An investigation of the English language proficiency and academic and clinical performance of University of Adelaide Medical School undergraduates.

UNIVERSITY OF ADELAIDE LIBRARY
13·11·98

Anna Chur-Hansen

Department of Psychiatry

Faculty of Medicine

Division of Health Sciences

University of Adelaide

South Australia

Thesis submitted in fulfilment of the requirements for

the Degree of Doctor of Philosophy.

Date submitted: July, 1998.

For JVR

Publications arising from the material presented in this thesis. (Appendix I)

Appendix I.I: Vernon-Roberts, J. and Chur-Hansen, A. (1995)
Letter to the Editor. Communication skills of interns in New South Wales.
Medical Journal of Australia, **163**, 112.

Appendix I.II: Farnill, D., Hayes, S.C. and Chur-Hansen, A. (1995)
Interrater reliability of the scoring of the Screening Test of Adolescent Language.
Psychological Reports, **76**, 1027-1032.

Appendix I.III: Chur-Hansen, A. (1997)
Language background, proficiency in English, and selection for language development.
Medical Education, **31**, 312-319.

Appendix I.IV: Chur-Hansen, A., Vernon-Roberts, J. and Clark, S. (1997)
Language background, English language proficiency and medical communication skills of medical students.
Medical Education, **31**, 259-263.

Appendix I.V: Chur-Hansen, A. & Vernon-Roberts, J. (in press)
Clinical teachers' perceptions of medical students' English language proficiency.
Medical Education.

Appendix I.VI: Chur-Hansen, A. (in press)
Teaching support in the behavioural sciences for non-English speaking background medical undergraduates.
Teaching and Learning in Medicine.

Grant received to support the thesis. (Appendix II)

Appendix II.I: University of Adelaide Teaching Development Grant (1995)
"The assessment of spoken language in a clinical setting" with Dr Jane Vernon-Roberts.
For the sum of $3,500.00

Presentations to learned societies and review groups.
(Appendix III)

Chur-Hansen, A. (1995) English language proficiency and medical communication skills. Department of Psychiatry Postgraduate Seminar, Royal Adelaide Hospital, Adelaide, South Australia, October 27th.

Chur-Hansen, A. (1995) The use of informal English in health interactions. Advisory Centre for University Education, University of Adelaide, Adelaide, South Australia, April 11.

Chur-Hansen, A. (1995) Overview of a study investigating the academic performance of non-English speaking medical students at the University of Adelaide. Evaluation Unit, Women and Children's Hospital, Adelaide, South Australia, August 18th.

Appendix III.I: Chur-Hansen, A. and Vernon-Roberts, J. (1996) Paper and Poster Sessions - "English language proficiency and performance in medical communication skills." Teaching Communication in Medicine Conference, St Catherine's College, University of Oxford, United Kingdom, July 24th-26th.

Appendix III.II: Chur-Hansen, A. (1997) Presentation and Poster Session - "Teaching Support for Non-English Speaking Background Medical Students." Australasian and New Zealand Association for Medical Education (ANZAME) Silver Anniversary Conference: Communication - Art and Science, Ormond College, University of Melbourne, Victoria, July 6th-9th.

Publications associated with this thesis.
(Appendix IV)

Appendix IV.I: O'Hanlon, A., Winefield, H., Hejka, E. & Chur-Hansen, A. (1995) Initial responses of first-year medical students to problem-based learning in a behavioural science course: role of language background and course content.
Medical Education, **29**, 198-204.

Appendix IV.II: Chur-Hansen, A. and Barrett, R.J. (1996)
Teaching colloquial Australian English to medical students from non-English speaking backgrounds.
Medical Education, **30**, 412-417.

Appendix IV.III: Barrett, R.J. & Chur-Hansen, A. (in press)
Getting off at Redfern: Gathering Colloquial "Aussie" English.
TESOL Journal.

Peer Acknowledgements.
(Appendix V)

Appendix V.I: Barrett, R. J. (1994)
Personal Communication (Letter).
Department of Psychiatry at the Royal Adelaide Hospital, University of Adelaide.

Appendix V.II: Farnill, D. & Hayes, S.C. (1996)
A User's Guide to the Australian Tertiary English Screening Test: AUSTEST.
Behavioural Sciences in Medicine, University Of Sydney, Glensdale Press.

Table of Contents

List of Appendices

Abstract.

This thesis has two major aims. The first is to investigate the English language proficiency of several cohorts of undergraduate medical students enrolled at the University of Adelaide, South Australia, employing both quantitative data and qualitative measures. The second is to consider the influence of students' English language proficiency upon their academic and clinical performance. For the purpose of this thesis "English language proficiency" is defined according to the quantitative and qualitative measures with which it has been assessed. "Academic performance", as conceptualised in the thesis, refers to the summative outcome measures of an examination in medical communication skills and performance in a one-year behavioural science course. "Clinical performance" has been taken in this thesis to refer to interactions between students and patients, both "real" and standardized.

The thesis begins with a description of the students who enrol in Medicine at the University of Adelaide Medical School, and two of the teaching interventions that have been implemented by the School in order to assist students experiencing difficulties with the course due to limitations in English language ability.

For the sake of cohesion and clarity a select literature review follows, based primarily upon writings from the field of medical education as opposed to the body of work in language teaching and learning. The latter has been largely omitted because the discipline employs its own terminology and philosophy, which is outside the author's area of knowledge and expertise.

The literature review commences with a consideration of the importance of language proficiency in medical training and practice, with a distinction made between the terms "language" and "communication". A discussion of the limited research that has been conducted regarding the English language proficiency of Australian medical students follows, including the predictive validity of language in the areas of academic and clinical performance.

A comprehensive review of the use of standardized patients to assess clinical skills is presented, with particular focus upon their employment in an Australian context and issues of reliability and validity. The literature review concludes with a discussion of United States research which has explored the use of standardized patients to evaluate the spoken English language proficiency of Foreign Medical Graduates in that country.

Following the literature review seven studies are presented which deal with one or both of the major aims of the thesis. Each study comprises one chapter, and takes the format of a scientific journal article, with its own introduction, method, results, discussion and conclusion sections.

The thesis ends with a general conclusion concerning the main findings gleaned from the seven studies conducted and puts forward recommendations for future teaching and learning strategies and research opportunities in the medical education arena.

Author's Statement

This thesis contains no material which has been accepted for the award of any other degree or diploma at any university. To the best of my knowledge and belief it contains no material published or written by any other person except where due reference is made in the text. If accepted for the award of the degree I consent to the thesis being made available for photocopying and loan.

Signed

Anna Chur-Hansen

Acknowledgements.

This thesis is dedicated to my dear friend, mentor and patron, Dr Jane Vernon-Roberts ("JVR"), who first suggested that I pursue a higher degree based on my experience teaching medical students from non-English speaking backgrounds. Over the last four years she has provided me with considerable support, advice and inspiration as well as plying me with unwooded chardonnay and other assorted alcoholic beverages. I look forward to continued collaboration with her.

I am greatly indebted to my supervisors, all of whom have been exemplary in their role.

I am particularly grateful to Professor Derek Frewin, Dean of Medicine and the University of Adelaide Faculty of Health Sciences - I am indeed another very satisfied customer!

My close mate Dr Mary Katsikitis has listened patiently to my moaning and groaning (both professional and personal) not only over the duration of this research but for six years prior and probably will for another sixty years hence. Her statistical knowledge and guidance have been invaluable and I have appreciated her willingness to read both drafts of this thesis and of the papers published from it.

I thank Kate Cadman for her tutorage and advice on English Language and Learning issues. We have had many memorable discussions regarding "the document" and our coming from different "worlds" - whilst I don't think we are entirely convinced of each others' arguments, we have learned a great deal from them.

Professor Issy Pilowsky and Associate Professor Uschi Felix, who both acted as supervisors for one year at the commencement of my candidature, whet my appetite for publishing and I thank them for their counsel.

I wish to give particular thanks to my colleague, Dr Rob "PBL" Barrett. He offered shrewd advice, made me laugh, and is an excellent academic role model. On his advice I have now read "The Masters" by C. P. Snow to address my political naiveté. I look forward to future work with him in anticipation.

I also want to make an especial mention of the kind help I have received from Dr Douglas Farnill from the University of Sydney, who welcomed me always and was ready with an answer to my queries on the e-mail. He has spurred me on and provided another inimitable role model.

This thesis really began as a consequence of the introduction of the Language Development Programme at the University of Adelaide, and I wish to acknowledge the input I have received from the members of the Language Development Committee. In particular I say "thank you" to Mrs Helen Mullins and Ms Ursula McGowan, discussions and debates with whom stimulated my synapses and had me thinking of myriad new projects - they inadvertently charged my intellectual batteries when they were flagging.

Within the Department of Psychiatry I am grateful to Associate Professor Helen Winefield for organising the postgraduate presentations, Dr Marilyn Peay for her advice regarding the Malaysian Australian Tertiary Education Scheme (MATES), Susie Sullivan for having the best knowledge of the Australian vernacular, Professor Bob Kosky for getting me the coveted telephone, and Associate Professor Michael Sawyer for the postgraduate meetings and associated festivities.

A major focus of this thesis has been the use of standardized patients to evaluate students' spoken English language. I therefore thank those people who acted in this role. In the 1994 study these were Alex Ask, Natalie Beaumont-Smith, Dr Mary Katsikitis, Dr Jane Mortimer, Della Steen, Pieter Walker and Dr Robyn Young. In the 1995 study they were Julie Clifford, Dr Mary Katsikitis, Dr Julie Mattiske, Dr Jane Mortimer, Jan Scicchitano, Dr Niccola Spurrier and Dr Robyn Young.

A number of others have lent their support over the course of this research and I thank them for their help. Specifically;
Dr Sheila Clark allowed me access to her examination marks for the 1994 Medical Communication Skills OSCI.
Ann Francis assisted in co-ordinating on the days of the Language OSCIs.
David Burford and Ian Carman supplied me with information about students.
Margaret Gunn and Josephine Sando acted as expert English as a Second Language raters.
Pamela Griffiths and Radhika Naidu from the Royal Adelaide Hospital's Jepson Library who organised numerous inter-library loans.
Judy Liney and Susie Sullivan gave advice on page numbering and photocopying.
Dr Bill Venables, from the University of Adelaide Department of Statistics, who provided advice on interrater reliability.

When Dr Mary Katsikitis was not listening to my woes David "pub dog" Jacquier was, and I thank him very much for his friendship and support, and for carefully checking the appendices of this thesis for me.

My special friend, Brian Cartwright, patiently helped me with the referencing and I am grateful to him for his input. He has provided me with much appreciated support on many levels as well as reminding me that there is always more than one side to every story.

My dear mate Marion Kennedy proof-read the final draft of this thesis and I am indebted to her for her assistance, friendship and empathy.

I must also pay homage to friends and family who have contributed to this work in some way (tangible or otherwise), including my mother, David Jones, Verner Søngaard Sørensen and Rodney Oates.

Finally, I would like to give an especial acknowledgement of my debt to the many medical and dental students that I have taught and learned from since 1987, and in particular those from non-English speaking backgrounds. In real terms, it has not always been easy to distinguish which one of us was the "teacher" and which one was doing the "learning". For our mutual lessons and the memories I have of them, I will always be grateful.

UNIVERSITY OF ADELAIDE LIBRARY

Chapter I.

The University of Adelaide School of Medicine.

Outline of Chapter One.

This Chapter profiles cohorts of students who were enrolled in Medicine at the University of Adelaide Medical School between 1992 and 1996 with specific reference to their language backgrounds. The Medical School curriculum and its evolution during those years is outlined and two of the teaching interventions that were implemented by the School in order to assist students experiencing difficulties with the course due to limitations in English language ability are described and discussed.

I.I. The University of Adelaide Medical School Student Population.

In the late 1980's and into the 1990's the Commonwealth government has encouraged universities to solicit overseas fee-paying students (Webb, 1993). This has proved a successful endeavour, certainly in monetary terms (Phillips, 1990; Kennedy, 1992). Providing education to foreign students is currently one of Australia's most important export markets, and the International Development Program of the Australian Universities (IDP) which oversees the interests of Australian universities in foreign countries has become a major business enterprise (Harris and Jarrett, 1990; Commonwealth Universities Yearbook, 1995-1996). The exact value to the economy is difficult to pinpoint, and estimates differ. Sweetman (1996) has reported that primary, secondary and tertiary Asian students contribute about $3.6 billion to the Australian economy, with a projected value of $6 billion by the year 2000. In 1997 the Federal Government's Employment and Education Minister, Senator Amanda Vanstone, stated that the 140,000 international students in Australia at that time contributed approximately $3 billion to the country's economy (Daw and Penberthy, 1997). Klimidis, Minas, Stuart and Hayes (1997) have made a more conservative estimate, based on higher education only, and not including primary and secondary schooling, stating that as at 1997, educating students from overseas earned Australia approximately $1.7 billion, a figure also cited by the Commonwealth Universities Yearbook (1995-1996). Lloyd (1997) reported that

international tertiary students brought AU$90 million into the state of South Australia in 1997 alone. Clearly, despite variations in the reported figures, such students provide a lucrative source of direct income. In addition to the financial benefits immediately gained, in the longer term educating these students fosters business and research links between Australia and the students' country of origin (which is usually Asian). Consequently the economy may be further strengthened by such activities with our geographical neighbours. Substantial social, cultural and intellectual gains to the community also result from our participation in overseas student education (Webb, 1993; Daw and Penberthy, 1997), which may not necessarily be directly evident or easily quantified, such as racial tolerance, increases in tourism through the recommendations of students to family and friends to visit, as well as economic benefits through rental accommodation, dining and shopping (Tertiary Multicultural Education Committee, 1988; Lloyd, 1997).

The University of Adelaide 1995 Annual Report (1996) indicated that there were 1,034 International student enrolments in 1994, which increased to 1,151 in 1995. These figures represent approximately 8% of total student enrolments in both years. International students generated $14.9m and $15.3m in student fees for the University in 1994 and 1995 respectively. In 1997 the University of Adelaide Vice Chancellor, Professor Mary O'Kane indicated that one of the characteristics of a "great" university, a position to which Adelaide should aspire, is the intake of substantial numbers of overseas students (Edge, 1997a), thus heralding a new impetus to attract such enrolments. Ballard (1992) has also noted that an influx of international students is essential to any Australian university which aims to hold prestigious standing in the international community.

The Social Health Atlas of South Australia (Glover, Shand, Forster & Woollacott, 1996) which collates and publishes population demographics states that Adelaide is composed of approximately 50% of people born in Australia of Australian parents. The remaining 50% are either born overseas themselves, or have one or both parents who were. In addition to those who have English speaking origins, a substantial number of immigrants have come from Italy, Greece, Germany, the Netherlands, former Yugoslavia,

Poland and the former USSR, and more recently from Asia, the Middle East and Latin America. As a consequence around 3% of the population, or 25,200 people cannot speak English well or at all. Significant waves of immigration to Australia in the twentieth century have occurred over three periods, according to Thong (1990); after World War II, during the late 1960's and after the abolition of the White Australia Policy in 1973. Thus, the permanent resident Australian medical student is drawn from a diverse population base with regard to cultural and language backgrounds. It appears that non-English speaking background permanent residents are well represented in medical student cohorts. Aubert (1994) reported that immigrants from Malaysia, Hong Kong and Vietnam were up to ten times more likely to be accepted into an Australian medical school than Australian-born applicants and that 24% of the 6, 930 students enrolled in Australian medical schools in 1993 were Asian immigrants.

Australia began educating Asian students in 1950, with the commencement of the Colombo Plan (Chew, 1972). Since that time thousands of overseas students have studied at Australian tertiary institutions. The University of Adelaide Medical School offered 135 places each year from 1991 to 1996, which is the period in which the majority of students who are the subjects of this thesis enrolled in the course. One-hundred of these places were reserved for "local" students who were permanent residents of Australia and this cohort included students with special circumstances (such as mature age, rural or Aboriginal students). Around 25 places each year were allocated to students under the Malaysia Australia Tertiary Education Scheme (MATES). This scheme is described in detail in a following section of this Chapter. A further ten places were allocated to overseas students, who were primarily from Asia.

Thus, it is clear that the homogeneous medical student population of the past, which was primarily Anglo-Celtic and from English speaking backgrounds (Phillips, 1990), from urban, privileged families, commencing university immediately after completing (private) secondary schooling, is no longer. The current group of students is far more diverse, with representatives from "old Australians", whose families have lived

here for several generations, immigrants and the offspring of immigrants, students from various overseas locations who will return to their country of origin on graduation, mature age students and aboriginal students, all of whom come from a range of socio-economic backgrounds (Ballard, 1992). With this diversity of backgrounds comes a range of linguistic competencies in English (Martens, 1991).

I.II. The University of Adelaide Medicine Curriculum.

Newble and Hejka (1991) have distinguished between "traditional" and "problem-based" medical schools. They describe the traditional medical school as having a six year curriculum divided into premedical, pre-clinical and clinical components. Teaching within a traditional context relies primarily upon lectures, practical classes and tutorials with ward work in the clinical years. Assessment is largely based on objective, end of course examinations. Problem-based medical schools usually structure their curriculum around clinical and community problems and focus upon small group learning and individual, self-directed study. Achievement is assessed by considering self-development of knowledge rather than pass/fail assessments. Newble and Hejka (1991) have demonstrated through a comparison of the University of Adelaide, at that time a traditional school, with the University of Newcastle, a problem-based school, that the educational environment of the traditional institution was more conducive to a "surface" or rote approach to learning whilst the problem-based school encouraged and facilitated the "deep" approach, where problem solving skills are attained independent of concrete factual information.

In 1994, following several major reviews of the pre-clinical curriculum, the University of Adelaide Faculty of Medicine introduced a number of important changes to the Regulations, Schedules and Syllabuses of the degrees of Bachelor of Medicine, Bachelor of Surgery (MBBS) (Frewin, 1993c). These changes occurred as a result of several major factors including recurrent criticisms of first year by students, concern by academic staff about the rote learning style of students, a need to reduce the amount of factual information presented to students and an external review which called for more clinically-relevant pre-clinical teaching (Frewin, 1993c). Consequently, from 1994 the

revised curriculum was designed to; teach principles in place of factual information; integrate courses horizontally and vertically throughout the degree; make the clinical relevance of material clear to students; place an emphasis upon self-directed learning; introduce problem-based learning (PBL) systematically into the pre-clinical years of the course; improve the teaching of medical communication skills; allow for students to choose from electives where feasible; and to assess English language proficiency and introduce strategies for improving language skills where necessary (Frewin, 1993c). Appendix VI contains the required subjects for study from the University of Adelaide Calendars of 1992 to 1996.

In 1994 the University of Adelaide made explicit to students the need for a high level of English language proficiency in order to enrol in the medical course, even though formal English language qualifications for local students were not required. The University of Adelaide Undergraduate Course Prospectus for 1998 (1997) stated that an assumed knowledge for entry into the first year of the medical degree at the University of Adelaide is "a high proficiency in written and oral usage of the English language" (University of Adelaide Undergraduate Course Prospectus for 1998, 1997, p 60). Furthermore, the Prospectus advised that;

> "English is the language of instruction at The University of Adelaide and proficiency in speaking, listening to, reading and writing English is essential. . . . Students without an acceptable level of English will be required to satisfactorily complete an intensive course of English language before being allowed to study at The University of Adelaide. It may be possible to take an appropriate course before students leave their home country. Alternatively, English Language Intensive Courses for Overseas Students (ELICOS) programs are available in Adelaide." (p 18).

The two main tests used by Australian universities (whose language of instruction is English, including the University of Adelaide), to determine an acceptable level of language proficiency for applicants who wish to enter the institution are the International English Language Testing System (IELTS) and the Test of English as a Foreign Language (TOEFL) (Rusek, 1992; University of Adelaide Undergraduate Course Prospectus for 1998, 1997). Alternatively, a candidate must have successfully completed an Australian Year 12 examination.

The International English Language Testing System (IELTS) is the test favoured by the University of Adelaide. The IELTS assesses general English language proficiency across four subtests in listening, speaking, reading and writing (International English Language Testing System, 1989). These subtest scores are added together and a mean calculated to give an overall band score, which for candidates who wish to pursue an academic course, ranges from Band 1 (essentially a non-user of the language) to Band 9 (essentially a fluent, expert user of English) (International English Language Testing System, 1989). The minimum English language proficiency for entry to a University of Adelaide course on the IELTS is an overall band score of 6.0 or better (University of Adelaide Undergraduate Course Prospectus for 1998, 1997). Band 6 is described in the International English Language Testing System (1989) handbook as indicative of a "Competent User [who] has generally effective command of the language despite some inaccuracies, inappropriacies and misunderstandings. Can use and understand fairly complex language, particularly in familiar situations." (p 6).

The Test of English as a Foreign Language (TOEFL) is a multiple choice test and comprises three sections; Listening Comprehension, Structure and Written Expression and Reading Comprehension. As part of the TOEFL a candidate may also be assessed with the Test of Written English (TWE) to establish ability to write in English and the Test of Spoken English (TSE) which evaluates general spoken language proficiency (Educational Testing Service, 1997). A maximum score on the TOEFL is 677. The TWE is scored from 1 (incompetence in writing) to 6 (clear competence in writing), and the TSE is scored from 20 (no effective communication) to 60 (communication almost always effective) (Educational Testing Service, 1997). The University of Adelaide will consider candidates with a TOEFL score of 550 or better (including 4 on the TWE) (University of Adelaide Undergraduate Course Prospectus for 1998, 1997). A score of 4 on the TWE equates to "minimal competence in writing on both the rhetorical and syntactic levels" (Educational Testing Service, 1997, p 39).

I.III. The Malaysian Australian Tertiary Education Scheme (MATES).

In 1986 the Malaysian Government's Public Services Department commenced sponsorship of a scheme for students from the Matriculation Programme of the Universiti Sains Malaysia (USM) to undertake undergraduate studies at an Australian University. Originally the degrees offered under the scheme included Accounting, Computer Science, Engineering and Medicine and came to involve five institutions (the Universities of Adelaide, Flinders, Tasmania, Queensland and Western Australia) but since 1988 only Medicine and Dentistry have been covered by the Malaysian Australian Tertiary Education Scheme, or "MATES", an acronym used to refer to both the scheme itself and the students involved in it. By 1995 two institutions (Flinders and Queensland) were unable to continue with the scheme, as they commenced postgraduate entry requirements from that year for Medicine, meaning that the undergraduate Malaysian students were ineligible.

At the scheme's inception a quota of between 50 and 60 students for each annual intake was set. The University of Adelaide has provided places for at least half of these students each year since 1989, with twenty-five places available for the medical course and five for dentistry. The current quota as of 1997 is 45 places across the three participating universities.

At the time of applying for the Malaysian Australian Tertiary Education Scheme students were in the final year of a two year Matriculation Certificate, a course for which they had been selected on the basis of their secondary school performance. Their first year of matriculation was taught at secondary school, whilst the second year of study was completed at the Universiti Sains Malaysia (USM), in Penang. Applicants, who had already been screened by the Malaysian Government's Public Services Department and USM, were interviewed by a team of academic staff members from the participating Australian institutions, to globally assess a number of attributes including oral English skills, intellectual capacity, motivation and sense of purpose, and ability to adapt to life and study in Australia (Segaran, 1993). The success of applicants in terms of whether they were accepted into the scheme and whether they were offered their preferred course and

university was based upon performance during the interview and final Matriculation grades. Thus, these students arrived in Australia as a highly select group.

Prior to commencing studies in Australia students completed a two month pre-bridging programme at the University Sains Malaysia (USM) Medical School at Kota Bahru, in Kelantan, Malaysia. This was a preparatory course to introduce the students to fundamental components that were dealt with more extensively during a bridging programme held in Adelaide, South Australia, which involved three tertiary institutions. The University of South Australia's Centre for Applied Linguistics (CALUSA) provided an intensive two month course of English and familiarisation with Australian culture; the University of Adelaide offered seven weeks of teaching within the disciplines of Anatomy, Behavioural Science (with a psychological focus), Biology, Physics and Chemistry; and Flinders University provided two weeks of teaching in Behavioural Science (with an epidemiological and sociological emphasis).

The prebridging course, co-ordinated and taught by USM academic staff, was conducted for approximately eight weeks and was primarily devoted to improving English language skills. All MATES students had commenced formal English language studies at elementary school level, at about age 7, but the majority of students had not been exposed to native English speakers and may not have used their English outside of the scholastic forum. The overall aims of the prebridging programme were to strengthen students' writing, reading and comprehension and spoken language skills, with a particular emphasis on grammatical structures and formal as opposed to informal, non-academic language. Students were also exposed to lectures, practicals and tutorials based upon the academic disciplines of Anatomy, Behavioural Medicine and Psychology, and Community Medicine in Malaysia.

The general aims of the bridging programme conducted in Adelaide centred upon the overall goal of maximising the likelihood of academic success, with the understanding that this success would be dependent to some extent upon the ability to understand and

adapt to Australian society and culture. More specifically the five month programme aimed to improve English language proficiency; develop social skills; expose students to the study skills and styles of thinking expected in Australian tertiary institutions; enable students to experience different styles of teaching in subjects in which they had tuition in secondary school and in disciplines with which they had little previous knowledge, such as Anatomy and the Behavioural Sciences; and to provide information about the pragmatics of living in Australia, like shopping, accommodation and so forth.

Varki (1992) and others (McAdam, 1972; Klineberg and Hull, 1979; Biggs, 1997; Chalmers and Volet, 1997) have warned against the dangers of generalising about individuals who have been categorised on the basis of their country of origin, or the place in which they were educated. Similarly, the MATES students cannot be considered to be a homogeneous group; one must be wary of making statements about a student group which are oversimplistic labels that fail to acknowledge the diversity within that group and the disparity of the individuals of which it is comprised. Keeping this in mind, there were nevertheless several characteristics shared in common by the MATES students that distinguished them from their non-MATES peers.

Malaysia is home to a number of ethnic groups. The population can, however, be divided into two categories; the "bumiputra", (or "sons of the soil"), persons of Malay descent who are usually of the Islamic faith, and the non-bumiputra, which includes people of Chinese and Indian heritage (Moore, 1991). At the decree of the Malaysian Government, the Malaysian Australian Tertiary Education Scheme was available only to bumiputra students, since the Malaysian government requires that the indigenous, ethnic Malays be preferentially selected over other students (Razali, 1996). The overwhelming majority of students were therefore Muslim, and shared the first language of Bahasa Malaysia. They came from a variety of socio-economic backgrounds, with some students from urban areas with professional parents and others from rural, village backgrounds. As mentioned previously, all had undergone rigorous selection procedures to secure a place at an Australian university. All were required to sign a contractual agreement with their

government stating that they must be answerable to the Malaysian Consulate whilst studying in Australia and must repay their tuition fees in kind, by working in a Malaysian government hospital upon graduation for a period of ten years. Students who returned to Malaysia as a consequence of unsatisfactory performance in their course were required to repay their debt in cash, proportional to the number of years spent in Australia. With such a pressure, in addition to parental expectations and their own personal desire for academic attainment, it is reasonable to say that overall, the MATES students were highly motivated to succeed and extremely hard working[1].

I.IV. The Language Development Programme.

In 1991 a proposal for Faculty-based support for students from non-English speaking backgrounds was put forward by Dr Jane Vernon-Roberts, the Clinical Studies Adviser at the Royal Adelaide Hospital. Her proposal was based upon the collective voice of academic and clinical teaching staff, who had described to her through both formal and informal avenues their thoughts and experiences about the challenges faced by these students. In her proposal she summarised five common areas of difficulty that had been identified; (1) a reticence and lack of assertiveness in some students which resulted in a failure to respond to and ask questions in tutorial and clinical settings. Although this was attributed by staff to poor social skills, it was acknowledged that language skills and cultural restraints may have played a significant role in limiting participation; (2) an insufficient command of informal, idiomatic English; (3) a lack of experience with small group, interactive teaching and learning; (4) a reliance upon rote learning; and (5) the inability of some students to grasp the importance of a high level of English language proficiency for the study and practice of medicine at an Australian institution (Vernon-Roberts, 1991).

Prior to 1994, efforts to assist medical students with language difficulties by sending them to outside agencies or to the University of Adelaide's Language and

---

[1] I am grateful to Dr Marilyn Peay, Convenor of the MATES Bridging Programme Committee and Chair of the MATES Selection Committee for her assistance in compiling this section of the thesis.

Learning Service at the Advisory Centre for University Education (ACUE)[2] had generally resulted in poor attendance and ill-feeling on the part of the students concerned, who saw language tuition as irrelevant to their medical studies. Because referral for language assistance was not based on any systematic procedure, but solely on the suggestion of a teaching staff member, some students also considered the offering of it to them as insulting, especially where they felt that language proficiency per se was of limited relevance to the medical course.

In 1994, in response to the concerns of both staff and students, the Faculty of Medicine implemented a teaching programme to address the problems faced by non-English speaking students. This initiative involved careful screening of all incoming students, and was taught by an English as a Second language specialist employed as a member of the Medical faculty, who collaborated with Medical faculty academic and clinical teaching staff to ensure that the specific language needs of the students were addressed through the integration of language teaching "in context" (McGowan, 1995a). By embedding the language course into the medical course in these ways, it was hoped that the resistance to language teaching demonstrated by earlier groups of students would be reduced.

During 1993, in preparation for the introduction of this English language teaching programme Dr Jane Vernon-Roberts and Ms Ursula McGowan[3] consulted Dr Douglas Farnill, Senior Lecturer in the Department of Behavioural Sciences in Medicine at the University of Sydney at the request of Adelaide University's Dean of Medicine, Professor Derek Frewin. Dr Farnill and his colleague, Associate Professor Susan Hayes, had considerable experience in identifying Australian medical students low in English language proficiency, and were willing to share their expertise and knowledge with the

---

[2] The Advisory Centre for University Education (ACUE) at the University of Adelaide provides teaching support to academic staff and language and learning support to students (University of Adelaide Annual Report, 1995).

[3] Co-ordinator of the Language and Learning Unit, University of Adelaide Advisory Centre for University Education (ACUE).

University of Adelaide Faculty of Medicine. A review of the Farnill and Hayes literature is presented in Chapter II.

To oversee the process of identification of students with English language difficulties and the implementation of an English language teaching programme and assessment protocol, a management group was organised for the "Language Development Programme", as it became known (Frewin, 1993a). Membership initially comprised four members of the Faculty of Medicine and an applied linguist from the ACUE, with a sixth member later joining the "Language Development Committee"[4].

Based upon recommendations made by Farnill and Hayes, the ACUE, and by academic and clinical teachers, the University of Adelaide Faculty of Medicine Curriculum Committee recorded several proposals regarding the teaching of English for its students (Report of Meeting of 13th December, 1993). Specifically, it was agreed that all students commencing the first year of the MBBS from 1994 would be required to undertake the STAL (Screening Test for Adolescent Language) at the beginning of the academic year and that students who did not reach a benchmark standard would be required to attend an interview with the Language Development Committee.

Farnill had reported that many students who failed to reach the benchmark standard on the STAL at the University of Sydney expressed anger at this. He also noted that voluntary attendance at remedial English language programmes offered to students was poor, consistent with the experiences of the University of Adelaide (Report of University of Adelaide Faculty of Medicine Curriculum Committee Meeting 21st June, 1993). In addition, the ACUE advised that short-term interventions were unlikely to result in long-term language improvement (Frewin, 1993b), and that brief remedial interventions would be less effective than a systematic, long-term approach (McGowan, 1994). Accordingly, it

---

[4] Membership of the Language Development Committee in 1994 was: Dr Rob Barrett, Senior Lecturer, Department of Psychiatry (Convenor), Dr Jane Vernon-Roberts, Clinical Studies Adviser, Royal Adelaide Hospital, Mr Peter Devitt, Associate Dean, Faculty of Medicine, Ms Anna Chur-Hansen, Lecturer, Department of Psychiatry, Ms Ursula McGowan, Lecturer, Advisory Centre for University Education, Mrs Helen Mullins, Lecturer, Language Development Programme, and Dr Ted Cleary, Associate Dean, Curriculum Affairs, as an Ex-Officio member.

was decided that participation in the Language Development Programme would be compulsory for students identified as experiencing language difficulties by the screening test and an interview, and that inclusion in the programme would be for a period of two years in the first instance. Students who were not identified by the screening test as potentially in need of language tuition but who perceived that their language proficiency was inadequate were allowed the option to self-refer to the programme if they wished.

It was proposed by the Curriculum Committee that at the end of the first year of participation in the Language Development Programme students would be interviewed by a member of the Language Development Committee as part of its review of the progress of students, and that all students would be required to undertake an oral and written Clinical Communications examination at the end of the second year. It was acknowledged by the Curriculum Committee that English language proficiency and communication skills were not one and the same, and that an examination in one could not be used as an assessment of the other (Report of University of Adelaide Faculty of Medicine Curriculum Committee Meeting 5th November, 1993). However, it was suggested that an assessment of each student's English proficiency could be one component of the Clinical Communications assessment, providing that the teachers who assessed Communication Skills were not also responsible for the assessment of English language skills. Thus, it was decided that although the assessment of language could be conducted within the context of a Clinical Communications examination, it would be designed by separate teaching staff and administered at a different time from the Communication Skills examination (Barrett, 1993). It was further decided that any student who did not exhibit an adequate proficiency in oral and written English in the Clinical Communications examination would fail the subject "Doctor, Patient and Society II", notwithstanding his or her performance in other assessment tasks for the subject, and thereby would be required to repeat all studies for the Second Year Examination (Report of University of Adelaide Faculty of Medicine Curriculum Committee Meeting 13th December, 1993).

The Language Development Programme was designed and taught by Mrs Helen Mullins, a Lecturer with qualifications in applied linguistics and expertise in designing English language courses in tertiary institutions for students with professional backgrounds (Frewin, 1994).

The programme was conceptualised as long-term and "developmental," rather than short-term and "remedial" (Snowden, 1995), encouraging students to aim for "life-long language learning" (McGowan, 1995a). The programme sought to achieve this by promoting independent learning (Snowden, 1995) and by engendering a sense of responsibility for one's own continuing language development (McGowan, 1994; Mullins, 1995b). The class sizes were limited to between eight and 12 students where possible, to facilitate meaningful small group interactions and to allow for individual tuition. Students were encouraged to focus upon the language demands both of their long-term goal of becoming a medical practitioner and of their shorter-term goals, such as completing assignments and passing each year of the medical course (Mullins, 1995a). The programme adopted a "learner centred" philosophy, with students deciding upon the content and direction of the course as their needs dictated (Mullins, 1995b). In recognition of the fact that individual students had differing language needs, students were grouped by the Language Development Committee on the basis of performance on the STAL, an interview with two members of the Committee and a further interview with the programme's lecturer. A detailed description of this procedure is given in Chapter III. Students were either allocated to attend Language Development classes for one-and-a-half hours per week, or were allowed to choose sessions to attend which they felt would be most useful to them (Mullins, 1994a). It had been planned initially that all selected students would be committed to the programme for two years, but it became apparent that flexibility was required in this regard due to variability in language proficiency within the group. Consequently, after further evaluation by the programme's lecturer, some students were required to attend for the full two year programme, some attended for one year only, and others attended selected classes for six months. All students, regardless of their

allocation within the programme, were able to make appointments for individual assistance where necessary.

The primary objectives of the programme included gaining feedback from tutors and lecturers working with the first and second year students through questionnaires, comments on students' written work and direct contact with the Language Development Programme lecturer (Mullins, 1995b). A basis for the programme's content was provided by designing teaching material to address staff concerns including written and oral English language proficiency, cultural knowledge and social skills, in addition to responding to the students' own perceived language needs, determined through questionnaires, a written exercise and an interview (Mullins, 1995a, 1995b).

The programme aimed to provide students with "models" of language appropriate to the academic, clinical and social situations they would encounter (McGowan, 1994). For example, students were made aware of the need to modify their spoken and written language to be appropriate to the context of the situation (Snowden, 1995). Thus, students learned about the differing conventions for writing across academic disciplines for essays and reports, lecture note taking skills, tutorial participation (focusing on small group, problem-based learning sessions), the need for brevity and clarity in case-note writing, speaking to patients during an interview, and interacting with peers and clinical teachers both formally and informally. Suggestions for ways in which vocabulary could be expanded were also given (McGowan, 1994; Snowden, 1995). Students' attention was drawn to the cultural variation which exists not only in the wider community of Adelaide and Australia but also within the university and the Faculty of Medicine itself (Ballard and Clanchy, 1988; Ballard, 1992; Mullins, 1995; Snowden, 1995), and the programme sought to assist students in coping with the demands made upon them by the need to be sensitive and responsive to these cross-cultural and inter-cultural differences.

One aspect of cultural variation particularly stressed because of its importance to Australian tertiary education was the concept of "analytical" or "critical" thinking

(Mullins, 1995). The term "critical thinking" refers to a style of learning which involves weighing up the relative merit of an intellectual position through systematic analysis and questioning of the argument until a judgement or conclusion can be reached about its validity (Ballard and Clanchy, 1984; Ballard, 1987; Ballard, 1995). Being critical of what is taught and learned requires that students may need to question and challenge the concepts and ideas put forward by their tutors and peers; in some Asian cultures this may be deemed disrespectful and inappropriate (Ballard and Clanchy, 1991). The Language Development Programme sought to provide opportunities for students to grasp the concept of critical thought and become comfortable with its use.

An integral part of the Language Development Programme was provision of opportunities for theoretical knowledge about language and culture to be applied and practised, as these experiences were seen as instrumental in facilitating language development (McGowan, 1994).

Assessment in educational settings may be formative or summative. Formative assessment refers to a process which provides feedback about performance by drawing students' attention to their strengths and weaknesses, so that they can concentrate future studies in the appropriate areas. Summative feedback involves the assessor making decisions about the students' future studies on the basis of performance in an examination or similar assessment tool (Newble and Cannon, 1994). Formative assessment may be preferable to summative methods because it is more likely to encourage and improve learning for learning's sake and promote self directed learning (Prideaux, 1992) and "deep" learning, where material is understood and recalled and students are motivated by interest (Newble and Mullins, 1990; Newble and Hejka, 1991), although rigorous research on the reliability and validity of formative assessment is yet to be conducted (Rolfe and McPherson, 1995). Summative assessment, on the other hand, tends to encourage learning for extrinsic rewards like being academically 'better' than classmates, and may promote "surface" learning styles, where rote learning is relied upon and students are motivated by the desire to perform in examinations (Newble and Hejka, 1991). Assessment of student

progress within the Language Development Programme was primarily formative to allow for individual rates of progress (Mullins, 1995b).

Tools for assessment included feedback from questionnaires, interviews and small group discussions; feedback from peers and problem-based learning teaching staff; portfolios that documented progress in writing ability; periodic tape recordings of speech to monitor oral language progress; reading tasks which were rated on a standardized language rating instrument (the Australian Second Language Proficiency Ratings - ASLPR); self-assessment through each student keeping a journal to reflect upon their progress; and spoken and written presentations made to the Language Development Programme lecturer and other students participating in the programme (Mullins and McGowan, 1994; Mullins, 1994a; 1995a; 1995b). Evaluation methods of progress through the programme were designed to provide information to guide the direction of the course and to ensure that it remained receptive to students' needs by incorporating additional teaching into the course to address identified areas of weakness.

Formative assessments provided valuable information to both the Lecturer and the participants in the Language Development Programme. However, the Faculty of Medicine required a summative assessment, having agreed that a student who did not exhibit an adequate proficiency in oral and written English by the end of the second year would not proceed to the third year of the course, but would repeat their second year studies (Report of University of Adelaide Faculty of Medicine Curriculum Committee Meeting 13th December, 1993). For the purpose of equity it was also considered appropriate that all students be assessed for English language proficiency, not only those students who had participated in the Language Development Programme, since there was no guarantee that all students who had needed language assistance had actually received it.

The conflict regarding the appropriate mode of assessment resulted in a dilemma for the Language Development Committee, who aimed to encompass the philosophy of the Language Development Programme and at the same time satisfy the immediate needs of

the Faculty. A summative assessment, although necessary, was considered problematic because it was contrary to the notion of a developmental model of language, signalling to students that if their language was assessed as "satisfactory" during a barrier examination at the end of their second year, they would need to develop their language skills no further (Mullins, 1995a). It was argued that as language mastery involves "life-long learning" (McGowan, 1995a), allowing students to form such an impression would be pedagogically irresponsible.

After considerable deliberation and debate a method of assessment was agreed upon by the Language Development Committee which encompassed both formative and summative measures. All second year students were required to write six essays for the "Doctor, Patient and Society II" course in public health and epidemiology. The third of these six essays, submitted by students at the beginning of August 1995 was assessed for global written language proficiency by the faculty member who graded the assignment and also by an independent expert in teaching English to speakers of other languages using a more specific rating scale. Both rating scales were designed by the Language Development Committee and are in Appendix VII.I and Appendix VII.II. The English language teacher was blind to the student's identity in terms of name and allocation to the Language Development Programme. All language ratings were returned to students with their graded assignment to act as formative feedback, and students were urged to seek assistance with their written language where an "unsatisfactory" was recorded on their feedback sheet. This process was repeated when students submitted their fifth essay for "Doctor, Patient and Society II", in October 1995 to allow redemption for those students who had previously been rated as "unsatisfactory" in writing skills.

A methodology for the summative assessment of English language proficiency was negotiated by the Language Development Committee to parallel corresponding research studies which are the focus of Chapters VII and VIII of this thesis. As a comprehensive description of the procedure employed for the summative assessment of spoken and written language is provided in Chapter VIII, it will be covered only briefly here. In

November 1995 all second year students were assessed for spoken and written language proficiency during a twenty minute examination following the format of an Observed Structured Clinical Interview (OSCI). Students were required to interview a standardized patient for ten minutes whilst an expert in teaching English to speakers of other languages observed. The English language teacher then made an independent rating of either "satisfactory" or "unsatisfactory" in relation to the student's spoken language during the interview. The student was given a further ten minutes to record in writing an account of the interview that they had just conducted. This sample of writing was also rated by an English language teacher as "satisfactory" or "unsatisfactory". The rating scales for both spoken and written language, designed by the Language Development Committee, are in Appendix VIII.I and Appendix VIII.II. Students who were assessed as "unsatisfactory" in either spoken or written language on the basis of their performance during the OSCI were permitted an opportunity for redemption, by performing a second interview with two members of the Language Development Committee.

A student was deemed to have failed the requirements of the examination of English language proficiency if they were rated as "unsatisfactory" on any of the three components of the assessment (the fifth DPS II essay, the spoken language of the OSCI, or the written account of the OSCI). In the event of failure, the student's case was to be reviewed by the Language Development Committee, which would in turn forward a recommendation to the Faculty of Medicine's Board of Examiners as to whether or not that student should proceed to the third year of the course[5].

I.V. The Supplementary Programme.

In 1987 the then Dean of the University of Adelaide Faculty of Medicine, Professor Geoffrey Dahlenburg, allocated funds to the Department of Psychiatry to employ a full time member of academic staff whose brief was to assist non-English speaking background medical students through their first year subject, Behavioural Science (later to become

---

[5] I am indebted to Dr Jane Vernon-Roberts, Clinical Studies Adviser, Royal Adelaide Hospital, and Mrs Helen Mullins, Language Development Programme Lecturer, for their comments regarding this section of the thesis.

known as part of the discipline stream "Doctor, Patient and Society" or "DPS"). This initiative came as a result of suggestions made by the then Head of the Department of Psychiatry, Professor Issy Pilowsky and the Behavioural Science course co-ordinators, Dr Marilyn Peay and Dr Helen Winefield. It had been found that a number of students who did not have English as their first language experienced considerable difficulty studying in the behavioural sciences due to the demands made upon both their English language skills and their knowledge of Australian culture (University of Adelaide Department of Psychiatry, 1987). As with other Australian curricula in the health sciences (Hunter and Hayden, 1990) and psychology (Ballard, 1987) the Adelaide course is Western in cultural orientation, often taking a knowledge of Australian customs and beliefs for granted. With the commencement of the MATES programme it was anticipated that these Malaysian students, as with other non-English speaking background students, would find the course difficult and be at risk of failure. Thus, an equity programme was considered necessary (V. Beasley, 1990), to reduce the possibility that non-English speaking background students would experience any type of exclusion or academic failure in this first year subject. In order to provide subject based support for non-English speaking background students a tutor with qualifications in the behavioural sciences (the author) developed a tutorial programme which is described below. An academic staff member in DPS was considered preferable to a teacher of English as a Second Language for several reasons, including the fact that the former would be able to relate all learning back to the subject content and place language needs in context, an approach that has been advocated by Spack (1988). Budd (1995) has described a similar effort to promote literacy skills by academic staff in the context of their disciplines in the School of Biomedical Sciences at the Curtin University of Technology in Perth, Western Australia. Seedhom, Smeathers and Thompson (1985) also believed that they, rather than English language specialists, should be responsible for designing a course that used modified English, so that the information of their discipline would be accessible to Japanese orthopaedic surgeons.

C. Beasley (1990) has described a variant of this "subject-based" approach at Murdoch University in Western Australia, utilizing English as a Second Language teachers

who worked in collaboration with the discipline academics. This is known in the literature as the "adjunct model" (Brinton, Snow and Wesche, 1989; Gollin, 1990). C. Beasley (1990) noted, though, that an adjunct arrangement can only succeed "where there is a high level of goodwill and mutual interest and understanding" (p 18) between the two teaching parties.

The adjunct model has been put forward by a number of writers in the field of medical education. That the responsibility for teaching English should fall to individuals within the hospital system, without any necessary understanding of cross-cultural or language issues has been strongly criticised by De Sweemer (1972), who argued for more systematic training by qualified professionals. Levey (1992) noted that even though they have passed a language proficiency test, many foreign medical graduates may experience difficulties understanding English-speaking patients and their families on both language and cultural grounds. He suggested a comprehensive, preparatory course of lectures, reading assignments, language classes and tutorials on United States medical practices and re-orientation in clinical skills, taught collaboratively by medical faculty staff and language experts.

In order to teach language in the context of medical situations, a collaborative approach has also been expounded by Ferguson and Maclean (1988), who reported on a combined course in medical English and primary care conducted at the University of Edinburgh in Scotland between the Department of General Practice and the Institute for Applied Language Studies. Ferguson and Maclean (1988) reported that their course was more intellectually stimulating and more effective than traditional classroom language teaching.

Maher (1987) agreed that professionals who are equipped to recognize and respond to the needs of their students are the most effective language teachers, but he argued that teachers of English to medical students and graduates need not have medical knowledge themselves. This statement raises a number of issues. For example, one may question

whether students will consider teaching that is conducted by someone without medical knowledge as relevant to their requirements. Second, it may be difficult to teach "in context" if the teacher has a limited knowledge of the field. In fact, students may be given misleading information or advice that conflicts with their lecturers' or tutors' unless the English language expert works in close collaboration with subject co-ordinators (Maher, 1986b; Spack, 1988). Thus, although Maher's (1987) idea may be a reasonable one, it would not be acceptable for any English language tuition to occur within medicine without ensuring that it was relevant and responsive to the students' requirements (Maher, 1986b). Indeed, Maher (1986b) has acknowledged that "inter-doctor" teaching as opposed to "applied linguist-doctor" teaching may be advantageous, although he does not advocate an intra-professional approach in favour of an inter-professional one. The Language Development Programme which has been described in detail in the preceding section of this Chapter and in Chur-Hansen (1997) (Appendix I.III) has addressed these points by placing the English language expert within the Faculty as a staff member and by attempting to facilitate collaboration between her and academic and clinical teachers wherever possible, to ensure the relevance and appropriateness of the programme.

From its inception in 1987, the programme at the University of Adelaide Department of Psychiatry has aimed to assist non-English speaking background students in their study of behavioural science, without focusing on English language proficiency *per se* (Pilowsky, Winefield, Peay, Augoustinos and Chur-Hansen, 1989). It is therefore complementary to the Language Development Programme. Both programmes tackle the common problem of limitations in English language proficiency, but with different emphases and approaches. The Language Development Programme might be seen as generic, this programme could be labelled as subject specific. The former is taught by an English language expert who is not responsible for students' grades, the latter by a member of the lecturing staff who is one of the subject co-ordinators and who is responsible for assessments in that subject.

It is clear that language proficiency is only one aspect that needs to be addressed for non-English speaking background students experiencing difficulties (Cargill, 1996); cultural issues, expectations about learning and teaching and isolation from other students are also pertinent issues that must be dealt with. The programme described here was not conceptualised as "remedial", with the negative connotations that this label entails (Webb, 1993), but rather, was referred to as a "supplementary" programme, where attendance was voluntary and served as an adjunct to the DPS syllabus. The only stipulations for attendance were that the student identified him or herself as coming from a non-English speaking background and perceived that their English language skills adversely influenced their ability to benefit from and contribute to the DPS course. Placement for language development had no bearing upon eligibility for attendance in this programme.

During the Supplementary Programme the content of lectures was discussed, vocabulary explained, material placed into its cultural context where relevant, questions were asked and answered, and students had an opportunity to discuss many course-relevant issues that may have otherwise been overlooked. A major focus of the Supplementary Programme was to encourage students to reflect upon and question DPS lecture material, rather than simply accepting such information at face value, as is often the case in tertiary education (Biggs, 1996). The Supplementary Programme included workshops to assist students where they requested additional help, in the areas of essay writing, report writing, exam strategies, analytical thinking, how to interact with Australian peers, and how to become familiar with colloquial language used in the medical setting (Chur-Hansen and Barrett, 1996) (Appendix IV.II). Part of the rationale behind the programme was to provide a non-threatening environment for students. Often they would have liked to participate actively in the DPS tutorials they attended with their English speaking background course mates. However, due to feelings of embarrassment or inferiority because of their differences in language ability and/or understanding of Australian culture, many non-English speaking background students did not feel confident enough to voice their ideas in such a forum. Thus, in addition to the academic component, one of the aims

of the Supplementary Programme was to provide an opportunity to participate and to increase students' confidence and self-esteem.

Students who opted to attend the programme could elect to commit themselves to up to four hours of tuition per week. There was no summative assessment, but on a formative level, all students were given feedback during personal consultation regarding their performance in the DPS course. Students who were not progressing well had the opportunity to discuss the reasons for this and strategies were negotiated for improvement. Students were invited to discuss drafts of their written work for comment before submitting the final version for summative assessment as part of the requirements of DPS.

Chapter II.

Review of the Literature.

Outline of Chapter Two.

This Chapter comprises six parts. Part I reviews the literature that has emphasized the role of language in the training of medical practitioners and its importance in the doctor-patient relationship.

Specific aspects of language proficiency that can be considered important in medical training and practice are considered in Part II. These are register, including both the specialised language of the medical profession and informal, colloquial language, reading, writing, listening, conventions of polite address (speech acts) and non-verbal behaviours (including touch, eye contact, tone of voice, accent, rate of speech, physical distance norms, cue discrepancy and facial expression).

Part III draws a distinction between "language skills" and "communication skills" as these terms are used within a medical education framework.

Australian research on the measurement of medical students' English language proficiency is reviewed in Part IV, whilst Part V discusses studies that have investigated the relationships between English language proficiency and academic and clinical performance.

Part VI of this Chapter is a review of the use of standardized patients to assess medical students' and graduates clinical abilities, with particular emphasis upon their use within Australian settings. The concepts of reliability and validity as they are relevant to standardized patient ratings are explored. Part VI concludes with a review and discussion of research that has emanated from the United States of America which has employed standardized patients to assess the English language proficiency of Foreign Medical Graduates.

II.I. Language proficiency, medical training and the doctor-patient relationship.

"Verbal communication is an essential component of physician's practice, enabling them to gather information about patients, educate them and assure their compliance, follow up and receive feedback on the course of treatment and interact with the health care team. Furthermore, verbal communication with health insurance companies and other health agencies and organizations, contributes to the efficacy of the physician's work. Thus, the spoken English proficiency of physicians should be of prime concern to health organizations or programme directors who recruit graduates of foreign medical schools." (Friedman, Sutnick, Stillman, Regan and Norcini, 1992, p 303)

"(T)he penetration of English in medical communication has occasioned the need for a diversity of skills from the overseas doctor. Consider, for example, the skills required in the medical meeting: listening to a ten-minute lecture, understanding a fairly wide variety of regional accents and speech patterns/note taking, giving a presentation and, in discussion sessions, a knowledge of forms of questioning, discussion procedure, polite forms of address, and cultural norms in English-speaking societies. A whole range of skills is required urgently by doctors (from) overseas especially in the fields of reading and writing." (Maher, 1986a, p 217).

The above quotations refer to graduates of foreign medical schools and make the point that proficiency in the language of the country in which the doctor is to function is imperative. Similarly, an undergraduate, training to become a fully qualified medical practitioner needs to be proficient in the language of instruction. Even where the student is a sojourner, intending to return on graduation to their country of origin to practice medicine, a very good grasp of the language in which medicine is being taught is necessary, for the benefit of the student, his or her teachers, peers and the patients with whom he or she will come into contact during training. In a relatively early study on the function of language in the doctor-patient consultation Cassell and Skopek (1977) argued the importance of spoken language in medicine;

"The spoken language is the most important tool in medicine. Speech is the medium by which patients inform doctors of their symptoms and concerns and by which doctors elicit and respond to the patients' needs. At least in our sophisticated society, no operation, treatment, medication or even diagnostic test is carried out without pertinent speech. Language is the predominant device by which information is transmitted. Successful patient compliance, initial interviewing, preoperative and postoperative explanations, doctor-family meetings, and doctor-to-doctor discussions, to say nothing of effective reassurance and comforting, are indicative of successful verbal interaction. Therefore, physicians should have an understanding of language as a tool of the trade - knowing how it functions, how it is used, and how it can be used. (Cassell and Skopek, 1977, p 197).

Most medical educators would accept without debate the assertion that proficiency in the language in which a student is to learn and practice is a necessary (but of course not sufficient) prerequisite (Biggs, 1997) and the doctor's ability to understand the patient is an essential feature of the medical encounter (Poole and Sanson-Fisher, 1979) in addition to the patient being able to comprehend what the doctor says (Golden and Johnson, 1970). The following review of the literature demonstrates reasons why proficiency for medical practitioners in the language of their patients is desirable and goes on to identify the specific language skills that are important in medicine.

The majority of authors who have written about the role of language in the doctor-patient relationship have implicitly assumed that the doctor and patient share the same cultural and language background and equivalent spoken language proficiency (see for example Inui and Carter, 1985), though Shapiro and Lenahan (1996), writing from the United States noted that "there are encouraging indications of increasing ethnic diversity among residents in family practice residency programs, reminding us that cultural differences are a two-way street" (p 249). Cole-Kelly (1994) has similarly called for recognition of the fact that in the United States a number of medical residents as well as patients originate from diverse cultural and ethnic backgrounds. Collins, White and Mantell (1997) from the University of Auckland in New Zealand have reported on an affirmative action programme which actively promotes the entry of Maori and Pacific Island ancestry students to Medicine, to ensure that these minority groups are represented in the profession. Thus, situations will arise whereby the doctor and patient will not come from the same linguistic heritage.

When the literature describes interactions where the two parties are not from the same language background, the doctor is assumed to be fluent in the language of the dominant culture and the patient limited in language proficiency (see for example, Grassby, 1980). Most research that has investigated difficulties during the doctor-patient interaction have concentrated upon the characteristics of the patient (Schwenk, Marquez, Lefever and Cohen, 1989; Ley, 1982) rather than those of the medical practitioner,

although more recent research has focused upon the health care provider as well (Frederikson, 1995). Nonetheless, problems in the encounter that are attributed to language are usually assumed to be a result of the patient's lack of ability to express themselves, and it is often seen to be the patient's responsibility to "linguistically and culturally translate their concerns into information that will be meaningful to health professionals" (Zambrana, 1987, p 148 in Huttlinger, Krefting, Drevdahl, Tree, Baca and Benally, 1992). Rarely have studies considered the reverse situation, where the medical practitioner experiences difficulties in understanding and being understood because of his or her limitations in language skills. This is an important area for research, for one of the responsibilities of the doctor and the profession of Medicine, is to "assume the responsibility for knowing what common obstacles or barriers could prevent their successful communication of information" (Stone, 1979, p 53), and clearly a doctor's lack of proficiency in the patient's native language is one of these.

The following literature review covers research that has underscored the importance of the doctor and patient understanding each other. The review is not intended to be exhaustive but illustrative, and therefore only selected studies on patients with language difficulties have been included. Only a few studies could be located that have discussed situations where the health care professional is deemed to be the one with limited language skills. All of these are presented here.

In a comprehensive review of doctor-patient communication research written in the early 1980's, Pendelton (1983) proposed an "input-process-outcome model" to classify studies of doctor-patient communication which is still a useful way to organise the literature. In this model "input" refers to characteristics of the doctor and patient, including the patient's symptoms, his or her ideas and concerns about those symptoms, and the patient's expectations for the consultation. Process, which Pendelton stressed can only be understood in light of input and outcome, includes the language used by the doctor and patient, non-verbal behaviour, the doctor's diagnostic and management decision making, and the doctor's use of medical communication skills, which Pendelton calls "verbal

content". "Outcomes " may be physical or psychological, or both, and may be immediate, such as patient satisfaction and recall of the doctor's instructions. Intermediate outcomes include compliance (adherence) with the doctor's instructions, whilst the long term outcomes are a change or otherwise in the patient's health status and the patient's understanding and concern about their health.

There are many studies which have measured input variables pertaining to the patient's language or cultural background (for example Zola, 1963 in Pendelton, 1983; Uba, 1992; Chugh, Dillman, Kurtz, Lockyer and Parboosingh, 1993; Diehl, Westwick, Badgett, Sugarek and Todd, 1993; Matsuba, 1993; Grover, Berkowitz and Lewis, 1994; Shapiro and Lenahan, 1996; Like, Steiner and Rubel, 1996). Research has addressed a number of input variables related to doctor characteristics (Roter and Hall, 1993) such as gender (Ackerman-Ross and Sochat, 1980; Chur-Hansen, 1985), age (Murphy-Cullen and Larsen, 1984), attractiveness (Young, 1979), ethnicity (Murphy-Cullen and Larsen, 1984; Ahmad, Kernohan and Baker, 1989) and personality characteristics such as introversion versus extroversion (see Ong, DeHaes, Hoos and Lammes, 1995). However, comparatively speaking, studies on the impact of the doctor's language background, language proficiency or cultural background are few, and this is a fertile area for future research. Research which has addressed the impact of the doctor's language and cultural background and language proficiency are reviewed below.

For political reasons, a substantial number of doctors who cannot speak Arabic are currently working in Saudi Arabia. El Shabrawy Ali (1992) and El Shabrawy Ali and Ali Mahmoud (1993) conducted a survey of 900 patients to investigate satisfaction with primary health care services in Saudi Arabia and found that 19.4% of respondents complained about language barriers with staff. In both studies the authors indicated that to overcome this problem each consultation involved a translator, not only an expensive solution, as they noted, but also one with the potential to introduce further difficulties, as the potential problems associated with using interpreters in the medical setting are well recognised (Uba, 1992; Chugh, Dillmann, Kurtz and Lockyer, 1993; Shapiro and Lenahan,

1996; Pousada, 1995; Jenkins, Blank, Miller, Turner and Stanwick, 1996; Farnill, Todisco, Hayes and Bartlett, 1997; Medical Defence Association of South Australia, no date). Although the papers by El Shabrawy Ali (1992) and El Shabrawy Ali and Ali Mahmoud (1993) provided descriptive statistics regarding patient satisfaction, neither attempted to determine relationships between variables. For example, it would have been possible, with the data collected, to establish the extent to which language barriers were associated with patient dissatisfaction. This information would have been most informative but is not available to the reader.

Notzer, Weinbach and Lauden (1995) conducted an investigation into the Hebrew language proficiency of Russian immigrant doctors in Israel. The study found a low knowledge of medical Hebrew and Hebrew in general, and concluded that this must have implications for the doctor-patient relationship and result in critical errors in the doctor's decision making. The authors also reported a significant relationship between language proficiency and general medical knowledge. Unfortunately the details of their methodology are unavailable for evaluation, as only an abstract of their research has been translated into English.

In a study of the verbal behaviour of 53 general practitioners in the United Kingdom, Long (1985) distinguished between doctors of Anglo-Celtic, Jewish and Asian origins. He found that generally speaking, doctors of Jewish background were more often engaged in patient centred behaviour and had longer consultations compared with Asian doctors, who were more doctor centred in their consultations, which were also shorter in duration[6]. However, closer examination of the behaviour of individual doctors within each of these two groups revealed that some of the Asian doctors did hold longer consultations. Two of the Asian doctors had been in practice for seven months and Long (1985) noted that they were experiencing severe language difficulties resulting in lengthy and inefficient consultations. Long (1985) also investigated "sterotypicality" (using the same stock

---

[6] Long (1985) defines *"doctor centered behaviour"* as related to the doctor's need to know and to control, and *"patient centered behaviour"* "as stemming from the doctor's need to cause the patient to verbalize feelings, insights and understandings and generally to participate in the consultation on his (the patient's) terms." (p 12).

phrases for each patient) versus "flexibility" (adjusting language to suit individual patients) in the doctors' verbal behaviours. He found that in general Jewish doctors demonstrated the most flexibility in their speech with the patient, whilst Asian doctors tended to employ the same or very similar patterns of verbal behaviour for each patient regardless of the patient's age, gender and complaint. Closer inspection of this finding showed that Asian doctors who had lived in the United Kingdom for long periods were more flexible in their speech than other Asian doctors. Long (1985) proposed that Jewish doctors may have been attempting to demonstrate linguistic excellence as a reflection of their successful integration into their new country, whilst some Asian doctors were limited to stereotypical verbal behaviours because of their limited proficiency in the English language, including difficulties they experienced in comprehending British accents and dialects.

Long's (1985) study is an important one, because it has direct implications for the association between the doctor's language proficiency and patient satisfaction. Though there are methodological problems associated with process and outcome measures (Williams and Wilkinson, 1995; Winefield, Murrell and Clifford, 1995), a number of authors have advocated (and it is generally accepted) that patient centred behaviours result in higher levels of satisfaction for the patient (Smith and Hoppe, 1991; Winefield, Murrell, Clifford and Farmer, 1996; Winefield and Weinman, 1996) and may also increase satisfaction for the practitioners themselves (Smith and Hoppe, 1991; Winefield, Murrell, Clifford and Farmer, 1995; Winefield, 1996). Thus, it can be hypothesized that doctors who have limited language skills may have dissatisfied patients, and may themselves be unhappy with their performance and the outcomes of their consultations. Whilst support for this hypothesis can be inferred from indirect evidence, this has yet to be empirically tested through replication of Long's (1985) study in an Australian setting by measuring the satisfaction of non-English speaking background doctors of differing levels of language proficiency and the satisfaction of their English speaking background patients.

Research suggests that the verbal behaviour of patients and medical practitioners during consultations can affect outcomes (DiMatteo and DiNicola, 1982). There is

considerable literature on patient's understanding and recall of doctors' advice and instructions; however, corresponding studies of doctors' comprehension and recall of the information provided by patients, of which the doctor's language proficiency may be an important determinant, are lacking. Similarly, much attention has been paid to the intermediate outcome of patient compliance or adherence (for selective reviews of the earlier literature, see Ley, 1979; 1982; DiMatteo and DiNicola, 1982). The focus on patient adherence has tended towards identifying patient characteristics which mediate compliance or the lack of it; again, it would be instructive to research the influence of the doctor's language background and proficiency upon patient compliance. Ley (1982) concluded that amongst other factors, patient satisfaction and understanding of what has been said are likely to result in compliance with the doctor's advice. Recall of the instructions given has also been identified as an important precursor to compliance (Ley and Spelman, 1965). It is not unreasonable to suggest that health care professionals who have difficulty speaking in the native language of their patients may have less satisfied patients who may have problems comprehending what they have been told which consequently, may impact upon compliance rates.

Long-term outcomes such as an improvement in health have been found to be related to the immediate outcome variable of patient satisfaction (Pendelton, 1983). The roles of the doctor's language background and language proficiency are yet to be explored at this level. However, existing research can be used to generate hypotheses about the likely effects of the doctor's language proficiency on health outcomes. Orth, Stiles, Scherwitz, Hennrikus and Vallbona (1987) found evidence that allowing the patient to express themselves in their own words was related to a decrease in the systolic and diastolic blood pressure of patients with hypertension over a two week period. Additionally, a reduction in diastolic pressure was found for patients whose doctors had provided them with information during the consultation. Discussing non-medical problems with the doctor has been found to increase patient compliance and satisfaction with the consultation (Martin and Bass, 1989). Patients who perceived that they could speak to the doctor about their problems and receive all the information they desired from

the doctor were more satisfied with the encounter and more likely to comply with the doctor's instructions. The benefits of such a patient-centred approach have been supported by the work of Henbest and Stewart (1990) and Stewart (1984) and reported by DiMatteo and DiNicola (1982). In addition, Henbest and Fehrsen (1992) have demonstrated that patient-centredness is beneficial to the doctor-patient consultation in a non-Western context, as well as the Western cultures in which its utility is currently generally accepted (Fehrsen and Henbest, 1993).

Extrapolating from the results gleaned by Ley and Spelman (1965), Ley (1982), Stewart (1984), Orth, Stiles, Scherwitz, Hennrikus and Vallbona (1987), Martin and Bass (1989), Henbest and Stewart (1990) and Henbest and Fehrsen (1992) it can be postulated that medical practitioners who have difficulties with expressing themselves in the language of their patients may not provide clear explanations, and nor may their patients feel comfortable to speak freely about their concerns with the doctor. Consequently, the patient may be denied psychological and physiological health benefits as a result of the doctor's limited language skills. However, this has not been demonstrated empirically and requires further investigation.

Murphy-Cullen and Larsen (1984) investigated the interaction of selected socio-demographic characteristics of doctors and their patients, including ethnic background, to evaluate the impact of this upon patient satisfaction. The authors concluded that doctors' and patients' ethnic background had no impact upon patient satisfaction. However, these results should be questioned. Firstly, satisfaction was measured via a questionnaire that was distributed to 730 patients. With a response rate of only 29.7% and no information regarding the patients who were not included in the study, there is no possibility to discount the impact of ethnicity and also language background of patients (and possibly the doctor) upon satisfaction with the encounter. For example, it may have been that language difficulties (and thus ethnicity) precluded some patients from completing the questionnaire. Secondly, ethnicity for both the 19 doctors and the 217 patients was not

described in this study. These are substantial flaws of the research and therefore the authors' conclusions cannot be accepted.

A study conducted by Vu, Marcy, Verhulst and Barrows (1990) investigated standardized patients'[7] ratings of their satisfaction during a clinical encounter with fourth year medical students in the United States. Although specific details were not provided about the students, the authors indicated that they could be divided into "white" and "non-white". Factors which influenced satisfaction ratings included whether or not students "asked questions clearly or awkwardly, "shouted" questions at patients, and used technical language in the encounter" (p S30). Across two cohorts of students Vu, Marcy, Verhulst and Barrows (1990) found that for one group white students had higher ratings than non-whites, and in the other non-white males received the lowest satisfaction ratings of all students. The authors noted that a limitation of these findings was the small sample size, with only between 1 and 8 non-white students in total cohorts of 69 and 63 students. A further limitation is the fact that data concerning students' language proficiency were not collected in this study. Furthermore, the ethnicity of "white" and "non-white" was not given. Had it been available, information regarding language and cultural background may have contributed to an explanation of the lower ratings given to non-white students.

Ahmad, Kernohan and Baker (1989) conducted research in the United Kingdom into the preference of 241 Caucasian, Pakistani and Indian patients, according to their proficiency in English, for a doctor of a particular gender and ethnic background - either a Caucasian English female practitioner or an Asian doctor fluent in three Indian dialects. English language "fluency" (as it is referred to in this study) was determined by patients' own self reported perception of their ability to speak, read and write in English. A substantial number of Asian patients reported that they could not speak, read or write in English, or only poorly - for example, 69% of Pakistani women in the study had limited spoken ability. The authors found that patients preferred a general practitioner with a similar language background and cultural background to their own, and this was

[7] For an explanation of the term "standardized patient" see later in this chapter.

considered by patients to be more important that choosing a doctor on the basis of gender. Details about the Asian doctor's proficiency in English were not provided and it must therefore be assumed that his proficiency was high.

In a second study conducted in the United Kingdom, which did not take English language proficiency into account, Ahmad, Kernohan and Baker (1991) again found that Asian patients, mostly of Pakistani and Indian origin, were more likely to consult an Asian doctor in preference to a non-Asian doctor. This trend held even for female Asian patients, who chose to consult a male Asian doctor rather than a female non-Asian general practitioner. Given that a number of researchers have demonstrated a general patient preference for doctors of the same gender, especially where complaints are of an intimate nature (Ackerman-Ross and Sochat, 1980; Chur-Hansen, 1985), Ahmad, Kernohan and Baker (1991) speculated that the need for linguistic and cultural compatibility may override the issue of gender, when the doctor and the patient do not share the same language background or where the patient is not fluent in English. The authors concluded that considerable research remains to be conducted to understand the impact of patients' ethnicity and language proficiency upon the doctor-patient relationship; the converse situation similarly requires attention.

The preceding studies all give the impression that the doctor's lack of familiarity with the patient's language is likely to result in problematic encounters. A study by Elsass, Christensen, Falhof & Hvolby (1994) which failed to support this position was conducted with 50 Danish health care professionals (doctors, nurses, physiotherapists and auxiliary nursing staff) working in hospitals in Greenland and 50 of their patients. The majority of patients understood little or no Danish, whilst only five of the health care professionals reported having a "satisfactory" working knowledge of Greenlandic. Nevertheless, an interpreter was not ordinarily involved in consultations. The Greenlandic patients expressed high levels of satisfaction with their Danish practitioners, despite the fact that many patients had not understood explanations given to them about admission, diagnosis and treatment. Elsass, Christensen, Falhof & Hvolby (1994) acknowledged that the

Greenlanders' cultural conventions may have led to their concealment of dissatisfaction. Nonetheless, the authors refuted the notion that patient satisfaction and successful treatment are reliant upon the practitioner and patient understanding each other's verbal utterances, pointing to shaman healing rituals reported by anthropologists to be effective, where the patient cannot understand the words used by the healer.

A second study that also failed to support the notion that language difficulties result in patient dissatisfaction was conducted in the context of the nurse-client encounter by Butrin (1992). Interviews were conducted with 15 nurse-client dyads, where the nurses originated from North America and the majority of clients from South America. The fifteen clients ranged from fluent English speakers (n=2) to non-English speaking (n=6), with "fluency" judged by the author. Butrin (1992) reported that although almost all of the participants in her study considered language differences to be potentially problematic in terms of relaying information, mutual satisfaction was expressed in most of the encounters.

The studies by Elsass, Christensen, Falhof & Hvolby (1994) and Butrin (1992) might be explained in terms of the "placebo effect". The placebo effect refers to therapeutic effect of a substance (in the case of pharmaceuticals) or interaction (in the case of psychological interventions) which in itself is not curative for the condition being treated (Friedman, 1982). Thus, perhaps patients are satisfied with their encounter with the doctor, even if the language barriers between them cause misunderstanding, because of the attention paid or effort made by the health care practitioner, rather than the information that they are able to give and receive. Perhaps the doctor's most powerful and effective drug is in fact his or her personality (Helman, 1984), regardless of their language or cultural background. That is, perhaps "the doctor is the drug" (Balint, 1957 in Bensing, 1991, p 1301). A research design that could investigate this specific possibility would be extremely interesting.

A study which is included here for the purpose of drawing attention to an alternative way of interpreting difficulties where the health care provider is not proficient

in the language of the dominant culture was conducted by Vore (1991). She reported concerns that Spanish speaking nursing staff working in Washington may have clashed with their English speaking counterparts because of the potential for language and cultural difficulties to result in conflict at the workplace. Interestingly, barriers to patient care due to limited English language skills appear to have been of less concern. Whilst it is acknowledged that medical students and graduates who have language difficulties may similarly experience conflict with their peers and other members of the health care team as a result, this issue is beyond the scope of the present review.

Thus far all of the research presented has been based primarily upon quantitative research methodologies, including cross-sectional studies, interviews and survey questionnaires to consider input and process variables in the doctor-patient encounter. Qualitative research designs have also been employed to explore the role of language in the medical consultation.

Research on the input and process of the consultation includes sociolinguistic studies, which attempt to categorize types of spoken language used by the doctor and patient (Pendelton, 1983) after videotaping or audiotaping, transcribing and subjecting speech to qualitative procedures. One sociolingistic approach that has been applied to doctor-patient consultations is discourse analysis. Discourse analysis has been described by Nessa and Malterud (1990), who subjected a single consultation to analysis, as involving four steps. "Registration" refers to the method by which language is recorded, such as via observation, audiotape or videotape; "transcription" is the process of transferring the discourse to written text; "coding" where the discourse is broken down into units for analysis; and "interpretation" which requires the investigator to relate the data, acquired through transcription and coding, to a stated theoretical problem (Nessa and Malterud, 1990).

Erzinger (1991) conducted an ethnographic and discourse analytic study in the United States on the medical encounters of Spanish speaking Latino patients and their

medical practitioners, who had a range of Spanish language proficiencies and came from various cultural backgrounds. In the ethnographic phase of her research Erzinger (1991) observed and interviewed Spanish speaking patients to establish which aspects of conversation between doctor and patient required further exploration. Audiotapes of selected doctor-patient consultations were then subjected to discourse analysis. Erzinger (1991) concluded that the most effective encounters (in her estimation) were those where the doctor had far more than simply a knowledge of medical Spanish, but also a proficiency level whereby the doctor could actively listen, comprehend and encourage the patient to respond. Erzinger (1991) also found that the doctor-patient relationship was enhanced where the patient could assist the doctor by providing and teaching needed Spanish words, where the doctor's proficiency was limited.

In Erzinger's (1991) paper exerpts from the transcripts of four doctor-patient encounters are provided. Two of these transcripts demonstrated the difficulties which arose when the doctor was not fluent in the Spanish language. One doctor misunderstood an everyday word ("sneeze") and as a consequence discussed at length the state of the patient's ovaries. A second medical practitioner groped for her words, asking the patient frequently for the appropriate word or term in Spanish without acknowledging her appreciation to the patient for providing these, and resorted to using English medical terminology which the patient may not have understood. Clearly, such scenarios call into question the accuracy of the doctor's diagnosis and ability to make decisions under such circumstances. It is also doubtful that either the doctor or the patient would have been highly satisfied with the consultation process, although this can only be speculation since satisfaction as an outcome measure was not assessed in this study.

Rowland-Morin and Carroll (1990) conducted a discourse analytic study of five doctor's global style of interviewing across 52 consultations. They focused upon specific language variables that conveyed the doctor's level of involvement (rapport), expressiveness (a quantitative measure of silence vs speaking), communication dominance (interruptions and control of the interaction), and speech convergence (where one party

reciprocates the speech patterns of the other). This is a notion which parallels that of "stereotypicality" versus "flexibility" put forward by Long (1985) and discussed above. Rowland-Morin and Carroll (1990) found that language reciprocity was positively correlated to patient satisfaction, concluding that the more similar the use of words between doctor and patient, the more likely information could be effectively transmitted, with patients more likely to perceive that the doctor had understood them. Although not stated, it is assumed that the doctors and patients in this study were of the same language background, presumably English. It would be instructive to replicate the study with doctors and patients of different language backgrounds, to assess the impact that this may have both on reciprocity and on the doctor's and patient's satisfaction with the interaction.

Winefield and her colleagues have carried out a number of discourse analytic studies based upon a pool of 210 general practice consultations conducted in Adelaide, South Australia. This work, which is reviewed below, is based upon the methodology of Stiles (1978), who devised a method of discourse analysis based upon Verbal Response Modes, which he defined as a "category of language behaviour that implies a particular interpersonal intent or microrelationship between communicator and recipient" (Stiles, 1978, p 693). The purpose of categorizing speech according to Verbal Response Modes is to demonstrate aspects of interpersonal relationships, by considering the content of what is said. In this regard Stiles (1978) differentiates his method of content analysis from others which code the semantic meaning of speech or features of speech such as pauses, pitch and laughter.

Winefield and Murrell (1991) transcribed and coded verbal interactions between doctors and their patients in an attempt to identify speech patterns associated with satisfaction of the consultation for both parties. They coded the types of utterances made by the doctor (for example information seeking by virtue of open or closed questions and information giving) and the patient (for example information giving and seeking) and concluded that patients were most satisfied with consultations where they could relay their experiences and attitudes to the doctor, whilst the general practitioners felt most satisfied

with consultations where diagnostic instructions in the early phase of the interaction were later followed by predictions, reflecting confidence in their diagnosis and conclusions.

Winefield and Murrell (1992) conducted a similar study in which general practice consultations were transcribed and coded according to the categories of "emotional support", "informational support" and "diagnostic activity" for doctors and "relationship oriented" and "task oriented" for patients. Consistent with the findings reported by Winefield and Murrell (1991) doctors considered their most satisfactory consultations to be those in which the content of their speech implied confidence in their medical knowledge and problem solving. Patient satisfaction was not assessed in this study.

Winefield, Murrell and Clifford (1995) found only few, weak relationships between the verbal interaction of consultations and outcomes measures of the doctor's satisfaction with the consultation, patient satisfaction, quality of care (assessed by patient health change pre and post consultation) and patient compliance. A study distinguishing between different types of consultations (Winefield, Murrell, Clifford and Farmer, 1995) employed the Verbal Response Mode methodology to determine the doctor's patient-centredness and also quantified patient involvement during the consultation by calculating the percentage of patient speech. Patients were found to speak most in "Complex" consultations followed by "Psychosocial" consults, speaking least in "Straightforward" consultations[8]. In an attempt to measure patient-centredness Winefield, Murrell, Clifford and Farmer (1996) categorized speech according to the Verbal Response Modes of "Doctor Receptiveness" and "Patient Involvement". Doctor Receptiveness included speech where the doctor reflected upon what the patient had said, asked open-ended questions and acknowledged the patient's speech. Patient Involvement encompassed speech related to questioning the doctor, showing either a positive or negative attitude to treatment, reporting an unobservable symptom to the doctor, accounting actions or experiences to the doctor and expressing an opinion. Winefield, Murrell, Clifford and

[8] A full description of the characteristics of each of these three types of consultation can be found in Winefield et al (1995), page 404 and Winefield et al (1996) pages 815 to 816.

Farmer (1996) found that patients showed more active involvement in their speech when the doctors' verbal behaviour demonstrated receptiveness to the patient.

Although Winefield and Murrell (1991; 1992), Winefield, Murrell and Clifford (1995) and Winefield, Murrell, Clifford and Farmer (1995; 1996) have demonstrated the importance of certain features of the verbal interaction in the doctor-patient relationship and their influence upon outcomes including satisfaction with the consultation, their work is limited in its use for understanding the potential effects poor language proficiency of the non-English speaking background medical practitioner may have upon the consultation process. Winefield and her colleagues (1991; 1992; 1995; 1996) have explored the content of the speech, rather than the speech itself, as dictated by Stiles (1978) Verbal Response Mode methodology. In their research specific aspects of spoken language, such as tone of voice, use of specialized terminology, colloquial language and other components of language are not explored. Furthermore, the English language proficiency and language background of the doctor and patient were not described or considered. Thus the reader must assume that both parties were fluent English speakers, though this may not necessarily have been the case. Nevertheless, Winefield and her colleagues' work has important implications that should be explored in consultations where the doctor and patient do not share the same language background or speak with the same levels of language proficiency. For instance, do such linguistic variables affect doctor and patient satisfaction and patient-centredness across different types of consultations? Do patients speak quantitatively more or less with a doctor who is not proficient in their language as compared with one who is? If so, what effect does this have upon outcomes, such as quality of care and compliance? To date these questions have not been addressed in research studies. Winefield and associates' work is important in relation to this thesis because it involved the collecting of 210 comprehensive general practice consultations in Adelaide. Therefore, researchers interested in the language used by doctors and patients in this city at this particular historical point in time have a rich source of data from which to base their lines of enquiry. Indeed, the Winefield transcriptions have already been used in

a study on colloquial language by Chur-Hansen and Barrett (1996) (Appendix IV.II), described below.

Sociolinguistic research has been criticised by Pendelton (1983) on several methodological grounds. He argues that it involves expensive procedures and often employs small sample sizes, which threatens the generalizability of results. He also states that sociolinguistic studies allow considerable scope for the investigator to infer the meaning, intent, rules and roles of the participants in the interaction, limiting the objectivity, and thus the validity of the approach. Nessa and Malterud (1990), whilst defending discourse analysis in favour of the strict methodological rules of quantitative research, also acknowledge its drawbacks. With regard to registration, they note that audiotaping conversations does not allow important non-verbal behaviours to be taken into account. They also consider the problems associated with methods to capture non-verbal behaviours, such as having the researcher observe and note the interaction, which can result in bias of recall, in addition to the influence of the observer's presence in the encounter. In transcribing, Nessa and Malterud (1990) question the validity of a process which cannot adequately allow for the inclusion of non-verbal behaviours and aspects of speech that are not strictly verbal, such as pauses and interruptions. In coding, the meaning of each sentence and each word must be interpreted and therefore "the person coding thus becomes the instrument to the same extent as the coding system itself" (Nessa and Malterud, 1990, p 82), which leads to queries about the reliability and validity of discourse analysis as a research tool. As with coding, interpretation is open to the same limitations. Stiles (1978) has similarly critiqued the limitations of his categorization of Verbal Response Modes.

II.II. Specific aspects of language proficiency important in medical training and practice.

"Language proficiency" for the purpose of interaction with patients is thus clearly an important skill for medical practitioners, but this term is general and does not make clear exactly which characteristics of language are important for being "proficient".

In a profile of the opportunities available to foreign nationals in the United States Buchanan (1989) stated 'facility' in the English language to be a fundamental prerequisite for access to heath care education and training (p S13). It is unlikely that a potential student would be in a position to determine whether their personal proficiency would suffice, based on this vague directive. Similarly, in a discussion of medical school selection criteria Goldbeck-Wood (1996) included "intelligibility" as a prerequisite for being a "good doctor". Blacket (1990), reporting on the experience of the Australian Medical Examining Council (AMEC) and the Australian Medical Council (AMC) from 1978 to 1989 argued that the poor performance of some foreign medical graduates in examinations to qualify for medical practice in Australia could be attributed to an inadequate command of English. However, he did not investigate or further explain the role of English as a reason for failure in the AMEC and AMC examinations, a point acknowledged by himself, and also made by McGorry (1990) and Tran-Dinh (1990).

With such non-specific references to language skills it is necessary to determine exactly what aspects of English language proficiency are important for the training of medical students or overseas medical graduates.

Aspects of language that may result in problems within the doctor-patient interaction have been classified by Ford (1977) into two categories. The first includes those problems which stem from miscommunication due to differences in word usage, resulting as a consequence of the patient's language, educational or socio-economic class backgrounds. This category may encompass misunderstandings due to either party being a

foreign language speaker, the use of medical terminology by the doctor or the incorrect use of such terms by the patient, colloquial language usage by either party and accent. The second category involves the doctor-patient relationship, and includes communication difficulties which arise because of differences between the doctor and patient in assumptions, knowledge, attitudes, beliefs, expectations and emotional needs (Ford, 1977).

Crystal (1976) emphasised that difficulties may arise because either the doctor or the patient expresses themselves poorly, either the doctor or the patient misunderstands the other, or any combination of these four possibilities. The situation in which the doctor or medical student overestimates the patient's ability to understand or produce language has been well documented and researched (Crystal, 1976; Poole and Sanson-Fisher, 1979). The reverse situation appears to have attracted little attention in the literature.

Crystal (1976) distinguished between three aspects of the structure of language which are commonly investigated in the sociolinguistic field which may be viewed as possible impediments to the doctor-patient interaction; pronunciation, grammar and semantics. Pronunciation involves the use of different sounds (phonology) and qualities of the voice (phonetics). These two components of pronunciation can be subsumed under the general term of "accent". Grammar comprises the manner in which words are formed (morphology) and put into sentences (syntax). Semantics, which refers to the meaning behind the words (Nessa, 1995), includes the meaning of the words themselves (vocabulary) and the overall patterns of meanings the words form (discourse).

Maher (1986a) has called for a close examination of the actual language requirements considered necessary for foreign medical students and graduates. Many authors in the medical literature have discussed the need for language proficiency in medical practitioners and a close inspection of what they have written identifies common concerns related to specific areas of language usage. Thus, the following review of the literature discusses facets of language considered important in the medical encounter

which have been discussed by other authors. These are register, colloquial language, reading and writing, listening, speech acts and non-verbal behaviours.

(i). Register.

Register is defined by Bickley (1983, in Bochner, 1983), as the style of speech used by a particular group of people, who share special grammar, words or phrases. Professions such as medicine have their own specialised language (Anderson, 1991). Register also refers to the differences in the language used by an individual in different situations (Argyle, Furnham and Graham, 1981). Language may vary according to the setting, the role relationships of the participants in the interaction, the goals of the language user, the affective state of the language user, the channel (telephone or face-to-face for example) and the rules of the situation in which the language is being used (Argyle, Furnham and Graham, 1981). Isaac (1993) refers to the different registers required by university students for speaking with peers outside the classroom, participating in tutorials and seminars, giving formal oral presentations, writing an essay, report or assignment and taking lecture notes. In addition medical students must have an understanding of the appropriate manner of speaking to a patient and taking a medical history from them.

Van Naerssen (1985) has said that the doctor-patient relationship involves only one register which encompasses a range of variations, whilst language between the doctor and other health care professionals, such as nurses, paramedics, administrators and medical students, involves a second register, which again takes several different forms. Thomas and Steele (1966) and Maclean and Maher (1994) have also noted that doctors use different terms when speaking with colleagues than with patients. Bourhis, Roth and MacQueen (1989) consider doctors and nurses to be bilingual, with fluency in two registers - their "everyday" language (that they share with their patients) and their medical language (that they use when speaking to each other and members of related professions). Christy (1979) has taken this idea of a second register even further, arguing that in fact, by virtue of their training doctors speak a second language, which he has coined "Medspeak" (p 979).

In order to master the professional register, medical students undergo an apprenticeship which involves learning the technical language specific to the field of medicine (Scott and Weiner, 1984), the "jargon of medicine" (Shapiro, 1980, p 163), which is considered to be a necessary (Gastel, 1995a) but formidable task (Thomas and Steele, 1966; Lucas, Lenstrup, Prinz, Williamson, Yip and Tipoe, 1997). This jargon can facilitate discussions with fellow medical practitioners, although sometimes a reliance upon medical terminology at the expense of a straightforward word or phrase can cause confusion (Woods, 1981; Manning, 1989). A reliance upon jargon may also have a detrimental effect when speaking with patients who do not share the same vocabulary (Korsch, 1984; Farnill, Todisco, Hayes and Bartlett, 1997). Shapiro (1980) makes the assumption that patients may not understand the words used by the doctor, but to take an alternative view, it is feasible that students and practitioners from non-English speaking backgrounds may not always understand their patients or their fellow medical colleagues (Rush, 1972). Lending weight to the argument that doctors may not always understand specialised terminology, an interesting study by Hadlow and Pitts (1991) found that in a sample of doctors, nurses, health support workers and patients all groups of subjects, including the doctors had difficulty in identifying the correct definition of some terms. Doctors demonstrated low to medium comprehension of such words as "epilepsy", "neurosis", "schizophrenia" and "hysteria" (p 194). In an article about the "dilemma" of a learned profession, Pickering (1978) noted too that student doctors often make lists of technical terminology which they do not understand. Information about subjects' language background was not provided in Hadlow and Pitts' (1991) study; to have analysed the results according to this variable, if it were appropriate, would have been instructive. Shapiro (1980) has argued that patients who do not understand the doctor may not seek clarification. Similarly, it may be that some medical students or graduates do not ask the patient what he or she has tried to convey where the message is not understood.

Shapiro (1980) has suggested that some doctors deliberately use their specialised terminology to assert dominance and intellectual superiority over patients and other health care workers. Conversely, perhaps some non-English speaking background medical

students and doctors use such language as a replacement for more informal, conversational language with which they may be less familiar and therefore less comfortable in using. This would be consistent with the notions of stereotypicality reported by Long (1985) and speech convergence (Roland-Morin and Carroll, 1990), which are discussed above.

Korsch (1984) has stressed that communication in the doctor-patient interaction must be "two way" (p 917), arguing that just as the patient should relay their concerns as clearly as possible, it is the responsibility of the doctor to use language which is appropriate to the patient in terms of age, education and culture. This sentiment has also been expressed by Stone (1979). Doctors who insist upon employing their medical register of jargon terminology may have dissatisfied patients and in addition be unhappy with the encounter themselves (Korsch and Negrete, 1972).

In what was then a ground-breaking study, Korsch, Gozzi and Francis (1968) tape-recorded 800 paediatric outpatient visits to determine factors associated with patient satisfaction or otherwise. Their analysis was quite comprehensive, but of interest here were their findings regarding the use of medical jargon. The researchers reported that an "outstanding barrier" (p 862) to patients understanding was the paediatricians' reliance upon technical language, used in over 400 of the recorded interactions.

The differences in language usage between doctors and their patients can be understood in light of Giles (1977), Giles, Bourhis and Taylor (1977), and Bourhis, Roth and MacQueen (1989), who have described "accommodation theory" to explain how people use language when speaking to someone who uses a register that differs from theirs. "Convergence" refers to the two speakers' attempt to adopt similar patterns of speech. Bourhis, Roth and MacQueen (1989) speculated that doctors who make efforts to converge to the language of their patients may be more favourably viewed by their patients than doctors who opt to either employ their specialist register ("maintenance") or deliberately exaggerate and accentuate the differences between their register and the patient's ("divergence"). Giles (1977) theorised that the more similar one's speech is to

our own, the more we will be attracted to them. Following from this, then, higher patient satisfaction with the consultation may result from convergence. Consistent with this hypothesis is the possibility that doctors and medical students from non-English speaking backgrounds who are experiencing language difficulties may be unable to converge with their patients' speech and in turn may use maintenance and divergence strategies. As a consequence this may result in the patient and doctor liking each other less with resultant dissatisfaction.

A study by Bourhis, Roth and MacQueen (1989) could not investigate the influence of language background or English language fluency upon doctor, nurse and patient convergence, maintenance and divergence with each other's registers because all of the subjects reported English as their first language. Thus, whether language background and proficiency play a role in accommodation theory within medical settings has yet to be explored. Hui and Cheng (1987) conducted a laboratory experiment with Chinese undergraduate students which investigated the effects of the language proficiency of a listener upon both their perception of, and behavioural intention towards a speaker. Extrapolating this to the medical arena, we may ask whether a non-English speaking background medical student's language proficiency influences the way they view and consequently behave towards their English speaking patient. The student in this scenario can be viewed as the "listener" and the patient as the "speaker", although it is equally valid to argue the reverse. Hui and Cheng (1987) have put forward several hypotheses; that the listener's proficiency has no impact upon perception of the speaker, whose own language proficiency is of greater consequence; that the listener will view the speaker favourably because the interaction will be seen as a learning opportunity to improve language skills; conversely, the listener will feel threatened and therefore perceive negatively a speaker whose language proficiency is better than theirs; the listener will consider proficiency in the speaker's language as an act of loosing their own cultural heritage, and thus will have negative feelings toward the speaker; and finally, consistent with accommodation theory, the listener will have a more favourable perception of speakers who either have similar levels of language proficiency, or alter their speech to make it more so.

In their study Hui and Cheng (1987) employed self-rated ability in spoken English and performance in an English examination to assess language proficiency. They concluded that high proficiency speakers were viewed more favourably than speakers with poorer spoken language skills, that listeners prefer to engage in task-oriented social behaviours with proficient speakers and listeners like proficient speakers more than speakers with lower proficiency. Hui and Cheng (1987) also found that the listener's own language proficiency did not influence their perceptions of the speaker. An important methodological consideration which limits the extent to which these results can be generalised outside of the laboratory situation is the fact that the listeners rated the speakers from an audiotape. This is a very artificial situation and is not analogous to a medical encounter where the two parties may be evaluating each other on numerous other dimensions in addition to that of spoken language proficiency.

(ii). Colloquial Language.

One way in which medical students and doctors can employ their "second" register (Van Naerssen, 1985) and converge with their patients' speech is by using informal, colloquial language. Colloquial language refers to informal speech, including words, phrases, expressions and metaphors (Hughes, 1993). It has been demonstrated that Australian medical practitioners liberally "pepper" their consultations with colloquialisms when speaking to patients, who return with like speech (Chur-Hansen and Barrett, 1996) (Appendix IV.II).

It has long been recognized that a working knowledge of colloquial language is important for medical students and practitioners. In Australia, Posen (1968) expressed concern that some Asian students studying Medicine at the University of Sydney were unable to acquire a working command of idiomatic Australian English. Similarly, Lowry and MacPherson (1988) argued that some foreign medical graduates in London had an insufficient command of colloquial language to allow them to practice medicine in Great Britain. In the United States, the first symposium on the problems faced by foreign medical graduates was held in 1969 (Sutnick, 1970a). During this symposium participants

discussed the need for foreign medical graduates to be proficient in both speaking English and understanding spoken English, including colloquial medical and non-medical terminology, conversational English, and formal English. Sutnick (1970b) reported that a glossary of colloquialisms, in addition to a five week training programme in pronunciation, conversational English and medical terminology had been offered to foreign medical graduates in Philadelphia in the United States since 1964. The need for a working knowledge of colloquial language, both spoken and also informal language committed to paper on case notes by American doctors, was reiterated by Sutnick, Kelley and Knapp (1972).

In describing methods of orientation of foreign medical graduates to their new homes and places of work Sutnick, Reichard and Angelides (1971) outlined a methodology for teaching English, whereby the instructor acted as a patient whilst the foreign graduate took a medical history. During the interchange the "patient" used colloquial language and corrected the foreign graduate's grammar, pronunciation and use of words. Sutnick, Reichard and Angelides (1971) recommended tape-recording these exchanges so that the graduate's learning could be continued with concrete examples.

Millward (1970) stated that the taboo areas of "elimination, sex and death" (p 431) are often referred to euphemistically, and that the choice of term used by patients is mediated by gender, socio-economic status, knowledge of medicine, ethnic and cultural background, age and mood. She argued that foreign medical graduates are often ill prepared for the complexities of the English language that they encounter with patients, and that teaching programmes should address the issue of vernacular language. Sanders (1993) has discussed the importance of finding the right language to use when speaking with patients about sexual issues, and in particular the appropriateness of colloquial terms rather than clinical ones. Baumslag (1970) and Riederer (1970), in reply to Millward (1970), considered that the informal language barrier between doctors and patients of different socio-economic backgrounds was more problematic than doctors and patients of different language backgrounds. It has been demonstrated that general practitioners offer

more explanations to patients of higher social class than those from lower socio-economic status (Pendelton and Bochner, 1980). As well as replicating this finding, it has also been demonstrated by Waitzkin (1985) that doctors of higher social class origin provided more information to patients than those from lower-middle and working class families, although Freemon, Negrete, Davis and Korsch (1971) reported that paediatricians (whose class background was not specified) discussed the diagnosis more thoroughly with middle and lower class patients as compared with those of higher socio-economic status. None of these researchers have considered language background or English language proficiency in their investigations. Thus, the impact upon patients and their doctors who differ on both social class background and language background dimensions has still to be investigated empirically.

In a teaching programme designed to teach non-English speaking background Australian medical school undergraduates colloquial language, Chur-Hansen and Barrett (1996) (Appendix IV.II) drew upon a pool of general practice consultations conducted in Adelaide, South Australia, which had been audio-tape recorded and transcribed by Winefield and her colleagues (Winefield and Murrell, 1991, 1992; Winefield, Murrell and Clifford, 1995) (see above). In order to gauge the extent to which Adelaide general practitioners and their patients employed colloquialisms when interacting, 60 transcripts were randomly selected from a total of 210 and examined for examples of colloquial language. Of the 384 instances of informal language gleaned from this procedure, the most commonly used were "pop" (as in "pop up on the bed"), "tummy", "crook", "drop" (for example, "you can drop that slip at the front desk"), "crop up" and "not 100%" (Chur-Hansen and Barrett, 1996, p 414). Interestingly, the six doctors represented in these transcripts used informal language, including swearing, more often than did their patients.

A teaching project for non-English speaking background nursing students at the Curtin University of Technology in Perth, Western Australia (Brown, Chadwick, Kulski, Thompson, Palmer and Goldie, 1997) has similarly recognized that such students need to be exposed to Australian colloquial language for effective interaction with patients. The

project, called "Slanguage" is a novel one, in that it is accessed via the internet (www.curtin.edu.au/curtin/dept/nursing/ww/wickword.html) and illustrates the meaning of slang and euphemism via cartoon in addition to enabling students to hear the Australian pronunciation of the words and phrases, by means of a computer sound card.

Thus far this review has concentrated upon the need for doctors to know the colloquial language of their patients but it is also essential that they comprehend the colloquialisms used by their professional peers. Korsch, Gozzi and Francis (1968) reported that a major barrier to patient understanding was the doctor's use of "hospital shorthand" (p 863) to describe various procedures, such as referring to patient admission to the hospital as "admitting her for a workup". Shapiro (1980) has also noted that medical practitioners have their own forms of slang, often used when speaking to each other about patients in a derogatory or condescending manner.

Cohen-Cole and Friedman (1983) have provided an interesting example of the use of colloquial language by medical practitioners in the United States of America. They listed several pejoratives employed to describe "difficult" patients experiencing problems stemming from psychosocial as opposed to biological causes. In a one year research project they noted the incidence of disparaging slang when referring to troublesome patients, including "crocks", "turkeys", "trolls", "gomers" (p 52) "troll of the year", "in the head", "supratentorial" and "not real" (p 55). Christy (1979) has likewise noted the use of colloquial language between New England hospital-based medical practitioners, including "gorked out" for comatose, "squash" for "cranium" and "flight deck" for "neurosurgical intensive care unit" (p 980).

Similarly, in an ethnographic study by Barrett (1996) conducted in Adelaide, South Australia, around 40 metaphors were found to have been used by members of a psychiatric team, including doctors, nursing staff and social workers, to describe a person with schizophrenia in the psychotic stages of the illness. These colloquialisms included "out of his tree", "climbing up the walls" (p 147), "high as a kite" (p 148), "mad as a cut snake",

"silly as a wheel", "nutty as a fruitcake", "batty", "gone off his rocker", and "tipping over" (p 149). The ward in which severely psychotic patients were contained was referred to by staff as "the wombat pit" (p 149). That a reference to an Australian native animal is made in this slang illustrates the need for practitioners to learn the colloquial language of the geographical location in which they work. Even within the same language and country the colloquial language used may differ according to such factors as the age group and socio-economic status of the patients and whether the doctor works in a rural or urban setting (Chur-Hansen and Barrett, 1996).

Whilst there appears to be little recognition of the fact in the medical education literature, students from non-English speaking backgrounds are likely to have difficulty not only with what has been defined as "colloquial" language, but also in interpreting the meaning of the many metaphors and proverbs that abound in Australian English (McGowan, 1996b). Hughes (1993) refers to "figures of speech" as similes or metaphors that are used for special effect but not intended to be taken literally by the listener. Metaphors are especially prevalent in the media (McGowan, 1996b), which may be problematic for students who must analyse newspaper, radio and television reports on public health issues, for example. McGowan (1993) has studied the incidence of figurative, metaphorical and idiomatic language encountered by undergraduate medical students at the University of Adelaide by analysing audiotapes of a conference, lectures and doctor-patient interviews. She found a high proportion of metaphorical language was used by academic and clinical staff and patients. These were categorized as "cultural", reflecting the culture of the university and tertiary study, "ideational", involving reference to medical content, "interpersonal" including statements in the first and third person, "distortion", encompassing understatement, humour, irony, sarcasm, hyperbole and acronyms, and "breadth of usage", this being language requiring familiarity with South Australia, particularly in relation to local medical practice. An interesting future research project could further investigate the incidence of the use of these types of language in the doctor-patient encounter and also the level of understanding that all medical students have of this language, regardless of their language background.

(iii). Reading and Written Language.

Both medical students and practitioners are required to write for the purposes of scientific medical research and the practice of clinical medicine (Bloom, 1986). Medicine has been described as "the profession of writing" (Reynolds, in Reynolds, Mair & Fischer, 1992, p xv), with writing considered a fundamental skill for doctors (Holmes, Lin, Fath & Gray, 1992) who must relay information to colleagues and patients (Liu, Bassett and Sayre, 1994). Van Naerssen (1985) considers writing the medical record to be an example of one of the registers in which medical practitioners must be proficient.

In addition to writing, medical students and practitioners must read copious amounts of material, including the texts and notes of their undergraduate training to the case reports, letters from colleagues within and outside their own profession, journal articles and information from pharmaceutical companies. Yet surprisingly few studies have attempted to evaluate medical students' and graduates' reading abilities.

Yanoff and Burg (1988) surveyed 100 United States medical schools and found no fewer than 29 diverse and complex types of writing tasks were required of students and graduates. They reported that the most important writing tasks were deemed by medical educators to be the writing of case histories and physical examinations, progress notes, discharge summaries, peer-reviewed published papers, grant proposals, letters to referring doctors, outpatient records, consultation reports, admitting notes, student evaluations and communications to patients (Yanoff and Burg, 1988). Despite its recognized importance, little research has focused upon the case report (Maclean and Maher, 1994), and correspondingly, few efforts have been made by medical educators to teach students how to read efficiently and write clearly and concisely. A review follows here of the studies that report efforts to incorporate writing skills into the medical curriculum or to establish the need for reading and writing skills for medical students.

Kerr (1996) has referred to the "chaos" of case notes, noting the "cure" of this "disease" to be " . . . clear, accurate and contemporaneous patient records which report the

relevant clinical findings, the decisions made, information given to patients and any drugs or other treatment prescribed" (p 5).

Friedman, Sutnick, Stillman, Regan and Norcini (1993) piloted the use of a "patient note" written by foreign medical graduates after a simulated consultation with a standardized patient. This idea was expanded upon by Stillman, Regan, Haley, Norcini, Friedman and Sutnick (1992), who used the patient note to assess the accuracy of a written report of a standardized patient encounter[9]. Given that the primary function of the medical record is the documentation of clinical decision making in diagnosis, treatment, patient management and prognosis (Payne, 1979), it is important that medical educators encourage clarity and accuracy in medical students' and medical practitioners' writing skills. Levenson (in Reynolds, Mair & Fischer, 1992) stated that the medical record, especially within mental health, requires that the writer employ the jargon of several different professions and bureaucrats, as well as accurately recording the encounter between the doctor and patient. Thus, clarity and accuracy might become especially pertinent where the writer is not fluent in the language in which they are writing. Poor written expression may also be compounded by indecipherable script (Kerr, 1996; Payne, 1979), creating further difficulties for the reader. Van Naessen (1985) has pointed to some of the problems in preparing medical records faced by foreign medical graduates in the United States. These include limited vocabulary, the inability to express meaning unambiguously, interpretation of other practitioners' case notes, understanding the colloquial language of patients and translating it into medical terminology, the appropriate use of abbreviations and the passive voice, and the appropriate omission of verbs, subject nouns and articles. She recommends specific training in writing medical records for medical practitioners with English language difficulties, using real patient data in controlled situations.

Poirier & Brauner (1988) have discussed the fact that the case report is created by the medical practitioner by reducing and simplifying the disorganized and extensive amount of information presented by the patient to suit the traditional format of case

---

[9] These two studies are more fully described later in this literature review.

presentations. Some sections of a report will be primarily objective, such as the review of systems, physical examination and laboratory data, but other sections, such as the family and social history of the patient may be a combination of subjective and objective information (Van Naessen, 1985). Subjectivity could be as a result of the practitioner's personal biases regarding such characteristics as the patient's gender and racial background, or the practitioner's own adherence to a particular theory or hypothesis (Poirier and Brauner, 1990; Reynolds, Mair & Fischer, 1992). If a practitioner who shares the same language and similar sociocultural background as the patient can introduce subjectivity and thereby misrepresent the patient's narrative, perhaps a medical student who has far less in common with the patient may especially distort the story, in committing it to paper, particularly if they are not comfortable or skilled in writing in English.

Manning (1989) reported that medical students' and residents' discharge summaries are often incoherent. He attributes this to no training and states that; "medicalese" is the native tongue of most medical housestaff and that the English language proficiency of medical school graduates is borderline illiterate (and illegible when written)" (p 453). Unfortunately Manning did not provide further details about the graduates to whom he is referring, so it is unclear as to whether he is making a strong point about the poor writing skills of all medical graduates irrespective of language background, or whether he perceives particular problems with those from non-English speaking backgrounds.

Roland (1970) called for English speaking doctors to write more clearly for the benefit of foreign readers. He argued that if contributors to medical journals wrote for an imaginary third year medical student from a non-English speaking background not only would foreign readers be appreciative, but English language readers would also have increased comprehension. Crichton (1975) lamented the incidence of obfuscation in the writing style of medical journal articles. Similarly, R. Thompson (1990) concluded that all

undergraduate students, regardless of language background, benefited from a compulsory 12 hour course focusing on essay writing skills offered at the University of Florida.

Fox and Meijer (1980) have described a course in medical authorship and the oral defence of a doctoral dissertation for graduate medical students in Sweden. They found that whilst their students could write in English at a satisfactory level, many had limited experience in critically evaluating and revising written work. Fox and Meijer (1980) postulated that the ability to critically review one's own writing is a skill that should be taught to all medical professionals, regardless of the language in which they are working.

Greenberg and Jewett (1987) described a programme to improve undergraduate medical students' writing skills during a paediatric clerkship; although not stated, it would seem that their students all had a high degree of English language proficiency.

An important writing task for undergraduate medical students is the taking of lecture notes. Isaacs (1989) reports that virtually no literature within medical education exists on this aspect of writing. Drawing on research from other areas, he suggests that accurate note-taking is beneficial for recall and secondly, for facilitating active learning, by allowing students to decide what information they consider important to note.

Flaherty, Rezler & McGuire (1982) conducted a longitudinal study of the reading and writing skills of 231 incoming undergraduate medical students at the University of Illinois. They found that students' scores on standardized measures of reading and writing skills were predictive of these skills in clinical settings three years later. They also found that approximately 25% of their cohort were identified at initial testing and follow-up as experiencing problems with reading and writing. Unfortunately the demographic characteristics of the subjects were not provided, meaning that conclusions as to why such difficulties might have existed in this student group cannot be made. The authors suggested that at admission all medical students should be required to complete standardized reading and writing tests, whilst on entering medical school "at risk" students

could be offered course work in basic and clinical reading and writing. They stress that such innovations would benefit from longitudinal evaluations, and point out that knowledge in the areas of medical students' language proficiency and ways in which it can be improved are limited.

McGlinn and Jackson (1989) analysed the predictive validity of grade-point averages, an admission test and ratings of progress for the academic success of minority students at the Southern Illinois University School of Medicine. Details of these students in terms of their language background or residency status were not provided. McGlinn and Jackson (1989) found that the sole predictor of progress through the course was the Reading subtest of the admission test. This is notable, since it was the only language skill assessed in their study, and they suggested that perhaps admission tests need to consider these variables more closely when selecting students.

Lyons & Payne (1974) investigated the hypothesis that standards of medical recording and medical care are related, and concluded that they are. They quote Donabedian (1969);

> "Appropriate recording is itself an important dimension of the quality of care . . . because the record is the major instrument of communication in the management of care and, as such, an indispensable tool whenever two or more persons must cooperate in the provision of care. It is the major vehicle for the co-ordination of care during any one episode and for the continuity of care over time. . . . Most students of the field would agree that good recording is likely to be associated with good care mainly because the conditions that bring about good care are also responsible for bringing about good recording." (Lyons & Payne, 1974, p 714).

Fessel and Van Brunt (1972) have concluded that due in part to the inadequacy of medical records, such data may be incomplete and misleading if used as a measure of the quality of medical care. Such findings have most important implications for medical educators. For example, they beg two questions; first, "Do medical students who have difficulties with language keep poorer medical records than students who are proficient in English?", and second, "Do medical students who keep poor records (regardless of their language proficiency) graduate from medical school to offer their patients a lower quality of care?" Both of these questions require closer examination.

A search of the literature has failed to locate research specifically addressing Australian medical students' writing skills. Holbrook and Bourke (1989) assessed the writing skills of students from a range of courses (but not including Medicine) from the University of Newcastle, New South Wales, and the Universities of South Australia and Adelaide in South Australia. The language backgrounds of these students were not described. Holbrook and Bourke (1989) employed a standardized test (the English Skills Assessment) and their own checklist applied to a narrative writing task to measure aspects of written language, including spelling, punctuation and capitalization, sentence structure, usage and vocabulary. Frequent errors were made by students in each of these areas, and the authors indicated that their measure was useful in identifying areas for developing interventions to assist students with less than optimal written skills.

(iv). Listening.

> "Hearing the patient's message is the *sine qua non* of a great physician. To hear that message requires first the physician's interest, second his understanding of the meaning of the language, and third his sympathy towards and his knowledge and understanding of the circumstances of the patient's life; these again he hears best by listening to the patient" (Pickering, 1978, p 554).

It is well established in the literature that one of the primary tasks of the doctor is to listen to the patient (DiMatteo, Prince and Taranta, 1979; DiMatteo and DiNicola, 1982; Menahem, 1987; Egan, 1990; Presswell and Stanton, 1992). "Listening" in the medical literature on doctor-patient consultations usually refers to "active listening", when the doctor not only hears and comprehends what the patient says, but in addition is cognizant of the emotional meaning behind the spoken words (Presswell and Stanton, 1992). Some researchers have investigated how effectively the doctor actively listens to the patient, but none have considered the influence of either parties' language background or language proficiency upon listening. As no details about these aspects of the doctor or patient are provided in the majority of studies, it must be assumed that the two share a common language background and are both highly proficient.

Butler and Rollnick (1996) have outlined three potential impediments to understanding meaning when listening to a patient, which can also be applied to the patient's understanding of the doctor. The speaker (be it doctor or patient) may express themselves inaccurately; the listener (either party) may not hear what has been said; or the listener may incorrectly interpret the speaker's meaning. Whether medical students and practitioners are more or less likely to experience difficulties in listening according to their language background and level of language proficiency has yet to be explored.

(v). Speech acts.

Requests, apologies, thanks, invitations and other forms of socially conventional polite address are called speech acts (Richards, Platt and Weber, 1992). Such politeness differs across cultures (Furnham and Bochner, 1986). For example, cultural variations exist in terms of how direct or "blunt" people are in making and denying requests (Furnham and Bochner, 1986). Likewise, in some cultures a verbal "thank you" is obligatory, whilst in others it is not, or a non-verbal sign of thanks may suffice (Furnham and Bochner, 1986).

Very few authors have specifically considered the role of speech acts in the doctor-patient encounter. Robins and Wolf (1988) explored the occurrence of "redress and conversational repair [through] face preserving or polite linguistic gesture" (p 217) by examining the responses given by 172 first year University of Michigan medical students to the written prompt "You can talk to me all you want, but I will *not* follow that diet! I have had it with all of you telling me how to organize my life!" The forms of politeness that students wrote they would use in response to this hypothetical patient statement included apologizing, admitting an impingement upon the patient's rights, and giving reasons. The authors concluded that overall students were aware of the culturally sanctioned speech acts appropriate for such a situation. They argued that such linguistic gestures approximate the notion of respect, and are important for a therapeutic alliance. Whilst this study represents an important effort to investigate an otherwise neglected aspect of the doctor-patient encounter, it has several weaknesses. First, a written response

to a written prompt lacks the immediacy of a true interaction. Of greater value would be research that gauges actual spoken language in response to real patients. Second, no details were provided about the students' cultural or language background, and it could therefore be assumed that all were white, Anglo-Saxon Americans, although this may not have been the case. These demographic characteristics are important information, as it would be probable that a familiarity with the culture of the patient (who must also be assumed to be white and American) would result in more conventional speech acts along with other forms of response, such as management of the asymmetrical power base inherent in doctor-patient relationships (Lacoste, 1994; Erikson and Rittenberg, 1987).

(vi). Non-verbal behaviours.

Non-verbal behaviour is considered by Pendelton (1983) to be a process variable in the doctor-patient interaction. Such behaviour includes bodily contact, proximity, posture, physical appearance, facial expressions, gestural movements, direction of gaze and the non-verbal aspects of speech, such as timing, emotional tone (pitch and stress), speech errors and accent (Argyle, 1969). Pietroni (1976) categorises non-verbal behaviour into proxemics (clothing, spatial configuration of furniture, physical distance, conveying the amount of time that one can spend with another); kinesics (facial expression, head position, eye contact, body posture, hand gestures); paralanguage (rate of speech, fluency of speech, hesitations, interruptions); and touch (how, when and why it is used for communication). Non-verbal behaviour serves to "complement and illustrate aspects of the spoken word" (DiMatteo and Friedman 1982, p 91).

It is well recognised in the literature that the appropriate non-verbal signals between doctor and patient are essential in the building of rapport and a trusting relationship (DiMatteo and DiNicola, 1982; Drife, 1990; Cohen-Cole, 1991). Friedman (1982) reported that in 1979 the American Board of Internal Medicine required medical internists to demonstrate competencies in interpersonal skills including the ability to recognize and respond to the non-verbal cues from the patient. Friedman (1982) has argued the importance of this initiative in sensitizing medical students to these cues, given

that research indicates that patients not only express themselves through non-verbal channels, but are also sensitive to the non-verbal messages sent by their health care practitioners. Misunderstandings based upon the misreading of either the doctor's or patient's non-verbal cues can result in serious errors in diagnosis, can increase the likelihood of non-compliance and negatively impact upon the doctor-patient relationship (DiMatteo and Friedman, 1982; Harrigan and Rosenthal, 1986).

Many researchers have neglected to take non-verbal factors into account when examining the medical encounter (Larsen and Smith, 1981; Bensing, 1991). This is problematic given that it has been calculated that up to 75% of the affective component of a medical interview is conveyed through tone of voice, eye contact, position of the body, laughter, facial expression, touch and physical distance (Ong, DeHaes, Hoos and Lammes, 1995). There is very little empirical data on the role of non-verbal behaviour in medical settings, probably in part because of the methodological difficulties involved in recording and coding such behaviours (Stiles and Putnam, 1990). In 1980 Byrne and Heath commented that the non-verbal behaviour of doctors required far greater consideration in the literature, given its "immense importance" (p 327). However, to date research that has concentrated upon non-verbal behaviour in the doctor-patient interaction has often focused on the patients' expressiveness or otherwise, patients' decoding of the doctor's non-verbal behaviour and the effects of patients' non-verbal behaviours on the interaction (Pendelton, 1983) at the expense of studying the doctors' non-verbal behaviours. The following review of the relevant literature includes studies that have considered the health care practitioner's behaviour. Although not yet conducted, an investigation of the influence of the doctor's language background, language proficiency and cultural background upon non-verbal communication in the doctor-patient interaction would be an interesting and pertinent area of inquiry. It could be proposed that just as there may be misunderstandings between the doctor and patient based upon verbal factors, so too doctors may misperceive or inappropriately employ non-verbal behaviours where their language background or culture differs from their patient's. For example, in a study of discourse by Erikson and Rittenberg (1987), foreign medical graduates working in the United States have been

shown to fail to respond to the paralinguistic cues of their patients. Poyatos (1984) has postulated that an individual needs to master not only language (which he terms "linguistic fluency") but also non-verbal norms as well (for which he has coined the term "verbal-nonverbal cultural fluency"). In his words, "not only do we "speak" a language . . . but "move" it, and . . . people from other countries may display alien kinesic accents and not only audible ones" (Poyatos, 1984, p 456-457).

(i) Touch.

Touch, also referred to as "haptic behaviour" (Remland, Jones and Brinkman, 1991), has been identified as therapeutic and a means of relaying empathy (Friedman, 1979), although this has not been carefully researched. Friedman (1982) concluded that touch "may have symbolic value in healing, may create positive expectations, may have important physiological effects and . . . may affect the interpersonal nature of the practitioner-patient interaction" (p 59). Ong, DeHaes, Hoos and Lammes (1995) have also noted the role of touch in demonstrating affect. Blondis and Jackson (1977), writing about the nurse-patient encounter, suggested that in nursing touch may be the most important of all non-verbal behaviours. Despite the seeming importance of touch, there is a dearth of research in the area (Friedman, 1979).

In a comparatively early study conducted at the University of California, Aguilera (1967) investigated the relationship between mutually acceptable touching of psychiatric patients by nurses and increased verbal interaction. She concluded that touching by the nurses resulted in greater rapport with patients, more positive perceptions of nurses by patients and as hypothesized, increased verbal behaviours between patients and nursing staff. The author acknowledged that cultural and age differences might dictate the appropriateness of touching some patients, but concluded that touch is a powerful therapeutic tool in the health care setting, at least in Western cultures.

In a United States study Comstock, Hooper, Goodwin and Goodwin (1982) found no correlation between patient satisfaction and the amount of physical contact between the

doctor and patients from Anglo-American, Hispanic, American Indian and Negro ethnicity. In a study by Weinberger, Greene and Mamlin (1981) the use of touch was not found to be predictive of satisfaction for either the doctor or patient.

It is possible that barriers to touch, where the doctor and patient are from differing cultures, may result in dissatisfaction from the patient's perspective. Ong, DeHaes, Hoos and Lammes (1995) speculated that cultural differences may explain why patients in a study conducted in Washington in the United States by Larsen and Smith (1981) were less satisfied with the medical consultation and understood less of the information given to them where they had been touched by the doctor, whilst in a study conducted by Scarpaci (1988) Chilean patients evaluated good care as that which involved touching by the health care practitioner. In an Australian context, it could be hypothesized that Muslim students and doctors may touch their patients less than students and doctors from Christian backgrounds, particularly where the patient is of the opposite gender. It is important to stress, however, that there are no empirical data to support such a notion, although it would be an informative line of enquiry. Indeed, it would be interesting to consider whether opposite gender as compared to same-gender doctor-patient encounters result in more or less touching by the health care professional, regardless of cultural background.

(ii). Eye Contact.

Eye contact is considered as fundamental in eliciting information from patients (Cohen-Cole, 1991; Myerscough, 1992) and conveying interest and attention (Bensing, 1991). Active listening, described above, involves non-verbal signals such as the appropriate use of eye contact to relay to the patient that their meaning both literally and emotionally, is being heard and understood (Presswell and Stanton, 1992). Following an analysis of general practice consultations, Byrne and Heath (1980) concluded that eye contact is extremely important, and when used appropriately can facilitate the interaction between doctor and patient. They stated that inappropriate use of eye contact can hinder or prevent the likelihood of a satisfactory doctor-patient encounter, although Comstock, Hooper, Goodwin and Goodwin (1982) found no correlation between patient satisfaction

and the amount of eye contact between doctor and patient. Just what rules should be followed for eye contact in general practice were not made explicit by Byrne and Heath (1980), who considered that it should be the doctor's prerogative to judge what is "appropriate".

(iii). Tone of voice.

Tone of voice refers to the variations in pitch that are independent of the verbal content of speech (Friedman, 1982). Tone of voice is considered to one mechanism through which empathy (Zinn, 1993) and concern (Hall, Roter and Rand, 1981) can be relayed to the patient. In an early study by Milmoe, Rosenthal, Blane, Chafetz and Wolf (1967) doctors whose voice was judged to be low in the expression of anger were more successful in referring patients with alcoholism for further treatment. Friedman, DiMatteo and Taranta (1980) conducted a study in which they demonstrated a weak but statistically significant relationship between the doctor's ability to express positive affect through tone of voice and patient satisfaction. DiMatteo and Taranta (1979) reported that doctors skilled in reading the emotional messages relayed via their patients' tone of voice were rated by those patients as having a greater rapport with them in contrast to doctors lacking this ability. Friedman, DiMatteo and Taranta (1980) concluded that whilst emotion can be relayed through voice tone, individual differences exist in terms of one's ability to express emotion through tone of voice. Even though DiMatteo and Taranta (1979) indicated that over half of the 400 patients in their study were not born in the United States, whether two parties from different cultures will share a common understanding of and sensitivity to tone of voice has yet to be explored in the context of a medical encounter.

In a series of studies by Blanck, Rosenthal and Vannicelli (1986) it was found that psychotherapists were rated by their supervisors as being more competent in dealing with patients when their tone of voice was low in anxiety. The same psychotherapists were also found to alter their voice tone according to the patients' hospital status (outpatient vs inpatient), with more anxiety expressed when dealing with inpatients, and patient prognosis, and a more professional and competent tone used when speaking to patients for

whom a good outcome was expected. Female therapists were found to speak in a less professional and competent tone (as judged by independent raters) but also employed a more honest and warm tone than their male counterparts. An extension of this research would be an investigation of the effects of language background, language proficiency and cultural background upon tone of voice when interacting with patients who either share or differ from the doctor on these dimensions.

(iv). Accent.

Accent is a characteristic of speech that sends the listener clues about the speaker's background, such as the country in which they have lived, their socio-economic status, and whether they are a native speaker of the language in which they are conversing (Richards, Platt and Weber, 1992). One's accent is a non-verbal component of speech (Argyle, 1969), which reflects "national, regional, social class, educational or occupational background" (Argyle, 1969, p 114). All speakers have an accent; where that accent is different from the listener's the two parties may have little or no difficulty in understanding each other. Conversely, an accent can sometimes interfere with the comprehensibility of speech or can have detrimental effects upon interpersonal perceptions. For instance, referring to nursing students studying at an Australian University, Brown, Chadwick, Kulski, Thompson, Palmer and Goldie (1997) have advised that students from non-English speaking backgrounds may find that they are misunderstood by others because of pronunciation anomalies related to accent. Similarly, Colton, Heun and Link (1994) described a seemingly ethnocentric "accent reduction programme" (p 360) for a group of non-English speaking background students studying at a private college of osteopathic medicine in the United States.

Research upon the influence of accent upon the comprehensibility of speech was not located. However, several researchers have investigated the effects of accent upon the formation of impressions. Gallois and Callan (1981), writing from the University of Queensland, Australia, found support for the hypothesis that native speakers of English may react negatively to accented English, even where the non-native speaker's language is

fluent. In a study of British Asians, Elwell, Brown and Rutter (1984) concluded that the same person speaking with an Indian accent was rated less favourably on a number of personality measures by secondary school students from Kent, than when he spoke with a "standard" English accent. Bradac and Wisegaver (1984) have also found that "accent standardness" is associated with more favourable perceptions on the part of the listener.

(v). Rate of Speech.

"Rate" refers to the speed at which someone is speaking. It is possible to objectively measure the rate of spoken language by calculating the number of syllables uttered in a 60 second period (Street, Brady and Putman, 1983). It is also possible to gain a subjective impression of the rate of speech as seeming more rapid, or slower than the listener considers usual, although it is important to recognise that there is considerable individual and situational variation with regards to what is deemed as an acceptable rate of speaking (Street, Brady and Putman, 1983).

A fast rate of speech coupled with poor fluency and signs of anxiety tends to result in the formation of negative impressions about the speaker, whilst conversely a fast and fluent rate of speech is associated with positive judgements (Street, Brady and Putman, 1983; Bradac and Wisegarver, 1984). In addition, Street, Brady and Putman (1983) have demonstrated that a slow rate of speaking is not necessarily viewed negatively in situations where the speaker is in a context where they are to be evaluated, such as during an employment interview, or where the topic under discussion is unfamiliar or highly intimate.

In a teaching programme in interviewing skills conducted at the Faculty of Medicine at the University of Sydney in New South Wales, Australia (Farnill, Todisco, Hayes and Bartlett, 1997), undergraduate students' rate of speech was rated by a non-English speaking background interviewee on a five point scale from 1 (too slow) to 5 (too fast). Although the average rating on this measure was about 3, considered to be the

optimal, the interviewees stressed the importance of doctors speaking at a slower pace, pointing out that patients may not understand them if the rate of speech is too rapid.

(vi). Physical distance norms.

Surprisingly little research appears to have concentrated upon the spatial distances maintained between individuals, referred to as "personal space" or "physical distance norms" (Rustemli, 1986) or "interpersonal distance" (Sussman and Rosenfeld, 1982) within the doctor-patient encounter. Even though the doctor is exempt from many of the rules concerning where they can touch patients and how close they can get to them by virtue of their role (Winefield and Peay, 1991), it is the case that they must still abide by physical distance norms during many consultations, particularly during those where no physical examination occurs.

Research has demonstrated that there are cultural and subcultural differences in proxemic behaviours (Mehrabian, 1972; Bilmes and Boggs, 1979; Furnham and Bochner, 1986a; Remland, Jones and Brinkman, 1991, 1995). It has been claimed that if a doctor stands too close or too far from the patient problems can occur, as the patient may feel uncomfortable or distressed (DiMatteo and Friedman, 1982). However, this claim has not been carefully studied through research. A number of studies within social psychology have demonstrated differences in dyadic proxemic behaviour across cultures, including England, France, the Netherlands, Italy, Greece, Scotland and Ireland (Remland, Jones and Brinkman, 1995), the Netherlands, France and England (Remland, Jones and Brinkman, 1991), Japan, Venezuela and America (Sussman and Rosenfeld, 1982), and Fijians and Fiji-Indians (Thomas, 1974). Research is sorely needed which explores the impact of cultural norms upon the medical encounter, where theoretically these norms should be highly reduced in their impact, when the doctor and patient do not share the same cultural background. Remland, Jones and Brinkman (1995) have distinguished between "contact" cultures, such as those in Latin America and Southern Europe and "non-contact" cultures including those from Asia and Northern Europe. When individuals from two different

cultures meet and their preferences for proxemic and haptic behaviours are not similar, it is likely that misunderstandings may occur.

A study by Rustemli (1986) found that Muslim Turkish females allowed greater distances between themselves and males, and he concluded that females from Muslim cultures in general would be reserved in the presence of members of the opposite gender. Sussman and Rosenfeld (1982) found support for their hypothesis that where two people from a non-American culture interact in a shared second language, such as English, their proxemic behaviour becomes closer to American norms. To extrapolate from this finding, it may be that when two individuals from different cultures interact in a shared second language any variations in physical distance norms will be minimised. Whether this holds true for Muslim medical students interacting in English with Australian patients needs to be verified, along with a consideration of the impact of these differences in physical distance norms (if any) upon the doctor-patient relationship.

As well as cross-cultural differences regarding how close two people should be to each other, according to their relationship with one another, are the rules concerning proximity depending upon gender (Remland, Jones and Brinkman, 1991; 1995). Roter and Hall (1993) have speculated that since females (in Western cultures at least) have been found to show higher levels of eye contact, touching and closer physical proximity with other females than do males with males, this might be extrapolated to the doctor-patient encounter. Furthermore, Cline and Puhl (1984) noted that preferred seating arrangements differ within and between cultures as a function of gender. These are aspects of the doctor-patient encounter that require exploration, both in terms of the cultural backgrounds of the participants and their gender.

(vii). Cue discrepancy.

Cue discrepancy refers to the congruence between verbal and non-verbal cues (Friedman, 1979). Generally speaking, verbal and non-verbal components should be congruent (Argyle, 1969). It could be that doctors who make positive verbal statements

accompanied by a negative non-verbal cue or vice versa may elicit feelings of fear, anger or humiliation in the patient (Friedman, 1979). However, some research findings do not support this. Hall, Roter and Rand (1981) concluded that patients felt their doctor to be more concerned about them and were more likely to comply where negative affect was expressed through the tone of voice to relay positive affect expressed through words. Importantly, though, this study concentrated upon the paralinguistic and semantic features of speech employing audiotapes of doctor-patient consultations. Other important non-verbal behaviours such as facial expression or eye contact could not be considered. Studying the congruence between verbal and non-verbal cues apart from the paralinguistic aspects of speech would involve a number of methodological difficulties, which would possibly deter many researchers from working in this area (Bilmes and Boggs, 1979; Stiles and Putnam, 1990). Friedman (1982) has stated that considerable research is needed in order to understand the effects of cue discrepancies in health care interactions, and to date this is still the case.

(viii). Facial expression.

Historically, it has been argued that facial expressions are learned and therefore both similarities and variations in facial behaviour may be apparent across different cultures (Ekman, Friesen and Ellsworth, 1982). Similarly, evidence has been produced to support the notion of subcultural differences and similarities in facial expression (Davitz, 1964; Ekman, Friesen and Ellsworth, 1982). In a study by Ekman, Friesen, O'Sullivan, Chan, Diacoyanni-Tarlatzis, Heider, Krause, LeCompte, Pitcairne, Ricci-Bitti, Scherer, Tomita and Tzavaras (1987) it was demonstrated that there was considerable cross-cultural agreement in judging facial expressions across ten groups of subjects (from Estonia, Germany, Greece, Hong Kong, Italy, Japan, Scotland, Sumatra, Turkey and the United States). However, this research also found evidence for cultural differences in the judgement of the intensity of an emotion as portrayed by facial behaviours, with Asian subjects attributing less intense emotions to the expressions portrayed by Caucasians.

Matsumoto and Assar (1992) have argued that facial expressions of emotion are more accurately recognised by subjects in the English language as opposed to subjects who are non-English speaking. They posit that perhaps English is a more useful language for labelling and discussing emotion than some others (and here they suggest Chinese as an example), and provide data collected with Hindi subjects to partially support their argument. They concluded that a limited understanding of the relationship between language and judging facial expressions exists, calling for further research in the area.

A number of social psychologists researching the facial expression of emotion today would argue that there is a strong pan-cultural aspect to the encoding (or 'sending') and decoding (or 'reading') of facial expressions although this is by no means accepted by all scientists in the field (Katsikitis, 1992; Matsumoto and Assar, 1992). To clarify the debate between cultural universals and differences Ekman (1971) stressed the importance of differentiating between facial expressions of emotion, which could be viewed as innate and therefore universal across cultures and facial gestures, which he argued are learned and therefore may show differences across cultures. This distinction had in fact originally been made by Darwin in 1872 (Ekman, 1971). Facial gestures, which include winking, blinking and nodding are posited as being variable across cultures and may be independent of emotion (Ekman, 1971). In addition to expressions of emotion and facial gestures are "display rules", which refer to the norms that are learned through one's culture about which emotions are acceptable to show to whom and under what circumstances (Ekman, 1971). These rules may differ both between and within cultures (Ekman, 1971).

Phillips (1988), in a discussion of the ways in which Australian academics may need to adjust to the cultural mores of Asian students, indicated that whilst the facial gesture of nodding indicates understanding or agreement in an Australian context, for Indonesian students blinking and nodding are signals of respect. Martens (1991) has also stressed the need for tertiary educators to be culturally aware of and sensitive to such culturally specific non-verbal behaviours.

(ix). Non-verbal behaviours and patient satisfaction.

It is plausible that the skill with which doctors can encode and decode non-verbal behaviours may impact upon levels of patients' satisfaction with the medical encounter.

Harrigan and Rosenthal (1986) have stated that an important foundation for patient satisfaction with the medical encounter is empathy, a notion that has also been expressed by many other authors (for example Poole and Sanson-Fisher, 1980; Egan, 1990; Cohen-Cole, 1991; Myerscough, 1992; Ong, DeHaes, Hoos and Lammes, 1995). Truax and Carkhuff (1967) have defined empathy as "the ability to communicate [the perception that the listener knows and understands the thoughts and feelings of the other] in a language attuned to the client that allows him more clearly to sense and formulate his feelings" (Truax and Carkhuff, 1967, p 286). Harrigan and Rosenthal (1986) have argued that empathy is relayed not only verbally but non-verbally as well (see also Poole and Sanson-Fisher, 1979), through tone of voice, facial expression and the position of the body, including touch. They point out that because non-verbal behaviours are conveyed simultaneously through a number of channels including the face, eyes, head, arm, leg and body position and hands it is extremely difficult to code and study them. Researchers have attempted to gauge non-verbal behaviours in the medical encounter through standardized questionnaires, self-report from the doctor and patient and the analysis of videotaped consultations.

Harrigan and Rosenthal (1986) examined videotaped doctor-patient encounters and found significant relationships between doctors' body movement and patient ratings of rapport. Doctors were more favourably evaluated by patients when sitting face-to-face with them, nodding, leaning towards the patient and holding moderate levels of eye contact.

DiMatteo and Taranta (1979) have stated that empathy partially depends upon the ability to decode patients' facial and body cues and respond with the appropriate non-verbal and verbal cues. Rapport was also found by DiMatteo and Taranta (1979) and

DiMatteo, Friedman and Taranta (1979) to be significantly related to the doctor's ability to understand patients' emotional states through facial expression, body movement and posture. DiMatteo (1979) argued that doctors can improve the quality of their interactions with patients, including increased rapport and empathy, through the development of skill in reading non-verbal cues. In addition, she proposed that the ability to send messages of care, respect and concern to the patient through their own non-verbal signals such as touch, tone of voice and facial expression (Friedman, 1979) is an essential component of the doctor's expertise. The ability to decode and encode non-verbal behaviours is seen by DiMatteo (1979) as one of the doctor characteristics that may be predictive of increased patient satisfaction. Furthermore, DiMatteo, Prince and Hays (1986) provided evidence to support an association between doctors' popularity with patients, their ability to encode non-verbal signals and patients' likelihood of attending appointments with these doctors.

The relationship between the doctor's ability to send (encode) and read (decode) non-verbal messages of emotion with patient satisfaction was investigated by DiMatteo, Taranta, Friedman and Prince (1980) using standardized measures and patient ratings. They found that patients were more satisfied with doctors skilled in encoding and decoding non-verbal behaviours, though as the authors point out there were some methodological limitations which must be considered in light of this conclusion. First, because this was a correlational study, causal relationships could not be established and second, it was not possible to isolate non-verbal behaviours from other characteristics of the doctor, and therefore these other factors may have played a role in patient satisfaction. These qualifiers aside, the authors conclude that doctors who possess well developed non-verbal behavioural skills may serve to deliver higher quality care, shorter and therefore more cost effective consultations, increased patient compliance and greater patient satisfaction.

Immediacy refers to the degree of psychological "closeness" between individuals during an interaction (Mehrabian, 1972). Immediacy is a composite of liking, attention and openness to the encounter and is conveyed by touching, close proximity, forward lean and mutual eye contact (Mehrabian, 1972; Harrigan and Rosenthal, 1986; Larsen and

Smith, 1981). A study on non-verbal behaviour in the doctor-patient interview by Larsen and Smith (1981) found that immediacy was associated with both higher patient satisfaction and an increased comprehension of the information relayed by the doctor.

Weinberger, Greene and Mamlin (1981) observed and analysed 88 doctor-patient encounters and concluded that both parties reported greater satisfaction with the interaction where the doctor employed appropriate non-verbal behaviours. Doctors and patients were more satisfied when the former sat closer to the patient, rather than further from them, and demonstrated concern with nods and gestures.

An unusual study was conducted by Mendez, Shymansky and Wolraich (1986). It is unique in that it investigated the hand and speech behaviours[10] exhibited by medical students, residents and general practitioners, who were videotaped whilst informing a simulated patient that her baby had a genetic abnormality. Cultural and language background information was not provided about either party. Doctors rated as high on medical communication skills were found to exhibit facilitative and relaxed verbal and hand behaviours and to deal more effectively with the affective component of the interaction. Female practitioners were more skilled in these domains than were males.

---

[10] Speech and hand behaviours were categorized according to the Physician Nonverbal Communication Systems I & II (Mendez, 1983, in Mendez, Shymansky and Wolraich, 1986, p 28).
Speech behaviour categories were;
"Normal": doctor's oral expression does not exhibit any deviation from its normal flow and semantics.
"Sentence change": any interruption in the flow of an utterance.
"Repetition": any serial repetition of a word or words during the flow of a single utterance.
"Stutter": the act of stuttering.
"Sentence incompletion": any verbal expression that is left incomplete and is followed by a period of silence.
"Tongue slips": any substitution of an unintended word for an intended word.
"Incoherent sounds": any sound which is incoherent to the listener but does not alter the meaning of the utterance.
"Unclassifiable": any behaviour that does not fit the other categories.
Hand behaviour categories were:
"Neutral": doctor's hands and fingers are relaxed.
"Folded arms/hands": doctor's hands and arms are folded over the lap, chest or over the desk.
"Tense hands/fists": doctor's hands and/or fists are tightly clenched.
"Describing": the doctor uses hands to describe an object or event or simply makes gestures during verbalizations.
"Pointing": the doctor directs fingers towards the patient.
"Tapping/touching": the doctor taps or grabs patient's shoulders and/or hands.
"Manipulating": the doctor plays with an object or drums fingers over an object.
"Touch-self": the doctor touches, rubs chin, fingers, arms, and/or hands insistently.
"Unclassifiable": a behaviour that does not fit into the other categories.

The finding that being relaxed and being judged as demonstrating effective medical communication skills are associated variables might be explained by a third factor, that of confidence in one's abilities to interact with patients (Harrell, Kearl, Reed, Grigsby and Caudill, 1993). Thus, students and practitioners who are not fluent in the language of their patients may be hampered in their performance not only by their level of language skills but also by their self-confidence generally. This dimension of the non-English speaking background medical practitioners and students psychological experience and clinical and academic competence has not yet been considered in the literature.

The preceding studies examining non-verbal behaviours underscore their importance in the doctor-patient relationship, and lead to the premise that there is certainly a possibility for difficulties when the parties do not share similar cultural knowledge. Unfortunately though, no information was provided by any of the researchers regarding the language or cultural backgrounds of either the doctors or patients who acted as subjects. This is an area of investigation that has yet to be considered empirically within medicine.

Very few researchers have investigated the non-verbal behaviours of medical student populations. Evans, Stanley, Coman and Burrows (1989) conducted a study which investigated the non-verbal sensitivity scores on a standardized measure of fourth year undergraduate Australian medical students. The authors reported that the students fell within the normal range on the standardized test for decoding facial expressions, body movements and postures and voice tones. It was also found that non-verbal sensitivity scores did not increase after training in interpersonal skills, as had been hypothesised. Unfortunately, a description of the subjects' demographic characteristics, including language background and cultural origin were not provided by the authors. A replication of their research taking such variables into account would be worthwhile.

Hunter and Hayden (1990) have described the case of "Mary Tiang" a non-English speaking background speech therapy student at the Lincoln School of Health Sciences at La Trobe University in Victoria, who was unable to recognize and respond appropriately to

her clients' non-verbal behaviours. Clinicians at the University of Adelaide have observed inappropriate non-verbal behaviours in some of their undergraduate, English speaking and non-English speaking background medical students whilst both interacting with them, and with patients. Comments were most frequently made regarding Asian males and females. Clinical educators reported that some male students, on being corrected during teaching sessions would display a "smiling grimace" that was deemed highly inappropriate and considered provocative by some teachers who read it to mean disrespect on the student's part. Female students, particularly from Taiwan and Vietnam, were reported to avert their eyes when speaking to clinical teachers and some patients, a cultural norm that has been reported by Bradley and Bradley (1984). Again, this was argued to be problematic and frustrating[11]. Both the averted gaze and the grimace may indicate submission before a dominant other (Morris, 1982). Alternatively, there may be confusion and misunderstanding because of differences in culturally-based display rules between the educator and the student (Argyle, 1971). For example, some Asian female students may be uncomfortable in holding direct eye contact with a member of the opposite gender (Argyle, 1971). Perhaps, similarly, some Asian male students, on being reprimanded, have learned to mask embarrassment or shame by smiling.

[11] Personal communication.

II.III. Language skills distinct from Medical Communication Skills.

"There is an Australia-wide move to screen students before or shortly after entry to medical school to address English and communication skills. It remains unclear, however, what components of "English" (or language or communication) are important. Is it formal or informal English (or other language) competence and/or proficiency, verbal fluency (perhaps even glibness and plausibility) or an even higher order skill such as strength in the humanities? . . . Thus, if we are to swing enthusiastically to assessing future medical students in English and communication skills (and perhaps in enriching such competencies during the medical course), just what should be the focal points? . . (T)he need then is to define and refine the "thick" descriptors (e.g. "verbal competence") and variables (e.g. High School Certificate English mark) that are presently being considered as screening variables, which so far appear to have resisted clarification in the debate." (Parker, 1993, p 750).

This quote is notable on several counts. Parker calls for the need to define and clarify the idea that "language" plays a role in the performance of students in Australian medical schools. First, which aspects of language are of importance - are we referring to spoken English fluency as opposed to written or aural language skills, the ability to use English that is appropriate to the context of the interaction, or some other, unspecified dimension of language usage? Second, is it that "higher order skills", such as the critical, or analytical thinking style useful for constructing arguments and debates, that is more predictive of medical school success or otherwise, which require proficiency in the language of instruction, but which most certainly could not be argued to be the preferred style of thought for all highly proficient speakers of a language? Third, Parker makes the important distinction between "language" and "communication". Can these two terms be used interchangeably in the context of the Australian medical school system, where most curricula teach and assess "medical communication skills"? A number of authors have used the term "communication" interchangeably with the word "language" (Flaherty, Rezler & McGuire, 1982; Bloom, 1986). Can we validly make the assumption that proficiency in the English language is equivalent to the ability to demonstrate medical communication skills in clinical settings, such as this interchangeable use of terminology might imply?

The use of the terms "language" and "communication skills" in the medical arena to refer to the same skill is common. For example, a paper by Greenberg and Jewett (1987) entitled "The Case Presentation: teaching medical students writing and

communication skills" focused solely upon writing skills. Glendinning and Holmstrom (1991) entitled their textbook "English in Medicine. A Course in Communication Skills". Similarly, Love, Newcomb, Schiller, Wilding and Stone (1993) reported upon "A comparison of knowledge and communication skill evaluations by written essay and oral examinations in pre-clinical medical students". Love, Newcomb, Schiller, Wilding and Stone (1993) claimed to be investigating four communication skills by assessing students' written and spoken language; organization of the answer, the precise use of English, summarization and overall command of the subject matter. It could be argued that the way in which an answer is structured, how it is summarized and how familiar the student is with curriculum material may have little bearing upon English language skills *per se*. A number of other authors have also seemingly confounded "language" with "communication skills" (see for example, Maher, 1986a; Ferguson and Maclean, 1988; Holmes, Lin, Fath & Gray, 1992; Rolfe & Pearson, 1994; Torda, 1995; McGowan, 1996a). Lipton, Huxham and Hamilton (1988) took secondary school English to be indicative of "verbal communication skills" (p 383), which would seem to be an unwarranted interpretation, since matriculation English in their study would have been assessed through written essays and assignments. Rolfe, Pearson, Powis and Smith (1995) concluded in a study of the clinical performance of first year interns that;

> "We suggest that subject spread and the inclusion of a humanities background is important for effective medical practice, at least in the immediate postgraduate period. This view is not surprising since a large component of medical practice requires good communication skills in doctor-patient and inter-professional interactions." (p 1332).

It would seem that these authors are working on the stereotypical assumption that students who study in the sciences are bookish and lacking in interpersonal skills, presumably because their subject areas favour independent learning, whilst their humanities counterparts develop a number of interpersonal qualities that are desirable for doctor-patient interactions. However, many secondary school laboratory experiments in physics and chemistry require a collaborative effort, whereas the analysis of a Shakespearean play or a counterpoint exercise in music can be a solitary affair indeed. Or it may be that the authors consider the humanities, in comparison with the sciences, to require considerably greater ability in written language and reading, which may be so.

However, skills in these areas of linguistic ability do not necessarily translate into spoken language skills, and nor can they be seen to follow logically as akin to the skills of medical communication. Therefore it would be wise to debate and justify these assumptions before they become taken for granted as truths when they have such a limited basis.

In contrast, some authors have differentiated between language and communication. For example, Waitzkin (1985) drew attention to the fact that communication in medicine does not equate to the mere relaying of information, but encompasses the emotional bond and non-verbal elements between the doctor and the patient as well. In a consideration of the obstacles faced by overseas-trained doctors seeking registration to practice in Australia, Kidd and Zulman (1994) clearly distinguish English language proficiency from communication skills. Helman (1984) noted that communication is not comprised solely of "speech, non-verbal behaviour or the written word" (p 547) but also encompasses individual differences and the situation, or the "context". Stone (1979) clearly differentiated between the doctor's responsibility to take language barriers with patients into account as compared with the need for members of the medical profession to have an understanding of medical communication skills. Cummins (1980, in Light, Xu and Mossop, 1987) distinguished interpersonal communication skills from language proficiency. Reporting upon "Literacy Matters. Strategies for teaching *communication skills* to university students" (author's italics) Ingleton and Wake (1997) seemingly confound terminology, although they quote a report by the Federal Government Department of Employment, Education and Training (DEET) (1991) which stated that "literacy is the ability to read and use written information and to write appropriately in a range of contexts. Literacy involves the integration of speaking, listening and critical thinking with reading and writing" (Ingleton and Wake, 1997, p 6).

It is not difficult to see why "language" and "communication" might be used interchangeably; for instance, "communicative competency" is the term used in applied linguistics to refer to an individual's ability to "get their message across" using language in the context of the appropriate cultural and social knowledge (Richards, Platt and Weber,

1992; Maher, 1987; Bradley and Bradley, 1984). Bilmes and Boggs (1979) noted that; "It has become something of a truism that all of our behavior is, insofar as it is perceived by someone else, and whether intended or not, communicative" (p 48), and argued that this conception of communication is too broad. Neumann (1985) has captured the circularity involved in defining the two terms thus; "Language is means of communication of knowledge and effective communication depends on effective use of language" (p 194). Arguing that language proficiency and communicative competence are distinct skills Ingram (1984) quotes Sollenberger (1978);

> "The person's so called language proficiency, while it may have been quite accurate in technical skill terms, did not mean effectiveness in *communication*. In some cases, it may have enabled the person to misrepresent or foul up more effectively. This is to say that you can be a fool in any language; or, that you can put your foot in your mouth in any language. Nor does the fact of technical ability to use a foreign language without noticeable accent or grammatical errors mean that the person has something worth saying. I'm sure we all know people who talk nonsense fluently. On the other hand I know people who butcher the language, whose accents are atrocious and whose vocabularies are limited. For those reasons we give them a low proficiency rating. Yet, for some reason, some of them are effective communicators." (Sollenberger, 1978, p 8 in Ingram, 1984, p 17).

In medical education it is important to distinguish between language skills and communication skills, if they are not one and the same. To confuse them, if they are in fact independent, has implications for teaching, learning and research.

II.IV. Australian research on medical students' English language proficiency.

Very few studies have investigated the English language proficiency of medical students in Australia. Several researchers have reported demographic profiles of the countries of origin of their students (Tiller and Jones, 1984, 1985; Farnill, Hayes and Barrett, 1993). These give a descriptive account of the types of students enrolled in an Australian medical school, in terms of language background, but give no indication of levels of English language proficiency based on standardized testing. Whilst descriptive information about medical student groups is interesting, the questions that arise regarding the influence of language background and language proficiency upon performance in medical curricula cannot be addressed in such research.

In Australia the major contributors to our knowledge of medical students' English language proficiency and its relationship to academic performance are Dr Douglas Farnill and Associate Professor Susan Hayes, psychologists from the Department of Behavioural Sciences in Medicine, University of Sydney, New South Wales. A review of their findings and suggestions follows, including data from unpublished reports to the University of Sydney's Faculty of Medicine.

Hayes and Farnill commenced testing the language skills of incoming first year medical student groups in 1990. They assessed English language proficiency using the Screening Test of Adolescent Language (STAL) (Prather, Breecher, Stafford and Wallace, 1981), which is described in detail later in this Chapter. To 1994 they had consistently found that between approximately 12% and 19% of their students each year were identified by the STAL as having deficiencies in English language proficiency (Hayes and Farnill, 1993a, 1993b, 1994; Hayes, Farnill & Sefton, 1994).

Hayes and Farnill (1992) have demonstrated good concurrent validity between the STAL and another more extensive and reliable language test, the Woodcock Language Proficiency Battery (Woodcock, 1984 in Hayes and Farnill, 1992; Farnill and Hayes, 1995). Based on performances on the two instruments, Hayes and Farnill (1992) decided

upon a cut-off score of 19 out of a possible total score of 23 on the STAL for classification of students into those with difficulties in English language proficiency (scoring 19 or lower), and those without language difficulties (scores of 20 and above). They argued that although some misclassification may occur with a cut-off of 19 (false identification of students as having English language difficulties, and vice versa), the STAL is valuable nevertheless as a preliminary identifier of those who may need English language support programmes. They stressed, however, that the cut-off of 19 must serve as a rough guide for medical school staff from other Australian Universities, who must make their own decisions about sensitivity (correct identification of students with English language difficulties) over selectivity (correct identification of students with no English language difficulties) (Farnill, Hayes & Chur-Hansen, 1995) (Appendix I.II).

To further validate the STAL cut-off score of 19, Hayes and Farnill (1993a, 1993b) examined the performance of two cohorts of students in their Medical Communication Skills course.

In the second year assessment of the Medical Communications Skills course at Sydney University students conducted an interview with a staff member, and reported that interview in writing. The 13 students who failed in Medical Communication Skills had been identified by the STAL, administered one year earlier, as having language difficulties. However, a further 24 students who had also been identified by the test did pass the requirements of the Medical Communications course. In the third year assessment of Medical Communication Skills students were required to conduct six interviews with members of the public and conduct a clinical interview. Of the 30 students identified by the STAL as experiencing language difficulties two years previously, six failed to satisfy this course. Two students who failed the course had performed satisfactorily on the STAL. STAL data were unavailable for a further four students who also failed.

Hayes and Farnill (1993a, 1993b) argue that the predictive validity of the STAL is impressive, given consistent, although modest correlations between this measure of English language proficiency and academic grades for these two successive cohorts of students. The STAL and the Woodcock Language Proficiency Battery were found to be related to first and second year subject results in Anatomy, Behavioural Sciences, Biochemistry, Biomathematics, Chemistry, Histology and Embryology, Introduction to Medical Science, Pharmacology, Physics and Physiology. Hayes and Farnill (1993a, 1993b) caution that these correlations do not demonstrate a causal relationship between English language proficiency and academic performance, noting that confounding, unidentified variables associated with language difficulties may contribute to academic achievement.

On the basis of their validation studies Hayes and Farnill (1993a, 1993b) concluded that the STAL was successful in identifying those students who were experiencing difficulties in tasks which required satisfactory oral and written skills. They postulated that the discrepancy in the case of those students identified as having English language difficulties but who passed the second and third year Medical Communications course was due either to an improvement in language proficiency, or to the fact that the Medical Communications Skills examination was less rigorous in its standards for English language proficiency. However, they did not collect data to support either of these hypotheses. This throws into question the assumption that medical communication skills assessments can be used to validate measures of English language proficiency.

At the University of Sydney all students received a feedback sheet advising them of their performance in their language screening test; students who scored 19 or lower on the STAL were also invited to attend several workshops provided by the University's Learning Assistance Centre. Further, the Sydney University Counselling Service offered individual and small group support to students in need of language assistance. Over the course of the year a general lack of interest from students in the remedial workshops was demonstrated by poor attendance, even though the ratings given to the courses by those

who did attend were high. It appeared from the experiences initially reported by the University of Sydney that voluntary attendance at language-related workshops for undergraduate medical students was not effective in recruiting participants. However, Hayes and Farnill (1994) reported that the demand for language workshops from later cohorts of their students exceeded that which could be provided by staff, due to time and financial limitations.

Having administered the STAL to five successive cohorts of students, Farnill and Hayes (1996a) recognised the threats to reliability and validity associated with a brief testing instrument coupled with subjects who may be highly motivated to gain access to test items from students tested in previous years, so as to improve their performance. They developed an alternative test based upon the STAL and sharing its format, the AUSTEST, an acronym for the Australian Tertiary English Screening Test. The psychometric properties of the AUSTEST are described in Farnill and Hayes (1996b), including the test's criterion validity in terms of academic performance in first, second and third year medicine at the University of Sydney. Correlations between academic performance expressed as a weighted average mark and AUSTEST scores were found to be low to moderate (Farnill and Hayes, 1996a; Farnill and Hayes, 1996b). Further details about the AUSTEST can be found in Chapter IX which reports Study VII of this thesis.

II.V. English language proficiency and academic and clinical performance.

Farnill and Hayes, whose work has been reviewed in the previous section, have been the primary researchers in considering the relationship between English language proficiency and academic performance in an Australian Medical School. However, other researchers have also investigated how predictive English language proficiency is for academic success, although their study cohorts have not necessarily been Australian, nor exclusively students of Medicine. In addition, only Farnill and Hayes' research has used the Screening Test of Adolescent Language (STAL) (Prather, Breecher, Stafford and Wallace, 1981) or the Australian Tertiary English Screening Test (AUSTEST) (Farnill and Hayes, 1996a, 1996b) to screen for English language ability, which are two of the measures that have been used in this thesis.

A consideration of the methodological difficulties involved in research which attempts to predict academic outcomes by Graham (1987, p 506) identified four areas of concern; the criteria used to determine academic "success", which may not necessarily be valid indicators; the validity of the English language proficiency measures; the interpretation of statistical findings, which are sometimes dubious (for example, claiming that a significant correlation coefficient of 0.25 is sufficient to explain the relationship between two variables); and the numerous uncontrolled, confounding factors in addition to language proficiency that may account for academic success or failure. Thus in comparing research findings or attempting to generalize results, one must be cognizant of these methodological points. Indeed, the studies reviewed below highlight several of these issues, which if not taken into account by the researcher, can weaken the validity and generalizability of the findings. It should be noted that only selected studies have been included here as the intention is to be illustrative, rather than to give a comprehensive review of the literature in this area, as the main focus of this thesis is language proficiency as a predictor of medical students' academic and clinical performance in Australia (an area in which there is limited research).

Rusek (1992) conducted a study at CALUSA (Centre for Applied Linguistics in the University of South Australia) to investigate whether scores on the International English Language Testing System (IELTS) predicted academic success at South Australian universities. Based upon the first semester academic results of 63 students, (48 undergraduate, 15 postgraduate) including 24 undergraduate medical students, Rusek (1992) stated that all students had been "successful" in fulfilling the requirements of their course and therefore considered that the IELTS test scores (which ranged from "acceptable" [n=39] to "unacceptable" [n=24]) could not be used to predict academic achievement. This is a debatable conclusion, since success was defined as being permitted to continue their course in second semester. Students at South Australian universities are rarely asked to leave a full year course because of poor performance after having completed only one semester. Since Rusek (1992) indicates that four students withdrew from their course and three failed one or more of their course subjects in first semester, it is questionable whether these seven students can be regarded as being successful. In addition, a number of students may have barely passed examinations in first semester, or failed to satisfy the requirements of their course at the final end of year examinations. Unfortunately these data were not collected, with the author making the curious (and unsubstantiated) statement that "results beyond the first semester were not sought as the immediate influence of the IELTS would have well and truly been negligible by Semester 2" (p 2). Given the methodological limitations of her research, Rusek's (1992) findings appear wanting in validity and thus do not add to our understanding of the relationship between English language proficiency and tertiary performance.

Abdulrazzaq and Qayed (1993) evaluated the performance of medical students from the United Arab Emirates University to determine predictors of university performance in the first two years of the course. One of their measures was an English language proficiency test, the details of which were unfortunately not provided. The authors' conclusion that English language assessment did not correlate significantly with averaged grades at the end of the second year is a perplexing one, as they cite a Spearman correlation coefficient of 0.28 and a probability value of 0.046. This author would have

considered a weak but significant correlation to have been found on the basis of these statistics.

Ahmed, Ahmed and Al-Jouhari (1988) conducted research to investigate the factors predictive of the academic performance of 42 graduates from the University of Kuwait Faculty of Medicine. With the majority of students speaking Arabic as their first language and English being the language of instruction at the University of Kuwait, the authors assessed English language proficiency through performance on a 17 hour English subject taught early in the medical degree. English language proficiency was found to be moderately to highly correlated with premedical, pre-clinical and clinical components of the course, yielding coefficients of 0.90, 0.60 and 0.51 respectively. Though demonstrating language to be a predictor of academic and clinical performance, Ahmed, Ahmed and Al-Jouhari (1988) also interpreted their findings as being indicative of student problems associated with facility in the English language. They wisely suggested that a 17 hour course in English may not be sufficient in a medical school where the majority of students have English as a second language and who have not previously been educated in English.

Medical students in a study by Alfayez, Strand and Carline (1990) were required to study Medicine in English at King Abdulaziz University in Jeddah, Saudi Arabia although this was not their first language. This study sought to ascertain whether English language proficiency, assessed through secondary school English grades and self-reported ability to read, speak and understand English, was predictive of academic achievement in a cohort of 153 fifth and sixth year students. Females performed better in English at secondary school as compared to males, and also rated themselves as more competent in English than did their male counterparts. Secondary school English grades were not related to either basic sciences or overall performance at medical school, but were weakly related to performance in the clinical sciences ($r=-0.05$) when students were pooled together. When analysed according to gender, English grades accounted for more of the variance for male students than for females. The opposite pattern emerged when considering self-reported difficulties

with the English language, with low, negative and significant correlations on overall performance, and performance in basic and clinical sciences for females being stronger than the correlations for males. The authors do not explore this finding, but this author suggests that perhaps this is a reflection of more realistic perceptions of language ability on the part of female students, although this cannot be substantiated on the basis of the data presented in the article. The authors do conclude, however, that regardless of gender, difficulties with English adversely affected performance, and they put forward the suggestion that students' English skills may require greater attention either prior to entry to Medicine, or shortly after commencing university. They reject the idea of excluding students from their course on the basis of English language proficiency, considering it to be an unrealistic option in their country, where the language of instruction at tertiary levels of education is not the language of the general population.

Keats (1972) stated that the research on English language ability and academic success across a number of different courses resulted in "tenuous and contradictory findings" (p 106). Gray (1972, in Bochner and Wicks, 1972) also noted that screening for language ability did not provide a good indication of how well overseas students would cope with the academic demands of an Australian university. In 1987, Graham reviewed studies up to that year which found either no relationship, mixed or definite relationships between English language proficiency and academic success in various disciplines. She concluded that with so many conflicting findings, in addition to the number of methodological problems associated with making direct comparisons between studies, the area of investigation between language ability and teriary performance is complex and "murky indeed" (p 506). McGowan and Cargill (1994) writing from the University of Adelaide have also stated that controversy exists as to whether scores on language proficiency tests are sufficient or appropriate to predict academic success. Thus the summation made by Keats in 1972 seems still to apply;

> "whereas good ability in English will contribute to success [in Australian university courses], lack of it does not necessarily prevent success . . . ." (p 106).

(i). Studies that have originated from Australian medical schools.

A number of Australian researchers within medical education have been interested in considering the contribution of secondary school performance to the prediction of academic and clinical success. Lipton, Huxham and Hamilton (1984) compared the grades of students enrolled in Medicine at the University of Queensland with their Australian secondary school performance. They found that over the six years of the medical degree, secondary school English had significant but low correlations of 0.24 and 0.28 with subjects which could be grouped under the rubric of either "paraclinical" or "clinical" sciences respectively. The umbrella term "paraclinical sciences" included Psychology, Applied physiology and Biochemistry, Neuroanatomy, Neurobiology, Pharmacology, Microbiology, Parasitology, Human Growth and Development and Pathology (Huxham, Lipton and Hamilton, 1980). "Clinical sciences", as classified by Huxham, Lipton and Hamilton (1980) were Medicine, Obstetrics and Gynaecology, Surgery, Mental Health and Social and Preventative Medicine. Although the authors do not indicate this, it could be assumed that all the students were from English speaking backgrounds, since all completed secondary school in Australia and had taken English as a matriculation subject.

In a second study Lipton, Huxham and Hamilton (1988) collected data from four cohorts of incoming medical students to the University of Queensland, to investigate the utility of secondary school results as predictors of achievement during the six years of the medical degree. As in the previous research (Lipton, Huxham and Hamilton, 1984) all students included in the sample had matriculated in Australia and specifically in Queensland and all had studied English in their final year of secondary schooling. Students who had not completed their secondary education in that state, or who had come from outside of Australia, a group which totalled 13% of the incoming student cohorts, were excluded. Lipton, Huxham and Hamilton (1988) found that English was a significant predictor of success in clinical performance and in "psychosocial" subjects, such as Medical Psychology, Sociology, Psychiatry, Community Practice and Counselling. Thus, significant correlations between performance in Medicine and the English score for such a homogeneous group make these findings appear quite robust. In both papers the authors

stressed, however, that secondary school performance alone is not a sufficient predictor of academic or clinical success in medical school, and suggested the use of psychometric and personality trait measures in addition to cognitive assessments in selecting which students should be offered a place in the medical course, a position that has also been advocated by Huxham, Lipton and Hamilton (1980), Powis, Neame, Bristow and Murphy (1988) and Neame, Powis and Bristow (1992).

Tutton and Wigg (1990) examined the predictive validity for pre-clinical performance of the secondary education subjects taken and gender, of students enrolled in the Faculty of Medicine at Monash University in Victoria between the years 1979 and 1981. The subject "English Expression" was a compulsory one for consideration for entry into the medical course at that time, and so therefore it may be assumed that the sample was fairly homogeneous with regards to language ability. The authors analysed the data according to the subjects taken in the matriculation year, rather than according to performance in any of the subjects themselves. Since all of the students had taken English, no influence of this subject could be reported. Student gender was not found to be a factor in pre-clinical academic performance.

Neame, Powis and Bristow (1992) from the University of Newcastle in New South Wales investigated the criteria for medical school selection procedures in Australian medical schools as predictors of academic performance. In this study "performance" in the course was assessed by considering the extremes of failure to complete the course as opposed to an award of graduation with honours, in addition to fulfilling the requirements of the course to graduate. Whilst no subjects were found to be associated with outcome on graduation, failure to have studied English was related to premature withdrawal from the course.

A retrospective study (Parker, 1993) of the academic success of final year medical students at the University of New South Wales, including approximately 30% from a language background other than English, found that better performance (i.e. higher

averaged grade) was related to younger age, English as the first language, completing the final year of secondary schooling in Australia, having a higher High School Certificate (Matriculation) English score, leadership responsibilities in Year 12, and participation in a debating team in Year 12. The author cautioned that medical educators and administrators should curb any zeal for attributing medical school success primarily to language background and language proficiency at the expense of other variables.

Rolfe, Pearson, Powis and Smith (1995) attempted to identify the selection criteria for admission into the University of Newcastle's medical school that predicted intern performance in the first year after graduation. They found support for the hypothesis that interns whose academic background prior to entry to medicine was stronger in the humanities (including English) or in a combination of sciences and humanities, rather than merely in the sciences alone, were rated more favourably in terms of clinical performance by their supervisors.

An important methodological point should be made concerning the language backgrounds and English language proficiency of the students in Neame, Powis and Bristow's (1992) and Rolfe, Pearson, Powis and Smith's (1995) studies. English as a secondary school subject was not a prerequisite for entry into the course and although they do not mention these variables, it must be assumed that all students had a high level of English language proficiency, since all students had to fulfil several "language intensive" requirements in order to gain a place at the University of Newcastle medical school. These included not only achieving a minimum secondary school performance in the top 10% of New South Wales matriculants, or its equivalent, but also the satisfactory completion of a battery of psychometric tests and a structured interview (Powis, Neame, Bristow and Murphy, 1988). Some of the items on the psychometric tests are analogous to the items comprising the Graduate Australian Medical School Admissions Test (GAMSAT) (Aldous, Leeder, Price, Sefton and Teubner, 1997; Rolfe and Powis, 1997), which itself is demanding in terms of the sophisticated level of English language ability required (author's personal opinion). It may well be that a disproportionate number of students

accepted into the University of Newcastle medical school also came from English speaking background families.

All of the research from Australian medical schools reviewed thus far has been based on student populations with English language skills which can be assumed to be quite similar. Studies which are exceptions to this, in that they are based upon more heterogeneous groups are those of Farnill and Hayes (1996a, 1996b), Hayes, Farnill and Sefton (1994), Hayes and Farnill (1994), Hayes and Farnill (1993a, 1993b) and Rolfe, Andren, Pearson, Hensley and Gordon (1995).

As has been discussed in the preceding section, many students who have entered the University of Sydney medical school have done so with poor English language skills (Sefton, 1990). Hayes and Farnill (1993a, 1993b) and Hayes, Farnill and Sefton (1994) reported that Woodcock Language Proficiency Battery rankings and STAL scores showed low to moderate and significant correlations with almost all first and second year subjects across two cohorts of students. STAL correlations ranged from 0.12 for first year Chemistry to 0.46 for the first year subjects Introduction to Medical Science and History of Medicine. Hayes, Farnill and Sefton (1994) also reported that overall, the final secondary school English score was the best predictor of academic achievement in first year Medicine, although the range of correlations for each of the nine subjects was not provided. Farnill and Hayes (1996a) calculated weighted average grades for first year Medicine subjects, finding a significant correlation of 0.43 with the AUSTEST score. Weighted average grades for second and third year subjects and the AUSTEST score were also significantly correlated in the vicinity of 0.40 (Farnill and Hayes, 1996b).

Hayes and Farnill (1994) have stressed that these findings are correlational in nature and as such should not be used alone as conclusive evidence of the importance of the role of English language proficiency for academic success. Similarly, Farnill and Hayes (1996a) pointed out that low to moderate correlation coefficients do not account for high percentages of the variance, although they stated that this is to be expected, given the

multiple factors which may affect academic performance, including motivation, general ability and the unreliability of marking procedures.

In a cross-sectional survey Rolfe, Andren, Pearson, Hensley and Gordon (1995) evaluated the clinical performance of 486 interns in New South Wales via a team supervisor rating scale. Interns had graduated from a New South Wales University or other Australian institutions, from New Zealand, or from another overseas country. This latter group were rated significantly lower than any other group of interns on all of 14 criteria, including relationships with patients and families, relationships with other professionals and overall term performance. Rolfe, Andren, Pearson, Hensley and Gordon (1995) postulated that either the interns' English language competencies or racial discrimination on the part of the raters may have produced these findings, but were unable to draw firm conclusions. They called for both further investigation of the impact of overseas qualified doctors' English language proficiency upon clinical performance in Australia and also longitudinal research designs for the monitoring and evaluation of skills.

(ii). Studies which have considered English language proficiency and performance in specific subjects.

All of the research regarding the predictive validity of English language proficiency and academic performance reviewed thus far has in common a methodology that involves analysing performance in subjects as components of a whole course. That is, the final grade for the subject for a given year, or a classification on the basis of the achieved overall grade is used as the indicator of success or otherwise. Very few authors have attempted to consider more closely the individual units of the subjects that comprise an entire medical course. It may be that by refining the investigation in this way, less contradictory findings may emerge in research. For example, it might be relatively meaningless to compare the performance of students in two subjects which have completely different assessment methods. Students who fair poorly in subjects which rely heavily upon one end of year written examination but excel in weekly assignments or in practical work may do so not because of language skills, but for a number of reasons,

including past learning experiences and the level of tutorial assistance provided. Even within one subject, students may perform better on some assessment measures than others, which again, may or may not be contingent upon their level of proficiency in English. To conduct studies which look carefully at performance within each subject as well as between subjects may be prohibitive unless the investigator has an intimate knowledge of the subject or subjects in question. This may be one reason why so few have chosen this as an area of interest. Two studies which are known to have addressed the issue, one of which is, unfortunately, unpublished, and another recent publication, are discussed below.

In an unpublished report from the University of Adelaide Department of Psychiatry, Winefield (1993) administered the Australian Council for Educational Research Word Knowledge Test (WKT) to 124 medical, 48 dental and 10 health science first year students. One-hundred and eight (59.3%) of the cohort reported speaking a language other than English at home. These students performed significantly poorer on the test than students from English speaking backgrounds. No relationship was found between WKT scores and academic performance in the first year subject Behavioural Science, a result attributed to the efficacy of subject-based teaching interventions specifically designed for students from non-English speaking backgrounds in compensating for any possible disadvantages due to language difficulties. The role of this teaching intervention had not been empirically tested at that time, however, and is the focus of Study VII of this thesis. Winefield (1993) did not examine performance on the WKT as a predictor of subjects other than first year Behavioural Science. The concurrent validity of the WKT and the STAL as indicators of English language proficiency is the focus of Study II.

Lucas, Lenstrup, Prinz, Williamson, Yip and Tipoe (1997) considered language as a barrier to learning in first year gross anatomy at the University of Hong Kong, where all students must learn in English, usually their second, third or even fourth language. The authors assessed the quality of English language proficiency by having one rater categorize the scripts of a short essay in anatomy as either "good", "intermediate" or "poor", with the

emphasis upon clarity of meaning rather than upon grammatical correctness. Frequency distributions for these three categorizations were not provided. Quantity of writing was measured by using a ruler, i.e. the further down the page the words went, the greater the quantity of language proficiency. This seems an odd method, given firstly that it does not control for the size of the script or spacing of words, as the authors themselves acknowledge. It also fails to take into account the fact that students may simply repeat similar material several times in one answer and use the same words multiple times, making the rationale behind this type of quantification unsound. Nevertheless, Lucas, Lenstrup, Prinz, Williamson, Yip and Tipoe (1997) found significant correlations of 0.38 and 0.56 with a classroom anatomy test and the quality and quantity of English respectively. A significant correlation (the value of which is not given in the paper) was also found between students' grades in a 16 hour English course taken on admission to the medical course and the end of year final examination in anatomy (which included other subjects within anatomy as well as gross anatomy). The researchers concluded that "language is a crucial barrier to success" (p 85) in this particular subject. Although this appears to be so on the basis of this study, a more rigorous methodology and further details about their findings would make their argument more convincing.

II.VI. "Standardized Patient" research.

(i). Definition of the term "Standardized Patient".

A considerable literature exists in medical education on the use of "standardized patients", also referred to as "patient instructors", "simulated patients", or "programmed patients" (Swanson and Stillman, 1990; Stillman, Regan, Philbin and Haley, 1990; Barrows, 1993). Standardized patients have been defined as individuals who do not hold qualifications in medicine, either with or without "real" symptoms, who are taught by medical educators to act as patients, to assist in teaching, and to evaluate the clinical and interpersonal skills of medical students, (Stillman, Ruggill, Rutala and Sabers, 1980; Stillman, Burpeau-Di Gregorio, Nicholson, Sabers and Stillman, 1983), resident medical officers (Stillman, Swanson, Smee, Stillman, Ebert, Emmel, Caslowitz, Greene, Hamolsky, Hatem, Levenson, Levin, Levinson, Ley, Morgan, Parrino, Robinson and Willms, 1986), and foreign medical graduates (Friedman, Sutnick, Stillman, Norcini, Anderson, Williams, Henning and Reeves, 1991). Their use has been most popular in the United States of America, but they have also been employed in a number of other countries, including the People's Republic of China (Stillman and Sawyer, 1992), the Netherlands (Pieters, Touw-Otten and De Melker, 1994), the United Kingdom (McAvoy, 1988), Canada (Chugh, Dillmann, Kurtz, Lockyer and Parboosingh, 1993; Hodges, Turnbull, Cohen, Bienenstock and Norman, 1996), New Zealand (Thompson, 1990) and Australia (Newble, Hoare and Elmslie, 1981; and see later in this chapter).

Stillman, Regan, Philbin and Haley (1990) noted that a lack of consensus existed in the definition of the term "standardized patient", which has hindered a uniform approach to training across different institutions. Barrows (1993) attempted to clarify the meaning of terminology, commonly (and sometimes incorrectly), used in the literature. "Role-playing" he defines as an activity that requires an individual to improvise within a role without specific guidelines, apart from those dictated by the role itself. "Patient instructors", a term which Barrows attributes to Stillman, refers to actual patients who have been educated about their illness and trained to give medical students feedback about performance in history taking and physical examination techniques (Stillman, Ruggill,

Rutala and Sabers, 1980; Stillman, Rutala, Nicholson, Sabers and Stillman, 1982). Barrows (1993) differentiates a "standardized patient" from a "simulated patient", although he believes that the two terms can be used interchangeably in specific circumstances. He explains that a "simulated patient" is one who has been trained to accurately portray the history and physical symptoms of an actual patient, whereas a "standardized patient" is *either* a simulated or actual patient trained to present their own illness in a standardized manner that does not alter across consultations.

(ii). A history of the use of standardized patients in the United States.

Standardized patients have been employed in United States medical curricula since Barrows and Abrahamson (1964) first published the details of their teaching innovation. Later Barrows (1968) and Barrows, Williams and Moy (1987) outlined the advantages of using individuals to simulate symptoms, such as those typical of neurological diseases (Barrows, 1971), in place of true patients. The benefits and value of standardized patients in medical education were reiterated by him nearly 30 years later (Barrows, 1993). Howard Barrows can be considered the pioneer of the use of standardized patients for the teaching of gross anatomy, clinical neurology and neurosurgical principles (Barrows, 1968, 1991), and he lay the foundations for the use of such patients in the clinical interview situation.

Since Barrows first conceived of the standardized patient as an educational technique they have been most commonly used within the context of an objective structured clinical examination (OSCE), which refers to a framework for assessing the clinical skills of medical students and residents (Hart, 1986; Newble and Swanson, 1988), first introduced by Ronald Harden and his colleagues (Harden, Stevenson, Downie and Wilson, 1975). An OSCE is a structured clinical examination involving multiple stations around which students rotate to carry out a number of clinical tasks (Petrusa, Blackwell, Rogers, Saydjari, Parcel and Guckian, 1987). A station can be considered to be a single task, such as taking a patient history, performing a physical examination, writing of case-notes or interpreting an x-ray; stations may vary in duration from around four minutes to

an hour (Swanson and Stillman, 1990; van der Vleuten and Swanson, 1990). The OSCE has been judged as more valid, reliable and practical when compared to alternative, "traditional" subjective examination methods which involve several clinicians assessing a student's competence across several different patients (Harden and Gleeson, 1979). The OSCE is "objective", because numerous variables can be controlled, including the patients, their symptoms and the examiners, and is "structured" because all examinees can be scored on the same predetermined checklists and rating scales, allowing for meaningful comparisons between candidates to be made.

Although standardized patients have been employed in place of "real" patients during OSCEs to ensure objectivity, some researchers (Frazer and Miller, 1977; Anderson and Meyer, 1978; Kretzschmar, 1978, in Stillman, Ruggill, Rutala and Sabers, 1980) have included teaching as another function of the standardized patient, by training them to give feedback to examinees about their clinical performance and ways in which it can be improved. In 1978 Paula Stillman and her colleagues extended the role of standardized patient further still, to include that of evaluator, by training them in the use of objective assessment instruments (Stillman, Ruggill and Sabers, 1978), such as the evaluation of second and third year medical students' clinical skills in cardiovascular and pulmonary examinations (Stillman, Ruggill, Rutala and Sabers, 1980). Norcini, Stillman, Sutnick, Regan, Haley, Williams and Friedman (1993) summarised the advantages of using standardized patients as being three-fold; first, there are no difficulties in acquiring patients as may be the case if unwell patients from hospital wards are used, second, an objective rating system can be applied uniformly to all examinees, and third, all examinees can be exposed to the same experience, thereby enabling comparisons across performances to be more meaningful.

In 1983 Stillman and her colleagues reported on their six years of experience in the use of standardized patients to teach and evaluate interviewing skills at the University of Arizona College of Medicine (Stillman, Burpeau-Di Gregorio, Nicholson, Sabers and Stillman, 1983). They concluded that standardized patients were an effective method of

teaching and evaluating such skills in undergraduate medical students. Swanson and Stillman (1990) have argued that whilst standardized patients cannot replace real patients they should be viewed as an extremely valuable teaching and learning resource.

By 1993 interest in and research on the use of standardized patients in medical education in the United States was such that the Association of American Medical Colleges (AAMC) decided to publish a special issue of their journal *Academic Medicine*, devoted to the proceedings of a conference on the use of standardized patients in the teaching and evaluation of clinical skills. In his conference summary Miller (1993) concluded that participants were of the consensus that standardized patients should be further incorporated into medical education programmes, enriched by future research. Swartz (1992) reported on ASPIRE (Advancement of Standardized Patients In Research and Education), a body promoting collaboration and cooperation between educators and researchers interested in the use of standardized patients.

Anderson, Stillman and Wang (1994) replicated a survey originally conducted in 1989 (Stillman, Regan, Philbin and Haley, 1990) which sought to establish the incidence of standardized patient use for teaching and evaluating clinical skills in United States and Canadian medical schools. Anderson, Stillman and Wang (1994) found that the use of standardized patients had increased since the 1989 survey from approximately 66% of North American medical schools using such patients to approximately 80% of these schools using them by 1993.

(iii). The use of standardized patients in Australia.

The literature on the use of standardized patients in Australia is not as considerable as that generated from the United States. The main contributors to Australian research on standardized patients have been Robin Winkler, who was actively working in this area in the 1970's, and Jill Gordon and Robert Sanson-Fisher, who commenced their investigation of standardized patients as evaluators of medical students' and interns' clinical skills in the 1980's.

It is possible to divide the use of standardized patients into two main categories. In one group of studies the use of such patients is overt, where the medical student or graduate is aware that they are working with a standardized patient. Australian studies have favoured a covert methodology, where the student or graduate carries out various aspects of patient care without an awareness that the patient is not "real", in the tradition of Rosenhan's (1973) seminal study of "pseudopatients" conducted in the United States.

Rosenhan (1973) had eight research assistants gain admission into 12 psychiatric hospitals, playing the role of a patient, in order to observe and record their observations of the care received by themselves and fellow patients. The pseudopatient case histories were standardized to the extent that in seeking admission to the hospital all reported experiencing auditory hallucinations; following admission all pseudopatients behaved "normally", exhibiting no psychotic symptomatology. Information about the pseudopatient's life history and circumstances was consistent with their actual life, apart from a change of name and employment. Rosenhan (1973) reported that with an average length of stay of 19 days, none of the pseudopatients were identified as such by nursing or medical staff, although genuine patients were suspicious in some instances.

To further his investigation Rosenhan (1973) informed the staff of a psychiatric ward (not involved in the previous study) that in the subsequent three months one or more pseudopatients would attempt to gain admittance into the hospital. Although no pseudopatients actually approached the hospital, forty-one patients were alleged to have done so. Rosenhan's purpose was to illustrate the impact of type 1 (false negative) and type 2 (false positive) errors in diagnosis. However, his findings could also be interpreted as lending support for the premise that the hospital staff were unable to distinguish between "real" patients and pseudopatients.

Winkler (1974) replicated Rosenhan's (1973) study, using a pool of nine pseudopatients who gained admission to the psychiatric wards of suburban Sydney hospitals. He extended Rosenhan's (1973) work by providing the pseudopatients with a

UNIVERSITY OF ADELAIDE LIBRARY

standardized and extensive case history, and by employing trainee clinical psychologists to act as pseudopatients, with the aim of increasing empathy toward their future clients. At the time of publication, Winkler (1974) was in the process of contacting the hospitals involved in order to provide them with feedback, based on pseudopatient observations of hospital and staff functioning.

Owen and Winkler (1974) used ten pseudopatients to collect information during two consultations each with a sample of 25 general practitioners. Using a standardized history and life situation, exhibiting the same symptoms and provided with standard questions to ask the doctor and answers to give, the pseudopatients presented with mild depression of psychosocial origin. In none of the fifty consultations did the general practitioner suspect that the patient was not "real". The authors stressed that the observations made by pseudopatients were descriptions, and that the role of the pseudopatient did not involve making any evaluations about the doctors' quality of care. Evaluations were made by comparing the observations provided by the pseudopatients with the standards of care for psychotropic drug use recommended by the National Health and Medical Research Council (NHMRC), research statements on drug usage, and the sampled general practitioners own standards, which were collected by questionnaire.

In concluding their paper, Owen and Winkler (1974) called for the establishment of an Australian body to instigate, supervise and publish systematic pseudopatient observations on the provision of health care and the training of health professionals. To date, such a body is yet to be established.

In an Editorial for the Medical Journal of Australia McConaghy (1974) noted the potential value of employing pseudopatients as a means of evaluating the services provided by medical practitioners. However, he strongly criticised the work of Rosenhan (1973), and the two studies by Owen and Winkler (1974) and Winkler (1974), and argued that well designed research studies were sorely needed for the pseudopatient technique to be of scientific value and merit. McConaghy's primary concern stemmed from his

perception that these authors were negatively biased against medical practitioners and the medical profession, and in particular, psychiatrists and psychiatry. Over twenty years later Pololi (1995) and Stillman (1993) were both to echo the similar view that standardized patients (and by implication, the individuals who use them for teaching and research) should not hold unfavourable views of medicine. In an Editorial of the British Medical Journal (1974) the pseudopatient studies of Rosenhan (1973), Owen and Winkler (1974) and Winkler (1974) were strongly criticised, in part because they were seen to have misrepresented the conceptualisation of the standardized patient. It was argued that the pseudopatients in these three studies differed substantially from standardized patients because they had not been adequately trained in simulation, and also had lacked objectivity. Barrows (1975), the pioneer of the use of standardized patients, concurred, stating that;

> "simulated patients are carefully trained by health professionals to simulate accurately an actual patient with high fidelity for purposes of teaching and evaluation. They should not be confused with pseudopatients, who are untrained patient impostors who gain entrance into a health care system for purpose of evaluation" (p 625).

Following Winkler's work, but with a less controversial focus, is the work of Gordon, Sanson-Fisher and their colleagues, who have used covert standardized patients (which they term "simulated" patients), but with the informed consent of their subjects. Sanson-Fisher and Poole (1980) investigated whether second year undergraduate medical students at the University of Western Australia behaved differently when interacting with simulated patients as compared with other patients. They concluded that there were no significant differences in the levels of empathy expressed by students to simulated and "real" patients, and that students could not distinguish between the two types of patient. Gordon, Sanson-Fisher and Saunders (1988) conducted a similar, two year study in Newcastle, New South Wales, concluding that interns could not easily differentiate between simulated patients and "real" patients in a casualty department.

In the first of a further three studies in New South Wales, Gordon, Saunders and Sanson-Fisher (1989) conducted an observational study in which 61 interns were videotaped during consultations with three simulated patients to evaluate clinical

competence, communication skills and responses to opportunities for preventive care. A further observational study (Gordon, Sanson-Fisher and Saunders, 1990) reaffirmed the previous finding (Gordon, Sanson-Fisher and Saunders, 1988) that interns were unlikely to correctly identify simulated patients. In the third study Gordon, Saunders, Hennrikus and Sanson-Fisher (1992) measured changes in 28 interns' performance in clinical and communication skills, by comparing three videotaped consultations with covert simulated patients conducted at the commencement and conclusion of the intern year. In this study the interns' performances were rated by an expert panel which did not include the simulated patient. The simulated patient's role did not include any form of evaluation or the provision of feedback.

Other researchers in Australia have employed standardized patients without providing further details other than that they were incorporated into the methodology. Farmer and Prideaux (1994) utilized standardized patient encounters to assess postgraduate general practice trainees' patient centred medical communication skills for the South Australian branch of the Royal Australian College of General Practitioners. Newble, Hoare and Elmslie (1981) and Swanson and Norcini (1989) refer to the use of standardized patients in a University of Adelaide study investigating the reliability and validity of an objective structured clinical examination procedure to assess the clinical competence of final year medical students, but do not elaborate. Edwards and Devitt (1988) outlined the use of standardized patients for interdisciplinary clinical education at the Lincoln School of Health Sciences at La Trobe University in Melbourne, a teaching initiative introduced there in 1982.

(iv). Selection and training of standardized patients.

A survey conducted by Stillman, Regan, Philbin and Haley (1990) asked a number of questions about the ways in which standardized patients were used in the curriculum of American medical schools, including methods of training. It was found that training was conducted by both medical and non-medical faculty staff, occurring in either individual or group training sessions, in some instances with videotapes being the main teaching tool.

In some cases, trainees directly observed fully fledged standardized patients, or observed actual patients, whilst in others only teaching manuals were used. The average duration of training for standardized patients to evaluate communication skills, including feedback to the student and teaching, was ten hours. Although the authors expressed enthusiasm about the widespread use of standardized patients as a teaching resource, they noted that uniformity in training procedures and clarification of terms is required.

The duration of training varies considerably, depending upon the problem to be simulated and whether the standardized patient must provide feedback or evaluation. In their 1993 survey Anderson, Stillman and Wang (1994) reported that in Northern American medical schools times range from one to three hours, four to six hours and eight to twelve hours training standardized patients to teach and/or evaluate interviewing skills. Pololi (1995) reported that the training of a standardized patient took eight hours on average, with an additional ten hours of training (eighteen hours in total) if the standardized patient was required to provide written and verbal feedback as well as portraying symptoms. Ten hours of training was provided to persons with complicated chronic diseases to become standardized patients for assessing Internal Medicine House Officers' clinical and psychosocial skills by Calhoun, Woolliscroft and Ten Haken (1987). Stillman, Regan, Swanson, Case, McCahan, Feinblatt, Smith, Willms and Nelson (1990) taught standardized patients to complete checklists and rate students in about 15 hours, using both group and individual training sessions. Stillman and Swanson (1987) and Stillman, Ruggill and Sabers (1978) recommended 15 to 25 hours of intensive training for a standardized patient to fulfil the roles of patient, teacher and evaluator. Gordon, Sanson-Fisher and Saunders (1990) and Gordon, Sanson-Fisher and Saunders (1988) reported that the training of standardized patients to portray a clinical problem (with no requirement to provide feedback) was completed in four hours. Nieman and Thomas (1987) devoted two 90 minute sessions for training of standardized patients, who were not required to provide feedback. Tamblyn, Klass, Schnabl and Kopelow (1991) spent one to three hours divided across two or three sessions to train standardized patients with previous experience to present one clinical problem. Barrows (1993) argued that someone who has never been a

standardized patient before can be trained in two to three hours, although again, this referred to situations where the patient was not required to contribute to evaluations or complete checklists.

Barrows (1971) considered that almost anyone can be trained to act as a standardized patient, providing that they possess interest, intelligence, motivation and a flair for acting. Enthusiasm was identified by Nieman and Thomas (1987) as an important attribute for trainee standardized patients. Stillman (1993) has outlined her selection criteria for choosing individuals who would be amenable to training as a standardized patient. She states that they should possess the attributes of intelligence, an interest in medical education, good teaching skills and good communication skills. Pololi (1995) sought men and women with a good memory, physical stamina, and the capacity to articulate emotions. Attempts to select standardized patients through testing with personality inventories as an adjunct to the selection process have not been useful (O'Connell and Thayer-Doyle, 1993). Stillman (1993) and Pololi (1995) both concurred on the point first made by Barrows (1971) that potential standardized patients must not have negative attitudes toward medical practitioners or the medical profession. Davies (1989) has reiterated the need for standardized patients to be able to communicate well, adding that suitable candidates for training should have a genuine interest in medicine, enjoy working with students and have the ability to ad-lib where appropriate.

It has been recommended that the trainer of standardized patients have some background in medicine, or work very closely with someone who does (Stillman, 1993). Where the standardized patients are to be employed as part of a research project, Stillman (1993) advises that the trainer also have a suitable research background.

Kachur, Green, Bruun, Philbin, Addis, Simonton, Ward, Stillman and Sutnick (1992) surveyed 64 standardized patients involved in a study organised by the Educational Commission for Foreign Medical Graduates (ECFMG) (Sutnick, Stillman, Norcini, Friedman, Regan, Williams, Kachur, Haggerty and Wilson, 1992) to ascertain the

characteristics of these individuals, whose task included the evaluation of spoken English. They found that the majority held a tertiary degree, had prior teaching experience, had a health care professional in their family and believed that being a standardized patient should be viewed as a profession. Many had previous acting experience; some researchers have argued that professional actors are the most preferable candidates for standardized patient training (Cohen, Gromoff and Swartz, 1992), although cost is a consideration here, as the fees commanded by professional actors are relatively high by comparison with amateur actors (Davies, 1989). The primary motivations for acting as a standardized patient were reported as improving patient care, helping medical education and earning money.

(v). Reliability and validity of standardized patient ratings.

In considering any ratings or test scores the issues of reliability and validity must be addressed, and the ratings provided by standardized patients on medical students' performance are no exception in this regard (Ainsworth, Rogers, Markus, Dorsey, Blackwell and Petrusa, 1991). The following consideration of reliability and validity owes much to Anastasi (1988) and assumes that the reader has a basic knowledge of statistical principles.

Reliability is a measure of the consistency of the same subject's performance on the same or similar test or group of ratings on different occasions. As reliability considers the extent to which two sets of independent scores are consistent with one another, it can be expressed statistically in terms of a correlation coefficient (r). Correlation coefficients, which may also be referred to as reliability coefficients, demonstrate the degree of correspondence between two sets of scores and range from perfect correlations of +1.00 to -1.00.

Reliability takes several main forms. "Test-retest reliability" involves repeating the same test twice for a subject and calculating the reliability coefficient. Because the test taker may have recalled the items on the test, or improved with practice, thus artificially

inflating the reliability coefficient, "alternate-form reliability" may be considered. This involves testing the same person on two different occasions in close temporal proximity with parallel forms of the same test. However, equivalent versions of the same test are not always practical or available, and in this case "split-half reliability" allows the researcher to obtain two scores for each subject after administering a single test once. This is done by dividing the test into as equivalent halves as possible and correlating the resultant half scores. Since only one test is being considered, split-half reliability is also referred to as "internal consistency". It has been argued that internal consistency should be interpreted with care for criterion-referenced tests (designed to assess knowledge or skills), and that Cronbach's coefficient alpha be applied to the data (Newble, Hoare and Elmslie, 1981). "Inter-item consistency" similarly requires one test administration, and assesses subjects' consistency of responses to the test items. Inter-item consistency can be calculated with the Kuder-Richardson formulae, which examines performance on each item for tests which are scored as "yes" and "no" or "correct" and "incorrect", for example. Where items are multiple-scored (such as on a scale from "always", "sometimes" to "never") inter-item consistency is calculated with coefficient alpha, also known as Cronbach's alpha. Inter-item consistency can also be calculated by correlating items on the scale with each other. Finally, some tests are less objective than others in their scoring criteria and may be open to the judgements and biases of the scorer (Kazdin, 1977). In these cases, tests can be independently scored by two examiners and their results correlated to establish "inter-rater reliability". Kendall's coefficient of concordance *(W)* can be calculated to establish agreement between raters, whilst Pearson or Spearman *rho* can be used to establish rater consistency (Newble, Hoare and Elmslie, 1981).

Validity refers to whether the test is measuring what it purports to measure. A test cannot be considered valid in the absence of reliability (Newble, Hoare and Elmslie, 1981). Test validity is primarily assessed by examining relationships between performance on the test in question and independent, observable facts related to the skill or behaviour being studied. "Content-related validation", or "content validity" asks firstly whether the content of the test items comprises a representative sample of the skills or knowledge or behaviour

(or whatever) that the researcher has specified are to be investigated and secondly, if the subjects' performance on the test has been relatively unencumbered by irrelevant variables. Anastasi (1988) has stressed that content validity should not be confused with "face validity", which is not true validity at all in the methodological sense, but is the extent to which the test items seem relevant for the task to untrained observers. Face validity and true validity are not necessarily related to each other and one cannot be taken as an indicator of the other.

"Criterion-related validation" assesses the extent to which a test can predict a subject's performance on another, independent measure. Two types of criterion-related validation assessments can be made. "Concurrent validity" involves correlating the results of one test with those of another test where both claim to be measuring the same skill, knowledge or behaviour. The two test scores may have been obtained at approximately the same time or after a specified interval. Anastasi (1988) makes explicit that unless the second test is a shortened or simplified version of the first, correlations between the two tests cannot be used to determine criterion-related validation. Other authors consider that concurrent validity can be established in this manner providing the criterion test is regarded in the literature as an accepted measure (Newble, Hoare and Elmslie, 1981; Ingram, 1984; Winefield and Peay, 1991). "Predictive validity" considers the predictive power of the test score to any criterion situation that is argued as relevant by the researcher. For example, indices of academic achievement are frequently employed as criteria for determining the predictive validity of intelligence tests.

"Construct-related validation" involves determining the extent to which a test measures a theoretical construct, such as verbal fluency. If a test has construct validity its scores must enable inferences to be made about an underlying trait or traits (Skakun, 1986). Construct validity can be investigated via a number of techniques, including correlations with other, earlier tests of both the same and independent skills, knowledge or behaviours. If the test under examination has sound construct validity the correlations in the former case should be high and in the latter case low, although this procedure in itself

would not ensure validity. This method is consistent with convergent and discriminant validation techniques, which aim to show that the test scores not only correlate with variables with which it theoretically should, but also does not correlate with variables from which it theoretically should differ. To further explore construct validity of a test the researcher may conduct tests of internal consistency, to determine the homogeneity of the items or subtests. It is important to remember, however, that internal consistency is based upon the use of the test score(s) as the criterion and as such, uses no external data to validate the test. Therefore, although a useful procedure, alone it cannot be employed to understand what the test is measuring. Another means by which construct validity can be studied is by comparing pre and post test scores after a relevant intervention. Test scores should be found to be lower prior to the intervention and higher afterwards. If this is not the case, either the test lacks construct validity, or there is something amiss with the intervention. It is important that more than one of the above methods is used when exploring construct validity (Skakun, 1986) as an accumulation of evidence strengthens confidence in any conclusions that can be drawn if this is done, providing that the statistical procedures employed are sound (Skakun, 1986).

Van der Vleuten and Swanson (1990) compiled a comprehensive review of studies which have investigated the psychometric properties of objective structured clinical skills examinations employing standardized patients. They reported that interrater reliability is generally good, and therefore only one rater per station is generally necessary. They argued that either standardized patients or faculty staff are equally suitable to serve as raters providing that they have been given sufficient training.

The measurement consequences of having several standardized patients play the same role is an important issue (Swanson and Norcini, 1989). Van der Vleuten and Swanson (1990) argued that where multiple stations are involved, this is not problematic. However, in examinations with only a single station the effects of using different standardized patients who differ in the accuracy of their presentation must be considered.

Whether accuracy is a function of training or other factors has yet to be established and the authors considered this to be an area where research should be focused in the future.

Van der Vleuten and Swanson (1990) concluded their review by making a number of methodological recommendations. They ask that future researchers be more specific in the reporting of whether multiple standardized patients played the same role, how examinees were assigned to standardized patients and raters, the format and content of rating lists and checklists and scoring procedures. Statistically, Van der Vleuten and Swanson (1990) called for the reliability of criterion measures to be reported where appropriate.

A review of the psychometric properties of tests which utilize standardized patients for teaching and assessing clinical skills by Swanson and Stillman (1990) concluded that inter-rater reliability between standardized patients and faculty staff is generally high, although reliability from case to case is poor unless between 10 to 40 stations are tested. Reproducibility of pass/fail decisions has been demonstrated to be generally acceptable providing that the checklists or rating scales used are fairly short. Swanson and Stillman (1990) noted that very few studies have investigated the construct validity of standardized patients, by comparing the relative performances of different groups of students, such as first years compared with third years, for example and they called for further investigations in this regard. The concurrent validity of students' performances with standardized patients and other measures varies, although average coefficients have been found to be generally high. Swanson and Stillman (1990) point out that the content validity of standardized patient stations testing clinical skills may be debatable, if a clear definition of what is to be tested is not stated. Poorly designed checklists and rating scales and insufficient orientation of examinees to what is expected of them at stations are both serious threats to content validity.

Colliver and Williams (1993) summarised empirical evidence on the technicalities of the use of standardized patients in assessment. Referring substantially to the review by

Van der Vleuten and Swanson (1990), the authors concluded that evidence regarding the reliability and validity of standardized patient assessments for clinical skills would suggest that they are a most satisfactory tool for the purposes of medical education. It was noted by conference participants that reliability and validity issues are most important when the standardized patient assessments are used for summative purposes; they become less so when used solely for teaching, or for formative assessment (Miller, 1993).

A study on the use of standardized patients to evaluate the clinical skills of resident medical officers investigated the reliability and validity of such assessments for general internal medicine practice (Stillman, Swanson, Smee, Stillman, Ebert, Emmel, Caslowitz, Greene, Hamolsky, Hatem, Levenson, Levin, Levinson, Ley, Morgan, Parrino, Robinson and Willms, 1986). Interrater reliability was measured by correlating the standardized patient's ratings with those of an observer's. Intercase reliability was determined by measuring the degree to which assessments were reproduced between different standardized patient cases. Validity was measured by comparing ratings with years of residency, the academic reputation of the resident's training programme, the resident's self ratings of competence and researchers' ratings of the resident's competence, and General Internal Medicine Certifying Examination results. Stillman Swanson, Smee, Stillman, Ebert, Emmel, Caslowitz, Greene, Hamolsky, Hatem, Levenson, Levin, Levinson, Ley, Morgan, Parrino, Robinson and Willms (1986) found moderate to high interrater reliability. They reported that in order to achieve acceptable intercase reliability (a coefficient of 0.8) an examinee would need testing on six to ten standardized patient encounters in order to assess history and physical examination skills, 15 encounters for assessment of communication skills, over 25 encounters to assess differential diagnosis skills and approximately 40 standardized patient cases to assess the ability to use diagnostic studies. The authors suggested that it may be more practicable to reserve standardized patients for history taking and physical examination assessments, and examinations of communication skills, and to employ alternative methods of assessment for other skills.

Evidence of the validity of the standardized patient assessments was reported with regard to years of residency and the academic reputation of the resident's training. First year residents were consistently rated lower on all assessments than third year residents, with second year residents consistently scoring between these two groups. Similarly, residents from programmes with strong academic reputations were consistently assessed more favourably on all measures than were others. It is not clear from this paper as to whether the standardized patients were blind to the residents' level of study and place of training at the time of testing. If they were not, this would of course weaken the strength of this evidence as an argument for validity. No correlations were found when standardized patients assessments were compared to residents' self ratings. Both the researchers' ratings and Certification scores yielded low but statistically significant correlations with standardized patient assessments. Researchers' and resident self-assessments were not correlated; nor were the researchers' assessments correlated to months of training that the resident had completed. The authors concluded that these latter two findings were "troublesome" (p 769), and called for future research to scrutinise the validity of standardized patient assessments.

An aspect of validity that has been further investigated is that of "test security"; that is, whether students examined at a later testing session have an advantage over those tested earlier, by having access to knowledge about what took place in the examination from their peers (Stillman, Haley, Sutnick, Phibin, Smith, O'Donnell and Pohl, 1991). This preliminary study found that examination scores were not a function of testing date. However, the authors point out that the examinations in this study were not of a "pass/fail" nature and also that the students in their study may not have had the opportunity to converse with each other. They recommend that researchers need to consider the issues surrounding the possible influences of prior knowledge about the content and structure of the examination on test results, and of the motivation of students to obtain information about examinations which carry a heavy penalty for failure. Jolly, Newble and Chinner (1993) found some evidence for increments in students' performance through using stations in Objective Structured Clinical Examinations more than once. On the other hand

Rutala, Witzke, Leko, Fulginiti and Taylor (1991) concluded that skills demonstrated by students in Objective Structured Clinical Examinations should not change across successive testing sessions over several days, even with knowledge of patient complaints. This conclusion has also been drawn by Stillman, Regan and Swanson and the Fourth-year Performance Assessment Task Force Group (1987), Colliver, Barrows, Vu, Verhulst, Mast and Travis (1991) and Battles, Carpenter, Wagner and McIntire (1992).

A further aspect of validity that has been considered in the literature is rater bias according to the gender of the standardized patient, the student and the preceptor. Solomon, Speer, Ainsworth and DiPette (1993) found no evidence to support the hypothesis that the gender of preceptor and student might influence students' ratings. Similarly, Colliver, Vu, Marcy, Travis and Robbs (1993) found no influence of standardized patient gender upon their ratings of male and female students' interpersonal and communication skills, a conclusion further supported by Furman, Colliver and Galofre (1993). However, Rutala, Witzke, Leko and Fulginiti (1992) found evidence to show that females consistently achieved higher scores from standardized patients, a finding also supported by Stillman, Regan, Swanson and Haley (1991, in Rutala, Witzke, Leko and Fulginiti, 1992). Whether this was due to a preferential bias toward females, or a reflection of female students' better skills, could not be answered in either research study. Whilst Luketich, Colliver and Galofre (1992) found that in general no significant interaction effect was present between standardized patient and student gender, they did find that females received significantly higher ratings in communication skills than did males. Luketich, Colliver and Galofre (1992) also found that the male standardized patient in their study gave higher ratings to students regardless of gender than did the female standardized patient, although they conceded that this may have been due to the latter's greater experience and therefore confidence to award lower scores. Independent of gender, it is possible that two different standardized patients rating the same student may produce different scores (Abrahamowicz, Tamblyn, Ramsay, Klass and Kopelow, 1990). Thus, in any research which utilizes standardized patient ratings, the available demographic characteristics of all parties involved in the process should be considered, given these

inconsistent findings and the potential for bias as a result of such characteristics (van der Vleuten and Swanson, 1990).

It has been well established in the literature that medical practitioners cannot distinguish between genuine patients and standardized patients in the workplace (see for example O'Hagan, Davies and Pears, 1986; Stillman, 1986; Rethans and van Boven, 1987; Stillman, Regan, Swanson, Case, McCahan, Feinblatt, Smith, Willms and Nelson, 1990; Rethans, Drop, Sturmans and van der Vleuten, 1991; Rethans, Sturmans, Drop and van der Vleuten, 1991). Thus, standardized patients used as a teaching and research tool can be said to have high face validity (Tamblyn, Schnabl, Klass, Kopelow and Marcy, 1988; Hays, Jones, Adkins and McKain, 1990).

Pieters, Touw-Otten and De Melker (1994) explored the face and predictive validity of standardized patient encounters by comparing the consultation skills of General Practice trainees with real versus standardized patients. They concluded that behaviour differed between the two consultation types. In particular, they found that performance was usually higher in the simulated situation. They suggested that the trainees were demonstrating their maximum capabilities with standardized patients, rather than their everyday use of skills. They also found that trainees who performed poorly with standardized patients also performed less well with actual patients, although the reverse scenario, where a high performance with the simulated patient was predictive of a similar standard with the actual patient, did not hold true. Thus, the findings provided some support for the predictive validity of standardized patient encounters in inadequate cases.

Tamblyn, Schnabl, Klass, Kopelow and Marcy (1988) and Tamblyn, Klass, Schnabl and Kopelow (1991) reported evidence to support the validity of four assumptions that are made about standardized patient performances. These assumptions are reproducibility (that the same clinical problem is presented for each examinee); an absence of bias (whereby the standardized patient is able to provide an unbiased presentation of an actual patient case); replicability (that two or more standardized patients given the same

training can present the same clinical case); and portability (trainers in geographically different testing places can train their particular standardized patients to present the same case).

Thus far this review of the literature regarding standardized patients studies has involved their use in the assessment of clinical skills. The section which follows considers the use of standardized patients to assess the English language proficiency of foreign medical graduates in the United States within a clinical context.

(vi). The use of standardized patients to assess English language proficiency.

The main contributors to the literature concerning the use of standardized patients to evaluate spoken English proficiency who are active in the United States are Miriam Friedman, Paula Stillman and Alton Sutnick and their colleagues (Friedman, Sutnick, Stillman, Norcini, Anderson, Williams, Henning and Reeves, 1991; Friedman, Sutnick, Stillman, Regan and Norcini, 1993; Sutnick, Stillman, Norcini, Friedman, Williams, Trace, Schwartz, Wang and Wilson, 1994). They work under the auspices of the Educational Commission for Foreign Medical Graduates (ECFMG), which began in 1956 as the Educational Council for Foreign Medical Graduates, a body responsible for the assessing the suitability of foreign applicants for United States residency or fellowship programmes (Sutnick, Shafron and Wilson, 1992). Providing an applicant has fulfilled all of the necessary educational requirements, ECFMG certification is granted following satisfactory performance on examinations in the basic medical and clinical sciences, and passing a test of English language proficiency (Sutnick, Shafron and Wilson, 1992).

In the United States there has been concern for some time about the adequacy of the English language skills of some candidates entering the country to embark on graduate medical education programmes. In 1972, Sutnick, Kelley and Knapp reported on an examination procedure used by the ECFMG to assess English proficiency by testing the comprehension of spoken language. Examinees listened to a number of statements and a brief medical history about which they were then required to answer a series of multiple

choice questions. The authors concluded that their testing method was reliable and valid, but this was challenged by De Sweemer (1972), who queried the likelihood of Type I error with such testing (that is, rejection of the null hypothesis when it is actually true, which would mean finding statistically significant differences in the English language proficiency of candidates where none exists). De Sweemer (1972) also criticised the test for decontextualising language by ignoring the non-verbal components of speech, and argued that the administration of the test was poor, with sound being distorted by amplification and the tester's pronunciation unclear.

During the 1990's a more comprehensive procedure for assessing English language proficiency was funded by the ECFMG involving a project which piloted the viability of using standardized patients to assess foreign medical graduates in their ability to use spoken and written English adequately (Sutnick, Stillman, Norcini, Friedman, Regan, Williams, Kachur, Hagerty and Wilson, 1992). As has been discussed, standardized patients are commonly used in the United States to teach and assess clinical skills, but the authors of this paper claim to be the first to use such patients for the evaluation of language competence (Friedman, Sutnick, Stillman, Norcini, Anderson, Williams, Henning and Reeves, 1991). On the basis of their research they have concluded that standardized patients' ratings of English language proficiency have high reliability, high concurrent validity with standardized language tests and high predictive validity with clinical encounters (Friedman, Sutnick, Stillman, Regan and Norcini, 1992).

To establish the validity of the standardized patients' assessments of language proficiency Friedman, Sutnick, Stillman, Norcini, Anderson, Williams, Henning and Reeves (1991) compared their scores with scores from an established instrument, the Test of Spoken English (TSE), which has previously been validated as a measure of oral language proficiency in the health care professions (Powers and Stansfield, 1983). The Test of Spoken English evaluates spoken responses to questions, reading a passage aloud, completing sentences; telling a story both from a series of picture prompts and a single picture, and giving a short presentation (Educational Testing Service, 1993). Ratings can

be made on the TSE for comprehensibility, pronunciation, grammar and fluency (Clark and Swinton, 1980).

Friedman, Sutnick, Stillman, Norcini, Anderson, Williams, Henning and Reeves (1991) queried whether standardized patients could provide a reliable test of language proficiency, given that they cannot control the amount of spoken language generated during an interaction, and if rating a speaker would be too complex a cognitive task when other ratings of the graduate, such as clinical competencies, were to be made by the standardized patient from the same encounter. The study also investigated whether high concurrent validity existed between the standardized patient's rating of language as compared to the assessment of a professional language expert's. The study also aimed to ascertain if it was necessary to rate the graduate's spoken language proficiency within a clinical context.

The standardized patients employed to assess spoken language proficiency were trained in groups of four, for a total of four hours. They were made familiar with the criteria for each level of the Test of Spoken English, by listening to tapes and allocating scores on four samples of speech. A discussion following this allowed the trainees to clarify scoring discrepancies and ask questions.

In the study 122 foreign medical graduates and 24 United States graduates were involved in 8 clinical encounters with standardized patients, who rated them on history taking, physical examination, interpersonal skills and language proficiency. After the encounter the graduate completed a written multiple choice exercise regarding the findings from the history and physical examination, the diagnosis and recommendations. Spoken English was rated by the standardized patient on a four point scale, following the rating system of the Test of Spoken English; (Level 1 - comprehensibility very low even in the simplest forms of speech; Level 2 - generally not comprehensible; Level 3 - comprehensible with errors in grammar, pauses, rephrasing and errors in pronunciation; and Level 4 - completely comprehensible) (Clark and Swinton, 1980). The Test of Spoken

English was then administered by a ninth standardized patient at the conclusion of the eighth encounter, whereby the graduate was tape-recorded responding orally to subtests which were designed to assess comprehensibility, pronunciation, grammar and fluency. The tapes were then rated by professional scorers from an English language testing service

Results demonstrated that the standardized patients were able to carry out their assigned task, with a high correlation found between the ratings of the standardized patients and the professional English language scorers. The reliability of the standardized patient ratings was satisfactory. All of the graduates in this study were found to have acceptable levels of spoken language proficiency, and the authors identify this as a weakness in terms of the conclusions that can be drawn from the results, with a dispersion of language abilities needed in the sample to determine the necessity of rating language in a clinical setting.

A second study (Friedman, Sutnick, Stillman, Regan and Norcini, 1993) sought to investigate the reliability and validity of standardized patients' ratings of spoken English, including the predictive validity of these ratings with clinical performance, such as interpersonal skills. The influence of standardized patient training on the validity of ratings was also considered.

In this study eight standardized patients rated 382 graduates of non-US medical schools for spoken English skills, following a 15 minute encounter. Language was rated on a four-point Likert-type scale (0=low comprehensibility, 1=generally not comprehensible, 2=comprehensible with errors, 3=completely comprehensible), based on the Test of Spoken English scoring system used in the previous study (Friedman, Sutnick, Stillman, Norcini, Anderson, Williams, Henning, and Reeves, 1991). The students' resident programme director also rated the graduate on this four point scale, and a professional English rater assessed graduates on the Test of Spoken English. Thus each of the foreign graduates had three ratings on language, those of the standardized patients', the Test of Spoken English, and the program director's. Overall spoken English scores were

calculated for analyses by averaging ratings across the standardized patients and the professional English raters. The researchers also tested the relationship between the spoken language scores elicited from this study, and scores on standardized language tests, including some administered two and a half years previously when the graduate had applied for ECFMG certification, a requirement in the United States for foreigners to commence residency.

Of the 382 graduates only 107 could be included in analyses due to missing data on either the clinical or language measures. The ratings of spoken language between the standardized patients and the professional English raters correlated highly, whereas the ratings from the programme directors were modestly correlated to standardized patient and professional rater scores. Ratings of spoken language by the standardized patients and the professional raters and interpersonal skills were modestly correlated, with a low correlation found on program director ratings. Varying low to moderate correlations were reported between the spoken language ratings from this study and standardized language tests. The researchers found the predictive validity of the spoken English measures on clinical performance to be relatively high and more predictive than the standardized language test scores.

It is clear from Friedman, Sutnick, Stillman, Regan and Norcini (1993) that comparing the trained raters scores (i.e. the standardized patients) and the untrained (i.e. the programme directors) with the Test of Spoken English criterion measure clearly demonstrated the effect of training on the validity of the ratings, with the untrained raters consistently correlating less with the TSE score than the standardized patient. Thus, in order to rate English language skills in a medical education context the authors advised that standardized patients trained in the evaluation of English language proficiency be employed in favour of untrained staff.

This study acknowledges several weaknesses. As with the study by Friedman, Sutnick, Stillman, Norcini, Anderson, Williams, Henning and Reeves (1991) the range of

language abilities in the sample of graduates tested was skewed toward the upper levels. The authors do not indicate how many graduates were found to be unsatisfactory in English, but it would seem that the number was small. Findings from future studies would be enhanced by a range of language abilities, as truncated ranges in samples such as this weaken correlational findings. Secondly, the fact that some standardized measures of language had been administered several years previously may account for their weak correlations with current measures of language. Similarly, any improvement in language ability over this period cannot be accounted for with this method, and concurrent validity between the two instruments cannot be reliably established.

The authors suggest possible reasons for the low correlations found between all measures and the programme directors' ratings. They acknowledge that some doctors may be less sensitive to the psychosocial dimensions of medical care, including interpersonal communication and language. However, they assume that programme directors would be familiar with their students and their students' language capabilities. They do not indicate the number of students for which each director is responsible, or the length of time spent with each graduate. It may be necessary to control for this possible variation, which could influence the accuracy of ratings in a study such as this.

An overall concern is the considerable amount of missing data which precluded 275 subjects of the 382 from inclusion in analyses. With more than two-thirds of the subjects omitted from the results, it would be instructive to know on exactly which measures data were unable to be collected. If considerable data were missing from the spoken language ratings, one would need to know more about under what circumstances a graduate could not be rated, as Friedman, Sutnick, Stillman, Norcini, Anderson, Williams, Henning and Reeves (1991) have raised the point that perhaps standardized patients cannot control the interaction to the extent that sufficient language is generated for reliable ratings to be made. Further information about the amount of speech or variability of encounters in terms of the levels of discussion between the two parties may have explained if and why data was missing on this variable.

This study (Friedman, Sutnick, Stillman, Regan and Norcini, 1993) reported that graduates completed a "patient note", summarising findings of history taking, physical examination, differential diagnosis and management, to assess written communication with members of the health care team. This material does not appear to have been scored for written language proficiency, and is not reported upon separately in this paper.

To assess written English the "patient note" was investigated in a study of 265 residents by Stillman, Regan, Haley, Norcini, Friedman and Sutnick (1992). The study sought to establish whether a foreign medical school graduate's written account of a standardized patient encounter, completed in seven minutes under examination conditions, could be employed as a measure of ability to summarise subjective and objective findings, differential diagnosis and proposed management plan. In order to rate the patient note trained coders matched the degree of agreement between the content of the graduate's written description and the key words and phrases noted by eight standardized patients as important outcomes of the encounter, and calculated a weighted score for each note. No penalty was made for incorrect findings, inappropriate test orders or misdiagnosis; neither were quality of grammar, spelling or sentence construction taken into account. A range of agreement was found between the standardized patient and graduate accounts, which the authors attributed to a reflection of differing levels of difficulty across the eight cases and the variability of the graduates' skills. Stillman, Regan, Haley, Norcini, Friedman and Sutnick (1992) concluded that a correlation of 0.20 between the mean patient note score (across the eight cases) and the Test of Spoken English for 223 of the sample for which this data was available indicated a relationship between spoken English proficiency and the written material.

Stillman, Regan, Haley, Norcini, Friedman and Sutnick (1992) acknowledged that a limitation of their study was that scoring did not take inappropriate procedures or incorrect diagnoses into account. It is important to remember that the study's primary concern was with the level of accuracy with which graduates' recorded standardized patient encounters, and the ease with which these written accounts could be scored by

trained coders. The extent to which English language proficiency influenced the production of patient notes appears to have been of more peripheral interest. Nevertheless, it would have been useful had the authors given a description of the range of Test of Spoken English scores. It would be reasonable to assume that like other samples employed in ECFMG research, the scores for spoken language proficiency were negatively skewed. This may explain why the correlation between the patient note and measure of English language proficiency was low. Stillman and her colleagues call for replication studies, and any further research would need to take the range and dispersion of scores within the sample on language proficiency instruments into account. A future study might also determine whether English language proficiency has any bearing on inappropriate clinical conclusions. The written recording of inaccurate descriptions of the encounter and what took place within it could also be investigated. If finances permitted, a replication study would benefit from a videotape of each encounter, to ensure that the standardized patients' written account is accurate, before using that account as a criterion measure for the graduates' account.

In 1993 Sutnick, Stillman, Norcini, Friedman, Regan, Williams, Kachur, Haggerty and Wilson investigated the feasibility, reliability and validity of assessing spoken English proficiency for 525 first year residents in an ECFMG clinical competence study. As in previous ECFMG research, language was rated by eight individual standardized patients on a four point scale from 'low comprehensibility' (1) to 'completely comprehensible' (4). The Test of Spoken English was administered to assess comprehensibility, pronunciation, grammar and fluency, and employed as an external validation criterion of the standardized patients' language ratings. As in a previous study (Friedman, Sutnick, Stillman, Regan and Norcini, 1993) the sample means were high on both the standardized patient ratings and the Test of Spoken English. Standardized patient language ratings yielded high reliability coefficients, and moderately high correlations were found between the standardized patient and Test of Spoken English professional rater scores. The authors noted that both measures of language proficiency were highly correlated with residents' interviewing and interpersonal skills.

In 1994 Sutnick and his colleagues (Sutnick, Stillman, Norcini, Friedman, Williams, Trace, Schwartz, Wang and Wilson) conducted a pilot project of the assessment of clinical competence for the ECFMG, testing among other variables, the rating of spoken English proficiency. They used this clinical competence assessment as a formative method, suggesting that with knowledge of individual students' strengths and weaknesses remedial programmes can be implemented where appropriate. They implied that this would apply not only to clinical skills, but to English language skills as well.

Although their study does not specifically consider the English language proficiency of foreign medical graduates, the work of Stillman, Swanson, Regan, Philbin, Nelson, Ebert, Ley, Parrino, Shorey, Stillman, Alpert, Caslowitz, Clive, Florek, Hamolsky, Hatem, Kizirian, Kopelman, Levenson, Levinson, McCue, Pohl, Schiffman, Schwartz, Thane and Wolf (1991) should be included here. In an assessment of the reliability and validity of standardized patients' evaluations of clinical performance the authors concluded that since foreign medical graduates performed more poorly overall than United States and Canadian graduates this finding could be taken as an indication of construct validity. Omoigui (1991) criticised this argument, stating that foreign medical graduates may not have poorer clinical skills, but rather, may have been rated more poorly by standardized patients who themselves may not have been culturally and racially representative of their examinees, thus resulting in bias. Stillman, Swanson and Regan (1991) conceded this as a flaw in their methodology, but added that;

> "greater general familiarity with American culture and with medically relevant aspects specifically, as well as increased language fluency, is likely to improve interactions with standardized patients, just as they do with actual patients" (p 158).

Summary

The review of literature presented in Chapters I and II opens up a number of lines of enquiry for research which focus upon the English language proficiency and academic and clinical performance of University of Adelaide Medical School undergraduates. It is recognized that many other possible avenues for future studies have also arisen from the review, such as the relationship between medical students' English language proficiency and their patients' levels of satisfaction, adherence to the doctor's advice and health outcomes. However, the following areas have been chosen for inclusion in this thesis:

1. The profile of University of Adelaide Medical School students with regards to both their language background and level of English language proficiency.
2. Establishing whether any significant differences in students' academic and clinical performance can be attributed to the introduction of the Faculty's Language Development Programme.
3. Establishing a methodology whereby English language skills and medical communication skills can be clearly differentiated within a medical education framework.
4. Rating students' ability to employ appropriate English language skills when interacting with patients. These skills include register, colloquial language, speech acts and non-verbal behaviours.
5. The overt use of standardized patients within the context of a clinical encounter to assess the English language skills of undergraduate students with a range of proficiencies by means of a rating scale. Standardized patients have been chosen as a research tool in preference to a discourse analysis methodology following the consideration of the strengths and weaknesses of both methods.
6. An assessment of students' ability to convey clearly and accurately in writing a brief clinical encounter between themselves and a patient.
7. The predictive validity of measures of English language proficiency for academic performance in a specific subject within the curriculum.
8. Determining the role of a Supplementary Programme in the academic performance of students from non-English speaking backgrounds in a first year subject.

Aims and objectives of this thesis

In the light of the preceding summary, this thesis has two major aims. The first is to investigate the English language proficiency of several cohorts of undergraduate medical students at the University of Adelaide, South Australia, by employing both quantitative and qualitative measures. The second aim is to consider the role of English language proficiency upon academic and clinical performance. It should be noted that for the purpose of this thesis "English language proficiency" is defined according to the quantitative and qualitative measures with which it has been assessed. "Academic performance" refers to the summative outcome measures of an examination in medical communication skills and performance in a one-year behavioural science course. "Clinical performance" refers here to interactions between students and patients, both "real" and standardized.

To achieve these aims the thesis comprises seven interrelated studies. These are;

Study I (Chapter III) which describes a procedure whereby first year students were screened for aural (listening) English language proficiency utilizing a standardized paper-and-pencil instrument. The relationship between performance on this instrument and demographic characteristics is considered. Further information about English language proficiency was collected from a subset of these students, by means of a structured interview. Finally, this study outlines a methodology employed to allocate students to the Language Development Programme, (a course of English language tuition), where such an intervention was considered necessary.

Study II (Chapter IV) compares the performance of a cohort of first year students on a standardized aural English language proficiency measure with their performance on a second standardized aural English language proficiency measure two years later. The relationship between these measures and the students' demographic characteristics is explored. The issues of concurrent validity and the improvement of language proficiency over time without language tuition are examined.

Study III (Chapter V) has two main objectives. Firstly, it extends the previous two studies by investigating spoken English language skills in addition to listening skills by way of a simulated medical interview between the student and a standardized patient. The relationships between language background, English language proficiency (as rated by a clinician and as assessed by a standardized instrument) and the demonstration of medical communication skills are investigated. Secondly, this Study considers the proposition that English language proficiency and the satisfactory demonstration of medical communication skills may be independent of one another.

Study IV (Chapter VI) is an exploration of the concerns and comments made by University of Adelaide Medical School clinical teachers regarding students' English language proficiency and the impact of poor proficiency upon clinical performance.

Study V (Chapter VII) describes the development, reliability and validity of a Spoken Language Rating Scale designed in the light of both the review of literature and data collected in Study IV. The piloting of this scale in a clinical context using standardized patients to rate the language of third year undergraduate medical students is reported.

Study VI (Chapter VIII) is a replication of Study V, with some refinement of the Language Rating Scale and the development, reliability and validation of a corresponding pilot version of a Written Language Rating Scale, with second and third year undergraduate students. A comparison is made of performance on the Language Rating Scale of the students in Study V (who did not have the benefit of the Language Development Programme) with those in this Study (some of whom did have the benefit of this programme) on both spoken and written measures.

Study VII (Chapter IX) takes one Subject Stream in the medical curriculum and examines it closely, by describing a teaching programme offered in that Stream for students from non-English speaking backgrounds. Academic performance in the Stream

and the factors which may influence performance are then considered for a subset of students within the cohort.

Subjects

The students who acted as subjects for this thesis were four selected cohorts of undergraduates enrolled in the University of Adelaide Medical School over the years spanning 1991 to 1996. Information about their language spoken at home, place of birth, resident status, method of entry into the first year of the medical course, age in years and gender was extracted from the Student Information System at the University of Adelaide. This data is collected from all students who enrol at the university (see Appendix IX). The University of Adelaide Committee on the Ethics of Human Experimentation ruled that informed consent was not necessary to be obtained from the students who were the subjects for this thesis.

It should be noted that since the research presented in this thesis is somewhat longitudinal in nature, not all students commencing an academic year in each cohort are necessarily represented in a later year. This is due to attrition rates resulting from failure to fulfil course requirements or temporary or permanent withdrawal from the course. Course withdrawal may have resulted from either a Faculty directive or a personal decision.

Subject pool for Studies I and VI

Studies I and VI were conducted with the first year medical student cohort of 1994. The entire cohort's language spoken at home, place of birth, resident status, method of entry into the first year of the medical course, age in years and gender are listed in Tables I, II, III, IV, V and VI respectively.

Table I: Languages spoken at home by subjects.

Language	N	%
African	1	0.7
Bahasa Malaysia	26	18.2
Cantonese	5	3.5
English	89	62.2
Greek	1	0.7
Hindi	2	1.4
Hokkien	1	0.7
Hungarian	1	0.7
Italian	2	1.4
Mandarin	5	3.5
Other Chinese	4	2.8
Persian	1	0.7
Slovak	1	0.7
Vietnamese	4	2.8
	143	**100.00**

Table II: Subjects' place of birth.

Place of Birth	N	%
Australia	67	46.9
Austria	1	0.7
China	1	0.7
Fiji	2	1.4
Hong Kong	3	2.1
India	3	2.1
Iran	1	0.7
Kampuchea	1	0.7
Malawi	1	0.7
Malaysia	41	28.7
Other African	1	0.7
Other Asian	1	0.7
Other European	1	0.7
Pakistan	1	0.7
Papua New Guinea	1	0.7
Singapore	4	2.8
Sri Lanka	2	1.4
Taiwan	4	2.8
Thailand	1	0.7
United Kingdom	2	1.4
Vietnam	4	2.8
	143	**100.00**

Table III: Residential status of subjects.

Status	N	%
Australian resident	109	76.2
Overseas fee-paying	34	23.8
	143	**100.00**

Table IV: Type of entry to first year medicine.

Entry type	N	%
Matriculant	87	60.8
MATES*	21	14.7
Overseas Student	10	7.0
Repeat	8	5.6
Repeat MATES	3	2.1
Special Entry	3	2.1
Second year	1	0.7
Tertiary transfer	10	7.0
	143	**100.00**

* MATES = Malaysian Australian Tertiary Education Scheme.

Table V: Age in years at commencement of medical course.

Age in years	N	%
16	4	2.8
17	43	30.1
18	35	24.5
19	19	13.3
20	15	10.5
21	15	10.5
22	3	2.1
23	3	2.1
24	1	0.7
25	1	0.7
30	2	1.4
35	1	0.7
39	1	0.7
	143	**100.00**

mean=19.01 years, standard deviation=3.09.

Table VI: Gender of subjects.

Gender	N	%
Male	79	55.2
Female	64	44.8
	143	**100.00**

Subject pool for Studies II, III, V and VI

Studies II, III, V and VI were conducted with the third year medical student cohort of 1994. The entire cohort's language background, place of birth, resident status, method of entry into the third year of the medical course, age in years and gender are listed in Tables VII, VIII, IX, X, XI and XII respectively.

Table VII: Languages spoken at home by students.

Language	N	%
Bahasa Malaysia	21	14.1
Cantonese	12	8.1
English	61	40.9
Farsi	1	0.7
German	3	2.0
Greek	3	2.0
Hindi	2	1.3
Hokkien	1	0.7
Italian	5	3.4
Khmer	1	0.7
Malawi	7	4.7
Mandarin	6	4.0
Other Chinese	8	5.4
Punjabi	1	0.7
Sinhalese	2	1.3
Spanish	1	0.7
Tamil	4	2.7
Vietnamese	9	6.0
Yugoslav	1	0.7
	149	**100.00**

Table VIII: Students' place of birth.

Place of Birth	N	%
Australia	59	39.6
Chile	1	0.7
China	1	0.7
Cyprus	1	0.7
Hong Kong	4	2.7
India	3	2.0
Iran	1	0.7
Ireland	1	0.7
Kampuchea	2	1.3
Malawi	7	4.7
Malaysia	46	30.9
New Zealand	1	0.7
Singapore	3	2.0
Sri Lanka	4	2.7
Taiwan	1	0.7
United Kingdom	3	2.0
Vietnam	11	7.4
	149	**100.00**

Table IX: Residential status of subjects.

Status	N	%
Australian resident	112	75.2
Overseas fee-paying	37	24.8
	149	**100.00**

Table X: Type of entry to third year medicine.

Entry type	N	%
Matriculant	81	54.4
MATES*	19	12.8
Overseas Student	17	11.4
Repeat	13	8.7
Repeat MATES	1	0.7
Special Entry	2	1.3
Higher year	1	0.7
Tertiary transfer	15	10.1
	149	**100.00**

* MATES = Malaysian Australian Tertiary Education Scheme.

Table XI: Age in years in third year of medical course.

Age in years	N	%
19	25	16.8
20	37	24.8
21	22	14.8
22	25	16.8
23	19	12.8
24	9	6.0
25	2	1.3
26	1	0.7
27	1	0.7
28	1	0.7
29	2	1.3
30	1	0.7
32	2	1.3
33	1	0.7
35	1	0.7
	149	**100.00**

mean=21.69 years, standard deviation=2.83.

Table XII: Gender of subjects.

Gender	N	%
Male	76	53.5
Female	66	46.5
	149	**100.00**

Subject pool for Study IV

Study IV was conducted with a sample of the third, fourth and fifth year students from 1991 to 1994. Demographic data about students was not collected for this study. However, their cohort profiles in terms of gender, language spoken at home and so forth would be comparable to the other groups included as subjects in this thesis.

Subject pool for Study VII

Study VII was conducted with the first year medical student cohorts of 1995 and 1996. The language background, place of birth, resident status, method of entry into the medical course, age in years and gender for both cohorts are listed in Tables XIII to XXIV respectively.

1995 student cohort

Table XIII: Languages spoken at home by subjects.

Language	N	%
Africaans	1	0.7
Arabic	1	0.7
Bahasa Malaysia	29	20.7
Cantonese	4	2.9
English	83	59.3
Farsi	4	2.9
Greek	2	1.4
Hindi	3	2.1
Japanese	1	0.7
Mandarin	1	0.7
Other Chinese	3	2.1
Russian	2	1.4
Serbocroatian	1	0.7
Swedish	1	0.7
Turkish	1	0.7
Urdu	2	1.4
Vietnamese	1	0.7
	140	**100.00**

Table XIV: Subjects' place of birth.

Place of Birth	N	%
Australia	66	47.1
Brunei	1	0.7
Canada	1	0.7
Fiji	2	1.4
France	1	0.7
Germany	1	0.7
Greece	1	0.7
Hong Kong	2	1.4
India	3	2.1
Iran	4	2.9
Japan	1	0.7
Malaysia	37	26.4
Middle East	1	0.7
Pakistan	1	0.7
Philippines	1	0.7
Singapore	3	2.1
South Africa	1	0.7
Sweden	1	0.7
Taiwan	2	1.4
United Kingdom	1	0.7
USSR	2	1.4
Vietnam	6	4.3
Yugoslavia	1	0.7
	140	**100.00**

Table XV: Residential status of subjects.

Status	N	%
Australian resident	108	77.1
Overseas fee-paying	32	22.9
	140	**100.00**

Table XVI: Type of entry to first year medicine.

Entry type	N	%
Matriculant	83	59.3
MATES*	23	16.4
Overseas Student	4	2.9
Repeat	7	5.0
Repeat MATES	4	2.9
Special Entry	3	2.1
Tertiary transfer	16	11.4
	140	**100.00**

* MATES = Malaysian Australian Tertiary Education Scheme.

Table XVII: Age in years at commencement of medical course.

Age in years	N	%
16	5	3.6
17	46	32.9
18	40	28.6
19	10	7.1
20	27	19.3
21	3	2.1
22	4	2.9
24	1	0.7
26	1	0.7
28	1	0.7
31	1	0.7
35	1	0.7
	140	**100.00**

mean=18.62 years, standard deviation=2.51.

Table XVIII: Gender of subjects.

Gender	N	%
Male	65	46.4
Female	75	53.6
		100.00

1996 student cohort

Table XIX: Languages spoken at home by subjects.

Language	N	%
Arabic	1	0.8
Bahasa Malaysia	22	18.5
Cantonese	3	2.5
English	70	58.8
Farsi	3	2.5
German	1	0.8
Greek	2	1.7
Hindi	2	1.7
Mandarin	5	4.2
Other Chinese	2	1.7
Polish	2	1.7
Sinhalese	1	0.8
Vietnamese	5	4.2
	119	**100.00**

Table XX: Subjects' place of birth.

Place of Birth	N	%
Australia	48	40.3
China	2	1.7
Egypt	1	0.8
France	1	0.8
Hong Kong	5	4.2
India	3	2.5
Iran	2	1.7
Kenya	1	0.8
Malaysia	26	21.8
Middle East	1	0.8
Poland	2	1.7
Singapore	1	0.8
Spain	1	0.8
Sri Lanka	2	1.7
Sudan	1	0.8
Taiwan	6	5.0
United Kingdom	4	3.4
United States of America	3	2.5
Vietnam	8	6.7
Zambia	1	0.8
	119	**100.00**

Table XXI: Residential status of subjects.

Status	N	%
Australian resident	92	77.3
Overseas fee-paying	27	22.7
	119	**100.00**

Table XXII: Type of entry to first year medicine.

Entry type	N	%
Matriculant	78	65.5
MATES*	21	17.6
Overseas Student	6	5.0
Repeat	1	0.8
Repeat MATES	1	0.8
Special Entry	2	1.7
Tertiary transfer	10	8.4
	119	**100.00**

* MATES = Malaysian Australian Tertiary Education Scheme.

Table XXIII: Age in years at commencement of medical course.

Age in years	N	%
17	44	37.0
18	36	30.3
19	7	5.9
20	25	21.0
21	2	1.7
22	2	1.7
26	2	1.7
27	1	0.8
	119	**100.00**

mean=18.44 years, standard deviation=1.81.

Table XXIV: Gender of subjects.

Gender	N	%
Male	67	56.3
Female	52	43.7
	119	**100.00**

Chapter III.

Study I.

Testing the 1994 first year student cohort for English language proficiency employing the Screening Test of Adolescent Language (STAL).

Aims of the present study.

The study described in this chapter had four main objectives:

(i) to screen all 1994 first year medical students enrolled at the University of Adelaide for English language proficiency, using the STAL.

(ii) to construct a profile of the English language proficiency for this group of students, as assessed by the STAL.

(iii) to investigate the relationship between students' demographic characteristics and performance on the STAL.

(iv) to ascertain, through performance on the STAL, and by structured interview, which students from this group should be directed by the University of Adelaide Faculty of Medicine, to undertake compulsory English language intervention.

Method.

Subjects.

One hundred and thirty-nine students (97.2%) of the 1994 first year Medicine intake were tested on the STAL. They ranged in age from 16 to 39 years (mean=19 years); 55.4% of the group were male, 44.6% female.

Four students from the 1994 intake were not tested with the STAL due to late entry to the course (n=2), or illness (n=2).

The most common language spoken at home was English (61.9%), followed by Bahasa Malaysia (18.7%). The distribution of first languages, defined as the language reported by students to be spoken at home, are listed in Table XXV.

Table XXV: Languages spoken at home by students.

First Language	N	%
African	1	0.7
Bahasa Malaysia	26	18.7
Cantonese	5	3.6
English	86	61.9
Greek	1	0.7
Hindi	2	1.4
Hokkien	1	0.7
Hungarian	1	0.7
Italian	1	0.7
Mandarin	5	3.6
Other Chinese	4	2.9
Persian (Farsi)	1	0.7
Slovak	1	0.7
Vietnamese	4	2.9
	139	**100.00**

Students' country of birth are listed in Table XXVI. Most students were born in Australia (46.0%) or Malaysia (29.5%).

Table XXVI: Students' place of birth.

Place of Birth	N	%
Australia	64	46.0
Austria	1	0.7
China	1	0.7
Fiji	2	1.4
Hong Kong	3	2.2
India	3	2.2
Iran	1	0.7
Kampuchea	1	0.7
Malawi	1	0.7
Malaysia	41	29.5
Other Asian	1	0.7
Other European	1	0.7
Pakistan	1	0.7
Papua New Guinea	1	0.7
Singapore	4	2.9
Sri Lanka	2	1.4
Taiwan	4	2.9
Thailand	1	0.7
United Kingdom	2	1.4
Vietnam	4	2.9
	139	**100.00**

The majority of students (75.5%) were Australian residents, with the remainder (24.5%) residing in Australia for the duration of their studies, most probably to return to their country of origin on completion of their medical degree.

The majority of students (59.7%) had gained admission to the University of Adelaide medical school through their matriculation scores. Other avenues for admission are listed in Table XXVII.

Table XXVII: Type of entry to first year medicine.

Entry type	N	%
Matriculant	83	59.7
MATES*	21	15.1
Overseas Student	10	7.2
Repeat	8	5.8
Repeat MATES	3	2.2
Special Entry	3	2.2
Second year	1	0.7
Tertiary transfer	10	7.2
	139	**100.00**

* MATES = Malaysian Australian Tertiary Education Scheme.

Instruments.

The Screening Test of Adolescent Language (STAL) (Appendix X).

The Screening Test of Adolescent Language (STAL) is a twenty-three item test originally devised for individual administration to children of eleven years of age and older, to screen for language disorders (Prather, Breecher, Stafford and Wallace, 1981). The test was designed to be administered orally, with the tester reading each item, and recording verbatim the answer given by the student. Hayes and Farnill (1992) have made several modifications to the STAL for use with Australian University students. Several Americanisms in the test have been substituted with more familiar Australian terms[12]. They have also employed the STAL for group administration, requiring students to give

[12] In Subtest One , Item 2, the word "kettle" was removed, and substituted by the word "saucepan". In Subtest Two, Item 2, "The school on the west side of town" became "The school in the next suburb". Item 3 became " . . . the theatre that gives student discounts" instead of ". . . .the theatre that takes coupons." Item 2 of Subtest Three became ". . . combat boots to wear to the school dance" rather than ". . . combat boots to wear to the Junior Prom".

paper and pencil responses to all items, rather than the verbal answers specified in the test manual.

The STAL comprises four subtests:

Subtest One: "Vocabulary" has twelve items. Students are required to provide a synonym for a word, ensuring that the synonym can be used in a given sentence. For example, in the sentence *Carve the turkey*, "slice" or "cut" would be acceptable synonyms.

Subtest Two: "Auditory Memory Span" has three items. Administered as a paper-and-pencil test, these items essentially become dictation tasks. Students must listen to a sentence, which is read out only once, and write it down word-for-word. An example of such an item is the following: *The fire drill that we had last week turned out to be the real thing.*

Subtest Three: "Language processing" has five items. The student is given an illogical sentence, and must reason what it is about the sentence that is not logical, and why this is the case. For example, in response to the sentence *The sun was shining so brightly last week on Tuesday that I had to wear my sunglasses in the movie theatre* the student is required to explain that one does not wear sunglasses in a theatre, because the sun does not shine there.

Subtest Four: "Proverb Explanation" has three items. The meaning of a proverb must be explained, either generally, or through the use of a specific example. For instance, the student could state that *Practice makes perfect* means that the more you practice something, the better you become at that particular task.

Criticisms of the STAL.

At a theoretical level, the STAL has been criticised because it is not based on any specific developmental model of language (Lieberman, Heffron, West, Hutchinson and

Swem, 1987). This has two implications; firstly, the test authors cannot justify why they have chosen to include the items that they have, or why they chose one method of administering the test over another. Secondly, the test lacks construct validity in the absence of a theoretical foundation.

Sommers (1985) has queried the premise put forward by the authors of the STAL that the test has been designed to screen "adolescent" language, when normative information collected by the authors of the test included sixth grade children between the ages of 10 to 13 years, in addition to older ninth grade children. Ages for the ninth graders are not reported in the test manual (Prather, Breecher, Stafford and Wallace, 1981). Prather, Brenner and Hughes (1981) have designed a five-item mini-screening language test from the STAL, which they administered to 547 students from sixth to twelfth grades. The authors do not report ages for their subjects.

As with any psychometric test, when administering the STAL there is always a possibility that errors of classification will be made. Some students who do not have difficulties with the English language will be identified as requiring language assistance (false positives), and similarly, some students who do have difficulties with the English language will not be identified as such (false negatives). The authors of the STAL (Prather, Breecher, Stafford and Wallace, 1981) have stressed the importance of not relying upon a sole test score when making decisions about English language proficiency. Indeed, the primary purpose of the STAL is to identify students whose language skills warrant further evaluation (Lieberman, Heffron, West, Hutchinson and Swem, 1987). Stephens and Montgomery (1985) suggest interviewing the student, and people who would be familiar with his or her language skills, such as teachers, as well as viewing samples of written work, in addition to standardized testing. In this way, students who have been incorrectly identified as having difficulties with the English language, or who have failed to be identified by the STAL where a difficulty exists, can be reclassified.

Sommers (1985) has argued that misclassifications may occur because the STAL confounds language abilities with verbal intelligence, as measured by standardized IQ tests. Given that medical students in Australian universities have been selected from the highest academic achievers, it seems reasonable to assume that they would have well above the necessary intelligence to understand the test items (Hayes and Farnill, 1993a). A further threat to the reliability and validity of the STAL is the brevity of the test, coupled with highly motivated, intelligent subjects. Students can remember test items, and may pass this information down to successive generations of first year medical intakes. The consequence of this would be an increase in the number of false negatives (Hayes and Farnill, 1994). Test-retest reliability cannot be meaningfully established because most items on the test will be remembered. Therefore, the STAL is an instrument for "one-time-only" administration (Stephens and Montgomery, 1985).

Structured Interview. (Appendix XI).

The structured interview was based upon a similar interview designed by Hayes and Farnill (1993a; 1993b), which they used as one component of a two-year follow-up of a group of students tested using the STAL.

The interview was designed to be completed in no more than 15 minutes. All interviews were conducted by two staff members. One member of the Language Development Committee acted as the interviewer, whilst the other acted as scribe. All interviews were audio-tape recorded.

Questions relating to language use and perceived proficiency in language were included in the interview, requiring both closed and open-ended responses from students. Language related items were either adapted from the Hayes and Farnill interview, or were devised specifically for this study. Items about social supports were also covered, both to encourage the flow of discussion where questions about language skills might not, and also to investigate possible associations between English language proficiency and social supports. The socially oriented items were based upon questionnaires investigating

overseas student adaptation to university life (Mullins and Hancock, 1990; Hancock and Mullins, 1991; Felix, 1992).

The interview covered three broad areas and consisted of nineteen questions;

(i) *Demographics.*

Questions 1 to 3 dealt with demographic information about first language, age at which English learning commenced, and any formal studies of the English language. Questions 5 to 6 investigated the student's place of birth and period of residence in Australia. Questions 7 to 9 investigated schooling history, including the language primarily used for study before attending the University of Adelaide.

(ii) *Language Proficiency.*

Questions 12 and 13 inquired about the language the student used most often in Adelaide, and the language used most often at home with their family. Question 4 required the students to rate their perceptions of their spoken, listening, reading and writing language proficiencies, on a five-point scale, from 'very poor' (1) to 'excellent' (5). Questions 18(a) and 18(b) also probed perceived difficulties with the English language.

(iii) *Social Supports.*

Questions 10 to 11 and 14 to 17 elicited information about social supports, including with whom they lived, and in particular whether the student was interacting with Australian persons, and their use of the English language in informal settings as opposed to scholastic and institutional environments. Students were asked to evaluate the language proficiency of their friends, on the same five-point scale that they rated themselves, from 'excellent' to 'very poor'. Question 18(c) asked the student to identify any problems they may have experienced in dealing with aspects of Australian culture.

The final Question, "Why did you choose to study Medicine?", was included to elicit a lengthy sample of speech from the student, requiring reasonably complex language.

Procedure.

Two days before being tested, students were required to attend an information session[13], explaining the English Language Development Programme, and outlining the format of the STAL and the type of items it includes (Appendix XII). Students had an opportunity to ask questions and clarify any aspects of the process. It was made clear that attendance at the testing session was compulsory and that any student who did not attend would be interviewed. Students were also advised that if identified by the STAL as experiencing language difficulties, attendance at an interview would be necessary for decisions to be made about that student's suitability for the Language Development Programme.

On the day of testing, students were gathered in a large lecture theatre, and provided with a pro forma answer sheet (Appendix XIII). All items were administered by following the STAL Instruction Manual (Prather, Breecher, Stafford and Wallace, 1981) except for Item 2 of Subtest Three, which was modified to read "a pair of army boots to wear to the school dance", instead of "a pair of combat boots to the Junior Prom", as South Australians would not usually use the terms in the original test item. The Instruction Manual does not indicate how long an interval of time should pass between each test item. After a pilot test involving three staff members, it was decided that students should be given ten seconds for each item on Subtest One (Vocabulary), twenty seconds for each item of Subtest Two, (Auditory Memory Span), and forty seconds for each item on both Subtests Three (Language Processing) and Subtest Four (Proverb Explanation). Testing took approximately 10 minutes in total.

---

[13] I am grateful to Dr Rob Barrett, who, as Convener of the Language Development Committee, addressed students at the information session, and also administered the STAL on the day of testing.

Scoring.

The Instruction Manual allows for some subjectivity in scoring of items. For example, acceptable synonyms on Subtest One (Vocabulary) can be credited if they reflect regional or colloquial variations, although they may not appear in the original guide to correct responses. Judgement is also required on the part of the scorer for Subtests Three (Language Processing) and Four (Proverb Explanation), as decisions must be made as to the acceptability of the answer. Subtest Two (Auditory Memory Span) is the only truly objective component of the STAL in terms of scoring, as each omission, addition or substitution in the sentence can be counted as an error. Farnill, Hayes and Chur-Hansen (1995) have investigated the effects of subjectivity in scoring of the STAL for 299 subjects, and concluded that although discrepancies existed in the scoring of individual items, a significant Spearman rank-order correlation of 0.83 ($p<.001$, one tailed) demonstrated satisfactory interrater reliability of the test.

In order to investigate interrater reliability in this study, each completed test paper was duplicated. One set of papers was then scored by four raters: one rater per subtest[14]. For ease of description, these four raters will hereafter be referred to as "Rater 1". The other set of papers was scored by one independent rater (the researcher), hereafter referred to as "Rater 2".

Results.

Screening Test of Adolescent Language.

The means, standard deviations, difference between means and correlations for the STAL subscales and STAL total scale for Rater 1 and Rater 2 are presented in Table XXVIII. The distributions were negatively skewed, and therefore the nonparametric Spearman *rho* (one tailed) was employed in analyses.

---

[14] I would like to acknowledge the contribution of Mr Ian Carman (Subtest One), Mr David Burford (Subtest Two), Ms Ursula McGowan (Subtest Three), and Ms Helen Mullins (Subtest Four), who acted as raters in this study.

Table XXVIII: Means, standard deviations, differences between means and Spearman *rho* for the STAL subscales and the STAL total for Rater 1 and Rater 2.

Scale	Items	Range	$\bar{x}$ Rater 1	SD Rater 1	$\bar{x}$ Rater 2	SD Rater 2	$\bar{x}$ Diff.	*rho*
Vocabulary	n=12	3-12	10.42	1.59	10.43	1.87	-0.23	0.89*
Auditory Memory	n=3	0-3	2.65	0.75	2.69	0.70	-1.92	0.89*
Language Processing	n=5	0-5	4.41	1.15	4.45	1.07	-0.87	0.75*
Proverb Explanation	n=3	1-3	2.57	0.60	2.55	0.69	0.40	0.47*
Total Scale	n=23	5-23	20.06	3.24	20.12	3.57	-0.58	0.86*

* $p<0.001$.

The means for the four subscales and the total scale were not significantly different between raters. The Spearman rank-order correlations for the scores for the subscales and total scale were all significant at $p<0.001$. The correlation of the total scores, at 0.86, indicates good interrater reliability. The correlations of 0.89 for Vocabulary and Auditory Memory Span reflect the objectivity of the scoring criteria for these two subscales, whilst the weaker correlation of 0.47 for Proverb Explanation reflects the level of subjectivity required in scoring the items on this subtest. Examination of the three individual items on Proverb Explanation indicates that the two sets of raters most often disagreed on Item 1 "Practice makes perfect" (Rater 1 scoring 4 students incorrect, Rater 2 scoring 13 as incorrect), and Item 2 "Actions speak louder than words" (Rater 1 scoring 37 students as incorrect, Rater 2 scoring 31 as incorrect).

In terms of selectivity and sensitivity, if Rater 1 is used as the standard for comparison, Rater 2 achieved a sensitivity of 82.5%, correctly identifying 33 of the 40 students classified by Rater 1 as experiencing English language difficulties, and a selectivity of 97.9%, identifying 96 of the 98 students classified by Rater 1 as having no difficulties with language. If Rater 2 is adopted as the standard, Rater 1 achieved a sensitivity of 94.1%, identifying 32 of the 34 students classified by Rater 2, and a selectivity of 92.3%, identifying 94 of 104 students classified by Rater 2.

For purposes of classification of students into those with a possible English language difficulty and those without, the scores of Rater 1 have been employed, and these

are the scores referred to henceforth. Table XXIX gives the distribution of scores for the 139 students for Rater 1. Of the 139 students tested, 40 (28.8%) scored 19 or lower.

Table XXIX: Distribution of STAL scores, Rater 1.

Score	N	%
5	1	0.7
7	1	0.7
10	1	0.7
11	1	0.7
12	1	0.7
13	2	1.4
14	3	2.2
15	5	3.6
16	2	1.4
17	3	2.2
18	7	5.1
19	13	9.4
20	15	10.9
21	21	15.2
22	43	31.2
23	19	13.8
	139	**100**

Relationship between demographic variables and performance on the STAL.

Gender was not related to overall performance on the STAL. Age ($r=-.35$, $p<0.001$), language background, place of birth, method of entry into medical school and overseas fee paying versus permanent Australian resident status were all significantly related to STAL total scores.

Students over the group mean age of 19 years were more likely to have English language difficulties (x^2(continuity correction)=26.32, df=1, $p<0.001$); older students were also most likely to have entered Medicine through the MATE scheme ($x^2=90.62$, df=2, $p<0.001$). To investigate the nature of the association between the students' language background and STAL performance, students' reported first languages were categorised into "English", "Bahasa Malaysia" and "Other". Students from English speaking background were least likely to be identified as experiencing difficulties with language, and students from non-English speaking backgrounds the most likely, particularly those students whose first language was Bahasa Malaysia ($x^2=51.7$, df=2, $p<0.001$).

Categorising place of birth into "Australia", "Malaysia" and "Elsewhere" showed a significant association between students born in Malaysia and identification by the test as experiencing language difficulties (x^2=25.17, df=2, p<0.001). Permanent Australian residents were least often identified as having language difficulties (x^2(continuity correction)=30.32, df=1, p<0.001); MATES students were the most frequently identified group of students (x^2=48.47, df=2, p<0.001).

Structured Interview.

Following Hayes and Farnill (1992) it was decided that all students who scored 19 or below on the STAL should be interviewed individually, to collect data about their language abilities and also to identify any students who had been incorrectly classified as experiencing difficulties with English. These 40 students were all requested to attend an interview with a member of the Language Development Committee, as were the four students who had not attended on the day of testing. Over the following week 42 interviews were conducted; (one student decided to withdraw from the medical course, and a scoring discrepancy in one test between the two sets of raters resulted in the decision that the student should not be interviewed).

Four members of the Language Development Committee acted as interviewers[15]. Interviews ranged from 4 to 22 minutes in duration, ($\bar{x}$=12.9 minutes, sd=4.52).

Results of demographic items.

The 42 students interviewed stated their place of birth to be Malaysia (54.8%), Australia (21.4%), Hong Kong (4.8%), Taiwan (4.8%), Fiji (4.8%), Nigeria (2.4%), Vietnam (2.4%), Iran (2.4%) and India (2.4%). Students' first languages were Bahasa Malaysia (52.4%), followed by English (26.2%), Mandarin, (7.1%), Cantonese (4.8%), Hindi (4.8%), Persian (2.4%) and Vietnamese (2.4%). Students had commenced learning the English language from between 2 and 15 years of age ($\bar{x}$=7.16 years, sd=2.4). All 42 students had received formal instruction in English at secondary school level. The final

[15] Anna Chur-Hansen interviewed 13 students, Dr Rob Barrett 15, Dr Jane Vernon-Roberts 9, and Ursula McGowan 5.

year of secondary studies was completed either in Malaysia (50.0%), Australia (45.2%) or Fiji (4.8%). Before attending medical school, students had been studying in English (50.0%) or Bahasa Malaysia (50.0%).

The period of time that students had been living in Australia before the interview (if Australia was not their place of birth) varied from 2 months (6.1%), to over 17 years (3.0%). On average, students had resided in Australia for approximately 2 years (24.55 months, sd=39.99 months).

Results of language proficiency items.

When students were asked to identify the language most often used whilst they had been living in Adelaide, not all could state categorically that one language or another was favoured. English was the language used by 45.2% of interviewees, and 28.6% stated that they used a language other than English most of the time. However, 26.2% deliberated over this item, and concluded that they most often spoke a combination of English and another language. When at home with their families, whether their family lived in Australia, or elsewhere, 11.9% of those interviewed stated that they spoke with a combination of English and another language. Bahasa Malaysia was spoken in the family home by 42.9%, followed by English (31.0%), Cantonese and Mandarin (both 4.8%), and Hindi and Vietnamese (both 2.4%).

When students were asked to rate themselves for their ability to speak, listen, read and write in English on the five-point scale from 'very poor' to 'excellent', several students were unable to choose one rating, and elected between two points on the scale, such as between 'good' and 'excellent'. To analyse these responses, the five-point scale was expanded to nine-point. Table XXX lists the points on the resultant scale.

For speaking ability, responses ranged from 'between poor and fair' (2.4%), to 'excellent' (11.9%), $\bar{x}$=6.21, sd=1.44. Responses for ability to understand spoken English, or listening, ranged from 'fair' (23.8%) to 'excellent' (16.7%), $\bar{x}$=6.83, sd=1.32. Self-

evaluation of reading ability ranged from 'poor' (2.4%) to 'excellent' (16.7%), $\bar{x}$=6.71, sd=1.49. Writing ability was rated from 'between poor and fair' (2.4%) and 'excellent' (16.7%), $\bar{x}$=6.41, sd=1.55. Table XXX gives the frequency distributions for the four self-evaluated language skills.

Table XXX: Distributions of students' self-ratings (n=42).

	3 (%)*	4 (%)	5 (%)	6 (%)	7 (%)	8 (%)	9 (%)
Speaking		1 (2.4)	20 (47.6)	1 (2.4)	14 (33.3)	1 (2.4)	5 (11.9)
Listening			10 (23.8)	3 (7.1)	20 (47.6)	2 (4.8)	7 (16.7)
Reading	1 (2.4)	1 (2.4)	9 (21.4)	3 (7.1)	19 (45.2)	2 (4.8)	7 (16.7)
Writing		1 (2.4)	18 (42.9)	2 (4.8)	12 (28.6)	2 (4.8)	7 (16.7)

***key**
3 = 'poor',
4 = 'between poor and fair',
5 = 'fair'
6 = 'between fair and good',
7 = 'good',
8 = 'between good and excellent',
9 = 'excellent.

Items that invited an open-ended response about any difficulties in either understanding Australian speakers, or in making themselves understood to others did not generate substantial comment, although some students did elaborate briefly on the reasons why they had experienced difficulty. As with their self-evaluations of language abilities, students were asked to rate themselves for the degree of difficulty they had experienced for these two items. They were given a three point scale - 'a great deal of difficulty', 'some difficulty', 'no difficulty at all' - with the option of saying that they were 'unsure' about their difficulties, if appropriate. Several students were again unable to decide upon one rating on the scale, and the scale was converted to a five-point scale for the purpose of analyses, allowing for the categories of "between a great deal of difficulty and some difficulty" and "between some difficulty and no difficulty at all".

The majority of students (57.5%) reported no difficulty at all in understanding Australian speakers. Some difficulty was reported by 42.5% ($\bar{x}$=4.15, sd=1.00). The comments made by this latter group are instructive in establishing where those areas of

difficulty might lie. Of the eight comments made regarding this question, four mentioned problems comprehending Australian slang, three noted that the Australian rate of speech was too fast for them to easily understand, one felt that he could understand urban dwelling Australians, but not rural dwellers, one student stated that the Australian accent made speech unclear, and two students differentiated between the rate of speech of younger people as compared to "oldies", whom they said they found difficult to follow in conversation.

Some students (52.5%) were also likely to say that they had no difficulty at all making themselves understood when speaking in English to someone else in Australia. However, other students reported 'between a great deal and some difficulty' (2.5%), 'some difficulty' (42.5%), and 'between some and no difficulty at all' (2.5%). Six students gave some further comment in addition to rating their difficulties. Four of these comments reflected the reactions made by others when the student spoke. Examples included speaking loudly and slowly to the student, misunderstanding what the student had said, asking the student to repeat themselves, and saying "hey?". One student felt that he had difficulty expressing "big" ideas to others, such as summarising a journal article, and another thought that time spent translating from one language to another in her mind before responding in conversation, might make her difficult for others to understand at times.

Results of social support items.

Students lived either with friends, usually other students (54.8%), family (31.0%), at a residential college affiliated with the university (11.9%), or in private rental accommodation (2.4%).

All 42 students said that they had friends living in Adelaide, and spent time in recreational activities with them. The majority of students had a friend studying with them at medical school, and therefore reported seeing a friend every day (62.2%). Others saw a friend once per week (18.9%), twice per week (2.7%), or between every day and twice per

week, depending upon the friend in question (16.2%). Ninety-five percent believed that at least one of these friendships was "close". All friends spoke English, with 4.8% preferring to speak English mixed with another language. The students were asked to rate the spoken language ability of their friends, in general, on the five point scale, from 'excellent' to 'very poor'. Again, this rating was converted to a nine-point scale, to facilitate analyses. Ratings of friends' English ranged from 'fair' (20.5%), 'between fair and good' (5.1%), 'good' (25.6%), 'between good and excellent' (15.4%) to 'excellent' (33.3%) ($\bar{x}$=7.36, sd=1.51). When spending time with their friends, 57.9% of the students spoke in English, 28.9% in Bahasa Malaysia, 10.5% in a combination of English and another language, and 2.6% in Cantonese Ninety-five percent felt they had at least one friend whom they perceived to be an "Aussie".

The students were asked to give some indication about how they felt concerning getting acquainted with their Australian classmates. Some of the answers to this question were monosyllabic and to the point. However, this question elicited eighteen more lengthy answers, and gave rise to some emotional responses. Answers of more than one sentence were transcribed, and categorised as either reflecting positive feelings about getting to know the classmates, negative feelings, or mixed feelings. An independent rater[16] also categorised the responses, resulting in agreement in 14 of the 18 responses (77.78%). On the four transcripts where there was disagreement, the two raters discussed why this was so, and agreed upon a category. Representative examples from each category are given below, whilst all transcribed responses can be viewed in Appendix XIV.

Positive feelings about getting to know Australian classmates:

"Great. Learn about a new culture, their way of thinking, and feeling. So it's good. Quite easy [to make friends] at College, because everybody's an individual, so they're all open to friendships."

"I haven't actually gotten to close contact with them really. No, they're fine, friendly, really helpful. It will be interesting really, 'cause, well, ahh, since you have a different cultural background. It will be interesting."

[16] I would like to thank Sue Sullivan for acting as a rater for this task.

Negative feelings about getting to know Australian classmates:

"Sometimes quite difficult, because, like, aahh the, we don't really have that much in common, so you keep on talking the same topic over and over again, like every time you meet them. Yeah. And like 'cause sometimes it's quite hard because they have like, their own, their own clique, and it's quite hard to join them."

"Sometimes, umm, it's quite embarrass, you know, it's sort of when I talk with them, or when I talk with some guy, they don't or he doesn't understand me, and they say 'pardon' and I say it again, and then he still say 'pardon', and then teasing, umm, and doesn't want to listen again, so I feel embarrassed and sometimes I feel humiliating. So, but my close friend, she's an Australian, she's very good, and because she's very patient, she try to understand me and, first time my expression not very good, so I change another expression again, so she try to understand me, I ask her which part I wrong, she will help me, 'You should say this, this, this.' 'Usually it's said in this way', or something. So, that means when I talk to my close friend I feel very comfortable, sorry comfortably. But I mean when I talk to other Australian, they very impatient, I feel very embarrassing. And sometimes when I have dinner or lunch or breakfast [at College] and I sit on same table and I'm just only Chinese, they all speak English and they use some slangs or something I just try to understand. I don't understand what they talking about, because maybe they talk very fast, sometimes they may use some slang, don't understand."

Mixed feelings about getting to know Australian classmates:

"Well, I like to know them, but since they tend to be among themself, and you know, I feel, like, awkward, just to go and say 'hi', and , and also because of the language barrier, I feel quite uneasy, you know, like awkward, just go and mix with them, and they speak fast, and sometimes I don't understand them, so yeah, and I sometimes shy too, yeah, I would like to make friends with them."

". . . I am happy because they can accept me, but sometimes I feel quite little and tiny to them because they are good, and especially in the language we are not that good, so we feel belittle ourselves. But try, because sometimes they are very helpful at times."

Of the 42 students interviewed, 39 indicated their willingness or otherwise to become acquainted with the Australian students in their class. Positive feelings were expressed by 64.1%, negative feelings by 20.5%, and mixed feelings by 15.4%.

Students were asked about their perceived difficulties with understanding Australian customs, again rating themselves on the five-point scale of 'a great deal of difficulty' to 'no difficulty at all'. The majority of students said that they had no difficulty at all understanding Australian customs (87.2%), one said that she was unsure (2.6%), and the remaining four students had some difficulty in understanding the Australian way of

life. Three of these students made comments about their difficulties, each of them commenting about social aspects of the culture about which they were unsure, or bemused:

"I mean, what they expect us to say and how, what they expect us to answer, when they ask us something. When we are a good friends of them, what they expect, what will they expect of us."

"Social circle . . . compared with my culture. In my culture - I'm not saying my culture is good - there is a boundaries between males and females, so, compared with Australia, where the people are very open-minded, so I just, I sometimes I can connect it with some other problems, like the collapsing of moral problems . . "

"It's the difference . . . yeah, they drink a lot actually, alcohol, and they talk when they group they don't usually talk about the work, but they talk about something else, like sports, and I've been to quite a lot of parties, and that was one thing, I was quite surprised, the men in one corner and the women in the other. So yeah, and I found it quite interesting though, because everybody is friendly, you don't have to know anything about it."

The question about why the student had selected Medicine as their choice of study was included primarily to encourage free-flowing speech. Responses to this question have not been included in analyses.

Relationship between self-evaluations and the STAL.

The students' self-evaluations regarding spoken English ($r=0.57$, $p<0.001$), understanding others ($r=0.46$, $p<0.005$), reading ($r=0.46$, $p<0.005$) and writing ($r=0.54$, $p<0.001$) were all correlated with the STAL total score, as was the perception that the students were understood by others ($r=0.32$, $p<0.05$). Evaluation of the ability to comprehend Australian speakers and STAL score just failed to reach significance with a correlation of 0.27, $p<0.06$.

No significant correlations were found between the socially oriented items regarding attitudes towards getting acquainted with Australian classmates and understanding Australian customs and the STAL total score.

Allocation to the Language Development Programme.

As has been stressed by previous authors (Prather, Breecher, Stafford and Wallace, 1981; Stephens and Montgomery, 1985; Lieberman, Heffron, West, Hutchinson and

Swem, 1987) decisions about students' need for language interventions were not based solely upon the STAL score. Allocation to the Language Development Programme was based upon performance on the STAL, the Structured Interview conducted with a Language Development Committee member, and a possible subsequent interview with the English language specialist.

Allocations by the Language Development Committee.

Following the Structured Interviews, the Language Development Committee members met to discuss the possible language intervention that would be provided for each of the 42 students. Decisions were made collaboratively, on the basis of the student's performance on the STAL and the responses given during the interview. Six students were considered in no need of assistance with language. Four of these students had been interviewed because they had not been tested with the STAL due to late entry to the course or illness, and it was judged that they had no difficulties with language; the remaining two students were deemed false positive cases, incorrectly identified by the STAL as requiring language intervention.

The remaining 36 students who had been interviewed were allocated to one of three streams: those who needed a further interview by the English language specialist for a more comprehensive assessment of their needs (7 students); students who had reported specific areas of weakness, such as writing, and who could be offered modules to assist them in their specific area or areas of need (7 students); and students who were believed by the interviewer and the Committee as a whole, to require a full, two-year programme of language training (22 students).

Self-Referrals.

Four students who were not identified by the STAL as requiring language assistance requested permission to enter the Language Development Programme, believing that they would benefit from the course. These students were interviewed for their need

and suitability for the course by the English language specialist. Three of the students were allocated to the two-year programme, and one was offered specific modules.

Two of the self-referred students scored 20 on the STAL, two scored 22. One student was male, three female, and ranged in age from 19 to 21 at the time of testing. Three students gave their place of birth as Malaysia, and one as Thailand. English was recorded as the first language for two students, and Bahasa Malaysia for two. Two of the students were fee-paying students, here in Australia for the duration of their medical degree; two were Australian permanent residents.

Final allocations.

The English language specialist interviewed the 7 students directed by the Language Development Committee to attend another interview for further assessment. These students were not offered a placement in the Language Development Programme.

Following further assessment by the English language specialist, 25 students in total (22 identified by the STAL, three self-referred) were allocated to the full, two year programme. Eight students (seven identified by the STAL, one self-referred) were advised that they required specific modules for language intervention, and were requested to attend the Language Development Programme for a period of six months in the first instance. The 29 students who were selected as candidates for the Language Development Programme comprised 72.5% of the 40 students initially identified at testing as requiring language intervention.

The seven students who were not required to participate in the Language Development Programme scored between 16 and 19 on the STAL ($\bar{x}$=18.57, sd=1.13). The 25 students who entered the full Language Development Programme scored between five and 22 on the STAL ($\bar{x}$=15.2, sd=4.18), as compared with the eight students who were directed to participate in modules of the Programme (STAL range 17-20, $\bar{x}$=18.25, sd=1.04). This difference in performance on the STAL between students in the full

Programme compared to those allocated to modules was significant ($t=-3.34$, $p<0.005$). A comparison of the STAL performance between the 33 candidates in the overall Programme and the seven students who were not offered a placement showed a significant difference ($t=-3.29$, $p<0.005$).

Those in the full Programme ranged in age from 17 to 21 years at the time of taking the STAL ($\bar{x}=20.16$, sd=1.03). Nine of the students were male, 16 female. The majority of this group were born in Malaysia (n=19), with the remainder giving their place of birth as Australia (n=1), Hong Kong (n=2), Taiwan (n=1), Thailand (n=1) or Vietnam (n=1). Students' first languages were Bahasa Malaysia (n=18), English (n=3), Cantonese (n=2), Mandarin (n=1) or Vietnamese (n=1). Eighteen were overseas sponsored fee-paying students, seven Australian residents. Three of the students were undertaking their first year medical studies for the second time, having failed to reach a satisfactory standard of performance in the previous year.

Those students required to participate in modules of the Programme ranged in age from 18 to 22 years at the time of taking the STAL ($\bar{x}=19.88$, sd=1.36). Six of the students were male, two female. The majority of this group were also born in Malaysia, as were those in the full Programme (n=6), with the remainder giving their place of birth as Australia (n=1) or Fiji (n=1). Students' first languages were Bahasa Malaysia (n=5), English (n=1), Hindi (n=1) or Mandarin (n=1). Six were overseas sponsored fee-paying students, two were Australian residents. Two of the students were undertaking their first year medical studies for the second time, as a result of failing to pass their examinations in the preceding year.

Data collected at the Structured Interview showed a significant difference between the two groups of Language Development Programme participants on the self-evaluation rating of spoken language ability ($t=-2.38$, $p<0.05$). There were no differences between the two groups on the self-evaluations of understanding spoken English, reading, or writing. Comparisons excluded the four students who self-referred into the Programme, as

they were not interviewed by the Language Development Committee. A comparison of the 33 Language Development Programme candidates with the seven students who were not selected for the Programme showed significantly higher self-evaluations in the latter group on spoken ability ($t(pooled)=-5.77$, $p<0.001$), understanding spoken English ($t(pooled)=-3.80$, $p<0.01$), reading ($t=-3.55$, $p<0.01$) and writing ($t(pooled)=-3.97$, $p<0.001$).

Discussion.

The aims of this study were to test a first year medical student cohort using the STAL, and describe the group's language proficiency according to that test, as well as investigating the effects of demographic variables on STAL performance. The study also sought to identify students who should be required to attend additional classes to develop their English language skills, whilst studying in the medical course.

For the purpose of screening the entire first year medical student intake for language difficulties, the STAL proved to be an easily administered and reliably scored instrument. This supports the findings of Hayes and Farnill (1992), who commended the test for identification of students who may need language assistance, and is also consistent with other studies which have reported good interrater reliability (Farnill, Hayes and Chur-Hansen, 1995). The testing of students using a standardized instrument was useful in establishing a profile of the English language proficiency of a large group of students, in an economical manner. The STAL was administered with several modifications to the instructions specified in the original test manual, and it may be that these alterations influenced the reliability or validity of the test, although given the high test sensitivity found in this study, and reported in previous research (Farnill, Hayes and Chur-Hansen, 1995) (Appendix I.II), it would seem that this has not been the case.

Previous research has not reported on the relationship between performance on the STAL and age and gender. The finding that age was related to language ability in this study can be explained by the fact that MATES students were many of the candidates

identified by the test as experiencing language difficulties. MATES students enter the University of Adelaide several years older than most of their Australian born or Australian educated peers. The MATES students also account for the association between lower STAL scores and being born in Malaysia, and being a temporary resident in Australia. The finding that students from non-English speaking backgrounds were more likely to be identified on the STAL was not unexpected. However, it is noteworthy that three students who entered the Language Development Programme did report their first language to be English.

The primary purpose of the Structured Interview was to facilitate the decision making process of the Language Development Committee. It was necessary to ensure that students were not incorrectly classified as requiring assistance with language, and also to determine whether students might benefit from a long-term, or shorter term language intervention. As well as serving its purpose as an indicator of language proficiency, the interview generated interesting data about students at the lower end of the STAL score distribution. Self-evaluations of speaking, listening, reading and writing skills are useful. Although the Likert rating scale of "very poor" to "excellent" can be criticised for its subjectivity, students' ratings of themselves on these items were significantly, although modestly, correlated with their performance on the STAL, as was the perception that the student could be understood by others. Students' self-evaluation of ability to understand Australian speakers just failed to reach significance on this measure. That the standardized test results and the self-evaluations are related may indicate that this group of students had realistic perceptions of their abilities, and were willing to report these during the interview. To verify this, it would be necessary to compare the students' self-evaluations with evaluations of language ability made by the English language specialist. However, this finding is consistent with previous research, that has shown that self-ratings and interviewer ratings of English language ability can yield high correlations (Smith and Baldauf, 1982).

A methodological flaw in the interviewing process was that only students who fell below the STAL cut-off of 19 were interviewed, to identify the false classification of a student as experiencing English language difficulties. Students who scored over 19 were not interviewed, meaning that students who were not identified by the STAL, but who did have language difficulties, were not reclassified. Four students who had not been present for the testing session were also interviewed, to establish their language proficiency. The interviewer judged that no intervention was required, but without STAL scores for these students, no comment can be made as to whether this decision was consistent with the screening instrument. Four students decided independently that they had been incorrectly classified by the screening test, and that they would benefit from language interventions. It would be instructive to follow this entire group of first year students over their medical course, and to establish whether students not initially identified by the STAL were later identified through other mechanisms, such as self-referral to the Language Development Programme, or via recommendations by Faculty teaching staff to seek language support. On the basis of their interview, seven students were not offered a place in the Language Development Programme, even though they fell below the cut-off score of 19 on the STAL. With a longitudinal study it may be possible to determine whether this decision could have adversely affected these students' performance, or whether they self-referred for language intervention of their own accord later during the course of their studies.

The information that the students provided about specific areas of difficulty in understanding Australian customs, and speech in Australia, gives medical educators an indication of where to concentrate teaching efforts. For example, the fact that students reported difficulties with Australian slang would suggest that students from non-English speaking backgrounds may need specific instruction in this area, especially in the informal speech frequently encountered in medical or health-related situations. A response to this need has already been implemented for subsequent student cohorts, as a result of this finding (Chur-Hansen and Barrett, 1996) (Appendix IV.II). Teaching initiatives that focus on familiarising students with aspects of Australian culture may also be valuable.

Perceived difficulties in understanding Australian culture were not found to be related to students' STAL performance, in those students who were interviewed. Many students said that they had no difficulties whatsoever in this area. This may be an accurate reflection of their knowledge, or it may be that the question was ambiguous; it may be that some students were uncertain as to what the term "culture" referred. That only three students chose to give further comment to this item lends weight to this notion. Future interviews may need to probe more deeply into students' comprehension of this item, to ensure that valid answers are given.

Answers provided by the students concerning their friendships and feelings about their classmates were sometimes frank, and a rich source of information. That some of the students interviewed felt that they did not enjoy a close friendship in Adelaide may be of concern, if the student was socially or emotionally isolated as a consequence, though this cannot be established from the data collected during this one interview. Negative and mixed feelings towards becoming acquainted with Australian peers were evident during some interviews; this is consistent with the findings of other researchers working with Asian students studying in Australia (Choi, 1997). Given that these interviews were conducted in the first week of the first semester of classes, it may be that the students had not yet had the opportunity to meet class members. A lack of positive experiences and interaction with others from different language and cultural backgrounds may have contributed to students' feelings of inadequacy and shyness. Furnham and Bochner (1986) have reported that a lack of friendship networks has been found to be associated with academic failure and failure to complete a course. It has also been found that students who mix with locals and establish relationships with them report higher degrees of satisfaction with both academic and social experiences (Furnham and Bochner, 1986). A follow-up of the students in this research is necessary to establish whether their attitudes alter over time, and if not, whether reluctance to associate with Australian classmates impacts upon performance in the medical course. Attitudes toward getting to know Australian classmates were not associated with STAL performance in those students who were interviewed in this study. A more extensive interview with students from a range of

language backgrounds and abilities is required to identify factors that might predispose students to feel uncomfortable about mixing with those of different language and cultural backgrounds from themselves. However, from the information gleaned here it is clear that the majority of the non-English speaking background students wanted to mix with local English speaking background friends. This is inconsistent with the conception that overseas students "stick together" (Chalmers and Volet, 1997, p 92).

In their validation studies of the STAL, Hayes and Farnill (1993a; 1993b) and Hayes, Farnill & Sefton (1994) make the assumption that performance in Medical Communication Skills examinations is contingent upon English language proficiency. They note, however, that a number of students who were identified by the STAL as experiencing language difficulties performed to a satisfactory standard in the Medical Communication Skills examination, and postulate that this reflected an improvement in the group's language skills, or that the Medical Communication Skills examination was not assessing language proficiency as does the STAL. These are propositions which require further clarification and investigation. The relationship between medical communication skills and English language proficiency is the subject of a comprehensive research project described in Chapter V.

The correlations found by Hayes and Farnill (1993a; 1993b) between English language proficiency and academic performance in subjects other than Medical Communication Skills over first and second year needs closer investigation. It is necessary to isolate possible confounding factors, such as social, interpersonal, financial, and cultural that may contribute to academic achievement or the lack of it. It may be that these other factors, rather than language proficiency, account for non-English speaking background students' difficulties in the course. A comparison of students from English and non-English speaking backgrounds, of varying levels of language proficiencies, may help to isolate such factors and the role they play. Attitudes toward language intervention programmes should be evaluated, of both programme participants and non-participants, as one important determinant of student outcomes. Where such a programme meets with

non-acceptance, the reasons for this can be documented and addressed; similarly, where the programme is viewed positively, the factors that make it attractive to students can be emphasised and capitalised upon.

Farnill & Hayes (1994) have established norms for the STAL in an Australian medical school with 962 Sydney University first year Medicine students, enabling meaningful comparisons to be made between Adelaide and Sydney populations. This study demonstrates a considerably higher proportion of students identified by the STAL as experiencing difficulty with the English language (28.8%) enrolled in the first year at the University of Adelaide, compared with enrolments at the University of Sydney. This finding supports the wisdom of the University of Adelaide Faculty of Medicine to implement and fund a Language Development Programme to provide assistance for what can be considered to be a substantial number of students.

The results of the present study provided considerable data for the Language Development Committee's immediate use, in allocating students to a teaching programme, which as a finite resource, could only accommodate a limited number of students. The study also served to indicate a number of areas for further inquiry. Many of these can best be addressed through a longitudinal research design, to determine the role that English language proficiency plays in academic performance in an Australian medical school, and to isolate the relative contribution of factors that interact with language proficiency, such as cultural awareness, and the availability of social supports, to successful progression through the medical course. A comparison of this cohort of students, with a similar cohort who have not received substantial language intervention may indicate whether language programmes ameliorate any negative consequences of difficulties with language in terms of academic performance. This would be one way in which the Language Development Programme can be evaluated. Rush (1972) and Levey (1992) have argued that the implementation of training programmes requires rigorous evaluation, both before commencing the programme, at points during, and at the programme's conclusion. Evaluations may be subjective, such as feedback given by patients, medical and nursing

staff and the students themselves, or be more objective, such as examination results and successful completion of the course (Rush, 1972; McGlinn and Jackson, 1989; Levey, 1992). This Chapter has described a procedure for assessing students prior to their entry into an intervention programme. Chapters VIII and IX of this thesis are concerned with subjective and objective data as evaluative tools for assessing the utility of a language-based intervention to improve academic and clinical performance.

Chapter IV.

Study II.

Testing the 1994 third year student cohort for English language proficiency employing the Screening Test of Adolescent Language (STAL) and the Word Knowledge Test (WKT).

Aims of the present study.

The study described in this chapter had three main objectives;

(i) to screen all 1994 third year medical students enrolled at the University of Adelaide for English language proficiency, using the STAL.

(ii) to examine the performance of students on two standardized tests of English language proficiency taken at different testing times.

(ii) to explore the possibility of improvement in English language proficiency over time where none or minimal language tuition is provided.

Method.

Subjects.

Subjects were the 1994 third year medical students. In 1992 Associate Professor Helen Winefield, (Associate Professor in Clinical Psychology at the University of Adelaide Department of Psychiatry), administered the Australian Council for Education Research (ACER) Word Knowledge Test (WKT) to 174 first year students comprising 124 from Medicine, 40 from Dentistry and ten from Health Sciences[17]. Of the 124 first year medical students screened in 1992, WKT results were available for 101 (67.8%) of the 149 students in the 1994 third year cohort.

The WKT (Form F) was administered according to the Test Manual. Students were gathered in a lecture theatre and completed the test under examination conditions.

---

[17] Associate Professor Winefield generously provided all data related to the Word Knowledge Test.

Students were told that testing was compulsory, but that results on the test would not be counted toward any subject grades.

The STAL was administered as closely as possible to the way in which it was for Study I with the 1994 first year students, in the same location, and with the same tester[18].

In order to establish concurrent validity, comparisons are made between test scores obtained at the same time, or after a specified, brief time period (Anastasi, 1988). A substantial period of time (roughly two and a half years) had elapsed between the students' completion of the WKT and the STAL. It was therefore not possible to establish concurrent validity between the WKT and the STAL. Instead, a comparison was made between performance over the two and half year period for those students for whom scores on both measures were available. As there are no data on the concurrent validity between the two tests and considering the period of time between each testing session, findings in this regard must be interpreted with caution.

Instruments.

The Screening Test for Adolescent Language (STAL).

The Screening Test for Adolescent Language has been described in detail in Study I.

Word Knowledge Test (WKT) (Appendix XV).

The Australian Council for Educational Research (ACER) has developed a paper and pencil test, the Word Knowledge Test (WKT), which can be used for screening large groups of subjects at one administration time, to assess familiarity with, and understanding of a range of words. The test takes ten minutes to administer, and can be scored by hand or by machine. The test manual describes common applications of tests such as this, including the selection of applicants for courses that require a minimum level of language ability (de Lemos, 1990).

---

[18] I am again indebted to Dr Rob Barrett for administering the STAL.

The ACER Word Knowledge Test (WKT) (1986) is a revised version of a previous test, the Adult Form B (ACER, 1960, in de Lemos, 1990). The WKT has two parallel forms, E and F, both consisting of 72 items. Each item comprises a word in capital letters, such as "WOOD", followed by five options, lettered A to E, such as "door", "fire", "house", "timber", "tree". The subject must choose the option that is closest in meaning to the prompt word. In this example the subject would answer option D, "timber". Items are of varying difficulty.

Word knowledge tests are generally regarded as measures of acquired knowledge, and as such are included in standardized IQ tests. In this way they could be criticised, as has the STAL, for confounding verbal intelligence with language proficiency (Sommers, 1985).

The WKT manual includes normative data based on over 1,200 Year 9, 10 and 11 Australian secondary school students, and information about the reliability and validity of Forms E and F. The main language spoken at home and parental language background were investigated for their influence on performance on the WKT. Students from non-English speaking home backgrounds, or who had parents from non-English speaking backgrounds were found to perform significantly less well on the WKT than those who had English speaking parents and spoke English at home. The relationship between the parents' ethnic background and WKT performance was also investigated. Students from Anglo-Saxon and Northern European backgrounds were found to perform better than those from Eastern European, Southern European and Asian backgrounds, although the manual stresses that this finding must be viewed with caution, as the results are based upon some small numbers within ethnic categories.

The manual warns that although the WKT is valid for measuring current language skills, verbal reasoning and the adequacy of language skills for a particular purpose, this does not imply that the test can be used to make predictions about potential verbal ability. Thus, in the case of students from non-English speaking backgrounds, low scores may

reflect a lack of exposure to the English language, but do not provide information about the likelihood of improved mastery (or otherwise) with increased exposure.

Results.

The Screening Test for Adolescent Language (STAL).

Of the 149 students in the third year student cohort, 142 students (95.3%) were tested with the STAL. Ages of the tested students ranged from 19 to 35 (mean=21.65); 53.5% were male, 46.5% female. The most common language spoken at home was English (40.1%) followed by Bahasa Malaysia (14.8%). Most students were born in Australia (38.7%) or Malaysia (31.0%). The majority of students were Australian residents (73.2%), and had gained admission to the university's medical school through their matriculation scores (53.5%).

The STAL was scored by one rater (the author). Scores ranged from 7 to 23, ($\bar{x}$=20.01, sd=3.15), and were negatively skewed, with 70.4% of the tested students scoring 20 or over. The means, range and standard deviations for the STAL total and subscales are presented in Table XXXI. The distribution of scores can be found in Table XXXII.

Table XXXI: Means, ranges and standard deviations for the STAL total and subscales.

Scale	Items	Range	$\bar{x}$	SD
Vocabulary	n=12	4-12	10.44	1.56
Auditory Memory	n=3	0-3	2.56	0.74
Language Processing	n=5	0-5	4.48	0.96
Proverb Explanation	n=3	0-3	2.54	0.70
Total Scale	n=23	7-23	20.01	3.15

Table XXXII: Distribution of STAL scores.

Score	N	%
7	1	0.7
8	1	0.7
10	2	1.4
11	1	0.7
14	2	1.4
15	7	4.9
16	4	2.8
17	9	6.3
18	5	3.5
19	10	7.0
20	17	12.0
21	26	18.3
22	31	21.8
23	26	18.3
	142	**100**

Relationship between demographic variables and performance on the STAL.

Females scored significantly higher on the STAL than males (t=-2.68, df=134.11, $p<0.01$). Age was significantly related to the total STAL score (Spearman r=-0.36, $p<0.001$). To further investigate this, the group was divided into two age groups by rounding the mean age up, to 22 years; students over 22 years of age were more likely to be deemed as experiencing English language difficulties compared with those under 22 (x^2(continuity correction)=9.09, df=1, $p<0.005$).

The relationships between age, gender and performance on the STAL were more closely examined. Age and gender were not related. Method of entry into medical school and gender were also not related. A chi-squared analysis to test the association between age and method of entry was conducted by categorising students into four groups; matriculants, tertiary transfer and special entry students; MATES and repeat MATES; repeating students; and overseas students. Matriculants, tertiary transfer and special entry students were significantly younger than MATES students (who were all 22 years of age or older) and overseas students (x^2=58.97, df=3, $p<0.001$).

Language backgrounds were categorised into "English", "Bahasa Malaysia", "Chinese" and "Other". Students whose language backgrounds were Bahasa Malaysia and

Chinese were most likely to be identified by the STAL as experiencing language difficulties, with English speaking background students least often identified (x^2=36.45, df=3, p<0.001). Categorising place of birth into "Australia", "Malaysia" and "Elsewhere" yielded a significant association, with those born in Australia least often identified by the STAL as experiencing language difficulties, and those born in Malaysia and Elsewhere being most frequently identified (x^2=22.14, df=2, p<0.001).

To examine the relationship between STAL performance and the method of entry to medical school, students were divided into three categories; matriculants, tertiary transfers and special entry students; MATES and repeat MATES; and overseas students. Repeat students were excluded from analyses, as their inclusion rendered chi-square analyses as invalid. An association was found between poor performance on the STAL and entry to Medicine via the MATE scheme (x^2=19.21, df=2, p<0.001).

Word Knowledge Test (WKT).

Of the 149 students in the third year student cohort, 101 students (67.8%) had been tested with the WKT. Ages of the tested students in 1994 (approximately two and a half years after having completed the test) ranged from 19 to 33 (mean=21.07, sd=2.46); 51.5% were male, 48.5% female. The most common language spoken at home was English (43.6%) followed by Bahasa Malaysia (16.8%). Most students were born in Australia (38.6%) or Malaysia (37.6%). The majority of students were Australian residents (75.2%), and had gained admission to the university's medical school through their matriculation scores (54.4%).

The WKT was scored by hand by one rater[19]. Scores were normally distributed, and ranged from 19 to 70 out of a maximum possible score of 72 ($\bar{x}$=49.74, sd=11.6). These results were not significantly different to the results of the 124 students tested in 1992, (range=16-71, $\bar{x}$=49.8, sd=12.0), of which the 101 students in this study are a sample. In 1992, students who fell at approximately one standard deviation below the

[19] Ms Melissa Raven scored the Word Knowledge Test.

mean (a score of 35 or less) were advised to seek language related assistance from university student support services. There were 19 students included in this category in 1992. In 1994, 14 of these 19 students had progressed to third year. Of the five remaining students, four had failed in their first or second year studies, and one student had withdrawn from the medical course, to pursue a degree in engineering.

Relationship between demographic variables and performance on the WKT.

The WKT total score was not related to gender, age, language background, place of birth or method of entry. Performance on the WKT was related to overseas vs permanent Australian resident status, with permanent Australian residents more likely to score above the group mean of 49.7 out of a possible 72 (rounded to 50 for analyses) (x^2(continuity correction)=14.11, df=1, p<0.001).

Relationship between the STAL and WKT.

As has been stated earlier, it is not possible to establish concurrent validity between the WKT and the STAL. However, it is feasible to consider correlational data between the two measures, bearing in mind that the following results are based upon a subset 101 of the original 124 students tested with the WKT in their first year of Medicine, (in 1992) and 142 students tested with the STAL in their third year of Medicine, (in 1994).

STAL scores were not available for 26 students for whom there were WKT scores, either because the student had not attended the STAL testing session in 1994 (n=4), or because they had not proceeded to third year as would have been anticipated, by 1994, either due to academic failure (n=13), withdrawal (n=5) or intermission (n=2) from the course, or because the student had gained admission to the medical course through a special entry scheme, and had not been required to progress sequentially through the medical course (n=2).

WKT scores were not available for 45 students for whom there were STAL scores, either because they did not attend the WKT testing session in 1992 (n=3), or due to

academic failure the student was in third year, when chronologically a higher year level should have been achieved (n=35), or because the student had gained admission to the medical course through a special entry scheme, and had commenced studies at the second or third year level (n=7). Neither STAL nor WKT scores were available for three students.

Spearman correlation coefficients were significant for all STAL subtests and whether the students was identified by the STAL as experiencing English language difficulties (by scoring 19 out of 23 of less) with performance on the WKT (Table XXXIII).

Table XXXIII. Correlation coefficients between the STAL and the WKT.

STAL Subtests	r^2 with WKT (n=96)
Vocabulary	0.67*
Auditory Memory	0.58*
Language Processing	0.21†
Proverb Explanation	0.44*
Total Scale	0.66*
Identified Difficulty	-0.42*

*significant at $p<0.001$ †significant at $p<0.05$

As could be expected, of the four subtests the Vocabulary subtest of the STAL was most highly (although modestly) correlated with the WKT, which is itself a test of vocabulary. The STAL subtests of Proverb explanation and Auditory memory showed modest but significant correlations with the WKT score, with Language processing demonstrating a low but significant relationship. It may be that the cognitive tasks required by the STAL on the Language processing and Proverb explanation subtests require different abilities, as compared to the Vocabulary subtest, which is most akin to the WKT. Nevertheless, the STAL Total scale score and the WKT were significantly correlated at 0.66.

Discussion.

The STAL and WKT scores demonstrate a range of English language abilities in the 1994 third year medical student cohort. The data from the WKT are interesting, but the STAL data should be viewed as more reliable, given that the WKT is of historical importance for these students, and also because STAL data were available for 95.3% of the cohort, as compared with only 67.8% of the cohort on the WKT. Because concurrent validity between the STAL and the WKT cannot be established from this study, conclusions as to whether students had improved in their English language proficiency as measured by these tests over the two and a half year interval from their first to third year of medical school are dubious. However, Farnill and Hayes (1996c) found in a more controlled study that students whose English was poor on commencing a medical degree did not improve significantly over time as a result of simply studying within the course.

Older students were found to perform more poorly on the STAL than younger students; this can be explained as a function of the fact that MATES students and overseas students were older in general than the majority of their classmates. Why females should perform better than males is difficult to explain, but this may reflect a true difference in favour of the females in the cohort.

The correlations between the STAL and WKT measures should be considered in light of the fact that in 1992, 124 first year medical students were tested with the WKT; by 1994, 101 students (81.5%) had progressed to third year. It could be proposed that the 23 (18.5%) who had not reached third year by 1994 may have been hampered in their studies by difficulties with English; this would result in a truncated range of WKT scores by 1994, and could account for the lack of significant relationships found between the demographic variables of gender, age, language background, place of birth and method of entry and WKT scores. Similarly, it should be remembered that WKT scores were not available for 48 (32.2%) of the 1994 third year cohort.

Paper-and-pencil tests like the STAL and WKT are useful instruments for screening large numbers of subjects at one administration time (Hayes and Farnill, 1992). However, these tests, administered as they have been in this study, assess only listening (aural) and, to a limited extent, writing skills. They do not indicate proficiency in reading or speaking.

A number of researchers have stressed the importance of verbal communication in the medical encounter. As described in Chapter I, Vernon-Roberts (1991) also suggested that clinical educators were most concerned with students spoken English rather than their ability to listen, read or write. The following study sought to investigate more closely students' spoken language proficiency, in the context of a mock medical interview.

Chapter V.

Study III.

The relationship between language background, English language proficiency and medical communication skills.

Aims of the present study.

The study described in this chapter had one main objective:

(i) to investigate the relationship between language background, English language proficiency, and medical communication skills in the third year cohort. The aim of the study was to determine whether coming from a non-English speaking background is associated with lower proficiency in English, and if this lack of proficiency hampers the demonstration of medical communication skills in an examination setting.

Introduction.

Medical communication skills, which include attending, listening, responding, probing and challenging, open-ended, and closed questioning, demonstrating empathy and building rapport (Egan, 1990) are recognised as being fundamental to effective health care (Friedman and DiMatteo, 1982; Roter and Hall, 1993; Del Mar, 1994; Sanson-Fisher and Cockburn, 1997). Because of the importance of these skills, many medical schools have implemented training programs designed to develop medical communication skills in their students, as recommended by Doherty (1988). As an evaluation of the long-term effectiveness of such training, Rolfe and Pearson (1994) compared the communication skills of interns practising in New South Wales who had trained either in Australia, New Zealand, or an overseas institution. The authors concluded that graduates from the University of Newcastle, New South Wales, were significantly better than other interns in the demonstration of medical communication, and that the communication skills of foreign graduates were rated less favourably by comparison with the other graduates. Although language background was not controlled for, this finding was attributed to difficulties with language and a lack of training in interpersonal skills. Glover (1995) and Young (1995)

have argued that this is a serious flaw of the research. Torda (1995) has suggested that although it may be that foreign graduates perform poorly in communication skills in English, this may not be the case were they to be examined in their first language.

Method.

Subjects.

All 149 students (54.4% males, 45.6% females) enrolled in the third year of an undergraduate medical degree at the University of Adelaide in 1994 took part in this study. The demographic profile of this group is given in Tables VII to XII, above.

Procedure.

Students had been assessed under examination conditions for written and aural language proficiency with the Screening Test for Adolescent Language (STAL) (Prather, Breecher, Stafford and Wallace, 1981), as described in Study I. Six weeks later they were examined by means of an Objective Structured Clinical Interview (OSCI) to assess their medical communication skills, following a 15 hour interpersonal skills training programme.

The Instruments.

The Screening Test for Adolescent Language (STAL).

This measure has been described in Study I.

The Observed Structured Clinical Interview (OSCI).

The Observed Structured Clinical Interview (OSCI) took the form of an examination of medical communication skills, designed and co-ordinated by Dr Sheila Clark from the University of Adelaide's Department of General Practice[20]. The OSCI examination followed a fifteen hour training course, which involved two one hour lectures, three workshops of three hours each and two sessions of two hours working with a general practitioner and his or her patients. During the workshops students participated in role

[20] I am grateful for the assistance of Dr Sheila Clark, who generously made all data generated from the 1994 OSCI freely available to me.

plays to practice communication skills and to prepare for their general practice attachment and their OSCI examination.

Feedback for students' progress was formative, and utilised self-evaluations through comprehensive communication skills checklists. Videotapes of interviews involving the student and a simulated patient were reviewed by the student in conjunction with a course tutor, and rated on self-assessment proformas. During the attachment with the general practitioner the student interviewed a 'real' patient , and audiotape-recorded the interaction. The student was then required to self-assess their communication skills performance after listening to their tape-recording.

The nature and requirements of the OSCI examination were fully explained to students prior to the commencement of communication skills training. A sample marksheet for the examination was provided with the course handbook, and it was stressed that satisfactory performance in the OSCI examination was necessary for the student to proceed to the fourth year of the medical curriculum. Students whose standard was unsatisfactory in the OSCI examination were offered a second attempt following further training. Those who were unsuccessful in their second examination were required to repeat the third year of their medical degree.

In the OSCI examination, each student interviewed one of fifteen roleplaying standardized patients for a maximum of ten minutes, observed by one of a pool of seven medical practitioners. The cases, standardized for complexity, were emotional problems presented in a clinical context, such as a grieving situation being used to request a prescription for hypnotics. The examination was criteria scored by the clinician and standardized patient following examiner training.

Twelve facets of the interview were scored on Likert type scales ranging from 1 (poor) to 5 (excellent), with a midpoint of 3 (adequate) (Appendix XVI). These twelve skills were;

1. the student's introduction including explanation of the purpose of the interview and obtaining informed consent from the patient;
2. exploration of the patient's knowledge of the problem;
3. exploration of their concerns;
4. summarising, checking and finishing the interview;
5. rapport and attending;
6. listening;
7. appropriate use of questioning techniques;
8. empathy skills;
9. personal and professional qualities, such as warmth, concern, sincerity and respect;
10. the use of clear, unambiguous and fluent English;
11. the use of appropriate language to the person and context; and
12. the overall feeling of empathy, reflected in the creation of an atmosphere of trust, support and sensitivity to the patient's emotional state.

Results.

Screening Test of Adolescent Language (STAL).

The STAL results for this cohort of students are described in Study II.

Observed Structured Clinical Interview (OSCI).

When performance was expressed as a score out of a total of 60, outcomes ranged from 23 to 60. Adjusted scores for assessment purposes, with a pass grade of 24 out of 48 (mean=32.49, sd=9.39, range=11 to 48), resulted in 26.8% (n=40) of the cohort failing to satisfy the required standard for the examination.

Ratings of spoken language made by the standardized patient and clinician on a scale from 1 (poor) to 5 (excellent), with a midpoint of 3 (adequate), resulted in the majority of students (85.9%, n=128) assessed as being "adequate" to "excellent" in their use of "clear, unambiguous and fluent English", with the remainder scored as "poor" (3.4%, n=5) to "less than adequate" (10.7%, n=16) (mean=4.12, sd=1.12).

The relationship between language background, English language proficiency and medical communication skills.

Chi-square analyses were conducted employing continuity correction to test for associations between variables. Percentages are of the cohort totals.

Gender was not associated with language background, overseas student status, performance on the STAL, medical communication skills, or ratings of spoken language fluency.

One hundred and forty-nine students were examined in medical communication skills during the OSCI. Of the 40 students who failed the requirements of the OSCI, 14 (9.4%) were native English speakers. Twenty-six (17.4%) of the 88 students from a non-English speaking background did not satisfy the requirements of the examination. Chi-square analysis to test for an association between language background (English speaking vs non-English speaking) and performance on the OSCI (pass vs fail) was not significant.

Of the 42 students identified by the STAL as likely to be experiencing difficulties due to language, 16 (11.3%) failed to reach a satisfactory level in medical communication skills assessed during the OSCI. Similarly, 21 (14.8%) of the 100 students with no identified language difficulties did not pass. This association between written and aural language proficiency and the demonstration of medical communication skills approached, but failed to reach, significance (x^2=3.64, df=1, p=0.06).

Analyses on the standardized patient and clinician's rating made on the basis of speech during the OSCI of "clear, unambiguous, fluent language" (satisfactory vs unsatisfactory) rendered an association with language background (x^2=12.43, df=1, p=0.00), overseas student status (x^2=28.78, df=1, p=0.00), and performance on the medical communications examination (x^2=15.66, df=1, p=0.00), with students identified as having unsatisfactory fluency of speech more likely to fail the OSCI examination. The STAL and rating of the fluency of speech were also associated with each other: students identified as

experiencing written and aural language difficulties on the STAL also tended to be identified as having unsatisfactory spoken language as rated during the OSCI (x^2=20.88, df=1, p=0.00).

Spearman *rho* correlation coefficients for medical communication skills and language background, performance on the STAL, and the standardized patient and clinician ratings of "clear, unambiguous and fluent English" and the use of "appropriate language for the person and context" during the OSCI are listed in Table XXXIV.

Table XXXIV: Spearman *rho* correlation coefficients for medical communication skills and language background, STAL performance, and the ratings of "clear, unambiguous and fluent English" and the use of "appropriate language for the person and context" during the OSCI.

Medical Communication Skills demonstrated during OSCI	Language Background	STAL	Clear, unambiguous and fluent English during OSCI	Appropriate language for person and context during OSCI
Introduction	NS	.28*	.23*	.24*
Exploration of problem	NS	.18*	.32*	.35*
Exploration of concerns	NS	NS	.45*	.35*
Summary, checking, finishing	NS	NS	.24*	.20*
Rapport and attending	NS	.26*	.27*	.17*
Listening	NS	.17*	.43*	.30*
Questioning	NS	.21*	.30*	.29*
Empathy skills	NS	.21*	.37*	.32*
Personal and professional qualities	NS	NS	.16*	.18*
Clear, unambiguous, fluent English	NS	.57*		.65*
Appropriate language	-.22*	.31*	.65*	
Overall empathy	NS	.22*	.37*	.28*
Final OSCI mark	NS	.31*	.55*	.46*
Satisfactory/Unsatisfactory mark	NS	-.19*	-.32*	-.27*
STAL performance	-.34*		.57*	.31*
Language background		-.34*	-.22*	NS

*p<0.05

Discussion.

The purpose of this study was to investigate the relationship between language background, English language proficiency and medical communication skills in undergraduate medical students studying in Australia. Results indicate that, in this cohort,

coming from an English speaking background does result in better performance in a screening test of written and aural English language proficiency, but does not result in the more satisfactory demonstration of medical communication skills, as rated by clinicians and standardized patients in an examination setting. The strongest predictor for satisfactory performance in medical communication skills for this cohort was a rating of the fluency of spoken language.

Results on the STAL yielded a range of scores comparable with those of other Australian medical student cohorts (Hayes and Farnill, 1993a; 1993b). The majority of the cohort experienced no difficulties with aural and written English. Students identified as experiencing language difficulties, and especially those scoring at the lower end of the distribution may be functioning with substantial language problems.

The reliability and validity of the STAL (Hayes and Farnill, 1992) and OSCI examinations have been discussed elsewhere (Newble, Hoare and Elmslie, 1981). It is possible that students identified by the STAL as experiencing language difficulties may be false positive cases, although research on the selectivity and sensitivity of the test in comparable populations would suggest that this is not the case (Farnill, Hayes and Chur-Hansen, 1995) (Appendix I.II).

There are four aspects of language proficiency - reading, writing, listening and speaking (International English Language Testing System, 1989). The STAL is designed to screen writing and listening, or aural skills. Reading and speaking skills are not assessed. The association of the rating of fluency of speech made by the examiners during the OSCI examination with performance on medical communication skills and the STAL would suggest that this aspect should be taken into account in future research. Oral language proficiency may be an especially pertinent variable to consider because listening and reading are often referred to as "passive" language skills, and may be more highly developed in a non-English speaker compared to the "active" skills of speaking and writing (Bradley and Bradley, 1984, p 193). Thus, the active skills may be a more useful

indicator of proficiency. In the United States, Friedman and her colleagues (1993, 1991) have assessed spoken language proficiency with graduates from foreign institutions in a clinical setting using standardized patients, and conclude that the concurrent validity of such assessment with standardized language tests and predictive validity with clinical competence is high.

More extensive screening and testing of the active language skills are necessary in settings where non-English speakers comprise a significant proportion of the undergraduate medical population, as is the case in several Australian universities. Rather than assessing only the global fluency of speech, the specific components of speech that may hinder performance in medical communication skills require investigation. These include accent, rate of speech, the ability to use and understand colloquial language, and the use of appropriate grammatical constructs.

Inspection of the correlation coefficients in Table XXXIV demonstrates moderate correlations between the ratings of clear, unambiguous, fluent spoken English and both STAL scores ($r=0.57$, $p<0.05$), and the appropriate use of language for the person and context ($r=0.65$, $p<0.05$). All other correlations which are significant are also comparatively low, indicating that spoken language accounts for some, but not all, of the variance. Other, unexplained factors apart from spoken English must therefore also be implicated in the demonstration of satisfactory medical communication skills. Nevertheless, the role of language is obviously of importance given the number of significant correlations.

Table XXXIV also shows clearly that language background *per se* has very little impact upon ratings of medical communication skills, apart from the use of appropriate language for the person and context. It is plausible that language and cultural background may be confounded here, as possibly it is one's knowledge of Australian medical cultural conventions which mediates performance in this arena, rather than spoken English proficiency.

An important area of inquiry is the early identification of students who may experience difficulties in the acquisition and demonstration of medical communication skills as the result of their language skills, so that intervention strategies to assist these students can be designed and implemented. It would be instructive to determine the differences between those students from non-English speaking backgrounds who are hampered by language proficiency and those who are not. Of course, it would be reasonable to suggest that there are factors other than linguistic which may influence students' performance in medical communication skills which must be isolated and not confounded with language proficiency or background (Vernon-Roberts and Chur-Hansen, 1995) (Appendix I.I); Smith, Hamilton, Rolfe and Pearson, 1995). These might include cultural background, assimilation into the Australian culture, personality factors and attitudes toward studying medicine. Attitudinal influences might include whether the student is studying through choice or through familial expectation or other external pressures, and whether the student has a psychosocial or biomedical orientation; the latter may be less motivated to practice and improve their medical communication skills (Marteau, Humphrey, Matoon, Kidd, Lloyd and Horder, 1991).

This study demonstrates that whilst language background and English language proficiency are associated, unsatisfactory performance on medical communication skills examinations such as the OSCI cannot be attributed to language background alone. Written and aural English language proficiency are also poor indicators of performance in examinations of medical communication skills. In the present study, spoken fluency was most likely to be associated with satisfactory performance in examinations that assess these skills. The specific aspects of spoken language which might be considered important by practising clinicians and educators in the medical setting are further explored in the following chapter.

Chapter VI.

Study IV.

Clinicians' perceptions of students' language difficulties at the University of Adelaide Medical School.

Aims of the present study.

The study described in this chapter had one main objective:

(i) to collect written comments about students' language abilities, made by their clinical teachers, in order to establish the occurrence of individual difficulties and to obtain an overview of clinicians' perceptions of this issue.

Introduction.

There are a number of studies which have surveyed university students from non-English speaking backgrounds to assess their difficulties and learn of their experiences. However, very few researchers have been interested in the perceptions and opinions of the teaching staff concerning their non-English speaking background students.

Samuelowicz (1987) surveyed academic staff and overseas students from 50 departments including those in the Faculty of Medicine at the University of Queensland, Australia, to establish their ideas about the nature and extent to which educational problems existed. Both staff and students concurred that language problems were significantly hampering adjustment to tertiary education and to Australian culture.

Felix (1992) interviewed academics from several departments, including the Faculty of Medicine, to ascertain agreement between non-English speaking background students and their teachers regarding English language competency. She found that generally speaking, academic staff tended to rate students' ability lower than the students rated themselves, although both groups agreed that language difficulties were impacting upon academic, and in the case of Medicine, clinical performance. However, the results of

this study should be viewed in light of the limitations of the methodology. Felix (1992) asked staff to assess "the competence of a typical NESB (non-English speaking background) student in their course" (p 8) and these ratings were then compared with students' self-ratings. As she herself says, there is no such student as a "typical NESB", and thus, the validity of the answers given in response to this prompt may be questionable.

Concern at the University of Adelaide regarding students' English language proficiency expressed by academic staff and clinical tutors had been documented in the Minutes of Faculty Committees and letters to the Dean of Medicine from Heads of Departments as well as in internal Departmental reports (Winefield, 1988). In 1991, a proposal for Faculty-based support for students from non-English speaking backgrounds was put forward by Dr Jane Vernon-Roberts, the Clinical Studies Adviser for the Royal Adelaide Hospital. Her proposal was based upon the collective voice of academic and clinical teaching staff, who had described to her through both formal and informal avenues their thoughts and experiences about the challenges faced by these students. As described in Chapter I of this thesis, in her proposal she summarised five common areas of difficulty that had been identified; (1) a reticence and lack of assertiveness in some students which resulted in a failure to respond to and ask questions in tutorial and clinical settings. Although this was attributed by staff to poor social skills, it was acknowledged that language skills and cultural restraints may have played a significant role in limiting participation; (2) an insufficient command of informal, idiomatic English; (3) a lack of experience with small group, interactive teaching and learning; (4) a reliance upon rote learning; and (5) the inability of some students to grasp the importance of a high level of English language proficiency for the study and practice of medicine at an Australian institution (Vernon-Roberts, 1991).

However, the comments and concerns regarding students' English language proficiency were either anecdotal or did not specify the ways in which students' language skills were wanting. No quantifiable data had been collected to support the overall claim

that the English language skills of some students were inadequate for the tasks of studying at tertiary level and for interacting with other students, staff and patients.

No systematic method of commenting upon students' academic or clinical performance was implemented until the students reached the third year of their course. This was the year in which the students were first assessed by clinical teachers on the wards. Therefore, it was not possible to collect comments regarding first and second year students. Students were assessed in writing by their clinical teachers in their third, fourth, fifth and sixth years of the medical course. Clinical tutors were provided with an assessment sheet, designed by the Clinical Studies Adviser[21]. This sheet required that the tutor give a graded assessment, in which the student was scored from Distinction (75% or above) to Fail (below 50%), on the basis of their performance on the ward. In addition, tutors were asked to make comments about the students' performance on the assessment sheet where appropriate. The graded assessment and any comments made about third year students were based upon demonstrated competence in taking a history from patients, conducting physical examinations, and the presentation of a case to staff on the ward, as well as the student's overall attitude, knowledge, enthusiasm and attendance. Additionally, fourth, fifth and sixth year students were assessed on establishing rapport with patients, the ability to synthesize signs and symptoms, ward involvement and general knowledge of the patients on the ward. Commenting by tutors was not mandatory, and with the exception of one assessment sheet (the 1993 third year students' sheet), no prompt for specific types of comments was given. In 1993 the tutor was specifically invited to pass comment on individuals level of English language proficiency, if they wished to do so. Examples of the tutor assessment sheets are presented in Appendices XVII.I and XVII.II.

---

[21] The position of Clinical Studies Adviser was held by Dr C. Dearlove at the Queen Elizabeth Hospital, Adelaide, South Australia, from 1992 to 1994, and Dr J. Vernon-Roberts at the Royal Adelaide Hospital, Adelaide, South Australia, from 1991.

Method.

Subjects and Procedure.

The written comments concerning five groups of students who were assessed by their clinical teachers over a period of four years were examined for mention of language related difficulties[22]. Details of the five groups are summarised in Table XXXV. Group 1 comprised 69 1991 fifth year students; Group 2 were 114 1992 fifth years; Group 3 were 114 1992 fourth year students, whose comments were also available in 1993, when they were in fifth year; Group 4 were 127 1992 third year students, whose tutor comments were also examined in their fourth year (1993) and fifth year (1994); and Group 5 were 144 1993 third year students, whose tutor comments were also checked for their fourth year of the course (1994). Numbers for each group of students differ because assessment sheets were not available for inspection for all students over the 1991 to 1994 period. The assessment sheets for sixth year students were not available for inspection.

In total, the assessment sheets of 568 students were checked for comments about English language proficiency in this study.

Table XXXV: Students groups by Year, Level and Number.

Group	Year	Student Level	Number
1	1991	V	69
2	1992	V	114
3	1992-1993	IV-V	114
4	1992-1993-1994	III-IV-V	127
5	1993-1994	III-IV	144
			568

Two different tutors assessed each student, resulting in 1136 potential opportunities for comment. All comments that mentioned the student's language abilities were transcribed, verbatim, from the assessment sheets, without identifying the student, level of study, or year. A copy of the transcripts can be found in Appendix XVIII. These

[22] I am grateful to Dr Jane Vernon-Roberts for making the tutor assessment sheets available to me.

comments were categorised by "nature of comment" by the author. These coding categories were then applied by a second, independent rater[23].

Results.

Frequency of comments

Not all tutors took advantage of the opportunity to make written comments about their students, and chose only to mark the students' grades on their assessment sheets. A count of the number of tutors who did as opposed to did not comment in writing about students was not conducted. However, it was the author's impression that over 75% of the assessment sheets recorded a grade only.

The number of comments made specifically about English language proficiency, the year in which the comment was made (1991 to 1994), the level of the student (third to fifth), the number of students involved and the number of clinical tutors who elected to comment in writing on assessment sheets are summarised in Table XXXVI.

Table XXXVI. Number of comments made about students' English language proficiency by Year, student level, and tutors.

Year	Student Level	Comment (N)	Students (N)	Tutors (N)
1991	V	2	2	2
1992	III	10	9	9
1992	IV	7	7	6
1992	V	2	1	2
1993	III	15	13	14
1993	IV	10	9	9
1993	V	2	2	2
1994	IV	6	6	5
1994	V	2	1	1
		56	**47**	

The total number of comments made about English language proficiency was 56. These comments were made concerning 47 individual students, as seven students were commented upon twice by two different tutors, five in the same year, and three in

[23] Ms Anne Francis acted as rater for this task.

consecutive years. One student was commented upon three times by three different tutors over two consecutive years. In total, 36 different tutors were responsible for making the comments. The total number of tutors listed in Table XXXVI does not sum to 36 because sometimes the same individuals made comments about students' language in different years.

Nature of comments.

Comments were classified into thirteen categories; personality factors and language, lack of familiarity with colloquial language, the student's rate of speech, difficulties in differentiating when to use "professional" language and when to use informal speech, comprehensibility of speech, pronunciation, poor interpersonal skills and language, difficulties with patient interactions due to language skills, difficulties with staff due to language skills, language and cultural problems, language skills hampering the student's ability to relay knowledge, and difficulties in "conversational" speech. The final category included comments where language proficiency was commented upon as being problematic, but no particular aspect of the language was specified.

The author and the second rater agreed outright on 29 (51.8%) of the 56 comments. The main reason for this modest agreement was because 31 comments encompassed more than one aspect of language, and could therefore be coded on up to four categories; in total, 83 instances of coding were possible from the 56 comments. Taking into account partial agreement in addition to full agreement, the raters concurred on all but six of the comments, or in 89.3% of cases. Where there was any lack of agreement, the author re-examined the comments, and in light of the second raters codings, made a final judgement. The resultant categorisations are listed in Table XXXVII, whilst the transcribed comments are presented according to their coding categories in Appendix XVIII.

Table XXXVII: Frequency and Nature of Comments made by Clinical tutors about students' English language skills.

Nature of comments	Frequency	%
Colloquial language	5	6.02
Comprehensibility	2	2.41
Conversation difficulties	2	2.41
Cultural difficulties	2	2.41
Difficulties talking with patients	7	8.43
Difficulties talking with staff	1	1.20
Interpersonal skills and language	1	1.20
Personality factors	10	12.05
Pronunciation	1	1.20
Rate of speech	5	6.02
Relaying of knowledge	8	9.64
Register	2	2.41
Specific problem not mentioned	37	44.60
	83	**100.00**

Representative examples of the types of comments most frequently made include the following:

Colloquial language:

"Needs <u>much</u> more practice with idiomatic colloquial English in conjunction with patient contact. I think the University needs to address this issue; lack of familiarity with colloquial English is preventing this student from getting adequate patient exposure."

Difficulties talking with patients:

"She appears to have a significant language and possibly cultural block to communication with patients which severely inhibits her clinical ability."

Personality factors:

"Shyness seems to be more of a problem than language but difficult to judge."

Rate of speech:

"Language will be a problem as tends to speak very fast when under pressure."

Relaying of knowledge:

"Struggling with her conversational English. Knowledge better than she can express. A very keen student. She attends all clinic activities and tries very hard."

Specific problem with language not mentioned:

"Major communication problems with language. Reasonable theoretical knowledge/examination technique."

Discussion.

Although this study cannot lay claim to a rigorous methodology, and was intended as an exploration of the area, it was useful in determining the occurrence and nature of the comments being made about students' language skills.

Fifty-six comments from 1,136 potential comments would seem a somewhat unimpressive finding in this study. However, it is important to consider that a substantial number of tutors did not comment at all on the assessment sheets. The unsolicited nature of 41 of the total number of comments is also noteworthy. Since clinical teachers felt the need to pass comment on their students' language skills, this would indicate that a problem in this domain was perceived by them. Clinical teachers were not specifically requested to pass comment upon language, apart from one year. Where they were invited to comment on language, for the 1993 third year group, the rate of commenting increased, with more tutors offering their opinion. Interestingly, the 1993 fourth year students were also more frequently commented upon concerning their language, even though the invitation to comment in this area was not included on the assessment sheet. It may be that clinicians were more sensitised to students' language related performance after being invited to comment. Or, it may be that they felt that permission had been granted to make evaluations on what could be considered outside of their areas of expertise.

Medical training stresses objectivity and universalism, the notion put forward by Parsons (1952) that doctors should treat all patients as equal to each other, and not allow

non-medical details, such as race, sex or socio-economic status influence treatment. Thus, it is plausible that doctors might generalise the principles of universalism to their students to some extent, and fear that passing judgement on an individual's language proficiency could be interpreted by others as having racist connotations. Charges of discrimination within medicine towards foreign colleagues have been levelled by several authors, including Varki (1992) in the United States, Esmail, Nelson, Primarolo and Toma (1995) and Esmail and Everington (1993) in the United Kingdom and the Human Rights and Equal Opportunities Commission (reported by Gerber, 1991) in Australia. Some doctors may be reluctant, therefore, to make public statements about the adequacy of another's language proficiency, or to commit themselves to paper. An investigation in this sensitive area would be fruitful, but is beyond the scope of this thesis.

The distinction made in an ethnographic study by Barrett (1996) between "professional" and "unprofessional" communication is also useful in further exploring why many clinicians may not have felt comfortable in committing comments about students' language skills to paper. Professional behaviour in this context refers to formal, objective and non-judgmental evaluations which can be defended rationally in a public domain and which can be written; unprofessional refers to informal, subjective and judgmental opinions which are usually conveyed verbally to colleagues. Whilst clinicians in the present study were willing to formally express concern about groups of students language difficulties and informally discuss individuals, it appears that many were less confident in formally criticising any one particular student, who could later view the written comments and perhaps challenge these and the clinician who wrote them. This is a situation analogous to case notes, where clinicians may be reluctant to record their true impressions about patients, a point which was first made by Garfinkel (1967).

The nature of the comments that were written about student's language is valuable information, as it gives structure to the nebulous claim that a student has "language difficulties". Practitioners of medicine are not trained in the assessment or evaluation of spoken language, but as experts in their fields of clinical medicine it could be argued that

they are well qualified to decide whether a student's language ability is of the standard required to perform clinical tasks satisfactorily. Conversely, it could be suggested that without training in the assessment of language, tutors may make unsubstantiated judgements based upon fragmentary information, or upon factors not necessarily related to English language proficiency, such as personality or appearance. In order to explore the accuracy of clinical tutors' comments about individual students' language skills, a controlled research study would be required. Such a study could also explore the supposition that some tutors are more sensitive to language-related issues than others, and thus notice and comment upon this aspect of the students' performance than others. Similarly, some clinical teachers may believe that they are the only ones who think that a particular student has a problem, resulting in a failure to identify and deal with the issue in question (Hunt, Khalid, Shahabudin, Jaafar and Carline, 1995). Without surveying the tutors involved, this can only be speculation.

The types of comments made by tutors demonstrates that spoken language was considered inadequate on several different dimensions. Although 44.6% of the comments did not indicate the reason why the student's language was considered inadequate, the remainder identified areas of weakness. Of the twelve specific areas encompassed by the comments, several may reflect abilities other than language. Difficulties in carrying out a conversation, problems with cultural morés, poor interpersonal skills and personality factors may be problems that require social skills training rather than being a consequence of English language ability. Because of the brevity of most comments the meaning of "struggling with conversational English" for example, is not clear. It is possible that the clinical tutors perceived language to be weak where social skills were inadequate, or vice versa. A further explanation involves the possibility that clinicians were either not aware or unable to express the specific features of students' language that made it seem problematic. Fluency in any given language is dependent upon the sophisticated use of paralinguistic and non-verbal behaviours (Poyatos, 1984) as well as verbal proficiency. Where these features are absent there may be a sense that language skills are lacking, yet it may be difficult to identify and articulate what is missing. These are postulations which

cannot be tested from the information collected for this study, but would lend themselves well to studies of conversation analysis within an ethnographic framework (Heritage, 1984).

Pertinent to the findings of this study are those of Hunt, Khalid, Shahabudin, Jaafar and Carline (1995), who found that Malaysian students studying in Malaysian medical schools were more frequently rated as "shy" by their clinical teachers as compared to United States students studying in Washington. They also discovered that Malaysian students viewed "problem" students as those who challenge and question their teachers. Thus, it may be that in this Australian study, clinical teachers are confounding cultural behaviours with language difficulties. That is, perhaps Asian students are regarded as having "language problems" because they are not vocal and do not question their teachers, when in fact they are obeying cultural rules of respect.

Hunt, Khalid, Shahabudin, Jaafar and Carline (1995) make the most important point that often students who are labelled as problematic may simply have failed to have reached their teacher's expectations, rather than have a problem *per se*. For example, it may be that some clinicians perceive that a student has difficulties with language if they do not use the "Queen's English", whilst others may be more flexible in the standards that they deem acceptable. It is possible that some English speaking background clinicians may feel that "perfect" language skills should be taken for granted and a common sense prerequisite for any entering medical student who seeks membership within the ranks of their profession (Garfinkel, 1967, in Cuff and Payne, 1980; Schutz, 1976). Differences between clinical teachers, including their own language background and level of English language proficiency, and their experiences both in teaching and working with non-English speaking background students and colleagues should thus be considered before accepting any evaluations of students' English language skills.

The comments that specified problems with colloquial language, comprehensibility of speech, difficulties communicating with patients and staff members, pronunciation, the

rate of speech, confusion about the appropriate registers in a given situation, and the inability to relay clinical knowledge due to the student's command of English are aspects of language that require further exploration. That specific aspects of spoken English are identified is valuable knowledge when planning teaching programmes to improve students' performance. Programmes can be designed to address common areas of difficulty, or tailored to suit individual student needs.

All of the comments made by the clinical tutors concerned spoken language proficiency, consistent with the report made by Vernon-Roberts (1991). Reading, writing and listening skills were not mentioned. Given that the writing of case records is an important clinical task, it is surprising that no indication of the students' ability to record their interviews or patient observations was given. The two studies that follow sought to investigate students' written language proficiency in addition to their oral skills.

Chapter VII.

Study V.

Language Rating Scale (LRS) Pilot Study.

Aims of the present study.

The study described in this chapter had two main objectives:

(i) to develop a reliable and valid instrument that could be used by trained standardized patients to rate students' spoken English language in the context of a clinical encounter.

(ii) to administer that instrument as a pilot study to the 1994 third year student cohort.

Introduction.

A call was made in 1969 by participants of the first symposium held in the United States of America on the problems faced by foreign medical graduates, for research into the relationship between performance on standardized tests of English and the doctor's actual use of English in the hospital setting (Sutnick, 1970a). A review of the research that has investigated the English language proficiency of foreign medical graduates as rated by standardized patients and preceptors during clinical interviews has been discussed in Chapter II. To date, researchers have not considered assessing undergraduate medical students' English language proficiency with such a methodology and the present study and the study that follows it sought to address this.

This research followed closely the methodology employed in previous studies (Friedman, Sutnick, Stillman, Norcini, Anderson, Williams, Henning and Reeves, 1991; Stillman, Regan, Haley, Norcini, Friedman and Sutnick, 1992; Friedman, Sutnick, Stillman, Regan and Norcini, 1993; Sutnick, Stillman, Norcini, Friedman, Regan, Williams, Kachur, Haggerty and Wilson, 1993) which have evaluated English language proficiency through the ratings of standardized patients. A mock medical interview was

conducted in order to evaluate spoken and written language abilities because it has been argued strongly that "context" is very important if meaningful conclusions are to be drawn (Helman, 1984). Not only must a speaker be competent in the grammar, pronunciation and semantics of a language (Crystal, 1976), but must also have competency in knowing the rules about when to say what, where - that is, with regard to the context of the situation in which they use language (Bilmes and Boggs, 1979). Patterson (1976) argued that the language proficiency of medical professionals can be best assessed in a simulated hospital-like situation. Vu, Barrows, Marcy, Verhulst, Colliver and Travis (1992) have similarly argued that medical students' clinical, communication and interpersonal skills are best assessed in situations that approximate real clinical practice (that is, they should have high face validity).

There are a number of instruments for rating spoken language proficiency, such as the Australian Second Language Proficiency Ratings (ASLPR) (Ingram, 1984). The ASLPR and other such tests are not designed for use in specific contexts, and require that the assessor has extensive training in applied linguistics or an affiliated area of expertise. The rationale behind designing a new rating scale was to enable standardized patients with limited training, and no extensive knowledge of the assessment of spoken language, to rate students in the context of an Observed Structured Clinical Interview (OSCI). As described in Study IV of this thesis, clinical teachers at the University of Adelaide reported that some undergraduate medical students were experiencing difficulties in their dealings with staff and patients due to their language skills. Thus, a further rationale behind developing a rating scale was to explore the extent to which clinicians' perceptions of students' English language abilities concurred with the assessments of independent raters.

Examination of the nature of the comments made by clinical tutors showed that not all could be considered as strictly related to language skills. However, concerns about "familiarity with colloquial language", "comprehensibility of speech", "rate of speech" and "register" were seen to be valid items to include in a rating scale of spoken language proficiency. "Pronunciation" was also deemed relevant, but covered under

"comprehensibility of speech", thus negating the need for a separate evaluation of this aspect of language.

In addition to referring to the findings of Study IV, a list of specific aspects of language hypothesized as having the potential to influence a medical student-patient interaction was compiled in collaboration with a clinical colleague, Dr Jane Vernon-Roberts, whose experience and knowledge as a medical professional were invaluable. Three linguists[24] were consulted during the construction of the scale, for advice on the appropriateness of the inclusion and exclusion of items and to comment on the method of the scale's administration and scoring.

Based upon the information gleaned from the review of literature presented in Chapter II, the data collected in Study IV, the knowledge and experience of my clinical colleague, informal discussions with clinical teachers and the advice of the linguists, the final scale (the "Language Rating Scale" or "LRS") (Appendix XIX) comprised ten aspects of spoken language proficiency. These were;

1. Use of correct tense
2. Use of appropriate register
3. Comprehensibility of speech due to accent
4. Appropriate rate of speech
5. Appropriate use of non-verbal communication
6. Response to requests, apologies, and/or thanks (speech acts)
7. Understanding of informal language
8. Clarification where comprehension lacking
9. Fluency of speech
10. Overall impression of language proficiency.

---

[24] I am grateful for the advice and assistance of Associate Professor Uschi Felix, who supervised the work for this thesis during 1994, as well as Mrs Helen Mullins and Ms Ursula McGowan, who are both members of the Language Development Committee of the Faculty of Medicine, University of Adelaide.

Each of these items was included in the scale to address questions about that particular aspect of speech and its effect on an interaction between a student and a standardized patient. It was necessary to clearly define what was meant by each of the items and to specify the question or questions that were to be addressed by inclusion of the item in the scale.

"Tense" is defined as the relationship between the form of the verb and the time of the action or the state that it describes (Richards, Platt and Weber, 1992). The student was rated as to whether or not he or she employed the correct tense during the interview. It was hypothesised that students from non-English speaking backgrounds would be more likely to make errors in the use of tense than English speaking background students.

With regards to "Register", it was hypothesised that some students would be more flexible in their style of speech when interviewing a standardized patient. Anecdotal evidence, supported by clinician's feedback suggested that students from non-English speaking backgrounds are less able to modify their register to suit the situation, and tend to rely upon their knowledge of specialised registers where their English language skills are not adequate for the situation in which they find themselves. This is consistent with Long's (1985) hypothesis of stereotypicality.

A rating of the student's "Comprehensibility of speech due to accent" aimed to determine whether this aspect of speech was detrimental to the standardized patient's understanding of what was being said. It was hypothesized that some students' pronunciation of words may make them difficult to comprehend.

"Rate of speech" was assessed as clinicians noted in their comments that students from non-English speaking backgrounds sometimes spoke quickly to patients, and that this impeded understanding. A rating on this dimension of language allowed for an investigation of whether non-English speaking background students are perceived to speak

more quickly than other students, and whether this interfered with the standardized patient's understanding.

Anecdotal evidence suggested that clinicians believed that students from non-English speaking backgrounds were more likely to display incongruent "Non-verbal behaviours" whilst speaking to, or listening to patients. For example, it was stated that Asian females did not make eye contact with patients or authority figures, and that Asian males tended to use smiling where such a facial expression was not expected or appropriate. A rating of the student's non verbal behaviour sought to determine whether these were in accord with verbal behaviour.

"Response to requests, apologies and/or thanks" or "speech acts" was evaluated because anecdotal evidence suggested that some students employed inappropriate verbal responses to patients and peers where conventions required polite language, such as a word of thanks or an apology. Since it would be unlikely that a student would be in a position to demonstrate their own speech acts within a brief interview with a standardized patient, it was decided that the student's response to the patient's speech acts would be evaluated.

Since clinicians' had seen difficulties with Australian colloquialisms to be an important impediment in student-patient interactions "Understanding of informal language" was included to determine whether students did indeed have problems with such expressions.

"Clarification where comprehension lacking" was rated on the basis of anecdotal information from clinicians that students experiencing problems with English often failed to ensure that they had understood, by asking that something be repeated or restated. It was thus hypothesized that students who did not understand the standardized patient's meaning would attempt to cover this by continuing the conversation without requesting clarification.

"Fluency of speech" is a global indicator of the quality of speech (Richards, Platt and Weber, 1992). Fluent speech seems natural to the listener, with the appropriate use of pauses, rhythms, rate, intonations, stresses on syllables and words, interjections and interruptions. Richards, Platt and Weber (1992) further define fluency as reflecting a level of oral language proficiency which includes the ability to produce spoken language with ease, to speak with a good but not necessarily perfect command of intonation, vocabulary and grammar, to be able to express ideas effectively and to produce a continuous flow of speech devoid of comprehension difficulties or breakdowns. The standardized patient rated their perception of the fluency of a student's spoken English.

"Overall impression of language proficiency" was included as a global rating. Hui and Yam (1987) define "language proficiency" as a collective term for variations in syntax, grammar, vocabulary and discourse strategies. Thus, the standardized patient was required to make a final evaluation of how proficient in speaking English they considered the student to be, regardless of the ratings they had made on the previous nine items.

Finally, an opportunity for "Comments" was considered extremely important to be provided to the raters, since as a pilot study further refinements of the scale were to be made partially on the basis of the open-ended feedback provided by them.

Each of the ten items was rated on a five point scale, from 1 (poor) to 5 (excellent), with a mid-point of 3 (adequate). In order to decide upon what each point on the scale should represent, the Australian Second Language Proficiency Ratings (ASLPR) (Ingram, 1984) were employed as a guide for this pilot study. The following definitions for each of the five points of the scale were arrived upon. It should be noted that unless stated otherwise, "speaker" refers to the student.

1. Fluency is uneven. Poor grasp of the social conventions in conversation, such as responses to apologies, thanks and requests. Grammatical errors, such as the inappropriate use of tense create misunderstanding. Poor vocabulary means the language produced is

hesitant and time is spent searching for the 'right' words. Words may not be in their correct place within the sentence. May be difficulty in pronunciation, and intonation and rate will not be that of a native speaker and may interfere with the listener's comprehension. May rely on the use of the specialist register to compensate for poor grasp of informal, or lay language. May use titles inappropriately (for example, Miss Marilyn). Influences of the person's culture may adversely affect the interaction (for example, the inability to make appropriate eye contact). Both speakers (student and listener) may have to repeat themselves in order to be understood.

2. Frequent hesitations as the speaker searches for vocabulary and the correct grammar. Vocabulary should be sufficient to sustain the interaction. Accent may be present but does not significantly affect listener's comprehension. Speaker may have to repeat him or herself in order to be understood. Overall rate of speech not as a native speaker's. Speaker shows more difficulty in longer or more complex sentences than simple, briefer ones.

3. At this level the speaker is deemed 'adequate' in English language skills for the purpose of a medical student/role-playing patient interaction. Speaker appears comfortable using language. Rarely needs to search for the words required, and can fill in gaps with other speech whilst searching for appropriate vocabulary where necessary. Accent may be present but does not hamper listener comprehension. Fluency is good, and the appropriate registers can be employed as required. Speaker may have some difficulty with colloquial/informal language.

4. Ability to use language fluently and accurately, using wide vocabulary. Is comfortable with colloquial language. May obviously be a non-native speaker of English, and may have an accent, but is easily comprehended by the listener. Grammatical errors are few. Sensitive to register, and can modify language accordingly.

5. Uses language as would a native English speaker. Has complete fluency, accuracy and vocabulary. Rate of speech, paralinguistics, grammar, use of informal speech, cultural references are all as a native speaker's. May have an accent, although this does not influence comprehensibility.

Method.

In 1994 all third year medical students were examined for their ability to demonstrate Medical Communication Skills. This examination took the form of an Observed Structured Clinical Interview, referred to as an 'OSCI'. The 1994 Medical Communication Skills examination is described in detail in Chapter V.

This study followed the procedure employed in the Medical Communication Skills OSCI examination, and extended upon the methodology of studies conducted in the United States of America, which have utilised standardized patients to evaluate the spoken English proficiency of foreign medical graduates in that country as reviewed in Chapter II.

A standard script was written which outlined a presenting complaint (chronic upper abdominal pain) and the history of that complaint (Appendix XX)[25]. The script was such that the standardized patient had scope for improvisation. However, it was necessary that the standardized patient mention specific issues related to pain, past medical history and lifestyle factors. With regard to language content, the standardized patient was required to ensure the use of two 'trade' names, "Quickeze" and "Mylanta", both of which are antacids purchased 'over the counter', without need for a doctor's prescription. Standardized patients also incorporated three examples of colloquial language into their interview; 'take a sickie', 'feeling crook' and 'feeling under the weather'. These three examples of colloquial language are often used in Australia (Wilkes, 1993) and are commonly heard in Australian doctor-patient interactions (Chur-Hansen and Barrett, 1996) (Appendix IV.II). After the interview was completed, the standardized patient was required to rate the student's spoken language proficiency on the Language Rating Scale.

---

[25] I am indebted to Dr Jane Vernon-Roberts who wrote the standardized patient script for this study.

Standardized Patients.

Seven Psychology postgraduates from the University of Adelaide were employed to act as standardized patients for this study[26]. The study required six standardized patients; however, it was thought prudent to employ and train an additional person in the event of illness or other misfortune on the day of testing. Two of the standardized patients were male; all were native speakers of and fluent in English. They all had teaching experience with undergraduate students, and five had experience teaching medical students from non-English speaking backgrounds at the University of Adelaide. It should be stressed, however, that none of the seven individuals approached to act as standardized patients for the study had teaching responsibilities for any of the students in the 1994 third year cohort.

Six weeks prior to this study the seven standardized patients met for one hour to discuss the task and the purpose of the project. A videotape of a 1993 non-English speaking background third year medical student interviewing a role-playing patient was shown, to facilitate discussion regarding language proficiency in medical interviews. After agreeing to act as standardized patients for the study, all six were sent a copy of the Language Rating Scale, definitions of each of the ten items and the definitions corresponding to each of the five points of the rating scale, the script and the format of the procedure for the day of testing.

Four weeks prior to the study (two weeks after the initial meeting) the seven standardized patients took part in an intensive two hour training session. After clarifying questions related to the material that had been distributed after the first meeting, three videotapes of 1993 third year medical students practising their interviewing skills were shown. In each videotape one student role-played as doctor, the other as patient. The student roleplaying as the doctor had been selected by the researcher, in collaboration with a clinical colleague[27], to represent an example of either 'good', 'poor' or 'very poor'

---

[26] Alex Ask, Natalie Beaumont-Smith, Mary Katsikitis, Jane Mortimer, Della Steen, Pieter Walker and Robyn Young were employed to act as standardized patients for this study.
[27] Dr Jane Vernon-Roberts.

English language proficiency. After each of the videotapes had been viewed, the standardized patients rated the role-playing doctor on the Language Rating Scale. Ratings were then discussed as a group, and ambiguities were identified and clarified. In the final stage of training the researcher role-played as a standardized patient for the study, with her clinical colleague acting as the medical student. The trainee standardized patients then rehearsed some parts of the script with each other, before training concluded. During the four week lead-up to the study, the standardized patients were requested to practice their role-playing techniques.

Subjects and Procedure.

An information sheet was circulated to all 1994 third year students outlining the dates on which language would be tested and outlining the format of the STAL (Appendix XXI). On the day, but prior to testing on the STAL, all students were invited to attend an orientation lecture, to advise them of the purpose and procedure of the study, referred to as the "Language OSCI" (Appendix XXII). It was stressed to students that although participation was compulsory, the results of the Language OSCI did not count toward academic grades. It was also explained that the results of the OSCI were confidential and available only to the Dean of Medicine and the members of the Language Development Committee. Students were assured that those identified by the study as experiencing difficulties with spoken language proficiency would be offered an opportunity to discuss improvement strategies with a Language Development Committee member, and would be strongly encouraged to participate in a brief programme of workshops offered by the Faculty.

Students were told that on the day of testing they would interview a standardized patient for approximately ten minutes and then write an account of that interview. They were informed that the standardized patient was trained to listen to, and rate, their spoken English language skills. It was made clear that the standardized patient was not at liberty to rate their Medical Communication Skills, (although students were advised that they

should draw upon these skills as appropriate), or their Clinical Science Skills, which were to be assessed independently by other examiners.

The entire 1994 third year medical student cohort (n=149) participated in this study. During the examination period held at the end of their first semester (in June), students were assembled as a group after the completion of their Medical Communication Skills Observed Structured Clinical Interview (OSCI) examination. Students entered an outpatients area of the teaching hospital in cohorts of six. When instructed to do so, each student entered one of six rooms in which their allocated standardized patient was waiting. The interview then commenced, and was timed to allow for a maximum of 10 minutes. Due to financial and practical constraints it was not possible to videotape record the interviews; however, all interviews were audiotape recorded. At the conclusion of the interview the student left the room, and entered a second room, where they were allowed five minutes to write about the interview that they had just conducted with the standardized patient[28]. They were given a sheet of paper on which to record the information (Appendix XXIII), with the following instructions;

You have 5 minutes to record the interview you have just completed. Please write an account of this interview so that another health care professional would understand what took place and the decisions and conclusions drawn by yourself and the patient. Do not use medical jargon.

Whilst the student was writing the account of the interview, the standardized patient had five minutes to complete the ratings for spoken language for that student, and to comment on the student's performance where appropriate.

Interviews commenced at 9.00 am and ended at 5.15 pm, with a one hour lunch break, and two 30 minute morning and afternoon tea breaks.

---

[28] Whilst students were required to write an account of the interview, the data generated from this exercise will be presented in the following Chapter.

English Language Specialist Ratings.

Two expert English language teachers who had no personal knowledge of any of the students involved in the study independently rated spoken language from the audiotapes. Their ratings were made "blind", in that they had access only to the students' names, but no other information.

Researcher's Ratings.

The audiotaped student-standardized patient encounters were rated by the researcher, who had previously taught all students in the cohort for at least one year, two years prior to the study.

Results.

Standardized Patient Ratings.

Six of the pool of seven standardized patients were required on the day of testing. Five were female, one male. Five of the standardized patients rated 25 students, the other rated 24. Interviews ranged in duration from 1.9 minutes to 8.8 minutes with the average length of each interview being 4.53 minutes (sd=1.35 minutes).

Individual standardized patient ratings can be found in Appendix XXIV. Inspection of the collective ratings for each of the ten items on the Language Rating Scale show negatively skewed distributions (Table XXXVIII).

Table XXXVIII: Standardized patients' ratings.

n=149	1	2	3	4	5	missing*	$\bar{x}$	sd
tense	2 (1.3)	7 (4.7)	23 (15.4)	32 (21.5)	85 (57.0)	-	4.28	0.98
register	1 (0.7)	7 (4.7)	29 (19.5)	49 (32.9)	63 (42.3)	-	4.11	0.93
accent	3 (2.0)	4 (2.7)	29 (19.5)	30 (20.1)	70 (47.0)	13 (8.7)	4.18	1.01
rate	1 (0.7)	7 (4.7)	31 (20.8)	34 (22.8)	75 (50.3)	1 (0.7)	4.18	0.97
nvc	-	7 (4.7)	36 (24.2)	46 (30.9)	57 (38.3)	3 (2.0)	4.05	0.91
speech acts	-	-	18 (12.1)	51 (34.2)	71 (47.7)	9 (6.0)	4.38	0.70
informal	5 (3.4)	18 (12.1)	35 (23.5)	26 (17.4)	61 (40.9)	4 (2.7)	3.83	1.20
clarification	2 (1.3)	11 (7.4)	13 (8.7)	29 (19.5)	54 (36.2)	40 (26.8)	4.12	1.09
fluency	1 (0.7)	8 (5.4)	37 (24.8)	31 (20.8)	72 (48.3)	-	4.11	1.00
overall	1 (0.7)	4 (2.7)	37 (24.8)	42 (28.2)	65 (43.6)	-	4.11	0.92

percentages are in brackets.

* missing=missing data, whereby the standardized patient did not provide a rating.

Key: 1=poor
2=between poor and adequate
3=adequate
4=between adequate and excellent (good)
5=excellent

Comments.

Open-ended comments made by the six standardized patients can be found in Appendix XXV. Of the 149 rating sheets, comments were made on 95 (63.8%). The number of comments made by individual standardized patients ranged from 6 (24.0%) to 23 (92.0%).

A number of the comments can be generally subsumed under six main categories (which are listed in no particular order);

1. Difficulty arriving upon a rating (N=4)

   Not worth a 5 (for overall proficiency) but close to it i.e. 4.5 (Rater 1).

2. Difficulty keeping the student talking (N=10).

   This one needed to talk more but couldn't get him to. Reasonable. (Rater 3).

3. Nervousness/anxiety of the student (N=8)

   Very good. Nervous, so tended to stare at me a bit, but otherwise, competent. (Rater 2).

4. Failure of student to clarify (N=10)

   He would ask, but it was clear he still didn't understand. He wouldn't ask for further clarification. (Rater 4).

5. Non-verbal communication (N=15)

   His eyes darted back and forth quickly, which was a little off-putting. (Rater 6).

6. Use of jargon (N=17)

   Used "radiate"(jargon), "stools". (Rater 5).

In addition, four of the comments referred specifically to medical communication skills (N=1), including empathy (N=2) and counselling (N=1).

Table XXXIX: English Language Specialist I Ratings.

n=149	1	2	3	4	5	missing*	$\bar{x}$	sd
tense	-	6 (4.0)	20 (13.4)	32 (21.5)	88 (59.1)	3 (2.0)	4.38	0.87
register	-	1 (0.7)	27 (18.1)	70 (47.0)	42 (28.2)	9 (6.0)	4.09	0.72
accent	-	3 (2.0)	36 (24.2)	33 (22.1)	73 (49.0)	4 (2.7)	4.21	0.89
rate	-	1 (0.7)	28 (18.8)	70 (47.0)	46 (30.9)	4 (2.7)	4.11	0.73
nvc						149 (100)		
speech acts	-	-	6 (4.0)	17 (11.4)	12 (8.1)	114 (76.5)	4.17	0.71
informal	1 (0.7)	1 (0.7)	13 (8.7)	20 (13.4)	45 (30.2)	69 (46.3)	4.34	0.89
clarification	-	2 (1.3)	10 (6.7)	47 (31.5)	35 (23.5)	55 (36.9)	4.22	0.72
fluency	-	3 (2.0)	38 (25.5)	63 (42.3)	39 (26.2)	6 (4.0)	3.97	0.79
overall	-	2 (1.3)	24 (16.1)	54 (36.2)	56 (37.6)	13 (8.7)	4.21	0.78

percentages are in brackets.

* missing=missing data, whereby the English Language Specialist did not provide a rating.

Key: 1=poor
2=between poor and adequate
3=adequate
4=between adequate and excellent (good)
5=excellent

Comments.

The need to include an item on stress patterns (intonation) on the scale and the inclusion of an item to assess students' loudness of voice was suggested as existing items did not accommodate for this.

This rater also recommended better training for the patients in eliciting speech from the students and ensuring that all had the same opportunities to demonstrate their fluency of speech.

It was suggested that rapport be included on the scale.

Table XL: English Language Specialist II Ratings.

n=149	1	2	3	4	5	missing*	$\bar{x}$	sd
tense	4 (2.7)	4 (2.7)	7 (4.7)	31 (20.8)	100 (67.1)	3 (2.0)	4.5	0.92
register	1 (0.7)	2 (1.3)	26 (17.4)	70 (47.0)	48 (32.2)	2 (1.3)	4.10	0.78
accent	1 (0.7)	5 (3.4)	24 (16.1)	69 (46.3)	50 (33.6)	-	4.09	0.83
rate	-	2 (1.3)	11 (7.4)	87 (58.4)	49 (32.9)	-	4.23	0.64
nvc						149 (100)		
speech acts	1 (0.7)	5 (3.4)	18 (12.1)	33 (22.1)	35 (23.5)	57 (38.3)	4.04	0.95
informal	-	5 (3.4)	22 (14.8)	41 (27.5)	79 (53.0)	2 (1.3)	4.32	0.85
clarification	-	5 (3.4)	23 (15.4)	47 (31.5)	49 (32.9)	25 (16.8)	4.13	0.86
fluency	-	2 (1.3)	29 (19.5)	49 (32.9)	69 (46.3)	-	4.24	0.81
overall	-	8 (5.4)	25 (16.8)	55 (36.9)	57 (38.3)	4 (2.7)	4.11	0.88

percentages are in brackets.

* missing=missing data, whereby the English Language Specialist did not provide a rating.

Key: 1=poor
2=between poor and adequate
3=adequate
4=between adequate and excellent (good)
5=excellent

Comments.

This rater made quite extensive comments which have been transcribed here.

With regards to the use of appropriate tense she noted that objectivity in rating was high.

Appropriateness of register was determined by this rater according to;

- relevance of questions
- evidence of exercise of power in the exchange
- degree of that power - condescension, patronisation, impotence, peer relationship.

- distance between participants

  e.g. a) "cocky" initial statements (resembling arresting police officers) - reciting statements about confidentiality were not regarded favourably

  b) inappropriateness of students revealing their family's medical history
- use of fillers - extreme sympathetic noises
- indifferent responses
- non-engagement with the seriousness of the problem

Interference due to comprehensibility of accent was deemed to occur if the following were noted;

- broad Australian vowel sounds
- excessive rising pitch and other artificial tone patterns which created diversionary meanings - e.g. defensiveness
- elision/slurring due to excessive speed
- "clipped" phonemes in word endings
- misplaced stress in words or sentences

This rater commented that she encountered speech that was too rapid far more often than speech that was too slow. Speech rate contributed to the rater's impression of the level of the student's confidence as well as his or her attitude to the standardized patient. It was noted that speech that was too fast could indicate nervousness, self-consciousness, over-familiarity or disregard for the condition or the language origins of the standardized patient, whilst speech that was too slow could indicate a pedantic/patronising attitude. Speech that was too slow due to a lack of fluency in English was not encountered according to this rater.

Commenting on speech acts, this rater looked for;

- turn-taking, but particularly the student's capacity to participate fully in the interview - i.e. not permit the standardized patient's monologue to continue for too long
- appropriate fillers
- cohesion of interaction e.g. if the standardized patient discussed a relative catastrophe and received a detached "mmmm" followed by an unrelated question, this was rated low.

She commented that requests from standardized patients were extremely rare, and that apologies were non-existent and thanks occurred only at the end of the interviews.

This rater felt that she could not formulate a definitive judgement for the student's understanding of informal language, since she considered that the definition of "informal" given to raters did not encompass her definition of this term. She noted that most interviews were "relaxed" and both parties used language which was informal rather than "formal". That is, the language was not necessarily littered with colloquialisms, clichés or idioms but rather, familiar, "informal" speech patterns e.g. "What's it made of?" as opposed to "Of what is it made?"

With regard to informal speech it was also commented that the instructions to standardized patients to include set idioms did not result in uniform usage, although where the standardized patients did use them, and the students had to ask for an explanation, the rating was lowered.

This rater commented that only a few standardized patients asked students to explain what they meant by a certain term. She also noted that one of the standardized patients acted particularly obtusely, asking awkwardly for simplistic explanations. Here, this rater scored the student's responses according to "speech act" performances or "register" (whether the student lost their temper or not). She also indicated that she scored

most students as a 4 or 5 on this item because standardized patients rarely asked students to clarify as their comprehension of what the students were saying didn't appear to provoke uncertainty or confusion.

This rater felt that the definition of fluency of speech was clear and she had no difficulty arriving upon a rating for each student on this item.

In terms of her ratings of overall impression of language proficiency, she indicated that as she works professionally with people in the initial stages of English language learning, she considered the majority of students to be a 3 or above. She commented, however, that she did not award a 5 for this item if tense, register, accent, speech acts or fluency were less than 5, although the remaining criteria were not necessarily deemed to reduce a high judgement of overall proficiency.

Table XLI: Researcher's Ratings.

n=149	1	2	3	4	5	missing*	$\bar{x}$	sd
tense	2 (1.3)	7 (4.7)	8 (5.4)	9 (6.0)	123 (82.6)	-	4.64	0.89
register	-	1 (0.7)	9 (6.0)	39 (26.2)	99 (66.4)	1 (0.7)	4.60	0.64
accent	5 (3.4)	14 (9.4)	20 (13.4)	21 (14.1)	89 (59.7)	-	4.17	1.18
rate	-	8 (5.4)	16 (10.7)	32 (21.5)	93 (62.4)	-	4.41	0.89
nvc						149 (100)		
speech acts	-	1 (0.7)	7 (4.7)	8 (5.4)	18 (21.1)	115 (77.2)	4.27	0.90
informal	3 (2.0)	7 (4.7)	7 (4.7)	13 (8.7)	42 (28.2)	77 (51.7)	4.17	1.20
clarification	-	5 (3.4)	4 (2.7)	8 (5.4)	51 (34.2)	81 (54.4)	4.54	0.91
fluency	3 (2.0)	15 (10.1)	14 (9.4)	26 (17.4)	91 (61.1)	-	4.26	1.11
overall	3 (2.0)	14 (9.4)	15 (10.1)	26 (17.4)	91 (61.1)	-	4.26	1.10

percentages are in brackets.

* missing=missing data, whereby the researcher did not provide a rating.

Key: 1=poor
2=between poor and adequate
3=adequate
4=between adequate and excellent (good)
5=excellent

Researcher's comments.

On listening to the audiotapes it was noted that some standardized patients had difficulties in incorporating the colloquialisms into each of their interviews. There was a

tendency for several standardized patients to "self-select" those students for whom they thought such language might be problematic, and thus used it in these encounters but not in others. Several standardized patients also failed to ensure that they made a request, apology or form of thanks during the interview.

Some students used a medical register during the encounter. Terms noted included "stools", "exacerbate", "first degree relatives", "bowel movements", "epigastric" and "localized pain".

Also noted was the fact that several students were confused about the nature of the task and their role within the encounter, and were therefore quite apprehensive at the commencement of the interview. This apprehension seemed to have been minimised by the standardized patient's reassurance that they needed only to talk with them about a health problem they were experiencing.

Identification of students in need of English Language Intervention.

On the basis of the ratings made by the standardized patients, the researcher and her clinical colleague (Dr Jane Vernon-Roberts) identified thirteen students who they deemed were in considerable need of assistance to improve their English language skills. All thirteen students were rated as being significantly poorer on all ten items of the Language Rating Scale (employing one-tailed t-tests, $p<0.05$). These students were referred to an English Language Specialist, who in collaboration with Dr Vernon-Roberts, designed a short programme of tutorials for them.

RELIABILITY OF THE LANGUAGE RATING SCALE.

Split-half reliability (internal consistency).

Internal consistency can be considered a measure of the homogeneity of a scale, and as such gives further weight to construct validation (Anastasi, 1988). To first investigate the internal consistency of the Language Rating Scale the first 5 items were added and correlated with the second 5 items to construct two forms of the scale. Using

the standardized patients' collective scores and excluding missing data, 97 cases were included in analyses, yielding a correlation between the forms of 0.89 and a Guttman split-half reliability correlation coefficient of 0.94. When missing data were included so that all cases could be analysed, the correlation between the two forms was 0.72 and the Guttman split-half reliability correlation coefficient 0.83.

Because non-verbal communication could not be scored by the other three raters and the scale could therefore not be equally divided into two forms, their data were not used to calculate split-half reliability.

Leary (1991) points out that one limitation with the split-half method to explore reliability of a scale is that the researcher is at liberty to decide how the items should be divided. Cronbach's alpha coefficient compensates for this by providing a statistic equivalent to the average of all possible divisions of the scale (Leary, 1991).

Cronbach's alpha calculated using the standardized patients' collective ratings and excluding missing data was 0.97. With missing data included, the alpha level was 0.84.

Cronbach's alpha could not be calculated for the other three raters if missing data were excluded from analyses. Accounting for missing values, the Cronbach alpha reliability coefficients were 0.61, 0.68 and 0.67 for the English Language Specialists I & II and the researcher, respectively.

Inter-item consistency.

In order to determine the inter-item consistency of the Language Rating Scale, all items were correlated with each other employing the Spearman *rho* (two-tailed) test with Bonferroni correction (Galambos and Simonelli, 1996).

Table XLII: Spearman *rho* correlation coefficients for Language Rating Scale item scores, standardized patient ratings.

Note: Correlation coefficients for individual standardized patients can be found in Appendix XXVI.

register	0.65								
accent	0.84	0.60							
rate	0.86	0.61	0.82						
nvc	0.54	0.45	0.58	0.72					
speech acts	0.68	0.58	0.69	0.71	0.67				
informal	0.69	0.49	0.80	0.66	0.51	0.60			
clarify	0.60	0.51	0.62	0.59	0.48	0.57	0.67		
fluency	0.84	0.64	0.80	0.85	0.64	0.70	0.73	0.69	
overall	0.83	0.62	0.88	0.87	0.70	0.73	0.81	0.72	0.90
	tense	**register**	**accent**	**rate**	**nvc**	**speech a**	**informal**	**clarify**	**fluency**

All significant at $p<0.001$.

Table XLIII: Spearman *rho* correlation coefficients for Language Rating Scale item scores, English Language Specialist I.

register	0.55								
accent	0.84	0.58							
rate	0.53	0.54	0.64						
nvc	-	-	-	-					
speech acts	0.53	0.63	0.66	0.71	-				
informal	0.78	0.57	0.87	0.55	-	0.82			
clarify	0.62	0.67	0.70	0.67	-	0.82	0.76		
fluency	0.74	0.59	0.78	0.66	-	0.69	0.69	0.73	
overall	0.78	0.67	0.85	0.68	-	0.78	0.81	0.75	0.83
	tense	**register**	**accent**	**rate**	**nvc**	**speech a**	**informal**	**clarify**	**fluency**

All significant at $p<0.001$.

Table XLIV: Spearman *rho* correlation coefficients for Language Rating Scale item scores, English Language Specialist II.

register	0.49								
accent	0.65	0.51							
rate	0.35	0.54	0.55						
nvc	-	-	-	-					
speech acts	0.37	0.63	0.37	0.50	-				
informal	0.64	0.48	0.63	0.33	-	0.34			
clarify	0.56	0.50	0.52	0.39	-	0.52	0.64		
fluency	0.68	0.55	0.65	0.50	-	0.44	0.75	0.64	
overall	0.71	0.58	0.75	0.48	-	0.49	0.78	0.66	0.86
	tense	**register**	**accent**	**rate**	**nvc**	**speech a**	**informal**	**clarify**	**fluency**

All significant at $p\leq0.001$.

Table XLV: Spearman *rho* correlation coefficients for Language Rating Scale item scores, researcher.

register	NS								
accent	0.64	NS							
rate	0.65	NS	0.89						
nvc	-	-	-	-					
speech acts	NS	NS	NS	0.59	-				
informal	0.41	NS	0.81	0.73	-	NS			
clarify	NS	NS	0.44	0.41	-	NS	NS		
fluency	0.67	NS	0.95	0.90	-	NS	0.76	0.45	
overall	0.67	NS	0.95	0.90	-	NS	0.76	0.44	0.99
	tense	**register**	**accent**	**rate**	**nvc**	**speech a**	**informal**	**clarify**	**fluency**

All values significant at $p<0.001$.

Interrater reliability.

Spearman *rho* correlation coefficients were calculated in order to establish interrater consistency between the standardized patient ratings, the two English Language Specialists and the researcher.

Table XLVI: Spearman *rho* correlation coefficients, standardized patients (SPS) and the researcher and English Language Specialists I & II (ELS I & ELS II).

SPS	researcher	ELS I	ELS II
tense	0.53	0.62	0.53
register	NS	0.25	0.23
accent	0.77	0.68	0.67
rate	0.61	0.39	0.30
nvc	-	-	-
speech acts	NS	0.57	0.20‡
informal	0.74	0.56	0.57
clarify	0.40	0.39	0.24†
fluency	0.63	0.54	0.55
overall	0.73	0.56	0.62

All values significant at $p\leq0.001$
† Value significant at $p<0.01$
‡ Value significant at $p<0.05$

Table XLVII: Spearman *rho* correlation coefficients, the researcher and English Language Specialists I & II (ELS I & ELS II).

Researcher	ELS I	ELS II
tense	0.57	0.59
register	NS	NS
accent	0.80	0.71
rate	0.51	0.41
nvc	-	-
speech acts	0.48‡	0.38‡
informal	0.69	0.63
clarify	0.55	0.23‡
fluency	0.67	0.66
overall	0.65	0.72

All values significant at $p<0.001$
‡ Values significant at $p<0.05$

Table XLVIII: Spearman *rho* correlation coefficients, the English Language Specialists I & II (ELS I & ELS II).

ESL I	ELS II
tense	0.66
register	0.30
accent	0.63
rate	0.33
nvc	-
speech acts	0.48
informal	0.63
clarify	0.46
fluency	0.65
overall	0.68

All significant at $p<0.005$

To consider agreement between the four sets of raters Kendall's coefficient of concordance (*W*) was calculated for each of the 10 items on the scale. As the sample size for each item was greater than 7, the chi-square statistic was employed to ascertain inter-rater agreement (Siegal, 1956).

Table XLIX: Interrater agreement on the Language Rating Scale (LRS), Kendall's *W*.

LRS	cases	x^2	p
tense	143	35.99	0.0000
register	137	43.13	0.0000
accent	132	5.28	NS
rate	144	27.16	0.0000
nvc	-	-	-
speech acts	18	1.97	NS
informal	54	6.01	NS
clarify	39	38.67	0.0000
fluency	143	35.19	0.0000
overall	132	17.92	0.0005

To further explore differences between the raters' scores, a comparison of the six standardized patient's mean ratings on each of the ten items of the Language Rating Scale was conducted for each individual standardized patient using independent sample t-tests (two-tailed).

Standardized patient I rated his students significantly higher than standardized patient II on appropriate use of register (t(pooled)=3.24, df=47, p=0.002) and rate of speech (t=2.59, df=35.28, p=0.014), whilst standardized patent II rated her student's understanding of informal language higher than standardized patient I's (t=-2.49, df=33.51, p=0.018).

Standardized patient I rated his students significantly higher than standardized patient III on correct use of tense (t=2.38, df=30.64, p=0.024), register (t(pooled)=4.61, df=48, p=0.000), comprehensibility due to accent (t(pooled)=4.33, df=35, p=0.000) and rate of speech (t=2.43, df=33.64, p=0.021).

Standardized patient I rated his students significantly higher than standardized patient IV on correct use of tense (t=2.49, df=33.82, p=0.018) and register (t(pooled)=2.37, df=48, p=0.022).

Standardized patient I rated his students significantly higher than standardized patient V on correct use of tense (t(pooled)=5.88, df=48, p=0.000), register (t(pooled)=5.51, df=48, p=0.000), comprehensibility due to accent (t(pooled)=3.10, df=48, p=0.003), rate of speech (t(pooled)=4.89, df=48, p=0.000), speech acts (t(pooled)=5.90, df=47, p=0.000), fluency of speech (t(pooled)=4.31, df=48, p=0.000) and overall impression of language proficiency (t(pooled)=2.67, df=48, p=0.01).

Standardized patient I rated his students significantly higher than standardized patient VI on correct use of tense (t=3.39, df=31.19, p=0.002), register (t=4.43, df=38.02, p=0.000), comprehensibility due to accent (t=2.51, df=39.63, p=0.016), rate of speech (t=3.41, df=32.09, p=0.002), speech acts (t(pooled)=2.29, df=47, p=0.026) and fluency of speech (t=3.14, df=37.37, p=0.003).

Standardized patient II rated her students significantly higher than standardized patient III on comprehensibility due to accent (t(pooled)=6.65, df=34, p=0.000), understanding of informal language (t=3.54, df=32.31, p=0.001) and clarification where comprehension lacking (t(pooled)=2.93, df=33, p=0.006).

Standardized patient II rated her students significantly higher than standardized patient IV on understanding of informal language (t=2.18, df=34.0, p=0.036).

Standardized patient II rated her students significantly higher than standardized patient V on use of correct tense (t(pooled)=3.34, df=47, p=0.002), comprehensibility due to accent (t(pooled)=4.74, df=47, p=0.000), speech acts (t(pooled)=3.54, df=44, p=0.001), understanding of informal language (t(pooled)=5.55, df=45, p=0.000), fluency of speech (t(pooled)=2.76, df=47, p=0.008) and overall impression of language proficiency (t(pooled)=2.52, df=47, p=0.015).

Standardized patient II rated her students significantly higher than standardized patient VI on use of correct tense (t=2.06, df=39.47, p=0.046), comprehensibility due to

accent (t=3.22, df=31.20, p=0.003), understanding of informal language (t=4.42, df=31.63, p=0.000), clarification where comprehension lacking (t(pooled)=2.48, df=37, p=0.018) and fluency of speech (t=2.14, df=40.94, p=0.039).

Standardized patient III rated her students significantly lower than standardized patient IV on comprehensibility due to accent (t(pooled)=-3.74, df=35, p=0.001), appropriate non-verbal communication (t(pooled)=-2.83, df=47, p=0.007) and clarification where comprehension lacking (t(pooled)=-2.56, df=31, p=0.016).

Standardized patient III rated her students significantly higher than standardized patient V on speech acts (t(pooled)=3.63, df=47, p=0.001).

There were no significant differences in the mean ratings of standardized patients III and VI.

Standardized patient IV rated her students significantly higher than standardized patient V on use of appropriate register (t(pooled)=2.47, df=48, p=0.028), rate of speech (t(pooled)=2.95, df=48, p=0.005), appropriate non-verbal communication (t(pooled)=3.43, df=47, p=0.001), speech acts (t(pooled)=4.67, df=42, p=0.000) and overall impression of language proficiency (t(pooled)=2.76, df=48, p=0.008).

Standardized patient IV rated her students significantly higher than standardized patient VI on use of appropriate register (t(pooled)=2.11, df=48, p=0.04), comprehensibility due to accent (t=2.12, df=40.88, p=0.04), rate of speech (t=2.38, df=39.0, p=0.022), appropriate non-verbal communication (t(pooled)=2.47, df=45, p=0.017), understanding of informal language (t(pooled)=2.31, df=46, p=0.026) and clarification where comprehension lacking (t(pooled)=2.15, df=35, p=0.039).

Standardized patient V rated her students significantly lower than standardized patient VI on speech acts (t(pooled)=-3.25, df=46, p=0.002).

Using paired samples t-tests (two-tailed), a comparison of the mean ratings of the standardized patients' ratings for the student cohort with the two English language specialist and the researcher was conducted. The English Language Specialist I rated students as better able to understand informal language (t=-2.47, df=77, p=0.016) and provided lower ratings of fluency of speech (t=2.00, df=142, p=0.48). The English Language Specialist II rated students higher on use of correct tense (t=-2.52, df=145, p=0.013) and understanding of informal language (t=-6.03, df=142, p=0.000) and rated students as being poorer in speech acts as compared with the standardized patients' ratings (t=2.16, df=88, p=0.033). The researcher rated students significantly higher on use of correct tense (t=-5.23, df=148, p=0.000), register (t=-5.64, df=147, p=0.000), rate of speech (t=-3.44, df=147, p=0.001), understanding of informal language (t=-2.63, df=70, p=0.011), clarification where comprehension lacking (t=-3.27, df=60, p=0.002), and overall language proficiency (t=-2.38, df=148, p=0.019). The differences between the researcher's and the standardized patients' ratings of fluency of speech just failed to reach significance (t=-1.96, df=148, p=0.052).

A frequency distribution was conducted to investigate whether rater differences might be due to differences in the allocation of non-English speaking background students across the six standardized patients. Standardized patient I interviewed 11 English speaking background students and 14 non-English speaking background students; standardized patient II interviewed 14 English speaking background students and 10 non-English speaking background students; standardized patient III interviewed 10 English speaking background students and 15 non-English speaking background students; standardized patient IV interviewed 7 English speaking background students and 18 non-English speaking background students; standardized patient V interviewed 11 English speaking background students and 14 non-English speaking background students and standardized patient VI interviewed 8 English speaking background students and 17 non-English speaking background students.

A comparison of the mean ratings given by the two English language specialists and the researcher was conducted employing paired sample t-tests (two-tailed).

The English Language Specialist I gave significantly lower ratings than the researcher on the appropriate use of tense (t=4.09, df=145, p=0.000), register (t=6.63, df=138, p=0.000), rate of speech (t=5.15, df=144, p=0.000), clarification where comprehension lacking (t=4.67, df=50, p=0.000), fluency of speech (t=4.60, df=142, p=0.000) and overall impression of language proficiency (t=2.14, df=135, p=0.035).

The English Language Specialist II gave significantly lower ratings than the researcher on appropriate use of tense (t=2.28, df=145, p=0.024), register (t=5.89, df=145, p=0.000), rate of speech (t=2.66, df=148, p=0.009), clarification where comprehension lacking (t=4.68, df=56, p=0.000) and overall impression of language proficiency (t=2.38, df=144, p=0.018).

A comparison of the mean ratings of the two English Language Specialists showed that they differed significantly in their mean ratings on two items, appropriate rate of speech (t=-2.09. df=144, p=0.039) and fluency of speech (t=-5.22, df=142, p=0.000). Rater I was consistently lower in her ratings than Rater II on both these items.

VALIDITY.

Content validity.

In order to determine the inter-item consistency of the Language Rating Scale, all items were correlated with each other employing the Spearman *rho* (two-tailed) test, and this has been presented above. These data can also be examined to determine content validity of the Language Rating Scale.

Concurrent validity.

To establish the relationship between the Language Rating Scale ratings and the Screening Test of Adolescent Language (STAL) scores, Spearman correlation coefficients

(two-tailed) for the four subtests and total scores of the STAL (scored by the researcher) and the items on the Language Rating Scale for the four sets of raters (the standardized patients, the two English language specialists and the researcher) were calculated.

Table L: Spearman *rho* correlation coefficients for standardized patient ratings on the Language Rating Scale and Screening Test of Adolescent Language (STAL) scores.

	Vocabulary	Auditory Memory	Language Processing	Proverbs	Total Score
tense	0.37	0.51	0.26	0.27	0.44
register	NS	0.28	0.25	0.26	0.27
accent	0.51	0.52	0.35	0.33	0.58
rate	0.38	0.43	0.25	0.23	0.43
nvc	0.23	0.30	0.21	0.24	0.30
speech acts	0.32	0.35	0.23	0.31	0.36
informal	0.47	0.40	0.26	0.32	0.51
clarification	0.39	0.36	NS	0.22	0.42
fluency	0.37	0.47	0.23	0.23	0.41
overall	0.39	0.44	0.28	0.27	0.45

All values significant at $p<0.05$.

Table LI: Spearman *rho* correlation coefficients for English Language Specialist I's ratings on the Language Rating Scale and Screening Test of Adolescent Language (STAL) scores.

	Vocabulary	Auditory Memory	Language Processing	Proverbs	Total Score
tense	0.47	0.51	0.28	0.32	0.50
register	0.25	0.31	NS	NS	0.28
accent	0.41	0.52	0.23	0.34	0.47
rate	0.30	0.30	0.24	0.39	0.40
nvc	-	-	-	-	-
speech acts	NS	NS	NS	NS	NS
informal	0.33	0.50	0.31	0.26	0.43
clarification	0.28	0.38	0.21	0.38	0.38
fluency	0.45	0.47	0.29	0.39	0.52
overall	0.38	0.42	0.28	0.36	0.48

All values significant at $p<0.05$.

Table LII: Spearman *rho* correlation coefficients for English Language Specialist II's ratings on the Language Rating Scale and Screening Test of Adolescent Language (STAL) scores.

	Vocabulary	Auditory Memory	Language Processing	Proverbs	Total Score
tense	0.44	0.49	0.36	0.33	0.50
register	0.31	0.27	NS	0.22	0.27
accent	0.42	0.40	0.24	0.28	0.45
rate	0.23	0.23	NS	NS	0.24
nvc	-	-	-	-	-
speech acts	0.24	0.24	NS	0.24	0.23
informal	0.36	0.44	0.27	0.36	0.44
clarification	0.39	0.37	0.19	0.30	0.38
fluency	0.31	0.51	0.21	0.35	0.40
overall	0.38	0.46	0.25	0.37	0.44

All values significant at $p<0.05$.

Table LIII: Spearman *rho* correlation coefficients for the researcher's ratings on the Language Rating Scale and Screening Test of Adolescent Language (STAL) scores.

	Vocabulary	Auditory Memory	Language Processing	Proverbs	Total Score
tense	0.46	0.46	0.25	0.31	0.48
register	0.22	NS	NS	NS	0.19
accent	0.55	0.60	0.32	0.42	0.63
rate	0.57	0.56	0.34	0.40	0.61
nvc	-	-	-	-	-
speech acts	NS	0.56	0.40	0.43	0.50
informal	0.53	0.44	0.41	0.43	0.64
clarification	NS	0.47	NS	NS	0.27
fluency	0.58	0.56	0.32	0.39	0.63
overall	0.58	0.57	0.33	0.40	0.64

All values significant at $p<0.05$.

Word Knowledge Test (WKT) scores (see Chapter IV) were correlated (Spearman *rho*, two-tailed) with items on the Language Rating Scale for all four sets of raters to explore the concurrent validity of these two instruments.

Table LIV: Spearman *rho* correlation coefficients for standardized patient ratings on the Language Rating Scale and Word Knowledge Test (WKT) scores.

	tense	register	accent	rate	nvc	speech a	informal	clarify	fluency	overall
wkt	0.51	0.25	0.59	0.46	0.33	0.40	0.56	0.53	0.45	0.53

All values significant at $p<0.05$.

Table LV: Spearman *rho* correlation coefficients for the researcher's ratings on the Language Rating Scale and Word Knowledge Test (WKT) scores.

	tense	register	accent	rate	nvc	speech a	informal	clarify	fluency	overall
wkt	0.57	0.28	0.69	0.69	-	NS	0.57	0.39	0.71	0.71

All values significant at $p \leq 0.005$.

Table LVI: Spearman *rho* correlation coefficients for English Language Specialist I's ratings on the Language Rating Scale and Word Knowledge Test (WKT) scores.

	tense	register	accent	rate	nvc	speech a	informal	clarify	fluency	overall
wkt	0.61	0.25†	0.49	0.38	-	0.54	0.45	0.50	0.47	0.44

Significant at $p<0.005$.
†$p<0.05$

Table LVII: Spearman *rho* correlation coefficients for English Language Specialist II's ratings on the Language Rating Scale and Word Knowledge Test (WKT) scores.

	tense	register	accent	rate	nvc	speech a.	informal	clarify	fluency	overall
wkt	0.49	0.48	0.53	0.24†	-	0.37	0.60	0.39	0.50	0.58

Significant at $p<0.005$.
†$p<0.05$

Spearman correlation coefficients were calculated to explore the concurrent validity of the Language Rating Scale scores made by the standardized patients to the ratings of the "use of clear, unambiguous and fluent English" and the "use of appropriate language to the person and context" and "no medical jargon without explanation", as rated by the standardized patients examining in the Medical Communication Skills OSCI (described in Chapter V).

Table LVIII: Spearman *rho* correlation coefficients for standardized patient ratings on the Language Rating Scale and ratings of language made during the Medical Communication Skills OSCI.

	Clear, unambiguous, fluent English	Appropriate language for person & context	No medical jargon without explanation
tense	0.32	NS	0.20
register	NS	NS	0.21
accent	0.45	0.17	0.40
rate	0.36	NS	0.25
nvc	0.24	NS	0.21
speech acts	0.19	NS	0.21
informal	0.41	0.24	0.40
clarify	0.29	NS	0.24
fluency	0.34	NS	0.25
overall	0.40	0.21	0.31

All values significant at $p<0.05$.

Finally, the uniqueness of data (Andrew, 1977) was examined by correlating the scores made by standardized patients on the Language Rating Scale with students' performance on the Medical Communication Skills OSCI. Andrew (1977) stated that:

> "whenever a new assessment technique is introduced attempts should be made to evaluate whether the new technique assesses components of competence not already being assessed by existing examination instruments. An appropriate conclusion can be drawn from the extent to which performance on the new technique correlates with performance on existing measures. If these correlations are high, it is appropriate to conclude that the new technique is only measuring aspects of competency that are already being measured by existing instruments" (p 590).

Table LIX: Spearman *rho* correlations for standardized patient ratings on the Language Rating Scale and Medical Communication Skills OSCI items.

	1	2	3	4	5	6	7	8	9	10
tense	0.22	NS	NS	NS	0.23	NS	NS	NS	NS	NS
register	0.23	NS	NS	NS	0.19	NS	NS	NS	NS	NS
accent	0.30	NS	NS	NS	0.32	0.16	NS	0.15	NS	0.20
rate	0.23	NS	0.17	NS	0.31	0.20	NS	0.20	NS	0.23
nvc	0.20	NS	NS	NS	0.26	0.16	NS	0.16	NS	0.19
speech acts	0.21	NS	NS	NS	0.29	NS	0.18	NS	NS	NS
informal	0.22	NS	0.17	NS	0.29	NS	NS	0.19	NS	0.21
clarify	0.23	NS	0.19	NS	0.34	NS	0.17	0.20	NS	0.25
fluency	0.25	NS	0.16	NS	0.29	0.17	NS	0.18	NS	0.20
overall	0.25	NS	0.18	NS	0.30	0.18	0.15	0.21	NS	0.24

All values significant at $p<0.05$.

Key:
1. introductions, including explaining purpose of the interview and obtaining informed consent
2. exploration of the patient's knowledge of the problem
3. exploration of patient's concerns
4. summarizing, checking and finishing the interview
5. rapport and attending
6. listening
7. appropriate use of questioning techniques
8. empathy skills
9. personal and professional qualities such as warmth, concern, sincerity and respect
10. overall feeling of empathy, reflected in the creation of an atmosphere of trust, support and sensitivity to the patient's emotional state

Construct validity.

To first investigate the construct validity of the Language Rating Scale the internal consistency Cronbach's alphas were examined by comparing the two forms of the scale after employing the split-half method as described for the consideration of internal consistency. Only the standardized patients' ratings could be meaningfully analysed due to problems with missing data with other raters' scores. Where missing data were excluded from standardized patients' ratings, 97 cases could be included in analyses, yielding alphas of 0.90 and 0.90 for both halves of the scale. When missing data were included so that all 149 cases could be included in the analyses, the alphas were 0.77 and 0.70 for the first and second half of the scale respectively.

In order to further establish construct validity of the Language Rating Scale, t-tests were performed on all items comparing the mean scores of English speaking background students with students from non-English speaking backgrounds.

Using standardized patient ratings as the dependent variable, English speaking background students were rated as significantly higher than non-English speaking background students on use of correct tense (t=5.86, df=141.70, p=0.000), register (t(pooled)=2.38, df=147, p=0.019), comprehensibility due to accent (t=7.77, df=120.73, p=0.000), rate of speech (t=5.34, df=145.80, p=0.000), appropriate non-verbal communication (t(pooled)=2.84, df=144, p=0.005), speech acts (t(pooled)=4.13, df=138, p=0.000), understanding of informal language (t=8.62, df=141.54, p=0.000), clarification where comprehension lacking (t=5.01, df=105.30, p=0.000), fluency of speech (t=5.69, df=147, p=0.000) and overall impression of language proficiency (t=6.22, df=146.98, p=0.000).

The English language specialist I ratings showed significant differences in favour of the English speaking background students on use of correct tense (t=6.79, df=131.62, p=0.000), register (t(pooled)=5.13, df=138, p=0.000), comprehensibility due to accent (t=9.04, df=138.81, p=0.000), rate of speech (t(pooled)=4.50, df=143, p=0.000), speech acts (t(pooled)=2.64, df=33, p=0.013), understanding of informal language (t=6.27, df=54.32, p=0.000), clarification where comprehension lacking (t=5.48, df=92.00, p=0.000), fluency of speech (t(pooled)=6.96, df=141, p=0.000) and overall impression of language proficiency (t=7.57, df=133.63, p=0.000).

Similarly, The English language specialist II ratings showed significant differences on use of correct tense (t=7.06, df=87.31, p=0.000), register (t(pooled)=5.07, df=145, p=0.000), comprehensibility due to accent (t=6.38, df=146.80, p=0.000), rate of speech (t(pooled)=3.84, df=147, p=0.000), speech acts (t(pooled)=2.59, df=90, p=0.011), understanding of informal language (t=8.49, df=118.77, p=0.000), clarification where

comprehension lacking (t=6.76, df=121.97, p=0.000), fluency (t=7.02, df=147.00, p=0.000) and overall impression of language proficiency (t=8.75, df=137.06, p=0.000).

Where the researcher's ratings were used in analyses, English speaking background students scored higher on average on use of correct tense (t=5.29, df=87.0, p=0.000), comprehensibility due to accent (t=9.13, df=101.47, p=0.000), rate of speech (t=8.51, df=99.23, p=0.000), understanding of informal language (t=6.36, df=45.27, p=0.000), clarification where comprehension lacking (t=3.62, df=58.73, p=0.001), fluency of speech (t=8.86, df=94.91, p=0.000) and overall impression of language proficiency (t=8.86, df=95.07, p=0.000). No significant differences in mean ratings between the two groups of students were found for the scale items of use of appropriate register or speech acts.

Test security.

To investigate whether the time during the day at which the student was interviewed had any bearing upon ratings, Pearson correlation coefficients were calculated between the time of testing and each item on the Language Rating Scale as scored by the standardized patients. Significant correlations were found for "non-verbal communication" (Pearson r=0.17, $p<0.05$) and "clarification where comprehension lacking" (Pearson r=0.22, $p<0.05$).

To determine the relationship between the time at which the interview was taken and these two items on the Language Rating Scale chi-square analyses were conducted according to the session of the interview (9.00am to 10.45am, 11.15am to 1.15pm, 2.15pm to 3.30pm and 4.00pm to 5.15pm).

More students were rated lower on their non-verbal communication skills than would be expected between 2.15pm and 3.30pm, and were rated higher than expected from 9.00am to 10.45am. Students were rated higher than expected on "clarification where comprehension lacking" during the first two morning sessions of interviewing and lower in the final afternoon session, from 4.00pm to 5.15pm.

Rater bias.

Because of the fact that only one of the six standardized patients was male, it was not possible to conduct analyses to investigate whether the rater's gender had any influence upon their decisions. Unfortunately, other demographic data such as age, language background, knowledge of a second or third language and previous experience with people of non-English speaking background were not collected from raters.

The three other raters (the two English Language Specialists and the author) were also all female and no other demographic information about them was collected.

To examine whether raters scored students differentially according to gender t-tests comparing the mean ratings of males and females across all four sets of raters (standardized patients, the two English Language Specialists and the researcher) were conducted. There were no significant differences on the majority of Language Rating Scale items according to gender. However, both the standardized patients and the English Language Specialist II rated females as having a significantly more appropriate rate of speech ($t=-2.04$, $df=143.93$, $p=0.043$ and $t(pooled)=-2.22$, $df=147$, $p=0.028$ respectively). The researcher's scores just failed to reach significance on this item, also in favour of females ($t=-1.94$, $df=146.46$, $p=0.054$). The researcher also rated females as being significantly more appropriate in their speech acts than males ($t=-2.71$, $df=30.60$, $p=0.011$).

Validity of presentations.

From the comments made by raters and presented above it is clear that standardized patient presentations were not uniform.

Discussion.

This study aimed to develop an instrument that could be used by trained standardized patients to rate students' spoken English language in the context of a clinical encounter and to administer it to a group of third year medical undergraduates. As a pilot

study for Study VI which follows in Chapter VIII, a major purpose of this research was to identify methodological weaknesses and ways in which they could be rectified.

The present study was successful in terms of replicating the methodology of previous research which has utilized standardized patients to rate spoken English. The situation in which students interviewed a standardized patient approximated real clinical practice, and thus the procedure had good face validity.

The instrument itself was based upon information gleaned from clinicians who had experience with non-English speaking background students and was also approved by several experts in teaching English as a second language. On this basis, it can be argued that the Language Rating Scale also has good face validity.

The ratings made by all raters (the standardized patients, the two English Language Specialists and the researcher) indicate that although the majority of students could be considered proficient in their ability to use the English language appropriately with a patient, this group of students was far from homogeneous in their language skills (Tables XXXVIII, XXXIX, XL and XLI and Appendix XXIV). A number of students were rated as "poor" or "between poor and adequate", especially by the standardized patients, and thirteen students were considered by the researcher to be at considerable risk of academic failure without some form of language intervention. This concurs with clinicians' concerns, reported in Study IV, that some students were not equipped with the language skills necessary for interactions with patients.

The amount of missing data generated in this study is worrisome, and must be taken into account when interpreting the results of this research. For example, standardized patients were often unable to rate the item "clarification where comprehension lacking". This may have been because the students who did not comprehend were skilled at concealing their confusion, or it may have been that the standardized patients were unable to engineer situations where students were forced to ask

for further explanations. Some of the comments made by standardized patients (Appendix XXV) lend support to both of these propositions. Two standardized patients contributed substantially to the missing data on this item and also on the item regarding "accent". This may have reflected a misunderstanding on their part, in that they may have thought that no rating was to be provided if the student understood or if their accent did not interfere with comprehension. This would indicate that further training of standardized patients was in order.

Missing data was also problematic for both English Language Specialist ratings and the researcher's ratings. One of the English Language Specialists had a tendency to circle more than one point on the scale, or make midpoint markers, meaning that these judgements could not be used since they did not conform to the conventions used by other raters in the assessments. Also of concern for all three of these raters was the fact that since they listened to an audiotape of the interaction, non-verbal behaviours could not be rated. One English Language Specialist and the researcher were often unable to rate students' understanding of informal speech. Therefore it was deemed prudent to include an item on the second version of the Language Rating Scale which would compensate for this. "Speech acts" and "clarification where comprehension lacking" were also items which contributed to the overall levels of missing data. Comments made by the English Language Specialist II and the researcher indicated that this may have been partially due to a lack of opportunity for the standardized patients to engage students in these linguistic behaviours.

One consequence of the high incidence of missing data was that a meaningful global score of English language proficiency according to the scores on the Language Rating Scale (by adding the scores on each item and then dividing by the total number of items) could not be computed. Since the interviews between standardized patients and students were audiotaped and not videotaped, this must be borne in mind when making comparisons between raters who were present during the student-standardized patient interview and those who rated from the tape-recording. For example, interrater reliability

coefficients for non-verbal behaviours could not be calculated because of the unavoidable missing data on this item. Similarly, different ratings on items such as "understanding of informal language" may be due partly to the absence of non-verbal cues for the rater, where a student's facial expression may have indicated confusion. Videotape recordings would have avoided these problems, and have been used previously for evaluation of medical students' performance, particularly in the areas of medical communication and interviewing skills (Beckman and Frankel, 1994) employing both students and "real" patients (Block, Schaffner and Coulehan, 1985) and students and standardized patients (Schoonover, Bassuk, Smith and Gaskill, 1983). Unfortunately, the financial costs involved in providing the equipment needed to videotape in the present research was prohibitive. Where possible, videotaping is preferable to audiotaping because the data yielded are more complete. However, Thompson (1994) found that even with videotaped encounters, interrater reliabilities were low between scorers of the non-verbal behaviours of family practice examinees toward patients.

Previous researchers (Friedman, Sutnick, Stillman, Norcini, Anderson, Williams, Henning and Reeves, 1991) expressed concern that standardized patients may not be able to rate spoken language reliably since they cannot necessarily generate speech from the person interviewing them. This drawback was certainly relevant here, as evidenced by the comments made by standardized patients regarding difficulty in keeping students talking. Both English Language Specialists also noted that some standardized patients were more able than others to elicit speech from students. In the study which follows both the standardized patients and the students were informed that conversation from the student was an essential component of the process and thus both parties were responsible for facilitating this.

The comments made by all four sets of raters were extremely useful. The standardized patients had all been requested to give detailed comments where appropriate, and it was disappointing that no comments whatsoever were made for approximately 40% of the cohort. Standardized patients employed for the next study were directed to ensure

that they provided comments for as many cases as possible, particularly where such comments would elucidate the reasoning behind a given rating. The standardized patients' comments in the present study demonstrated some confusion over the distinction between language skills and medical communication skills. Again, this was deemed to be an artefact of inadequate training procedures.

The comments made by the English Language Specialist I also showed this confusion with medical communication skills and language skills. She suggested that rapport be included on the scale, but as it was deemed a medical communication skill as opposed to a linguistic capability, this was not accepted. Her suggestion to rate students on intonation was considered, but rejected on the grounds that it may be too difficult a cognitive task for standardized patients to perform during an interview. Her argument that students' audibility be assessed was accepted and incorporated into the second version of the Language Rating Scale.

The comments made by the English Language Specialist II were most comprehensive and invaluable in enabling the researcher to recognize areas of incongruity in the working definitions used to score students on the Language Rating Scale. For example, her conceptualization of "register" was far broader than the definition intended. This was also the case for "speech acts" and "clarification where comprehension lacking". Thus, it is apparent that raters may have been working with very different ideas about what each item on the scale represented than those provided. Indeed, Kazdin (1977) has referred to this as "observer drift", where raters tend to alter either definitions or the manner in which they apply them over time. Low reliabilities on the Language Rating Scale may therefore be due to items either requiring or allowing interpretation by the rater - higher reliabilities should be found for items where the items have clear behavioural criteria (Kraan, Crijnen, Zuidweg, van der Vleuten and Imbos, 1990; Vu, Marcy, Colliver, Verhulst, Travis and Barrows, 1992). This rater pointed out that she rejected outright the definition of informal speech provided, and instead used her own conceptualization of the way in which this item should be applied.

The researcher's observation from the audiotapes that many students seemed to be in a state of confusion and distress about their role and the task itself was important, in that some students may have been disadvantaged if their speech became faltering or hesitant as a result of emotional arousal. Furthermore, since this study was in no way related to their formal assessment, it was regrettable that the research project may have contributed to levels of stress for them.

The internal consistency of the Language Rating Scale as it was applied in this study was good to excellent. The inter-item consistency and content validity was also generally good to excellent across the scores of all four sets of raters. Inspection of the correlation coefficients on Tables XLII to XLV show the greatest consistency between items when students were rated by the standardized patients and the least when rated by the researcher. This may be explained by the phenomenon of "observer drift" discussed above. That is, perhaps the researcher was rating students according to different criteria than the other raters. It is important to note, however, that whilst inter-item consistency was high when collective standardized patients' ratings were considered, this was not necessarily the case when examining individual standardized patient ratings. For example, the inter-item consistency of standardized patient I was non-existent for "register" (Appendix XXVI). Thus, it is very probable that different standardized patients were working according to different definitions on at least some of the items on the scale.

Since register was moderately correlated with all of the items on the Language Rating Scale for three of the four raters (Tables XLII, XLIII and XLIV), it must be accepted as consistent with the scale in general. Of note, however, is the fact that the researcher's ratings of register (Table XLV) yielded non-significant correlations with all other items on the scale. Since the researcher was working according to the intended definition, it is perhaps questionable whether register, as it was conceptualised for this study, is a relevant aspect of language usage to assess. This is supported by the results of Table LVIII, which show a general lack of relationship between register as assessed on the Language Rating Scale and the Medical Communication Skills OSCI item of "appropriate

language for person and context" and a low correlation with "no medical jargon without explanation".

Interrater reliability between the four sets of raters differed according to the item in question and the method of analysis. Van der Vleuten and Swanson (1990) have argued that validity considerations are sometimes more important than interrater agreement. Considering Spearman *rho* correlation coefficients (Tables XLVI, XLVII and XXLVIII) register was generally poorly agreed upon by raters and accent highly agreed upon, with most correlations in the moderate range. However, Table XLIX shows that when Kendall's coefficient of concordance is applied to the data, interrater reliability is excellent on all items except accent, speech acts and use of informal language. The apparent incongruity of this finding for accent can be explained by the fact that the English Language Specialist II's ratings on this item were substantially lower than the other three raters, and as the Kendall statistic ranks scores, this harsher rating has been reflected in the *W* statistic. One must interpret the non-significant results of speech acts and use of informal language with caution, as the statistics have been calculated on comparatively small N's, due to missing data.

Inspection of the results of t-tests between the four sets of raters' assessments show that the researcher consistently rated students higher than the standardized patients on most items. It is possible that this resulted as a function of different conceptual frameworks (as has been discussed). Alternatively, perhaps this is an apt demonstration of why a person who develops a scale and who also knows the people to be rated with it quite well, should not administer that instrument. That is, ratings performed by examiners blind to as many aspects of the student as possible are preferable as they are less open to subjectivity and bias.

A further consideration which may explain some of the discrepancies within and between the ratings of the standardized patients and the other three raters, including the researcher, is the fact that the patients rated students "in vivo". Therefore, the issue of

"halo effects" when using rating scales (Andrew, 1977) must be considered. Although the standardized patients were required to evaluate specific components of the students' spoken English language, the rater's perception of extraneous characteristics of the student may have either positively or negatively influenced that rater's assessment. Such characteristics might include "the examinee's general appearance, likeability, manner of presentation, communication skills, demeanour and personality" (Harasym and Woloschuk, 1992, p 437). Thus, the halo effect can be defined as an inability to differentiate between the impression formed of a student on the basis of these extraneous characteristics and the specific skill to be assessed. For example, Wible and Hui (1985) have demonstrated that speakers considered highly fluent in the language of the listeners were rated as more competent, intelligent and ambitious than less proficient speakers. Harasym and Woloschuk (1992) have suggested that the halo effect can be minimised by dividing complex behaviours into simpler sub-tasks, evaluating performance from direct observation rather than from memory, piloting rating forms and responding to feedback for improvement from users and through statistical analyses and ensuring that raters are trained in the use of the form. Anastasi (1988) added defining the traits to be assessed in concrete terms and provision of training for raters as further ways in which the halo effect could be minimized. Thus, whilst precautions were taken to avoid this bias, it is plausible that it affected ratings in this study, and that the bias would be stronger where the candidate was seen rather than merely heard via audiotape.

Anastasi (1988) has also referred to common rating errors in addition to the halo effect, these being the error of central tendency (avoiding the scale's extremes) and the leniency error (whereby raters do not wish to be harsh in their judgements). She recommends that ranking of candidates be carried out in order to correct such biases, although this procedure was inappropriate for the present study, where subjects were evaluated by different raters, and therefore direct comparisons may not be meaningful (Anastasi, 1988).

Powis (1996) discussed a further possible bias in rating students, the contrast effect, whereby raters not only judge candidates according to their performance, but are influenced in their final ratings by trying to place the student in a ranking system, relative to other candidates that they have assessed. Powis (1996) also described a tendency for raters to respond positively to people of a similar background to their own. In this study such a bias may have operated in a manner consistent with Giles' (1977) hypothesis that the greater the similarity in speech, the more two people will be attracted to one another. As a consequence it is possible that more favourable ratings may be given by assessors who share the same language background as the student, although it was not possible in this study to ascertain whether this might be the case.

Powers and Stansfield (1983) listed a number of the characteristics of people used as judges of spoken English which they considered important to take into account when assessing the reliability and validity of their ratings. These were their country of birth, whether they had travelled abroad, whether they had studied a foreign language at secondary or tertiary level, whether they were proficient in another language apart from English, frequency and type of contact with non-native English speakers, whether they could be considered to be a rural, town or city dweller, and their age and gender. Thus, a future study might draw upon standardized patients from diverse language and cultural backgrounds in order to explore these factors as sources of possible bias when making assessments of language skills.

One possible bias which was avoided in this study was the primacy effect, which has implications for validity (Albanese, Prucha, Barnet and Gjerde, 1997). The primacy effect is observed where ratings on a Likert scale tend towards the positive when the positive end of the response scale is on the left rather than on the right. The Language Rating Scale moved from negative ratings on the left to positive on the right, in an effort to avoid this effect.

Some students were rated as "between poor and adequate" and "adequate" by the standardized patients on their non-verbal behaviours. This indicates room for improvement which can be addressed by training them in both the encoding and decoding of facial expressions, body posture, tone of voice and so on (DiMatteo, Friedman and Taranta, 1979; DiMatteo, 1979). DiMatteo (1979) has reported that evidence exists for the possibility of increasing a person's ability to decode others' non-verbal cues. Although an undergraduate medical student training programme conducted in Australia by Evans, Stanley, Coman and Burrows (1989) which sought to increase sensitivity to non-verbal cues was unsuccessful, a replication of their work would be in order, employing University of Adelaide medical students. These students, who come from heterogeneous language and cultural backgrounds may consequently show greater benefits from such training.

Were training in non-verbal behaviours to be conducted, the Language Rating Scale item of "non-verbal communication" would need to be considerably refined, however. That is, this category is too broad to be useful in assessing pre- and post-training differences. One would need to specifically rate students on touch, eye contact, tone of voice, physical distance norms, cue discrepancies and facial expressions. To have rated students in such detail in this study would have been too arduous a task for the standardized patients, with so many other aspects of the students' behaviour to consider. Again, had the interviews been videotape recorded, it may have been possible to conduct a more fine-grained analysis of non-verbal communication.

The reproducibility of scores across clinical cases (or test-retest reliability) presupposes that more than one station is involved in testing. Van der Vleuten and Swanson (1990) recommended 3 to 4 hours of testing for minimal reproducibility. As only two stations (the interview and the written task) were used for a maximum of twenty minutes in this study one could query the reproducibility of results. A more comprehensive research design would allow students' English language proficiency to be rated across several different standardized patients and for a number of different tasks. Future research would benefit from a consideration of the issue of case specificity

(Swanson and Norcini, 1989), meaning that the performance of a student with one standardized patient does not necessarily predict future performance with other cases. Thus, multiple stations which rate language proficiency would allow greater generalizability of results across students than could occur with this study.

The time of day at which interviewing was conducted had negligible impact upon students' performance. However, some differences were found in ratings of non-verbal communication and clarification where comprehension lacking as scored by the standardized patients. Since higher ratings were allocated during the morning on these two items, these findings cannot be attributed to students passing on information to later examinees about what possible criteria might be viewed positively by raters. Thus, it may be that these results are purely coincidental. Had more than one station been employed, the reasons for differences in scores across time could have been further explored, but this is not possible with the use of a single station as occurred here.

Concurrent validity of the Language Rating Scale with the Screening Test of Adolescent Language (STAL) and the Word Knowledge Test was generally moderate across all four sets of raters (Tables L to LIII and Tables LIV to LVII respectively). Thus, it can be argued that the Language Rating Scale is tapping into similar skills as those evaluated by these two instruments. Low to moderate but significant correlations were found between the Language Rating Scale and ratings of language made during the Medical Communication Skills OSCI (Table LVIII). That the correlations were so low between these ratings and register are of interest, and have been discussed above.

In support of the argument posited in Study III of Chapter V, that language skills and medical communication skills are independent, are the results presented in Table LIX. The Language Rating Scale and most of the non-language items scored during the Medical Communications Skills OSCI yielded low correlations, or were not significant. The students' introduction and their rapport and attending were most consistently related to the Language Rating Scale. The introduction may have been rated by the examiners with

students' language proficiency in mind because the candidate had not had any opportunity to compensate for language difficulties at that point in the interview. Rapport and attending is a problematic item to explain methodologically because it is double-barrelled. Were the examiners responding to rapport, or to attending? Possibly rapport may be hampered if students struggle with language, although there is no evidence to suggest that this is the case.

That three of the ten non-language items on the Medical Communications Skills OSCI (exploration of the patient's problem, summarizing, checking and finishing the interview and personal and professional qualities) yielded no significant correlations indicates that the Language Rating Scale is certainly assessing an independent skill. It is especially noteworthy to see that students' personal and professional qualities were not viewed negatively even if their language skills were considered wanting.

Students from non-English speaking backgrounds were consistently rated lower on items of the Language Rating Scale as compared with English speaking background students. Therefore, the construct validity of the instrument is high. Overall, there was no effect of performance during the interview according to the time at which it was conducted. Where differences were found (on "non-verbal communication" and "clarification where comprehension lacking") these were fairly minor, and must be explained as occurring by chance. Given that the ratings were of language skills that could not be rehearsed or learned over an hour or two, it is doubtful that test security need be considered a threat to the reliability or validity of these results.

Females were found in this study to be rated as significantly more appropriate in their rate of speech and speech acts. The latter finding might be explained in terms of social conventions and the possibility that females are more sensitive to the requests, apologies and thanks of others, although this open to conjecture. Conversely, only the researcher's ratings reflected this trend, and as a female herself who knew all of the students well, this result may simply reflect a bias on her part. Why females should be

considered by all four sets of raters to have better rates of speech than males is difficult to explain. Interestingly, it is consistent with the finding in Study II, whereby females scored higher on the STAL than did male students.

The consistency and therefore validity of the standardized patient presentations must be questioned. Clearly, as a pilot study, this research demonstrated that a more intensive, comprehensive and directive training programme was required. Norman, Muzzin, Williams and Swanson (1985) have said that although a useful tool, one must consider the costs associated with the training and use of standardized patients. Fortunately, this research was supported by a University of Adelaide Teaching Development Grant for $3, 500 supplemented by funding from the Faculty of Medicine, without which it could not have been conducted (Appendix II.I). Nevertheless, extensive training incurs additional expense that must be borne in mind when utilizing standardized patients.

The study that follows in Chapter VIII aimed to utilize the information gathered in this study and replicate it, after addressing methodological flaws where possible.

Chapter VIII.

Study VI.

Language Rating Scale (LRS) Study and Written Language Rating Scale (WLRS) Pilot Study .

Aims of the present study.

The study described in this chapter had six main objectives:

(i) to replicate the study presented in the preceding chapter, with some refinements.

(ii) to further explore the reliability and validity of the Language Rating Scale.

(iii) to compare the performance of the 1994 third year students with the 1995 second year students on the Language Rating Scale.

(iv) to develop a reliable and valid instrument that could be used to rate the written English language proficiency of a student's account of an interview with a standardized patient.

(v) to administer that instrument as a pilot study to the 1994 third year and 1994 first year (1995 second year student) cohorts.

(vi) to compare the performance of the 1994 third year students with the 1995 second year students in written English language proficiency as assessed by that instrument.

Introduction.

On the basis of the information gleaned from the Language Rating Scale pilot study reported in Chapter VII, two additional items were included in the scale.

Since standardized patients had sometimes found it difficult to evaluate students' understanding of informal language an item was included on the scale to assess students' use of such language. It was supposed that usage of colloquial speech could be used as a

marker of familiarity with such language where the standardized patient was unable to otherwise gauge this.

Furnham and Bochner (1986) noted that across cultures there are differences in the intensity of the sounds of speech - they give the example of Americans louder speech sounding brash to the more quietly spoken English person's ears. It was considered prudent to evaluate the audibility of students' speech, as it was hypothesized that non-English speaking background students may be more quietly spoken than their English speaking background counterparts, and that this may be detrimental to listener comprehension. The suggestion of the English Language Specialist rater in the preceding study that students' loudness of voice be assessed was instrumental in the inclusion of this item in the Language Rating Scale.

One other minor change was made to the overall scale. The term "correct" was altered to "appropriate" for the item dealing with the use of tense, to be consistent with the rest of the items.

Thus, the final scale (Appendix XXVII) for this study, which again utilized a five point rating scale, comprised 12 items;

1. Use of appropriate tense
2. Use of appropriate register
3. Comprehensibility of speech due to accent
4. Appropriate rate of speech
5. Appropriate use of non-verbal communication
6. Response to requests, apologies and/or thanks (speech acts)
7. Understanding of informal language
8. Use of informal language
9. Clarification where comprehension lacking
10. Audibility of speech

11. Fluency of speech

12. Overall impression of language proficiency

An opportunity to make open-ended comments was again provided.

To ensure that standardized patients were fully conversant with the meanings of each of the items on the Language Rating Scale the definitions and rating criteria were more clearly specified. In addition to defining each item on the scale, standardized patients were provided with a prompt question to ask themselves when arriving upon a score and what each score from 1 through to 5 represented. Thus, the following information was provided;

Use of appropriate tense.

"Tense" is the relationship between the form of the verb and the time of the action or state that it describes.

Does the student employ the correct tense during the interview, in terms of relaying the intended meaning?

Rating

1. inappropriate use of tense creates considerable misunderstanding
2. inappropriate use of tense creates some misunderstanding
3. inappropriate use of tense creates unimportant misunderstanding
4. inappropriate use of tense creates no misunderstanding
5. no inappropriate use of tense

Use of appropriate register.

"Register" is the style of speech that is used by a particular group of people, with its own special words or phrases, and sometimes with its own grammar, as in legal registers. Usually people who share a common register come from the same occupation, or have the same interests. Therefore doctors have their own register, and they use this when interacting on a professional level. It is inappropriate for a doctor to use a medical register when consulting with a patient. For example, the use of medical terminology ("jargon") without clear explanation of the meaning, is inappropriate.

Does the student use the appropriate register throughout the interview?

Rating

1. student often relies on specialist register to relay meaning
2. sometimes relies on specialist register to relay meaning
3. rarely relies on specialist register to relay meaning
4. sensitive to register, can modify speech to suit patient, minor errors
5 sensitive to register, can modify speech to suit patient, no errors

Comprehensibility of speech due to accent.

"Accent" is defined as a way of speaking which tells the listener something about the speaker's background, such as the country from which they come, their social class, and whether the speaker is a native speaker of the language. Sometimes accent can interfere with the listener's understanding of the speaker's words.

Does the student's accent influence the understanding of the meaning of the speech in any negative manner?

Rating

1. frequently pronunciation of words is poor, and speech difficult to understand
2. sometimes pronunciation of words is poor, and speech difficult to understand
3. accent is present, but does not significantly affect patient's comprehension
4. accent present, does not interfere with patient's comprehension
5. accent, if present, in no way hampers patient comprehension

Appropriate rate of speech.

"Rate" refers to the rate or speed at which a person is speaking. Rate of speech can be objectively measured by counting the number of syllables per minute. However, one can have a subjective impression of the rate of speech as seeming more rapid or slower than one feels is usual.

What is the student's rate of speech? Does it interfere with the listener's understanding of the speaker?

Rating

1. speech is notably too fast or too slow, interferes with listener's comprehension
2. rate of speech sometimes interferes with listener's comprehension
3. rate of speech rarely interferes with listener's comprehension
4. rate of speech is not of a native speaker's, but does not affect comprehension
5. rate of speech is as a native speaker's

Appropriate use of non-verbal communication

"Non-verbal communication" (nvc) includes facial expressions, head and eye movements, body and hand gestures that support, emphasize or alter the meaning of the spoken word.

Are the student's non-verbal behaviours in accord with what he or she is saying?

Rating

1. nvc significantly inappropriate to interaction e.g. laughing, poor eye contact
2. nvc sometimes inappropriate
3. nvc rarely inappropriate
4. nvc almost always congruent with interaction
5. nvc always congruent with interaction

Response to requests, apologies, thanks.

Requests, apologies, thanks and invitations are called "speech acts".

Does the student respond appropriately to the patient's speech acts?

Rating

1. poor grasp of social conventions, interferes with interaction
2. some difficulty with speech acts, interferes with interaction
3. some difficulty with speech acts, but does not hamper interaction
4. rarely inappropriate in response to speech acts
5. always responds appropriately to speech acts

Understanding of informal language.

Includes figures of speech, colloquialisms, slang. In a "figure of speech" most commonly the speaker uses a simile or metaphor for special effect, but does not intend the literal meaning to be taken by the listener. "Colloquialisms" and "slang" are informal forms of speech, and include the examples you will use in the standardized interview script, such as "crook", "take a sickie" and "under the weather".

Does the student show by his or her response that they understand the use of these kinds of speech by the patient?

Rating

1. clearly does not comprehend informal language used by the patient
2. does not comprehend some informal language
3. does not always comprehend, but this does not significantly affect interaction
4. usually does understand informal language
5. understanding of informal speech is as native speaker's

Use of informal language.

Does the student demonstrate familiarity with informal language by using it in their own speech?

Rating

1. never uses informal speech
2. attempts at informal speech are awkward, or incorrect
3. attempts may not be correct, but do not hamper interaction
4. may not be completely at ease with informal speech, but is appropriate
5. is clearly comfortable with informal speech and uses it appropriately

Clarification where comprehension is lacking.

"Clarification" refers to the skill of ensuring that one understands what the speaker has said, where there is doubt. For example, where the patient uses unfamiliar informal speech ("under the weather"), or the name of a drug ("Mylanta"), the student should clarify this by rephrasing, or asking the patient to explain.

Does the student clarify what they have not fully understood?

Rating

1. clearly misunderstands and does not ever clarify
2. rarely clarifies, obviously has misunderstood some information
3. sometimes clarifies information which is not understood
4. always clarifies that which is not understood
5. may not clarify, but this is because the student has clearly understood

Audibility of speech.

"Audibility" refers to the projection of speech.

Is the student's speech inappropriately loud or quiet?

Rating

1. speech is inappropriately loud/soft, interferes with interaction
2. audibility is inappropriate, sometimes interferes with interaction
3. audibility is inappropriate, but does not interfere with interaction
4. audibility is usually appropriate, does not influence interaction
5. student's projection is like that of a native speaker's

Fluency of speech.

"Fluency" refers to features of speech that make it seem natural to the listener, including the use of pauses, rhythms, intonations, stress on syllables and words, the rate of speaking, and the use of interjections and interruptions. Thus, fluency is a global indicator

of the quality of speech. It can be described as a level of proficiency in language which includes the ability to produce spoken language with ease; to speak with a good, but not necessarily perfect command of intonation, vocabulary and grammar; to communicate ideas effectively; and to produce continuous speech without causing comprehension difficulties or a breakdown of communication.

Does the student appear to be speaking in English without effort, or does the student appear to struggle? Does the student have difficulty finding the right words and in putting them together into a sentence? Are there difficulties in comprehension for either party? Does communication "break down" at any point, where "break down" is defined as a track of thought being abandoned due to a lack of understanding, the student asking a question/giving an answer that is irrelevant to the patient's initial response?

Rating

1. fluency is poor, uneven, limited vocabulary means time spent searching for right words, words are not in correct place in sentence, patient must repeat self to be understood
2. hesitations as student searches for right words and grammatical constructions, patient may need to repeat self to be understood, student shows more difficulty in longer or more complex sentences than simple, briefer ones
3. appears comfortable with speaking, rarely needs to search for words, and can fill gaps with other speech whilst searching for words where necessary, fluency is good
4. can use language fluently and accurately, using wide vocabulary
5. uses language as would a native speaker, with complete fluency, accuracy and vocabulary

Overall impression of language proficiency.

All things considered, where would you rate this student on the scale from 1 to 5, based on their performance during the interview?

Do you think this student has a level of language sufficient to negotiate an interview with a patient in a medical interview setting?

Rating

1. language is extremely poor, unacceptable standard for interaction with patients
2. language is poor, and is unacceptable for patient interactions
3. language is acceptable for patient interaction, but requires improvement
4. completely acceptable standard of language for patient interactions
5. faultless performance, completely acceptable standard of language

Comments

The items on this scale take into account the **frequency** of a response, and also the **impact** that a response has upon the interaction. You may need to comment upon why you have rated a student as "poor" (1), if for example, they used tense incorrectly only once, but you feel this error was an important one that significantly impacted upon the course of the interview.

Sometimes you will be unable to rate an item - explain why this was so here.

You may also find that you want to say something about the student's language that has not been included on the rating scale, and this should be noted here.

Your open-ended comments are extremely valuable. If there is not enough space on the front of the rating sheet, please feel free to write on the back, which is blank.

Written Language Rating Scale.

In addition to administering the Language Rating Scale, in this study a Written Language Rating Scale (Appendix XXVIII) was designed and then administered to the written accounts of the 1994 third year cohort (who were the subjects of the preceding Chapter) and the present subjects of this Study.

The Written Language Rating Scale was designed on the basis of the author's own experiences in assessing students' written work. To ensure face validity two English Language Specialists were also asked to comment and contribute to the scale[29].

The resultant scale comprised 11 items rated on a Likert-type scale from 1 (poor) to 5 (excellent) with a midpoint of 3 (adequate). The items were;

---

[29] Kate Cadman and Margaret Gunn.

1. Appropriate content
2. Use of appropriate register (jargon)
3. Use of appropriate register (value-judgements)
4. Use of appropriate vocabulary
5. Appropriate use of tense
6. Appropriate use of articles, pronouns, prepositions
7. Appropriate use of spelling, punctuation, capitals
8. Legibility of handwriting
9. Appropriate use of conventions
10. Fluency of written expression
11. Overall impression of written proficiency

As with the Language Rating Scale an opportunity for the rater to provide comments was allowed, as this feedback was considered valuable.

To ensure that the rater used the definitions as conceptualised by the researcher, each item and the criteria for arriving upon a rating were clearly described. Examples taken from actual passages written by students were provided to illustrate the intention of the researcher in including each item.

Content of the passage

Does the passage of text relay exactly what took place between the student and the standardized patient?

Each time the student does not report accurately, deduct one point from the scale.

Examples:

"The patient is afraid her stomach will burst" if the patient said "I am afraid the ulcer will burst".

"Patient takes medication from the doctor for the pain" if the patient said "I have taken something from the chemist".

Use of appropriate register

The student has been instructed to write without using any jargon (1994 third years) or avoiding the use of jargon where possible (1995 second years). Does the student do this?

Each time the student employs jargon, deduct one point from the scale. Do not count instances where the student defines jargon that is used.

Examples:

"Epigastric pain", "Retrosternal pain".

The student has learned that he or she should not make value judgements about patients. Does the student avoid making such judgements in the passage?

For each value judgement, deduct one point from the scale.

Examples;

"She has a pleasant disposition"

"She doesn't seem sincere"

Use of appropriate vocabulary

How sophisticated is the vocabulary in this passage? Do the words chosen clearly communicate the intended message? Are words inappropriately used?

For each instance of inappropriate word use, deduct one point from the scale.

Examples;

"this initiated her to come to the doctor" instead of "this prompted her"

". . . used to be active . . but since the pain she cannot do it" instead of ". . . cannot do what she used to do"

Grammatical construction

Has the student employed the appropriate tense throughout the passage?

For each instance of inappropriate tense, deduct one point from the scale.

Examples;

"the Mylanta milk which help" instead of "the Mylanta milk which helps" or "helped" or "has helped"

"pain that have started" instead of "pain that has started" or "that started"

Has the student used the appropriate articles, pronouns, prepositions?

For each inappropriate article, pronoun, preposition, deduct one point from the scale.

Examples;

"she said" instead of "he said"

"an gastric ulcer" instead of "a gastric ulcer"

Mechanics

Does the student demonstrate an ability to use the appropriate spelling, punctuation, capitalization?

For each error deduct one point from the scale. Do not penalise mis-spellings of trade names, such as "Mylanta" or "Quick-eze".

Examples;

"happilly" instead of "happily"

"a mild annoying gnawing pain" instead of "a mild, annoying, gnawing pain"

Is the handwriting legible?

Give your global impression of legibility on the scale.

Conventions

Does the student demonstrate a knowledge of Western cultural conventions, for example in terms of forms of address?

For each example of inappropriateness, deduct one point from the scale.

Examples;

"Miss Marry" instead of "Mary"

"Ms Sue" instead of "Ms Sue Blake" or "Sue Blake"

Fluency of written expression

Does the passage flow clearly? Does it communicate clearly to the reader the sequence of events during the interview? Does the account end suddenly?

Give your global impression of fluency on the scale.

Overall impression of written proficiency

All things considered, how able is this student to write an account of a short interview, so that another person, without a medical background, can clearly comprehend what took place?

Give your global impression of this passage on the scale.

Method and Procedure.

In 1995 all second year medical students were examined for their ability to demonstrate Medical Communication Skills during an Observed Structured Clinical Interview (OSCI), as has been previously described. Immediately following this examination students proceeded down the corridor to interview a standardized patient during a "Language OSCI" in order that an assessment of their spoken English language could be made (as discussed in Chapter I in relation to the Language Development Programme). The study described in this Chapter paralleled the preceding study in every way possible, so that comparisons could be made across the two cohorts of students. However, there are several important differences between the methodologies of the two studies.

In the previous study feedback about students' language skills was purely formative and in no way contributed to their academic results. In this study, at the direction of the Language Development Committee, the Language OSCI was one

component of an overall summative assessment of language proficiency, and as such acted as a barrier examination. The implication of this was that unsatisfactory performance during the examination could result in a student failing to proceed to the third year of their medical course.

Since the stakes were so high for this Language OSCI the Language Development Committee agreed that the Language Rating Scale should not be used as the instrument to determine satisfactory performance, since it had been designed primarily to serve as a research tool and did not have established reliability or validity. It was also considered inappropriate for standardized patients to be held solely responsible for faculty assessments. Thus, during the interview an English Language Specialist was present with the standardized patient, and it was this person who made a decision about the student's spoken language skills, using an assessment sheet designed by the Language Development Committee with particular guidance from a member of the Committee with Teaching English as a Second Language (ESL) training (Appendix VIII.I).

The Committee's "Oral Language Skills Assessment" sheet covered four areas of ability; 1) to express one's self without being misunderstood (with particular reference to the ability to respond to the standardized patient's opening statements, move from point to point and to interview conclusion, and control vocabulary and grammar); 2) to retrieve communication if the patient misunderstands (especially related to an ability to recognize when the standardized patient did not understand, to rephrase and communicate successfully and use language appropriate to the patient); 3) to clarify when the student was unable to understand the patient (by asking appropriate clarifying questions leading to success in understanding the standardized patient's meaning); and 4) to speak intelligibly (according to appropriate voice projection and modification of enunciation, rhythm and intonation where needed). Each of these four items was rated by the English Language Specialist on a scale from 5 (severe problem) to 1 (no problem). The English Language Specialist then made an overall assessment as to whether the student's spoken English language could be considered as either "satisfactory" or "unsatisfactory".

As with the previous study, following the interview students were required to proceed to another room to provide a written account of what had taken place. However, for this examination students were allowed 10 minutes rather than 5 and their instructions for recording were slightly different (Appendix XXIX). This time, they were advised;

"You have 10 minutes to record the interview you have just completed. Write an account of this interview so that another health care professional would understand what took place, and the decisions and conclusions drawn by yourself and the patient. You should write in complete sentences, and in paragraphs. Your writing must be legible. Please avoid the use of specialized medical terminology. If it is used, ensure that it is explained ."

The Written Language Rating Scale developed by the researcher was applied to the passages generated from this exercise. However, as with the Language Rating Scale, the Language Development Committee developed a corresponding assessment sheet to score the written account of the OSCI interview (Appendix VIII.II). This sheet encompassed three areas, and was rated on a scale from 5 (severe problem) to 1 (no problem). The aspects of written proficiency scored were overall communication of information (whereby the rater considered the extent to which the account communicated clearly to another health care professional of what took place); accuracy in the choice of words and grammatical constructions (where the rater decided whether a health care professional would be likely to misunderstand the events being recounted because of incorrect or misleading word choices, unexplained specialist vocabulary, omissions, incorrect links, inappropriate use of tense, prepositions or articles; and presentation and editing (where the rater considered legibility, lack of editing in punctuation, spelling, omissions of final -d or -s). The English Language Specialist then made an overall assessment as to whether the student's written English language could be considered as either "satisfactory" or "unsatisfactory". It should be noted that the English Language Specialist who rated the student's written work was not necessarily present at that student's interview with the standardized patient.

Standardized Patients.

Seven postgraduates were recruited to act as standardized patients for this study, although only six people were needed for the day of testing[30]. A stand-by patient was trained in the event of mishap, as in the previous study. All of the standardized patients were female and were native English speakers fluent in English. Six of the seven had teaching experience with undergraduate students, and four had taught non-English speaking background medical students at the University of Adelaide, although they had not had teaching responsibilities for the 1994 first years/1995 second years. One of the standardized patients had a qualification in medicine, one in public health, one in nursing and the others had degrees in psychology.

Three of the standardized patients in this study had participated in the previous study, whilst the other four had not acted in this role before. On the day of testing three of the "experienced" and three of the "novice" standardized patients were employed to take part in the study.

Training for standardized patients was conducted over two three-hour sessions. In the first session, held three weeks prior to the day of language testing (on November 2nd, 1995), trainees were oriented to the task and the format of the examination explained. Of particular importance was the need to stress that ratings made by the standardized patients would not be used in the formal assessment of students. It was also stressed during this session that an English Language Specialist would be in the room with the standardized patient and the student during the interview, and that it would be their role to provide language assessments for the faculty. However, standardized patients were advised that it was imperative that no colluding between themselves and the other rater took place when deciding upon ratings. Standardized patients were requested to avoid as far as practicable, any form of non-verbal contact or discussion with the English Language Specialist present in their interview room.

---

[30] The standardized patients were Julie Clifford, Mary Katsikitis, Julie Mattiske, Jane Mortimer, Jan Scicchitano, Niccola Spurrier and Robyn Young.

During the first training session the criteria for the Language Rating Scale were outlined and discussed. Difficulties that had arisen in the pilot study of the Language Rating Scale were considered, and standardized patients who had participated in the previous study were invaluable here in providing examples of problems they had encountered and how they had negotiated them.

Finally in this session the author and her clinical colleague, Dr Jane Vernon-Roberts demonstrated a mock student-standardized patient interview, as required on the day of testing.

Standardized patients left the session with a training satchel which included the standardized patient script (Appendix XXX), a copy of the Language Rating Scale (Appendix XXVII), definitions of the items on the scale and rating criteria (Appendix XXXI), instructions that were to be given to students prior to entering the examination (Appendix XXXII), the researcher's contact telephone numbers and a map of the hospital rooms in which the interviews were to be conducted on the day of the examination.

The second session of training was conducted on the 14th of November, 1995, one week before the examination. For the first hour the contents of the satchel were discussed and questions answered. The three videotapes of the 1993 third year student interviews utilized in the training session for the pilot study were viewed and evaluated on the Language Rating Scale. Discrepancies were discussed. During the second half of this training session the standardized patients were taken to the interview rooms in which they would be working in the following week. The room lay-out, including where the student, the standardized patient and the English Language Specialist would be seated during the interview was demonstrated. The optimal position for the audio-tape recorder was also shown. Several role-plays with the trainee standardized patients then took place in an interview room, which were audio-taped to allow the standardized patients to become familiar with the recording equipment. The role-plays were all observed by the researcher to ensure that a satisfactory standard had been achieved.

Standardized patients were strongly encouraged to practice before arriving for the examination, and asked to seek assistance if they felt the instructions were ambiguous or problematic.

Written Language Rating Scale.

An English Language Specialist was provided with photocopies of all of the written accounts of each students' interview. She listened to the audiotape of the interview and then scored the account according to the criteria of the Written Language Rating Scale. She was blind to all students' details apart from their name. This rater also acted as a rater on the Language Rating Scale in the previous study and for the current study[31].

Results.

One hundred and twenty-seven of the 143 students (88.8%) who commenced first year medicine in 1994 progressed to their second year of the course and were rated on the Language Rating Scale in this study. A further five students who were repeating their second year of the course were also involved in this study, but because they were not members of the original cohort they have not been included in the analyses. The 16 students of the 1994 cohort who did not proceed to second year failed to do so because of withdrawal from the course (N=6) or due to unsatisfactory academic performance (N=10).

Six of the pool of seven standardized patients were required on the day of testing. All were female. Students had been randomly allocated to each standardized patient, with the proviso that there was to be a roughly equal distribution of non-English speaking background students amongst them. Thus, Standardized patient I interviewed 22 students, 15 of whom were from a non-English speaking background; Standardized patient II interviewed 13 non-English speaking background students from a total of 22 students; Standardized patient III interviewed 11 (of 18 students); Standardized patient IV interviewed 22 students, 13 of whom had a non-English speaking background; Standardized patient V interviewed 13 non-English speaking background students out of a

[31] In the previous Chapter she has been referred to as English Language Specialist II.

total of 21; and 15 of the 22 students interviewed by Standardized patient VI were from a non-English speaking background.

Inspection of the collective standardized patient ratings for each of the twelve items on the Language Rating Scale show negatively skewed distributions (Table LX). Ratings for individual standardized patients can be viewed in Appendix XXXIII.

Table LX: Standardized patient ratings, Language Rating Scale.

n=127	1	2	3	4	5	missing*	$\bar{x}$	sd
tense	-	3 (2.4)	8 (6.3)	22 (17.3)	94 (74.0)	-	4.63	0.71
register	-	-	12 (9.4)	46 (36.2)	69 (54.3)	-	4.45	0.66
accent	-	-	14 (11.0)	25 (19.7)	88 (69.3)	-	4.58	0.68
rate	-	-	11 (8.7)	26 (20.5)	90 (70.9)	-	4.62	0.64
nvc	-	2 (1.6)	6 (4.7)	20 (15.7)	99 (78.0)	-	4.70	0.63
speech acts	1 (0.8)	2 (1.6)	2 (1.6)	16 (12.6)	105 (82.7)	1 (0.8)	4.62	0.64
informal 1	-	5 (3.9)	18 (14.2)	20 (15.7)	78 (61.4)	6 (4.7)	4.41	0.89
informal 2	1 (0.8)	1 (0.8)	3 (2.4)	27 (21.3)	54 (42.5)	41 (32.3)	4.54	0.73
clarification	1 (0.8)	2 (1.6)	10 (7.9)	46 (36.2)	66 (52.0)	2 (1.6)	4.39	0.77
audibility	-	2 (1.6)	4 (3.1)	21 (16.5)	100 (78.7)	-	4.72	0.60
fluency	-	8 (6.3)	21 (16.5)	38 (30.0)	59 (46.5)	1 (0.8)	4.17	0.93
overall	-	-	20 (15.7)	51 (40.2)	54 (42.5)	2 (1.6)	4.27	0.72

percentages are in brackets.

* missing=missing data, whereby the standardized patient did not provide a rating.

Key: 1=poor
2=between poor and adequate
3=adequate
4=between adequate and excellent (good)
5=excellent

informal 1 = understanding of informal language
informal 2 = use of informal language

Comments.

Open-ended comments made by the six standardized patients can be viewed in Appendix XXXIV. The number of comments ranged from 8 (44.4% of the interviews conducted by that patient) to 22 (100%). The remaining four standardized patients commented upon all but one or two of the students that they rated, resulting in a total of 127 comments (87.4%) made overall.

The comments could be generally categorized under the following areas. It should be noted that many of the comments were quite comprehensive and encompassed more than one area, and thus the total N does not sum to the total number of students about whom comments were made. Also note that these comments are not listed in any particular order.

1. Use of jargon (N=25)

   Used jargon - asked if I had brought up any sputum and if I had blood in my stools. (Rater 1).

2. Tense (N=5)

   Some problems with tense. (Rater 1).

3. Non-verbal communication (N=35)

   No problems with language but he laughed when I said my parents had died. (Rater 2).

4. Audibility (N=15)

   Spoke a little too softly. (Rater 5).

5. Comprehensibility (N=7)

   He was rather soft and tended to mumble at times but overall was very easy to comprehend. (Rater 1).

6. Colloquial language - positive comments (N=24)

   Used informal language - "the terrible twos," "kill two birds with one stone". (Rater 1).

7. Colloquial language - negative comments (N=26)

   Student clarified because she did not understand any of the informal speech. (Rater 5).

8. Accent (N=10)

   "Aussie" accent tends to interfere with comprehensibility of speech. (Rater 1).

9. Vocabulary (N=2)

   Although this student has an accent she has very fluent speech with no hesitations and she uses a wide vocabulary. (Rater 1).

10. Nervousness (N=27)

Excellent - but very nervous. (Rater 4)

11. Fluency (N=24)

Hesitations in speech (possible worsened by feeling nervous). (Rater 3).

12 Failure to clarify (N=9)

She did not clarify phrases, (but) this student summed up the interview very well and accurately. (Rater 6).

13. Rate of speech (N=23)

Spoke a little fast. (Rater 4).

14. Pronunciation (N=1)

Rather sloppy pronunciation but does not really interfere with comprehension. (Rater 1)

15. Difficulty keeping the student talking (N=3)

This girl was very nervous and not very talkative. (Rater 4).

16. Difficulty arriving upon a rating (N=10)

A bit difficult to assess understanding of informal language. (Rater 6).

Table LXI: English Language Specialist Ratings, Language Rating Scale.

n=127	1	2	3	4	5	missing*	$\bar{x}$	sd
tense	-	9 (7.1)	14 (11.0)	24 (18.9)	78 (61.4)	2 (1.6)	4.37	0.95
register	-	3 (2.4)	22 (17.3)	67 (52.8)	31 (24.4)	4 (3.1)	4.02	0.73
accent	-	-	29 (22.8)	50 (39.4)	46 (36.2)	2 (1.6)	4.14	0.77
rate	-	1 (0.8)	8 (6.3)	65 (51.2)	51 (40.2)	2 (1.6)	4.33	0.63
nvc						127 (100)		
speech acts	1 (0.8)	4 (3.1)	19 (15.0)	55 (43.3)	47 (37.0)	1 (0.8)	4.14	0.84
informal 1	3 (2.4)	4 (3.1)	14 (11.0)	32 (25.2)	72 (56.7)	2 (1.6)	4.33	0.97
informal 2	-	8 (6.3)	14 (11.0)	41 (32.3)	61 (48.0)	3 (2.4)	4.25	0.90
clarification	-	-	7 (5.5)	27 (21.3)	87 (68.5)	6 (4.7)	4.66	0.59
audibility	-	5 (3.9)	15 (11.8)	55 (43.3)	51 (40.2)	1 (0.8)	4.21	0.80
fluency	-	3 (2.4)	21 (16.5)	46 (36.2)	55 (43.3)	2 (1.6)	4.22	0.81
overall	-	2 (1.6)	22 (17.3)	58 (45.7)	38 (29.9)	7 (5.5)	4.10	0.75

percentages are in brackets.

* missing=missing data, whereby the English Language Specialist did not provide a rating.

Key: 1=poor
2=between poor and adequate
3=adequate
4=between adequate and excellent (good)
5=excellent

informal 1 = understanding of informal language
informal 2 = use of informal language

Comments.

The comments made by this rater, who also acted as a rater for the pilot study, can be found in the preceding Chapter. However, she also made comments pertaining to the additional two items included on the revised version of the Language Rating Scale.

With regards to students' use of informal language, she reiterated that the definition given for "informal" language was not satisfactory. Therefore, she did not rate students according to whether they mirrored the use of idiom or colloquial speech where it was used by the standardized patient, but instead, considered whether the style of students' comments were compatible to the conversational style of the standardized patient. If colloquialisms and idioms were used successfully, the grading was made more confidently by this rater, and she indicates that for this reason her scoring is not valid.

Commenting upon audibility of speech, the rater attempted, as far as possible, to focus on the quality of the student's voice production, as defined in the rating criteria. However, she noted audibility also had an effect on the standardized patient e.g. whether the student projected confidence in language use or knowledge of material. There was, she said, usually a change in the audibility of the native-speaker standardized patients if the student's voices were less boisterous than their own.

This rater also commented that clearly, some students have less out-going personalities than others or are naturally quiet. She also indicated that frequently the recordings were poor, with interference from footsteps, drilling, sirens, other context noises and poor audiotape recording equipment.

RELIABILITY OF THE LANGUAGE RATING SCALE.

Split-half reliability (internal consistency).

Employing standardized patients' scores and excluding missing data, 80 cases were included in analyses, resulting in a correlation between the two forms (divided equally into the first and last six items) of 0.78 and a Guttman split-half reliability correlation coefficient of 0.80. When missing data were included so that all cases could be accounted for, the correlation between the two forms was 0.91 and the Guttman split-half reliability correlation coefficient 0.95.

Because non-verbal communication could not be scored by the English Language Specialist, her data were not employed to calculate split-half reliability.

Cronbach's alpha calculated using the standardized patients' collective ratings and excluding missing data was 0.85. With missing data included, the alpha level was 0.97. Cronbach's alpha employing the English Language Specialist's ratings was 0.97 (including missing data). Analyses could not be conducted on these ratings if missing data were excluded.

Inter-item consistency.

In order to determine the inter-item consistency of the Language Rating Scale, all items were correlated with each other employing the Spearman rho (two-tailed) test.

Table LXII: Spearman *rho* correlation coefficients for Language Rating Scale item scores, standardized patient ratings.

Note: Correlation coefficients for individual standardized patients can be found in Appendix XXXV.

	tense	register	accent	rate	nvc	speech acts	informal under.	informal use.	clarify	audible	fluency
register	NS										
accent	0.59	NS									
rate	NS	NS	0.35								
nvc	NS	NS	NS	NS							
speech acts	NS	NS	NS	NS	NS						
informal under.	0.42	NS	0.65	NS	NS	0.28					
informal use.	0.30	NS	0.33	NS	NS	NS	0.52				
clarify	0.45	NS	0.50	NS	NS	NS	0.66	0.47			
audible	NS	NS	0.27	0.27	NS	NS	0.29	NS	NS		
fluency	0.52	NS	0.64	0.26	NS	NS	0.62	0.37	0.57	0.28	
overall	0.52	NS	0.65	0.29	NS	NS	0.67	0.49	0.63	0.27	0.82

All values significant at $p \leq 0.005$

Table LXIII: Spearman *rho* correlation coefficients for Language Rating Scale item scores, English Language Specialist.

	tense	register	accent	rate	nvc	speech acts	informal under	informal use	clarify	audible	fluency
register	0.39										
accent	0.70	0.33									
rate	NS	0.28	0.27								
nvc	-	-	-	-							
speech acts	0.40	0.54	0.34	0.25	-						
informal under.	0.71	0.35	0.60	NS	-	0.32					
informal use.	0.76	0.35	0.65	NS	-	0.39	0.75				
clarify	0.57	0.52	0.41	0.27	-	0.52	0.53	0.55			
audible	0.25	0.43	0.25	NS	-	0.47	0.27	NS	NS		
fluency	0.71	0.45	0.59	NS	-	0.53	0.58	0.60	0.45	0.47	
overall	0.69	0.55	0.68	0.29	-	0.56	0.58	0.58	0.52	0.43	0.79

All values significant at $p \leq 0.005$

Interrater reliability.

Spearman *rho* correlation coefficients were calculated to establish interrater consistency between the standardized patient ratings and those of the English Language Specialist.

Table LXIV: Spearman *rho* correlation coefficients for standardized patient and the English Language Specialist ratings on the Language Rating Scale.

tense	0.54
register	NS
accent	0.45
rate	0.29
nvc	-
speech acts	NS
informal use	0.56
informal under.	0.38
clarify	0.37
audible	0.19
fluency	0.42
overall	0.43

All values significant at $p<0.05$.

Because only two sets of raters were involved in this study, the application of Kendall's coefficient of concordance to the data was inappropriate. However, to further explore differences between raters' scores a comparison of the six standardized patients' mean ratings on each of the twelve items of the Language Rating Scale was conducted for each individual standardized patient using independent sample t-tests (two-tailed).

Standardized patient I rated her students as significantly lower on appropriate rate of speech than did standardized patient II (t=-2.63, df=21.00, p=0.016). There were otherwise no significant differences between these two patients' ratings.

Standardized patient I gave higher ratings of clarification where comprehension lacking, fluency of speech and overall impression of language proficiency to her students than did standardized patient III (t(pooled)=2.52, df=38, p=0.016), (t(pooled)=2.93, df=37, p=0.006) and ((t(pooled)=4.31, df=36, p=0.000) respectively.

Standardized patient I and IV did not differ significantly on their ratings for any of the Language Rating Scale items.

Standardized patient I scored her students lower than standardized patients V and VI on use of appropriate tense (t=-2.12, df=29.23, p=0.042) and (t=-2.32, df=31.19, p=0.027) respectively and register (t=-2.4, df=41, p=0.019) and (t(pooled)=2.05, df=42, p=0.047). Standardized patient I scored her students higher than standardized patient V on comprehensibility due to accent (t(pooled)=2.86, df=41, p=0.007), speech acts (t=3.10, df=41, p=0.003) and overall impression of language proficiency (t(pooled)=2.37, df=39, p=0.023). She scored her students higher than did standardized patient VI on clarification where comprehension lacking (t=2.18, df=33.35, p=0.036).

Standardized patient II rated her students as using a significantly more appropriate rate of speech than standardized patient III did in the evaluation of her students (t=2.75, df=17.00, p=0.014). These two patients also differed in the same direction on appropriate

non-verbal communication (t=2.09, df=20.90, p=0.049), understanding of informal language (t(pooled)=2.24, df=37, p=0.031), fluency of speech (t(pooled)=3.50, df=37, p=0.001) and overall impression of language proficiency (t=4.07, df=34.47, p=0.000).

Standardized patient II awarded her students significantly higher scores on appropriate rate of speech than standardized patient IV (t=3.58, df=21.00, p=0.002). Otherwise these two patients did not differ in their ratings.

Standardized patient II's ratings of comprehensibility due to accent were significantly higher than standardized patient V's (t(pooled)=3.11, df=41, p=0.003), as were her evaluations of rate of speech (t=3.83, df=20.00, p=0.001), speech acts (t(pooled)=2.21, df=40, p=0.033), understanding of informal language (t(pooled)=2.32, df=39, p=0.026), use of informal language (t(pooled)=3.94, df=30, p=0.000) and overall impression of language proficiency (t(pooled)=2.15, df=41, p=0.037).

Standardized patient II also awarded higher scores than patient VI on register (t(pooled)=2.60, df=42, p=0.013), rate of speech (t=2.59, df=21.00, p=0.017) and use of informal language (t=3.90, df=24.44, p=0.001).

Standardized patients III and IV differed in their ratings of clarification where comprehension lacking and overall impression of language proficiency, with the former giving significantly lower scores (t=-2.32, df=26.55, p=0.028) and (t=-3.24, df=34.56, p=0.002).

Standardized patients III and V differed in their ratings of register (t=-3.27, df=27.03, p=0.003) and overall impression of language proficiency (t(pooled)=-2.30, df=37, p=0.027). Standardized patient III gave the lower ratings of these two items. However, she was more lenient in her ratings than patient V on students' use of informal language (t=4.31, df=13.97, p=0.001).

Standardized patient IV rated students as significantly poorer than standardized patient V on register (t=-2.76, df=32.54, p=0.009) and higher than patient V on comprehensibility due to accent (t=3.94, df=32.74, p=0.000).

Standardized patient IV rated her students as being better in their ability to clarify where comprehension lacking than did patient VI (t=2.08, df=31.26, p=0.046). There were no other significant differences between these two raters' scores.

Standardized patient V gave higher ratings than standardized patient VI on register (t=5.08, df=32.91, p=0.000), and lower ratings on comprehensibility due to accent (t(pooled)=-2.08, df=41, p=0.044) and speech acts (t=-3.42, df=23.16, p=0.002).

Using paired samples t-tests (two-tailed), a comparison of the mean ratings of the standardized patients' ratings with the English Language Specialist's ratings was conducted. The English Language Specialist gave significantly lower ratings than the standardized patients on the appropriate use of tense (t=-3.61, df=124, p=0.000), register (t=-4.80, df=122, p=0.000), comprehensibility due to accent (t=-6.72, df=124, p=0.000), rate of speech (t=-4.68, df=124, p=0.000), speech acts (t=-8.07, df=124, p<0.000), use of informal language (t=-2.86, df=85, p=0.005), audibility of speech (t=-6.66, df=125, p=0.000) and overall impression of language proficiency (t=-2.34, df=117, p=0.021). The English Language Specialist rated students higher on clarification where comprehension lacking than did the standardized patients (t=4.27, df=118, p=0.000). There were no significant differences between the two sets of raters on understanding of informal language and fluency of speech.

VALIDITY.

Content validity.

In order to determine the inter-item consistency of the Language Rating Scale all items were correlated with each other (Spearman *rho*, two-tailed, Bonferroni correction).

These data have been presented above, and can be inspected in order to determine the content validity of the scale.

Concurrent validity.

First, Spearman *rho* correlation coefficients (two-tailed) were calculated for the four subtests of the STAL and the total scores of the STAL (scored by the researcher) and the items on the Language Rating Scale for the standardized patients' collective ratings and those of the English Language Specialist.

Table LXV: Spearman *rho* correlation coefficients for standardized patient ratings on the Language Rating Scale and Screening Test of Adolescent Language (STAL) scores.

	Vocabulary	Auditory Memory	Language Processing	Proverbs	Total Score
tense	0.39	0.41	0.24	0.30	0.41
register	NS	NS	NS	NS	NS
accent	0.38	0.46	0.31	0.24	0.40
rate	NS	0.21	NS	NS	NS
nvc	NS	NS	NS	NS	NS
speech acts	NS	NS	NS	NS	NS
informal 1	0.37	0.45	0.32	0.23	0.40
informal 2	NS	0.26	NS	NS	NS
clarification	0.31	0.37	0.25	NS	0.31
audibility	NS	NS	NS	NS	NS
fluency	0.29	0.40	0.33	0.23	0.34
overall	0.24	0.37	0.33	0.24	0.32

All values significant at $p<0.05$.

informal 1= understanding of informal language
informal 2=use of informal language

Table LXVI: Spearman *rho* correlation coefficients for the English Language Specialist's ratings on the Language Rating Scale and Screening Test of Adolescent Language (STAL) scores.

	Vocabulary	Auditory Memory	Language Processing	Proverbs	Total Score
tense	0.39	0.35	0.35	0.28	0.44
register	NS	NS	NS	NS	NS
accent	0.32	0.31	NS	NS	0.28
rate	NS	NS	NS	NS	NS
nvc	-	-	-	-	-
speech acts	NS	NS	NS	NS	NS
informal 1	0.30	0.33	0.26	0.24	0.36
informal 2	0.29	0.40	0.26	0.18	0.32
clarification	0.21	0.28	0.32	0.23	0.24
audibility	NS	NS	NS	NS	NS
fluency	0.33	0.36	0.34	0.20	0.40
overall	0.30	0.35	0.29	0.23	0.32

Significant at $p<0.05$.

informal 1= understanding of informal language
informal 2=use of informal language

Concurrent validity of the Word Knowledge Test (WKT) with the Language Rating Scale could not be considered for this cohort of students, since they were not tested on the WKT.

Unfortunately, the data from the Medical Communication Skills examination were unavailable for this cohort of students, and therefore both concurrent validity with the language relevant items of that examination and the uniqueness of data cannot be reported.

Although an unstandardized instrument, the Language Development Committee's Oral Language Skills Assessment criteria were considered as possible indicators of the validity of the Language Rating Scale. The frequency distribution for the four criteria are listed in Table LXVII.

Table LXVII: Frequencies, Oral Language Skills Assessment.

	5 severe problem	4	3 adequate	2	1 no problem
1	-	2 (1.6)	7 (5.5)	33 (26.0)	85 (66.9)
2	-	-	6 (4.7)	35 (27.6)	86 (67.7)
3	-	2 (1.6)	3 (2.5)	29 (23.8)	88 (72.1)
4	-	1 (0.8)	8 (6.3)	31 (24.4)	87 (68.5)

percentages are in brackets

Key: **1**=ability to express oneself without being misunderstood
2=ability to retrieve information if patient misunderstands
3=ability to clarify when unable to understand the patient
4=ability to speak intelligibly

Two students of the 127 in the cohort were identified as "unsatisfactory" by the English Language Specialists with regard to overall evaluation of spoken language on the Oral Language Skills Assessment instrument. A further eleven students were identified as in need of formative feedback about their spoken English language skills, but were considered to be satisfactory for the purpose of this examination nonetheless.

To ascertain concurrent validity between the Language Development Committee's Oral Language Skills Assessment and the Language Rating Scale, Spearman correlation coefficients (two-tailed and with Bonferroni correction) were calculated employing both sets of raters scores. (Note that for analyses involving the Oral Language Skills Assessment, scores were reversed to be consistent with those of the Language Rating Scale).

Table LXVIII: Spearman *rho* correlation coefficients for standardized patient ratings on the Language Rating Scale and the Oral Language Skills Assessment.

	1	2	3	4	5
tense	0.34	NS	0.35	0.42	NS
register	NS	NS	NS	NS	NS
accent	0.47	NS	0.32	0.61	NS
rate	NS	NS	NS	0.29	NS
nvc	NS	NS	NS	NS	NS
speech acts	0.29	NS	NS	0.33	NS
informal 1	0.41	NS	NS	0.49	NS
informal 2	NS	NS	NS	NS	NS
clarification	0.33	NS	0.36	0.36	NS
audibility	0.31	NS	NS	NS	NS
fluency	0.31	NS	0.30	0.41	NS
overall	0.43	NS	0.39	0.45	NS

All values significant at $p \leq 0.001$.

informal 1= understanding of informal language
informal 2=use of informal language

Key: **1**=ability to express oneself without being misunderstood
2=ability to retrieve information if patient misunderstands
3=ability to clarify when unable to understand the patient
4=ability to speak intelligibly
5=overall assessment of satisfactory/unsatisfactory oral language.

Table LXVIX: Spearman *rho* correlation coefficients for the English Language Specialist's ratings on the Language Rating Scale and the Oral Language Skills Assessment.

	1	2	3	4	5
tense	-0.34	NS	-0.41	-0.46	NS
register	-0.29	NS	-0.36	-0.37	NS
accent	-0.29	NS	NS	-0.36	NS
rate	NS	NS	NS	NS	NS
nvc	-	-	-	-	-
speech acts	NS	NS	NS	NS	NS
informal 1	-0.36	NS	-0.38	-0.38	NS
informal 2	-0.31	NS	-0.42	-0.41	NS
clarification	NS	NS	-0.44	-0.38	NS
audibility	NS	NS	NS	NS	NS
fluency	-0.40	NS	-0.34	-0.43	NS
overall	-0.45	NS	-0.37	0.47	NS

All values significant at $p \leq 0.001$.

informal 1= understanding of informal language
informal 2=use of informal language

Key: **1**=ability to express oneself without being misunderstood
2=ability to retrieve information if patient misunderstands
3=ability to clarify when unable to understand the patient
4=ability to speak intelligibly
5=overall assessment of satisfactory/unsatisfactory oral language.

To further explore the validity of the Oral Language Skills Assessment the instrument was correlated with the Screening Test of Adolescent Language (STAL) scores. There were no significant correlations between the two instruments on ability to express oneself without being misunderstood, ability to retrieve information if patient misunderstands, ability to clarify when unable to understand the patient or overall assessment of satisfactory/unsatisfactory oral language. However, ability to speak intelligibly was correlated with three of the four subtests and the total score of the STAL (Spearman *rho*, with Bonferroni correction) with coefficients ranging from 0.28 to 0.33, $p \leq 0.002$. Proverbs was the fourth STAL subtest, with which this item did not correlate.

Construct validity.

To first investigate the construct validity of the Language Rating Scale the alpha correlation coefficients were examined comparing the two forms of the scale after employing the split-half method as described above (see internal consistency). Only the standardized patient ratings were employed in analyses, as the English Language Specialist's ratings included too many missing data (partially as a result of the inability to rate non-verbal communication). Where missing data were excluded from the standardized patients' ratings, 80 cases yielded alphas of 0.49 and 0.84 for the two forms. When missing data were included so that all cases could be included in analyses, the alphas were 0.97 and 0.90.

T-tests (one-tailed) were conducted to ascertain differences in the mean ratings of English speaking background versus non-English speaking background students. Standardized patient ratings were consistently higher for the former group of students on use of correct tense ($t=4.88$, $df=52.16$, $p=0.000$), comprehensibility due to accent ($t=4.08$, $df=74.24$, $p=0.000$), understanding of informal language ($t=6.19$, $df=58.02$, $p=0.000$), use of informal language ($t=2.63$, $df=37.41$, $p=0.012$), clarification where comprehension lacking ($t(pooled)=5.00$, $df=123$, $p=0.000$), fluency of speech ($t=3.56$, $df=78.99$, $p=0.001$) and overall impression of language proficiency ($t(pooled)=5.23$, $df=123$, $p=0.000$). There were no significant differences between English speaking background and non-English

speaking background students on use of appropriate register, rate of speech, appropriate non-verbal communication, speech acts or audibility of speech.

When the English Language Specialist's ratings were employed to test for differences between English speaking and non-English speaking background students, the former group were rated more highly on use of correct tense (t=6.72, df=65.05, p=0.000), register (t(pooled)=2.65, df=121, p=0.009), comprehensibility due to accent (t(pooled)=5.46, df=123, p=0.000), speech acts (t(pooled)=2.30, df=124, p=0.023), understanding of informal language (t=6.03, df=59.12, p=0.000), use of informal language (t=6.68, df=62.56, p=0.000), clarification where comprehension lacking (t=3.76, df=63.90, p=0.000), audibility (t(pooled)=2.26, df=124, p=0.026), fluency of speech (t(pooled)=6.08, df=123, p=0.000) and overall impression of language proficiency (t(pooled)=6.21, df=118, p=0.000). There were no significant differences in mean ratings between the two groups on rate of speech.

Test security.

No significant correlations were found between Language Rating Scale item scores and the time of day at which students' interviews were conducted, apart from "audibility" (r=-0.26, p=0.003). Students spoke significantly more audibly in the afternoon as compared to the interviews conducted in the morning (t=-2.30, df=76.35, p=0.024).

Rater bias.

As in the previous study, reported in Chapter VII, it was not possible to test for bias according to rater's gender, as in this study, all standardized patients and the English Language Specialist were female. However, it was possible to ascertain whether students' gender was related to scores on the Language Rating Scale.

T-tests were conducted to compare the mean ratings made by the standardized patients according to student gender. Males were rated as significantly higher than females on use of correct tense (t=2.26, df=92.89, p=0.26), whilst females were judged to have

more appropriate rates of speech (t=-2.09, df=123.53, p=0.039). The English Language Specialist rated males as significantly better than females in their use of informal language (t(pooled)=2.35, df=122, p=0.021).

In order to determine whether there might be "practice effects" related to whether or not an individual had previously acted as a standardized patient, the mean ratings of the three individuals who had not been involved in the previous study, reported in Chapter X were compared with the three who had. T-tests indicated that the latter group rated students as significantly higher on comprehensibility due to accent (t=-2.08, df=120.12, p=0.039) and use of informal language (t=-4.48, df=57.76, p=0.000). There were no other differences in the ratings between the two groups of standardized patients.

To further explore possible effects due to experience, the amount of missing data generated by each standardized patient was examined (Appendix XXXIII). Generally speaking, missing data from the standardized patients in this study were negligible. However, the two who contributed most to missing data (raters 4 and 5) were an experienced patient and a novice respectively.

Validity of presentations.

The six English Language Specialists who used the Language Development Committee's Oral Language Skills Assessment criteria made the following comments, as reported by McGowan (1995b);

> "While the exam assessment team members were impressed by the high standard of language for communication in most of the students, they also expressed some reservations about the constraints under which the language examination was performed and the limitations these imposed on their ability to judge the depth of students' understanding and their ability to communicate in English." (p 2).

Written Language Rating Scale

Inspection of the ratings made by the English Language Specialist for each of the eleven items on the Written Language Rating Scale shows negatively skewed distributions

for both cohorts of students (the 1994 third years and the 1994 first years) (Tables LXX and LXXI).

Table LXX: Frequencies, Written Language Rating Scale items scored by the English Language Specialist, 1994 third year cohort.

n=147	1	2	3	4	5	missing	$\bar{x}$	sd
content	3 (2.0)	18 (12.1)	48 (32.2)	57 (38.3)	21 (14.1)	2 (1.3)	3.51	0.95
jargon	1 (0.7)	8 (5.4)	54 (36.2)	64 (43.0)	21 (14.1)	1 (0.7)	3.65	0.82
values	1 (0.7)	5 (3.4)	47 (31.5)	69 (46.3)	27 (18.1)	-	3.78	0.80
vocabulary	1 (0.7)	13 (8.7)	46 (30.9)	52 (34.9)	37 (24.8)	-	3.75	0.95
tense	9 (6.0)	13 (8.7)	16 (10.7)	24 (16.1)	87 (58.4)	-	4.12	1.26
articles	3 (2.0)	14 (9.4)	39 (26.2)	30 (20.1)	62 (41.6)	1 (0.7)	3.91	1.11
spelling	1 (0.7)	8 (5.4)	25 (16.8)	38 (25.5)	77 (51.7)	-	4.22	0.96
legibility	1 (0.7)	3 (2.0)	13 (8.7)	81 (54.4)	50 (33.6)	1 (0.7)	4.19	0.73
conventions	2 (1.3)	5 (3.4)	26 (17.4)	85 (57.0)	31 (20.8)	-	3.93	0.79
fluency	2 (1.3)	11 (7.4)	36 (24.2)	55 (36.9)	45 (30.2)	-	3.87	0.98
overall	2 (1.3)	13 (8.7)	37 (24.8)	59 (39.6)	37 (24.8)	1 (0.7)	3.78	0.97

percentages are in brackets.

Key: 1=poor
2=between poor and adequate
3=adequate
4=between adequate and excellent (good)
5=excellent

Table LXXI: Frequencies, Written Language Rating Scale items scored by the English Language Specialist, 1994 first year cohort.

n=127	1	2	3	4	5	missing	$\bar{x}$	sd
content	-	7 (5.5)	26 (20.5)	59 (46.5)	35 (27.6)	-	3.96	0.84
jargon	-	4 (3.1)	31 (24.4)	70 (55.1)	22 (17.3)	-	3.87	0.73
values	-	7 (5.5)	22 (17.3)	76 (59.8)	22 (17.3)	-	3.89	0.75
vocabulary	-	4 (3.1)	32 (25.2)	52 (36.4)	39 (27.3)	-	3.99	0.83
tense	2 (1.6)	7 (5.5)	17 (13.4)	24 (18.9)	77 (60.6)	-	4.31	1.01
articles	1 (0.8)	4 (3.1)	20 (15.7)	32 (25.2)	70 (55.1)	-	4.31	0.90
spelling	-	1 (0.8)	9 (7.1)	41 (32.3)	76 (59.8)	-	4.51	0.67
legibility	1 (0.8)	-	9 (7.1)	74 (58.3)	43 (33.9)	-	4.24	0.65
conventions	-	3 (2.4)	32 (25.2)	88 (69.3)	4 (3.1)	-	3.73	0.56
fluency	-	6 (4.7)	30 (23.6)	59 (46.5)	32 (25.2)	-	3.92	0.82
overall	-	6 (4.7)	29 (22.8)	64 (50.4)	28 (22.0)	-	3.90	0.80

percentages are in brackets.

Key: 1=poor
2=between poor and adequate
3=adequate
4=between adequate and excellent (good)
5=excellent

Comments.

The English Language Specialist again made most comprehensive comments pertaining to each of the items on the Written Language Rating Scale, which are transcribed here.

1. **Appropriate content**

This rater sought "exact" recounts. Judgements were more realistically made about the "substantive" content of the written account. Incorrect details were penalised, as were obvious omissions, although where time ran out for the written account, a judgement was made about the proportion of detail not recorded - whether the quality of detail given to that point was likely to continue. Judgements also took account of whether the relative lengths of time spent discussing a certain issue were reflected in the proportion given to that sequence in the written account.

2. **Use of appropriate register**

This rater considered that few students failed to write as if it were a medical report - though many lost points for imprecision, incoherent sequencing, failure to categorize ideas or colloquial writing. The rater felt that the definition of "jargon" given by the researcher was inconclusive e.g. are "abdomen", "stomach", "below the ribs" or "tummy" acceptable? All of these were in fact accepted by the rater (except "tummy" which was considered childish and too colloquial).

3. **Value judgements**

She noticed that a large proportion of third year students (compared to the second year students) failed to "intuit" that the standardized patients' reference to the friend who had a burst ulcer was to indicate that "prevention was better than cure". Many indicated in writing that they thought the patient (being female) was neurotic or easily swayed, hypochondriacal or jealous. These "value judgements" prompted the rater to judge them more severely.

Instructions given by the researcher here were deemed clear. So, patronising comments, or gratuitous remarks like "a cheerful young lady" or " a middle aged woman of Caucasian descent" were deemed to be "value-laden". Also, over-use of the patient's name was marked down.

4. **Vocabulary**

The rater thought instructions were reasonably clear, and gave scope to penalise inappropriate use of words whether singly or in phrases/clauses, or in poor constructions. However, this rater perceived considerable overlap between this criterion and the intention of subsequent ones e.g. articles, pronouns, prepositions (see below - item 6).

5. **Appropriate use of tense**

This was easily detected and judged strictly.

6. **Appropriate use of articles, pronouns, prepositions**

Literal judgements were made about errors in these parts of speech, but these alone did not account adequately for the quantity of written errors. For example, errors in adjectival and adverbial constructions, relative and subordinate clause structures etc. were judged "globally" under item 4 (vocabulary).

7. **Appropriate use of spelling, punctuation and capitals**

Judgements were relatively straightforward for the rater.

8. **Legibility of handwriting**

Subjective judgements only were possible here. This rater feared that the way doctors' writing is caricatured as uniformly illegible may have prompted this criterion. There was a wide variety of the writing styles judged as a "4" or a "5". Reports with extensive cross-outs were penalised, as were those where the writing was very small (and the photocopy was faint - indicating that the original was less than bold).

9. **Appropriate use of conventions**

The rater was concerned that this criterion was rated almost uniformly at "4" - not because students failed in the example given (forms of address - which she included under the construct "register") - but because these were not authentic medical reports. No matter how approximate they may have been to actual reports, she felt that she could not vouch for their total conformity with the requirements of Western cultural medical practice.

10. **Fluency of written expression**

There were only a few outstanding examples where written accounts were highly competent, succinct, accurate, cohesive and stylish. This rater felt that it was easier to grade spoken fluency than written fluency.

11. **Overall impression of written proficiency**

Some native speakers scored "5" but not all, just as some non-native speakers/writers scored the maximum. The disparity between the two macro-skills is not unexpected. A noticeable case in point is student "X". When this rater heard his tape, she recalled that he had been reticent at interview (which she attributed to an over-talkative standardized patient and his "quietish" personality). However, his written report was clear, succinct, accurate and highly proficient.

RELIABILITY OF THE WRITTEN LANGUAGE RATING SCALE.

Split-half reliability (internal consistency).

The Written Language Rating Scale was divided into two forms (the first six items versus the last five). This resulted in a correlation between the forms of 0.86 and a Guttman split-half reliability correlation coefficient of 0.91 where the data from the 1994 third year cohort were used in the analyses. A correlation of 0.80 and a Guttman split-half reliability correlation coefficient of 0.85 was found using the 1994 first year cohort data. Because the amount of missing data for the Written Language Rating Scale was negligible, they were not included in these analyses.

Cronbach's alpha on this scale was 0.92 and 0.90 for the 1994 third year and 1994 first year cohorts respectively.

Inter-item consistency.

To determine the inter-item consistency of the Written Language Rating Scale all items were correlated with each other (Spearman *rho*, two-tailed test with Bonferroni correction).

Table LXXII: Spearman *rho* correlation coefficients for Written Language Rating Scale item scores, 1994 third year cohort.

jargon	0.65									
values	0.53	0.62								
vocab	0.71	0.66	0.63							
tense	0.47	0.43	0.39	0.64						
articles	0.43	0.42	0.34	0.58	0.63					
spelling	0.35	0.36	NS	0.55	0.64	0.62				
legibility	NS	0.35	0.38	0.40	0.44	0.30	0.44			
convent	0.65	0.60	0.64	0.70	0.56	0.53	0.48	0.40		
fluency	0.61	0.60	0.58	0.76	0.73	0.63	0.56	0.49	0.71	
overall	0.67	0.62	0.61	0.79	0.73	0.66	0.60	0.49	0.72	0.90
	content	**jargon**	**values**	**vocab**	**tense**	**articles**	**spelling**	**legibility**	**convent**	**fluency**

All values significant at p=0.000.

Table LXXIII: Spearman *rho* correlation coefficients for Written Language Rating Scale item scores, 1994 first year cohort.

jargon	0.59									
values	0.56	0.59								
vocab	0.61	0.62	0.47							
tense	0.32	0.43	NS	0.56						
articles	0.37	0.48	NS	0.56	0.75					
spelling	0.31	NS	NS	0.40	0.46	0.40				
legibility	NS	NS	NS	NS	0.31	NS	0.33			
convent	0.46	0.63	0.50	0.57	0.50	0.47	0.32	NS		
fluency	0.49	0.59	0.34	0.61	0.64	0.68	0.41	0.37	0.57	
overall	0.46	0.54	0.37	0.64	0.73	0.73	0.48	0.37	0.53	0.82
	content	**jargon**	**values**	**vocab**	**tense**	**articles**	**spelling**	**legibility**	**convent**	**fluency**

All values significant at p=0.000.

VALIDITY.

Content validity.

To determine the inter-item consistency of the Written Language Rating Scale all items were correlated with each other. The data presented above can also be used to determine the content validity of the scale.

Concurrent validity.

To establish the relationship of the Language Rating Scale items to the scores on the Screening Test of Adolescent Language (STAL), Spearman *rho* correlation coefficients (two-tailed, Bonferroni correction) for the four subtests and total score of the STAL were calculated.

Table LXXIV: Spearman *rho* correlation coefficients for Written Language Rating Scale items and STAL scores, 1994 third year cohort.

	Vocabulary	Auditory Memory	Language Processing	Proverbs	Total Score
content	0.43	0.38	NS	NS	0.42
jargon	0.30	NS	NS	NS	NS
values	NS	NS	NS	NS	NS
vocabulary	0.46	0.42	NS	0.29	0.45
tense	0.52	0.54	NS	0.30	0.52
articles	0.41	0.43	NS	NS	0.39
spelling	0.44	0.37	NS	NS	0.40
legibility	0.30	NS	NS	0.34	0.31
conventions	0.41	0.35	NS	NS	0.37
fluency	0.50	0.49	NS	0.32	0.49
overall	0.53	0.43	NS	NS	0.48

All values significant at p=0.000.

Table LXXV: Spearman *rho* correlation coefficients for Written Language Rating Scale items and STAL scores, 1994 third year cohort.

	Vocabulary	Auditory Memory	Language Processing	Proverbs	Total Score
content	NS	NS	0.42	NS	0.40
jargon	NS	NS	NS	NS	NS
values	NS	NS	NS	NS	NS
vocabulary	0.34	NS	0.33	NS	0.41
tense	0.42	0.48	0.40	NS	0.48
articles	0.45	0.40	0.38	NS	0.50
spelling	NS	0.31	NS	NS	0.36
legibility	NS	NS	NS	NS	NS
conventions	NS	NS	NS	NS	NS
fluency	NS	0.36	0.37	NS	0.42
overall	0.47	0.49	0.40	NS	0.54

All values significant at p=0.000.

Word Knowledge Test (WKT) scores (see Chapter IV) were correlated (Spearman *rho*, two-tailed) with items on the Written Language Rating Scale to explore the concurrent validity of these two instruments (Table LXXVI).

Table LXXVI: Spearman *rho* correlation coefficients, Written Language Rating Scale and Word Knowledge Test (WKT) scores.

	WKT
content	0.40
jargon	0.40
values	0.37
vocabulary	0.62
tense	0.60
articles	0.43
spelling	0.47
legibility	0.35
conventions	0.42
fluency	0.57
overall	0.57

All values significant at p=0.000.

Spearman correlation coefficients (two-tailed with Bonferroni correction) presented in Tables LXXVII, LXXVIII, LXXIX and LXXX show the concurrent validity of the Written Language Rating Scale with the Language Rating Scale for both cohorts of

students, using the standardized patients' and English Language specialists' ratings on the latter instrument.

Table LXXVII: Spearman *rho* correlation coefficients for the Written Language Rating Scale scored by the English Language Specialist and Language Rating Scale scored by standardized patients, 1994 third year student cohort.

content	0.32	NS	0.39	0.30	NS	0.34	0.35	NS	0.38	0.35
jargon	NS	NS	NS	NS	NS	0.31	NS	NS	NS	NS
values	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
vocabulary	0.43	NS	0.50	0.34	NS	0.37	0.46	NS	0.38	0.40
tense	0.43	NS	0.50	0.35	NS	0.34	0.51	0.43	0.42	0.44
articles	0.31	NS	0.41	NS	NS	0.32	0.35	0.36	0.29	0.32
spelling	0.29	NS	0.34	NS	NS	NS	0.32	NS	NS	0.30
legibility	NS	NS	NS	NS	NS	NS	NS	NS	NS	0.30
convent.	0.31	NS	0.35	NS	NS	0.31	0.37	NS	0.35	0.30
fluency	0.44	NS	0.46	0.36	NS	0.39	0.47	NS	0.43	0.41
overall	0.41	NS	0.50	0.36	NS	0.39	0.52	NS	0.45	0.46
WRS/LRS	**tense**	**register**	**accent**	**rate**	**nvc**	**speech acts**	**informal**	**clarify**	**fluency**	**overall**

All values significant at p=0.000.

Table LXXVIII: Spearman *rho* correlation coefficients for the Written Language Rating Scale and Language Rating Scale scored by the English Language Specialist, 1994 third year student cohort.

content	0.44	0.42	0.43	0.30	-	NS	0.42	0.49	0.36	0.47
jargon	NS	0.36	0.30	0.29	-	NS	0.29	0.38	0.30	0.41
values	NS	0.39	NS	NS	-	0.40	0.32	0.45	0.30	0.33
vocabulary	0.49	0.49	0.42	0.30	-	0.37	0.54	0.46	0.49	0.57
tense	0.47	0.42	0.48	0.33	-	NS	0.54	0.46	0.48	0.59
articles	0.38	NS	0.34	0.28	-	NS	0.31	0.33	0.33	0.43
spelling	0.32	NS	0.38	NS	-	NS	0.36	NS	0.29	0.39
legibility	NS	0.36	0.31	NS	-	NS	0.31	NS	0.29	0.30
convent.	0.43	0.52	0.35	0.34	-	0.42	0.44	0.45	0.38	0.47
fluency	0.50	0.45	0.47	0.32	-	0.39	0.53	0.51	0.49	0.58
overall	0.55	0.51	0.53	0.36	-	0.48	0.58	0.57	0.53	0.65
WRS/LRS	**tense**	**register**	**accent**	**rate**	**nvc**	**speech acts**	**informal**	**clarify**	**fluency**	**overall**

All values significant at p=0.000.

Table LXXIX: Spearman *rho* correlation coefficients for the Written Language Rating Scale scored by the English Language Specialist and Language Rating Scale scored by the standardized patients, 1994 first year student cohort.

content	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
jargon	NS	NS	NS	NS	NS	NS	0.35	NS	0.39	NS	0.33	0.37
values	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
vocabulary	NS	NS	0.34	NS	NS	NS	0.43	NS	0.49	NS	0.40	0.44
tense	0.52	NS	0.47	NS	NS	NS	0.54	NS	0.50	NS	0.47	0.41
articles	0.43	NS	0.50	NS	NS	NS	0.56	NS	0.51	NS	0.45	0.43
spelling	NS	NS	NS	NS	NS	NS	0.33	NS	NS	NS	0.34	NS
legibility	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
convent.	NS	NS	NS	NS	NS	NS	0.41	NS	NS	NS	0.34	0.35
fluency	0.37	NS	0.44	NS	NS	NS	0.41	NS	0.38	NS	0.47	0.38
overall	0.53	NS	0.55	NS	NS	NS	0.52	NS	0.49	NS	0.52	0.46
WRS/LRS	**tense**	**register**	**accent**	**rate**	**nvc**	**speech acts**	**informal 1**	**informal 2**	**clarify**	**audibility**	**fluency**	**overall**

All values significant at p=0.000.

Key: informal 1=understanding of informal language
informal 2=use of informal language

Table LXXX: Spearman *rho* correlation coefficients for the Written Language Rating Scale and Language Rating Scale scored by the English Language Specialist, 1994 third year student cohort.

content	0.31	NS	NS	NS	-	NS	NS	NS	0.36	NS	NS	NS
jargon	0.34	NS	0.31	NS	-	0.31	0.33	0.35	0.32	NS	NS	NS
values	NS	NS	NS	NS	-	NS	NS	NS	NS	NS	NS	NS
vocabulary	0.50	NS	0.39	NS	-	NS	0.42	0.41	0.34	NS	0.40	0.34
tense	0.64	NS	0.48	NS	-	NS	0.65	0.62	0.39	NS	0.54	0.46
articles	0.55	NS	0.39	NS	-	NS	0.58	0.51	0.33	NS	0.49	0.41
spelling	0.36	NS	NS	NS	-	NS	0.38	0.32	NS	NS	NS	NS
legibility	NS	NS	NS	NS	-	NS	NS	NS	NS	NS	NS	NS
convent.	0.38	NS	0.32	NS	-	0.33	0.33	0.33	0.32	NS	0.35	0.35
fluency	0.60	NS	0.42	NS	-	NS	0.51	0.48	NS	NS	0.44	0.39
overall	0.63	NS	0.48	NS	-	NS	0.58	0.51	0.32	NS	0.46	0.43
WRS/LRS	**tense**	**register**	**accent**	**rate**	**nvc**	**speech acts**	**informal 1**	**informal 2**	**clarify**	**audibility**	**fluency**	**overall**

All values significant at p=0.000.

Key informal 1=understanding of informal language
informal 2=use of informal language

Although an unstandardized instrument, the Language Development Committee's Written Language Skills Assessment criteria were considered as possible indicators of the validity of the Written Language Rating Scale. The frequency distribution for the three criteria are listed in Table LXXXI.

Table LXXXI: Frequencies, Written Language Skills Assessment.

	5 severe problem	4	3 adequate	2	1 no problem
1	-	-	1 (0.8)	1 (0.8)	125 (98.4)
2	-	-	6 (4.7)	30 (23.6)	91 (71.1)
3	-	-	1 (0.8)	5 (3.9)	121 (95.3)

percentages are in brackets

Key: **1**=overall communication of information
2=accuracy in the choice of words and grammatical constructions
3=presentation and editing

All of the 127 in the cohort were identified as "satisfactory" by the English Language Specialists with regard to overall evaluation of written language on the Written Language Skills Assessment instrument. However, twelve students were identified as being in need of formative feedback about their written English language skills.

To ascertain concurrent validity between the Language Development Committee's Written Language Skills Assessment and the Written Language Rating Scale, Spearman correlation coefficients (two-tailed and with Bonferroni correction) were calculated. Prior to conducting analyses, scores on the Written Language Skills Assessment were reversed to be consistent with the scoring of the Written Language Rating Scale.

Table LXXXII: Spearman *rho* correlation coefficients for the Written Language Skills Assessment and Written Language Rating Scale.

	1	2	3
content	NS	NS	NS
jargon	NS	0.31	NS
values	NS	NS	NS
vocabulary	NS	0.47	NS
tense	NS	0.75	NS
articles	NS	0.64	NS
spelling	NS	0.33	NS
legibility	NS	NS	NS
convent.	NS	0.34	NS
fluency	NS	0.47	NS
overall	NS	0.62	NS

Significant at p=0.000.

Key: **1**=overall communication of information
2=accuracy in the choice of words and grammatical constructions
3=presentation and editing

To further explore the validity of the Written Language Skills Assessment the instrument was correlated with the Screening Test of Adolescent Language (STAL) scores. There were no significant correlations between the two instruments on presentation and editing, whilst correlation coefficients for overall written language could not be computed, since all of the cohort were deemed "satisfactory". Accuracy in the choice of words and grammatical constructions correlated significantly with all four subtests and the total score of the STAL (with coefficients ranging from 0.20 to 0.45, $p<0.05$). Overall communication of information was significantly correlated with the Vocabulary and Auditory Memory subtests and the total of the STAL (with correlation coefficients of 0.19, 0.32 and 0.20, $p<0.05$, respectively).

Construct validity.

To first consider the construct validity of the Written Language Rating Scale the internal consistency Cronbach alphas were examined comparing the two forms of the scale after applying the split-half method as described for internal consistency computations. Using the 1994 third year student cohort data the alphas were 0.86 for both halves of the

scale. Where the 1994 first year student cohort data was utilized the alphas were 0.85 and 0.80 for the first and second halves respectively.

Second, t-tests were performed comparing the performance of English speaking background students with their non-English speaking background counterparts.

With regards the 1994 third year cohort, English speaking students were rated as significantly higher than non-English speaking background students on appropriate content (t(pooled)=4.45, df=145, p=0.000), appropriate register (jargon) (t(pooled)=4.37, df=146, p=0.000), appropriate register (value-judgements) (t(pooled)=4.03, df=147, p=0.000), vocabulary (t(pooled)=6.94, df=147, p=0.000), tense (t=7.79, df=114.79, p=0.000), appropriate use of articles, prepositions and pronouns (t(pooled)=4.78, df=146, p=0.000), appropriate use of spelling, punctuation and capitals (t=3.91, df=145.85, p=0.000), legibility of handwriting (t(pooled)=3.68, df=146, p=0.000), appropriate use of conventions (t=5.92, df=146.21, p=0.000), fluency of written expression (t=7.82, df=146.99, p=0.000) and overall impression of written language proficiency (t=7.98, df=145.78, p=0.000).

For the 1994 first year cohort, English speaking students were rated as significantly higher than non-English speaking background students on appropriate content (t(pooled)=3.71, df=125, p=0.000), appropriate register (jargon) (t(pooled)=3.63, df=125, p=0.000), vocabulary (t=5.50, df=78.13, p=0.000), tense (t=6.56, df=67.30, p=0.000), appropriate use of articles, prepositions and pronouns (t=6.27, df=67.39 p=0.000), appropriate use of conventions (t(pooled)=3.99, df=125, p=0.000), fluency of written expression (t(pooled)=6.18, df=125, p=0.000) and overall impression of written language proficiency (t(pooled)=7.55, df=125, p=0.000). There were no significant differences between the two groups on the use of value-judgements, appropriate spelling, punctuation and capitals or legibility of handwriting.

Test security.

To ascertain whether the time of day at which the student was interviewed and subsequently wrote the account of that interview were related to item scores, Pearson correlation coefficients were calculated between time of the examination and all of the items on the Written Language Rating Scale for both cohorts of students. The time at which the examination was taken was not relevant to any of the Written Language Rating Scale scores.

Rater bias.

To consider whether males and females were scored differently on the Written Language Rating Scale, t-tests were conducted. Females in the 1994 third year student cohort were rated as having significantly more legible handwriting than males (t(pooled)=-3.54, df=146, p=0.001). Otherwise, there were no significant differences according to gender on any of the items.

More differences between males and females were found in the 1994 first year cohort. Males were rated as having significantly better use of appropriate tense (t=2.77, df=103.78, p=0.007), appropriate use of articles, pronouns and prepositions (t(pooled)=2.01, df=125, p=0.46) and a higher level of overall written language proficiency (t(pooled)=2.50, df=125, p=0.014). In order to investigate these differences a comparison of the number of males and females from non-English speaking background in the cohort was made. Of the 54 students who fell into this category 26 (48.1%) were male and 28 (51.9%) were female, thus making language background less likely as a possible explanation for the results.

Comparison of the two study cohorts.

Screening Test of Adolescent Language (STAL) Scores.

A comparison of the means on the STAL (scored by the researcher) for the 1994 first year students and the 1994 third year group showed no significant differences on any of the four subtests or total score.

Language Rating Scale scores.

T-tests were conducted comparing the mean scores of the 1994 first year cohort of students with the 1994 third year students on the Language Rating Scale. Since "use of informal language" and "audibility" had not been items on the pilot version of the Language Rating Scale, these could not be considered in analyses.

Where the standardized patients' scores were used, students from the 1994 first year cohort were rated as significantly higher on appropriate use of tense ($t=3.41$, $df=267.38$, $p=0.001$), register ($t=3.49$, $df=266.35$, $p=0.001$), comprehensibility due to accent ($t=3.84$, $df=238.44$, $p=0.000$), appropriate rate of speech ($t=4.49$, $df=257.21$, $p=0.000$), non-verbal communication ($t=6.93$, $df=258.87$, $p=0.000$), speech acts ($t=4.66$, $df=263.99$, $p=0.000$), understanding of informal language ($t=4.55$, $df=260.46$, $p=0.000$) and clarification where comprehension lacking ($t=2.18$, $df=191.60$, $p=0.030$). Interestingly, although all of the items on the Language Rating Scale thus far demonstrated significant differences in favour of the 1994 first year cohort, the final two items, fluency and overall impression of language proficiency failed to show such discrepancies.

Analyses utilizing the English Language Specialist's ratings showed a significant difference on clarification where comprehension lacking ($t=5.70$, $df=217.96$, $p=0.000$) in favour of the 1994 first year cohort. No other differences between the two cohorts on the Language Rating Scale were found for this rater.

To further consider differences across the two cohorts, t-tests were again conducted, this time employing the data pertaining to students from non-English speaking backgrounds only. Findings were generally consistent with those reported for the cohorts as a whole. Where standardized patients' ratings were used in analyses, students from the 1994 first year cohort were rated as better on register ($t=2.60$, $df=199.91$, $p=0.010$), comprehensibility due to accent ($t=2.91$, $df=120.98$, $p=0.004$), rate ($t=4.35$, $df=122.93$, $p=0.000$), non-verbal communication (t (pooled)$=5.16$, $df=130$, $p=0.000$), speech acts (t(pooled)$=4.24$, $df=127$, $p=0.000$) and understanding of informal language

(t(pooled)=2.43, df=129, p=0.017). No significant differences were found on appropriate use of tense, clarification where comprehension lacking, fluency, or overall impression of language proficiency. The English Language Specialist saw the 1994 third year students from non-English speaking background as significantly more appropriate in their use of tense (t(pooled)=-2.46, df=130, p=0.015), and their 1994 first year counterparts as having more appropriate rates of speech (t(pooled)=2.63, df=132, p=0.010) and better use of clarification where comprehension lacking (t(pooled)=4.05, df=118, p=0.000). No other differences between the non-English speaking background students of each cohort emerged for this rater's scores.

Written Language Rating Scale Scores.

T-tests were conducted comparing the mean scores of the 1994 first year cohort of students with the 1994 third year students on the Language Rating Scale.

The 1994 first year student cohort were rated as significantly higher than the 1994 third year student cohort on appropriate content (t(pooled)=4.12, df=272, p=0.000), appropriate register (jargon) (t(pooled)=2.32, df=273, p=0.021), vocabulary (t(pooled)=2.28, df=274, p=0.024), appropriate use of articles, pronouns and prepositions (t=3.30, df=272.17, p=0.001) and appropriate spelling, punctuation and capitals (t=2.96, df=263.72, p=0.003).

The 1994 third year cohort were rated as having significantly more appropriate conventions in their writing than the 1994 first year students (t=-2.37, df=263.97, p=0.019).

The two cohorts did not differ in their use of value-judgements, use of appropriate tense, legibility of handwriting, fluency of written expression or overall written language proficiency.

Role of the Language Development Programme in Language Rating Scale and Written Rating Scale Scores.

In order to determine the role of the Language Development Programme in Language Rating Scale and Written Rating Scale Scores separate one-way analyses of covariance with STAL total score acting as the covariate were performed for items where an improvement was found in ratings of the 1994 first year cohort (in comparison to the 1994 third years). The analyses tested whether higher scores for this group could be attributed to participation in the Language Development Programme or whether English language proficiency as measured by the STAL was central to ratings.

The results showed that the higher ratings of the 1994 first year cohort on the Language Rating Scale items of tense ($F=66.15$, (1, 117), $p<0.001$), accent ($F=45.48$, (1, 117), $p<0.001$), appropriate rate of speech ($F=5.88$, (1, 117), $p<0.05$), speech acts ($F=4.45$, (1, 116), $p<0.05$), understanding of informal language ($F=43.61$, (1, 111), $p<0.001$) and clarification where comprehension lacking ($F=19.24$, (1, 116), $p<0.001$) were all attributable to students' STAL scores. Therefore, for these items, improvement between the two student cohorts cannot be attributed exclusively to participation in the Language Development Programme. However, the STAL did not have a significant covariate effect on the items of use of appropriate register or non-verbal communication and thus improvements on these items may have been due to participation in the Programme.

The higher ratings of the 1994 first year cohort on the Written Language Rating Scale items of appropriate content ($F=21.11$, (1, 117), $p<0.001$), appropriate register (jargon) ($F=11.92$, (1, 117), $p=0.001$), vocabulary ($F=30.27$, (1, 117), $p<0.001$), appropriate use of articles, pronouns and prepositions ($F=66.11$, (1, 117), $p<0.001$) and appropriate spelling, punctuation and capitals ($F=17.68$, (1, 117), $p<0.001$) were all attributable to STAL scores.

Discussion.

This study aimed in part to replicate the research reported in the previous Chapter, with some refinements to the methodology. It would appear that in some respects improvements were achieved, whilst some weaknesses identified in the pilot study were further highlighted.

More comprehensive and specific definitions and instructions provided to standardized patients during training meant that their scoring may have been more reliable than in the previous study. Certainly, the number of open-ended comments made increased in comparison with the pilot study, and these comments were also far more comprehensive. Overall, the amount of missing data generated by standardized patients did not differ greatly between the two studies, as had been hoped. This was due primarily to one of the new items included in the scale, "use of informal language". However, item by item, there were substantial decreases in missing data across the two versions of the Language Rating Scale on "comprehensibility due to accent", "speech acts" and "clarification where comprehension lacking".

Further improvements between the pilot and this study can be seen by inspecting the content of the comments made. Whilst 10 comments were made in the first study pertaining to difficulties in keeping the students talking, only 3 such comments were made for this study. This improvement can be partially attributed to better standardized patient training, but may also have been a function of the fact that this study was incorporated into a barrier examination. Thus, students had been instructed that they were also responsible for maintaining the flow of conversation. Because students were acutely aware that a fail in this examination may have resulted in their inability to proceed to the next year of the medical course, many were understandably nervous. This was reflected in the increased number of comments made by standardized patients about anxiety and nerves, as compared with the previous study's students, who saw the exercise as purely formative and were consequently more relaxed. A regrettable and problematic consequence of

nervousness was its impact upon the fluency of some students' speech, as noted by some standardized patients in their open-ended comments.

More comments were made by standardized patients in this study than the previous one about difficulties in arriving upon a rating. This can be interpreted as a weakness, in that they required perhaps still more training. Alternatively, it could be speculated that this showed greater vigilance on the part of raters to be certain of a score before awarding it. Many comments that pertained to problems in rating involved either the understanding or use of informal language. It would appear that a number of students simply did not use informal language at all during their interview, and therefore could not be rated as "poor" in their use, since this would be misleading. It is quite reasonable to assume that such speech may not be part of everyone's repertoire during interactions, particularly with strangers. Thus, standardized patients found themselves in a dilemma, in that they had to evaluate whether a lack of informal speech indicated a lack of ability to use such language, or simply no desire to use it. This item had been included on the second version of the Language Rating Scale to increase the validity of evaluations of students understanding of informal language. However, it seems that it was only marginally successful in doing so. In any future version of the Language Rating Scale, this item should be deleted, as its usefulness is questionable.

The open-ended comments were notable in that none of them demonstrated a confusion between medical communication skills and English language proficiency. This may be directly attributed to improved standardized patient training procedures.

Despite extended training, certain aspects of the performance of the standardized patients in this study could have been improved. In some areas their behaviour may have been detrimental to the performance of the students that they interviewed. For example, the English Language Specialist commented that some standardized patients compensated for less audible students by speaking more loudly themselves. This may have been quite intimidating to a shy or anxious person. An interesting extension to both this study and the

pilot study would have been to collect information from the students via questionnaire about their perceptions of the process. As this was not done, any perceived impact, unfairness or positive experience for students can only be speculated upon.

A further standardized patient weakness commented upon tangentially by both the English Language Specialist and McGowan (1995b) concerned the fact that in some instances it appeared that the patient did most of the talking. This was definitely inappropriate, and whilst it may have been because the student was either shy, nervous or experiencing difficulties producing fluent English speech, more effective standardized patients would have engaged in encouraging behaviours rather than compensatory ones. Therefore, future training would need to ensure that standardized patients were armed with such strategies to facilitate talking. That standardized patients did not always generate speech is consistent with reports made by Friedman, Sutnick, Stillman, Norcini, Anderson, Williams, Henning and Reeves (1991).

As has already been discussed in the previous Chapter, once again financial constraints resulted in poor audiotape recordings and no access to videotape recordings. As already stated in relation to the pilot study, video-taping would have been preferable but was not possible.

The heterogeneity of the ratings on the Language Rating Scale for the students involved in this study was comparable to their 1994 third year student counterparts. Fewer students in this cohort were rated as "poor" but more were rated as "between poor and adequate" by standardized patients and the English Language Specialist. However, for the purposes of summative assessment only two students were considered to be "unsatisfactory" in their overall performance in oral English according to the Language Development Committee's Oral Language Skills Assessment, and in fact after reassessment, these two students were considered to be "satisfactory". Nevertheless, it could be argued that in this cohort some students would benefit from further language based interventions.

The internal consistency of the Language Rating Scale was comparable with the pilot version of the instrument, again being good to excellent. Inter-item consistency and content validity was generally good to high for the English Language Specialist's ratings. However, inter-item consistency and content validity was quite poor where the standardized patients' ratings were concerned. In particular, register and non-verbal communication did not correlate with any other items on the scale, and speech acts correlated with only one item (understanding of informal language). This discrepancy with the finding of the previous study may be due to a reduction in "observer drift" (Kazdin, 1977) in the group of standardized patients, as a function of improved training. Since the English Language Specialist did not receive such additional advice, it would not be expected that the inter-item consistency from her ratings would differ substantially across the two studies. Comparison of inter-item consistency between this and the pilot study shows that the standardized patients in this study were more consistent with the researcher's ratings in the pilot study. If the inter-item consistency results of the current study are taken to be accurate, then whether the items concerning register, non-verbal communication and speech acts should remain on the scale is questionable. The hypothesis that the standardized patients and English Language Specialist employed different definitions for register and speech acts is reinforced by the fact that inter-rater reliability on these two items was non-existent. Unfortunately, inter-rater reliability for non-verbal communication could not be ascertained without videotape recordings of the student-standardized patient interviews. Otherwise, inter-rater reliability in this study was fair.

Martin, Reznick, Rothman, Tamblyn and Regehr (1996) have explored the question as to whether a clinician or a non-clinician, such as a standardized patient, should rate students during a clinical skills OSCE. This question has at its core the issues of reliability and validity. Martin, Reznick, Rothman, Tamblyn and Regehr (1996) reported that neither clinicians nor non-clinicians are more reliable than the other, citing seven other studies which support this claim. The relative validity of the two types of raters was investigated by comparing the evaluations of students' clinical performance by a medical practitioner, a

standardized patient and an observer of the standardized patient-student encounter with "gold standard" encounters. Medical practitioners were found to rate closest to the "gold standard" and therefore be most valid, followed by standardized patients and observers. Martin, Reznick, Rothman, Tamblyn and Regehr's (1996) initial question and subsequent conclusions can be extrapolated to this research; i.e. "Is students' English language proficiency most validly rated by an English Language Specialist or by a trained standardized patient?" In this study, the English Language Specialist tended to be more strict in her ratings on a number of items than were the standardized patients as group. This was not as obvious a pattern in the previous study.

Comparison of the ratings of the English Language Specialists who comprised the assessment team for the Language Development Committee's Oral Language Skills Assessment with the English Language Specialist rater employed by the researcher for this study shows that the former group were quite lenient in their judgements. This may have been a result of the different rating instrument, or may have been a reflection of the desire to ensure that no student was unfairly deemed as unsatisfactory in spoken English, given the serious consequences such an evaluation might bring. That is, they erred toward false negatives rather than false positives. The rater employed for research purposes would have been less influenced by such qualms, since her evaluations carried no implications for any individual students. Thus, as to who serves as a better rater probably depends upon the purpose of the assessment. If accuracy is required it might be better that raters are made to feel less responsible for the fate of individuals on the basis of their assessments. The English Language Specialists knew that theirs was but one component of an overall evaluation, but it may have seemed to them that their assessments carried considerable weight.

Both the pilot version and the final Language Rating Scale were kept to a minimum number of items. Vu, Marcy, Colliver, Verhulst, Travis and Barrows (1992) have shown that standardized patients are good to very good at accurately recording clinical performance checklists, providing that they are not too lengthy. Vu, Marcy, Colliver,

Verhulst, Travis and Barrows (1992) calculated the predicted percentages of accuracy for different lengths of checklist, reporting 81.44% accuracy for lists of up to 10 items, falling to 80.12% for 15 items. Although this facet of the standardized patients' performance has not been assessed in the present research, since "accuracy" is a problematic concept in a study such as this, it could be postulated that the Language Rating Scale and its pilot were of a reasonable length and should not have presented a formidable cognitive task to trained persons. Vu, Marcy, Colliver, Verhulst, Travis and Barrows (1992) also noted that standardized patients in their study tended to check a student as having demonstrated a skill when they had not - an error of commission as opposed to omission. However, it is important to differentiate between scales and checklists here. Van der Vleuten and van Luyk (1986) stated that;

> "Scoring by means of a criteria-list is simply checking and marking of a behaviour shown. In its strictest sense there can be no personal influence on the part of the person scoring. Like multiple-choice items, checklists are unequivocal and leave no room for dispute about the scoring procedure. This method can be called purely analytical. By contrast, judgements or ratings by observers, whether we try to make them more objective (e.g. Likert scales) or not, never exclude ambiguity. It is an established fact that if interpretations are needed, variations occur and judgements become prone to halo-effects, leniency influences, or any other response style. As opposed to the analytical method, this scoring procedure can be called a global evaluation method." (p 120)

Thus, all of the biases discussed in Chapter VII should be considered in light of this study also. The halo effect (Andrew, 1977), the error of central tendency (Anastasi, 1988), the leniency error (Anastasi, 1988) and the contrast effect (Powis, 1996) are of relevance in this study. The findings by Hui and Cheng (1987) that high proficiency speakers are viewed more favourably than speakers with poorer spoken language skills, that listeners prefer to engage in task-oriented social behaviours with proficient speakers and listeners like proficient speakers more than speakers with lower proficiency all have important implications for this present research and that reported in Chapter VII. For instance, if the standardized patient is considered to be the listener, it is possible that ratings for more proficient students might be artificially boosted across all items on the scale because of general favourable impressions of them, whilst similarly, students with poorer language skills will be rated down on all items. Hui and Cheng's (1987) finding that the listener's own language proficiency did not influence their perceptions of the speaker is not

altogether relevant here, as all of the standardized patients and the English Language Specialists in both studies were operating with homogeneous, high levels of proficiency.

In a similar vein, Hui and Yam (1987) investigated the influence of both language proficiency and physical attractiveness on person perception. They concluded that positive perceptions of a speaker were more likely where language proficiency was higher, but that physical attractiveness could compensate for weaker language skills. Thus, speakers with poorer language proficiency who were deemed physically attractive were perceived more favourably than less attractive speakers with poor spoken ability. These findings have extremely important implications for standardized patient ratings for English language proficiency. Students who are perceived as more attractive to the rater may be attributed with better spoken language than less attractive students. If approved by an ethics committee, a future study could have the students in both the 1994 third year and 1994 first year cohorts rated for attractiveness from their photographs and statistical tests performed to explore any relationship between physical appearance and the scores made by the standardized patients on the Language Rating Scale as compared to the ratings made by others from audiotape (who could not be influenced by appearance).

The concurrent validity of the Language Rating Scale with the Screening Test of Adolescent Language (STAL) was generally moderate. Interestingly, using both the standardized patients' and English Language Specialist's scores, register, speech acts and audibility were not correlated with the STAL. Non-verbal communication ratings made by the standardized patients did not correlate with the STAL and nor did the item of rate for all but one subtest. Rate did not correlate with any STAL subtests or the total where the English Language Specialist's ratings were employed. Whether these items should remain on the scale is thus open to conjecture.

The concurrent validity of the Language Rating Scale with the Language Development Committee's Oral Language Skills Assessment was considered, but a number of methodological flaws of the latter instrument mean that results must be

interpreted with caution. The first item on this instrument required a rating that encompassed students' ability to respond to the patient's opening statements, move from point to point and to the conclusion of the interview and to control grammar and vocabulary. This is a very problematic item on two counts. First, it confounds five distinct and independent skills and second, at least two of those skills (moving from point to point and concluding the interview) could be argued as medical communication skills rather than markers of English language proficiency. The second item also combined three skills under one general rubric (recognize where the patient did not comprehend, rephrase and use language appropriate to the patient). A further complication of the Oral Language Skills Assessment was its vulnerability to be influenced by the primacy effect (Albanese, Prucha, Barnet and Gjerde, 1997), since it was scored from positive to negative. Finally, it would seem that the Oral Language Skills Assessment items may not all be measuring language skills, at least not in the same manner as the Screening Test of Adolescent Language, since only one of the five items correlated with it.

Where both the standardized patients' and the English Language Specialist's ratings were employed in analyses, the Oral Language Skills Assessment items "ability to retrieve information if the patient misunderstands" and the "overall assessment of satisfactory/unsatisfactory language" were not correlated with any item on the Language Rating Scale. The Oral Language Skills Assessment items did not correlate with register, non-verbal communication or use of informal language for standardized patients' ratings, and was not related to scores made by the English Language Specialist on rate of speech, speech acts or audibility. Where the Oral Language Skills Assessment did correlate with the Language Rating Scale, coefficients were generally moderate. Spearman correlation coefficients between the Oral Language Skills Assessment and the Language Rating Scale were negative in the case of the English Language Specialist ratings. Thus, students who were rated higher on the Oral Language Skills Assessment were rated lower on the Language Rating Scale by the English Language Specialist.

The construct validity of the Language Rating Scale in this study was again very good to excellent. The split-half alpha correlation coefficient of 0.49 for the first form of the scale can be explained by the items of register and speech acts, again reinforcing the conclusion that in future versions of the scale these items should probably be excluded. Overall, English speaking background students performed better than non-English speaking background students on the Language Rating Scale as scored by the English Language Specialist, apart from rate of speech, where there was no difference. Standardized patients did not find differences between the two groups of students on register, rate, non-verbal communication, speech acts or audibility. This provides further evidence that these items may not be appropriate to maintain on the scale.

In terms of test security, the time at which the interview was conducted did not impact greatly upon students' ratings apart from their audibility scores. Students interviewed in the afternoon were rated more favourably than those interviewed in the mornings. It is possible that students became aware that this was one of the criteria being examined, and therefore passed this information on to subsequent examinees. Increasing the loudness of one's voice should be fairly easily manipulated and therefore this hypothesis is feasible although it cannot be confirmed.

Females in this study were found to have more appropriate rates of speech than males, as was found in the pilot study. Males were judged to use tense more appropriately and use informal language better than females. Why these gender differences were found is difficult to explain.

Whether a person had acted as a standardized patient or not in the previous study did not appear to greatly influence the ratings given in this study or the amount of missing data generated, contrary to the speculation made by van der Vleuten and Swanson (1990), who queried whether the accuracy of standardized patient ratings might be a function of training. Although it was not measured, it may be the case that with experience comes increased confidence in interviewing or in allocating ratings, or both. It is possible,

though, that the quality and quantity of training rather than experience may be better indicators of effectiveness for a standardized patient who needs to act as an assessor. A more controlled study, designed specifically to address this question, would need to be conducted in order to establish this.

Grammar, spelling and basic written literacy skills are essential in medicine. For example, recording "a pain" may mean something very different from "the pain". The importance of adequate writing skills in the medical professions was highlighted at the inquest into the deaths of two people from Legionnaire's disease in South Australia (Clark, 1997). The newspaper headline "Spelling error: 2 people died" (*The Advertiser,* 29/11/97) underscored the fact that because of a misspelling, laboratory staff failed to diagnose and identify a carrier of the infection (see Appendix XXXVI).

In view of the forgoing, it is therefore worrying that the English Language Specialist who scored students' written accounts of their interview with the standardized patient noted that very few examples of writing were "highly competent, succinct, accurate, cohesive and stylish". Tables LXX and LXXI clearly show that within both cohorts of students a range of abilities in relaying information through writing existed. However, this is quite consistent with the concerns expressed by medical educators all over the world, as has been discussed in the review of literature in Chapter II. As has occurred elsewhere, in other universities, it would seem prudent that the University of Adelaide medical school implement courses in medical writing, regardless of the student's language background or English proficiency. Inspection of the means and frequency distributions in Tables LXX and LXXI show that whilst the 1994 first year cohort had a generally higher standard of writing skills than the 1994 third year cohort, with fewer students evaluated as "poor", a substantial number would still benefit from further tuition in this domain.

Reynolds, Mair & Fischer (1992) reported that in the context of mental health report writing medical practitioners are offered very little formal training in how to read and write a case record. They expressed concern about this, given that these are extremely

important tasks for psychiatrists and psychiatric registrars. Reynolds, Mair and Fischer (1992) also noted that since at least 1946 poor standards for mental health record keeping have been reported in the literature, with similar problems still existing at the time of their research. Thus, it must be considered that poor written proficiency as defined by accuracy and clarity may be as much a function of inadequate student education from the medical curriculum as a reflection of difficulties with English language proficiency. The findings of this study would suggest that comprehensive training in writing case records should be compulsory for all students. This echoes the call for record reading and writing courses made by Reynolds, Mair and Fischer (1992). That no clinicians in Study IV of this thesis commented upon students' written skills may be an indication of the low priority placed upon the teaching of medical writing at the University of Adelaide at that time.

As an instrument, the Written Language Rating Scale appears to have good face validity, based upon the comments made by the English Language Specialist. There were few difficulties encountered by the rater in arriving at a decision using this scale, although she felt that it was easier to evaluate spoken language fluency as opposed to written fluency. In an excellent example of Kazdin's (1977) observer drift, this rater entirely altered the rationale behind one item on the scale, "appropriate use of conventions". The researcher was interested in whether the student knew how to relay the record according to Western cultural rules of address, such as referring to the patient as Mrs Blake rather than as Miss Marilyn. The English Language Specialist instead attempted to rate students according to whether or not their writing conformed to Western medical practice. This was not the intention, since expertise in medical writing would then have been a prerequisite for the rater's evaluations to be valid.

The internal consistency of the Written Language Rating Scale was good to excellent, and inter-item consistency moderate to good (Tables LXXII and LXXIII). Inter-rater reliability could not be established as only one person utilized the scale in this study. Future studies would need to explore this aspect of the scale's reliability.

Content validity of the Written Language Rating Scale was good to excellent (Tables LXXII and LXXIII). Concurrent validity of the scale with the Screening Test of Adolescent Language (STAL) was moderate on most of the items on the scale and subtests of the STAL. Of note though, is the fact that the two items pertaining to register were not correlated with the STAL for either cohort of students (apart from one instance). This provides yet further evidence that register is perhaps a skill unrelated to language proficiency and therefore does not belong on this scale or the Language Rating Scale. Also of note is the pattern of correlations across the two cohorts and STAL subtests. The "language processing" subtest was not correlated with any of the items on the Written Language Rating Scale for the 1994 third student cohort data. Fewer significant correlations were found between the two instruments when the 1994 first year student cohort data were analysed. For this cohort, the subtest "proverbs" failed to correlate with any of the items on the Written Language Rating Scale. A more truncated range of Written Language Rating Scale scores may partially account for the decrease in the number of significant correlations between the two measures across the two cohorts of students.

Correlation coefficients between the Written Language Rating Scale and the Word Knowledge Test (WKT) were moderate to good. In fact overall, the WKT was a better predictor of performance on the Written Language Rating Scale than was the STAL, with higher correlation coefficients for several of the items, including vocabulary, tense, fluency and overall impression of written language proficiency.

Concurrent validity between the Written Language Rating Scale and the Language Rating Scale differed depending upon the raters involved in the Language Rating Scale and the cohort. Concurrent validity was greater where the English Language Specialist rated both instruments, with greater discrepancies found when the standardized patients scored students on the Language Rating Scale. Inspection of Tables LXXVII, LXXIX and LXXX again led to the conclusion that the ability to employ an appropriate register is probably not a language skill, as Long's (1985) study of stereotypicallity versus flexibility

might suggest. Whether speech acts, audibility and non-verbal communication belong in an assessment of language proficiency is also questionable from considering these results. The item which deals with rate of speech may also be of less relevance than other items on the scale.

The concurrent validity between the Written Language Rating Scale and the Language Development Committee's Written Language Skills Assessment was poor. This result was not unexpected, due to logistical flaws related to the latter tool. Item I assessed the extent to which the written account relayed to another health care professional what took place during the interview. However, the English Language Specialist who rated this item did not know what had occurred during the interview in question as they did not listen to a recording of the session, and often it was the case that the rater had not been present at any of the interviews and so could not have relied upon memory to assess written accounts. Thus, the validity of this item is low. The second item, related to grammar, whilst an objective assessment, confounded eight separate variables (incorrect words, misleading words, unexplained specialist vocabulary, omissions, incorrect links, tense, prepositions and articles). Similarly, item III related to seemingly unrelated skills (legibility of handwriting, punctuation, spelling and again omissions). Neither items I or III were related to the Written Language Rating Scale, whereas item II frequently was, particularly for the items of tense, articles and overall impression of written proficiency. The Written Language Skills Assessment did demonstrate some concurrent validity with the STAL, and again, it was the second item that was most significantly related to that test.

The construct validity of the Written Language Rating Scale was good to excellent. The 1994 third year students from non-English speaking backgrounds were consistently rated as poorer on the scale than were their English speaking background counterparts. The 1994 first years also showed significant differences between the two groups, in the expected direction, but on fewer items. In this cohort of students the discrepancy in writing skills between students on the basis of language background was reduced. This

provides some evidence that the Language Development Programme addressed and improved the writing skills of disadvantaged students.

Females in the 1994 third year cohort were judged to have more legible handwriting, a finding not inconsistent with the stereotypical assumption that females are more "neat". Males in the 1994 first year cohort were deemed to have a generally higher level of written language proficiency than females, including better use of tense, articles, pronouns and prepositions. Why this might be so is unclear. An attempt to discover the reason for this finding, by examining the language background of students according to gender, failed to provide an explanation.

One of the major aims of this Chapter was to draw comparisons between two cohorts of students; the 1994 third years, who had no access to the Language Development Programme and limited language assistance of any sort, and the 1994 first year cohort, members of which had been identified as requiring language intervention (see Chapter X) and had the advantage of the Language Development Programme for two years at the time this study was conducted.

According to the results of the Screening Test of Adolescent Language (STAL) there were no significant differences between the two cohorts of students. Similarly, the English Language Specialist's ratings on the Language Rating Scale showed no significant differences between the two cohorts apart from one item, whereby the 1994 first year group were judged to clarify where they did not comprehend to a greater extent than the 1994 third years. It would not be expected that the STAL scores should necessarily be different between the two groups, since the 1994 first years were tested before any language intervention commenced, and the 1994 third years had not received tuition in this area either. Medicine has been classified by the International Testing System (IELTS) Handbook (1966) as a linguistically demanding academic course. As such, it suggests that with an IELTS band score of 6.0, which is the minimum accepted score for entry into the University of Adelaide (University of Adelaide Undergraduate Course Prospectus for

1998, 1997), a student would require up to 300 hours of language tuition before a noticeable improvement in language skills would be registered by the IELTS measure. The Language Development Programme was offered to students for a maximum of 78 hours over two years. Farnill and Hayes (1996c) have demonstrated that students who simply participate in mainstream studies do not show increments in their language skills if these are poor at the commencement of their studies. Thus, the results of this study would lend support to their findings.

It would, however, be expected that the English Language Specialist's ratings would be significantly higher for the 1994 first years, given that eligible students had been provided with access to the Language Development Programme. This was not the case and on the basis of her ratings, it must be concluded that the Language Development Programme was not successful in improving students' proficiency in English, as assessed by the Language Rating Scale. Indeed, the 1994 third year students from non-English speaking backgrounds were deemed by her to use tense more appropriately than the 1994 first year cohort of non-English speaking background students. Where the standardized patients' ratings are employed as the benchmark, quite a different picture emerges. Although fluency of speech and overall impression of language proficiency did not differ across the two cohorts, all other items on the Language Rating Scale showed that the 1994 first year cohort performed significantly better than the 1994 third years. Thus, using standardized patient ratings, it could be argued that the Language Development Programme was beneficial to students. However, interpretation of these results begs the question as to whose ratings are the most valid - the standardized patients' or the English Language Specialist's? This is a debatable point, given the biases that both sets of raters may have been operating under, as discussed above. An improvement upon the methodology which could have helped establish which raters' scores were more valid would have involved assessing students' spoken language proficiency on a standardized test such as the Test of Spoken English (Clark and Swinton, 1980) and then determining the concurrent validity between that measure and the two sets of raters. Future investigations of the Language Rating Scale should thus ensure that another, standardized

test is also administered, specifically designed to assess spoken language (as opposed to aural and written skills as assessed by the STAL).

It would appear that students from the 1994 third year cohort were generally weaker in their writing skills than the 1994 first year group, based upon the ratings of the English Language Specialist on the Written Language Rating Scale. That she considered the former group to be more appropriate in their use of conventions must be carefully considered, since her interpretation of this item on the scale was not as intended by the researcher. Interestingly, as with the Language Rating Scale, differences were evident on individual items of the scale. However, these were not reflected in terms of fluency of writing or the rater's overall impression of written language proficiency.

Analyses investigating the relative roles of participation in the Language Development Programme and English language proficiency as assessed by the Screening Test of Adolescent Language (STAL) showed that generally any improvement between the two cohorts could not be attributed to the Programme alone, since language proficiency overrode this. Thus, students disadvantaged by poor English language proficiency did not improve significantly in their spoken or written English after up to two years in the Language Development Programme, at least as assessed by the instruments used in this study. Interestingly, the two items where improvement can be considered to have been a function of the Language Development Programme, use of appropriate register and non-verbal communication, have been speculated here as possibly being more relevant as medical communication skills rather than as language skills *per se*. It is possible, therefore, that the Language Development Programme has greater utility in teaching communication skills than language skills, although this depends upon one's definition of these two terms (as explored and discussed in Chapter V of this thesis).

Despite some questions of validity regarding the ratings, this study was useful in exploring methods of assessing students' spoken and written English language proficiency. It also enabled an objective comparison to be made between two cohorts of students, in an

attempt to establish the effectiveness of a language intervention, based upon the scores of several standardized and unstandardized instruments. Whilst the psychometric properties of these measures was carefully explored, predictive validity in terms of academic and clinical performance was not considered. This is an avenue for future research. Indeed, as has been discussed in the review of literature in Chapter II, a number of researchers have considered whether language can be used as a predictor of performance in medical school. The majority of investigators have taken overall performance in the course as their outcome measure; few studies have carefully examined students' performance in one subject. The study that is presented in the following chapter considers the predictive validity of students' language background, participation in the Language Development Programme and an updated version of the Screening Test of Adolescent Language (STAL), the Australian Tertiary English Screening Test (AUSTEST) for academic performance in a single first year subject. In addition, the effectiveness of the Supplementary Programme, designed to assist non-English speaking background students and described in Chapter I, is considered.

Chapter IX.

Study VII.

The predictive validity of language background, English language proficiency and participation in language and subject based interventions for performance in a first year behavioural science course.

Aims of the present study.

The study described in this chapter had four main objectives:

i) to describe the English language proficiency profile of 1995 and 1996 first year University of Adelaide medical students in terms of a standardized measure (the AUSTEST)

(ii) to consider the predictive validity of the AUSTEST for academic performance in a behavioural sciences subject

(iii) to determine the performance in this subject of students attending the Language Development Programme

(iv) to evaluate the effectiveness of a teaching programme in facilitating non-English speaking background students' learning in that subject as reflected by summative assessment and student perceptions.

Introduction

The University's curriculum for Medicine aims to integrate traditional pre-clinical disciplines with clinically oriented, problem-based frameworks. Subjects of study for first year students are Cell and Molecular Biology, Human Structure and Function, Introductory Medicine and Doctor, Patient and Society (University of Adelaide Calendar, 1996) (Appendix VI). The latter subject (hereafter referred to as "DPS") draws upon material from the behavioural sciences, including psychology, sociology and anthropology, taught in lecture, tutorial and problem-based formats (Pilowsky & Winefield, 1976; O'Hanlon, Winefield, Hejka & Chur-Hansen, 1995) (Appendix IV.I). The subject is assessed through

two end of semester written essay examinations, three reports of practical work and contributions during tutorials. This subject presents a particular challenge to students experiencing difficulties with English language skills, because it places an emphasis on essay writing, analytical thinking and tutorial participation in the form of debate and discussion. Thus, poor written or verbal skills may hamper some students in their ability to understand course material and learn in an active, self-directed way. In addition, weakness in language may be compounded by a lack of familiarity with Australian culture, further disadvantaging students. For example, in 1995 and 1996 approximately 22% of first year students were temporary residents, intending to return to their country of origin on graduation.

Method

Subjects

In 1995, 140 students were enrolled in first year Medicine; of these 57 (40.7%) were of non-English speaking background and 54 (38.6%) attended the Supplementary Programme. During 1996, the Faculty admitted 119 first year students, of whom 49 (41.2%) were from a non-English speaking background and 33 (27.7%) attended the programme.

The 1995 Supplementary Programme group comprised 23 males (42.6%) and 31 females (57.4%). Thirty-three students (61.1%) were temporary residents in Australia for the duration of their studies, being primarily from Malaysia. The language students reported that they spoke at home is listed in Table LXXXIII.

In 1996, 17 males (51.5%) and 16 females (48.5%) attended the Supplementary Programme. Twenty-three students (69.7%) in this group were temporary residents, again mostly Malaysian. The language spoken at home appears in Table LXXXIII.

Table LXXXIII. Languages spoken at home by supplementary programme attenders in 1995 and 1996.

LANGUAGE	1995	1996
Arabic	1 (1.9%)	-
Bahasa Malaysia	28 (51.9%)	21 (63.6%)
Cantonese	3 (5.6%)	2 (6.1%)
English	7 (13.0%)	3 (9.1%)
Farsi	4 (7.4%)	1 (3.0%)
Greek	1 (1.9%)	-
Hindi	2 (3.7%)	1 (3.0%)
Japanese	1 (1.9%)	-
Mandarin	1 (1.9%)	4 (12.1%)
Other Chinese	2 (3.7%)	-
Russian	1 (1.9%)	-
Serbocroatian	1 (1.9%)	-
Urdu	1 (1.9%)	-
Vietnamese	1 (1.9%)	1 (3.0%)
totals	54 (100%)	33 (100%)

Procedure

English Language Proficiency Screening.

In 1995 and 1996, all enrolling students were screened for potential English language difficulties with the Australian Tertiary English Screening Test (AUSTEST), a twenty item instrument that has been validated with Australian undergraduate medical students (Farnill and Hayes, 1996a; 1996b) (Appendix XXXVII). The AUSTEST is based upon another standardized language screening test, the Screening Test of Adolescent Language (STAL) (Prather, Breecher, Stafford and Wallace, 1980) which has been described and discussed in Chapter III. Like the STAL, the AUSTEST can be administered to large groups of students, rather than by conducting individual testing sessions; the examiner reads each item, and the students record their answers on a proforma sheet which can be expediently and reliably scored (Farnill & Hayes, 1996b). The AUSTEST is composed of four subtests: Subtest 1: Vocabulary (10 items), where an alternative word must be provided to the prompt word, which is presented in a sentence; Subtest 2: Dictation (3 items), where a sentence must be transcribed verbatim after having been read out only once; Subtest 3: Absurdities (4 items), where a nonsense sentence must be explained in terms of what does not make sense and why; and Subtest 4: Proverbs (3 items), where the meaning of the proverb must be given.

Allocation to the Language Development Programme.

On the basis of performance on the AUSTEST and an assessment of English language skills during two interviews with a Faculty staff member, students who were deemed to have potential difficulties with the English language were directed to undergo up to two years of language development. The Language Development Programme, designed and taught by an English language specialist, was taken concurrently with the medical course and required a maximum time commitment of one-and-a-half hours per week (for further details see Chur-Hansen, 1997 (Appendix I.III) and Chapter I).

The Supplementary Programme.

The University of Adelaide pre-clinical academic year is divided into two 13 week semesters. In 1995, 45 hours of supplementary tuition were offered to students in Semester I and 40 hours were available in Semester II. The same person (the author) taught students throughout the year. In 1996 Semester I was taken by a second person, who offered only one hour per week, totalling 12 hours of additional tuition to students from non-English speaking backgrounds. The author resumed teaching in Semester II, increasing the available time for classes to 43 hours.

DPS Summative Assessment.

Assessment for DPS in both 1995 and 1996 was a composite mark out of 100. Two three-hour examinations, where students were required to answer four questions in essay format each accounted for 30% of the total; three practical reports and participation in tutorials were each worth 10%. In the end of semester examinations students wrote on material they had learned through lectures, readings and problem-based learning format. Answers which reflected rote learning were usually less highly scored than more reflective, integrated answers. All three practical reports demanded a high level of independent learning from students. In one practical, students designed and conducted a questionnaire survey of public attitudes toward an aspect of medicine and reported their findings, including statistical analyses of their data; another required students to become familiar with and demonstrate medical communication skills during an interview about

psychological development within a family, while in the third report students integrated their observations of clinical practice with concepts from the DPS course (Peay, 1977). The tutorial programme extended the material presented in lectures and practical sessions and stressed group interaction. In 1995, grading was roughly equally shared between four staff; in 1996, five staff shared this task (three of whom had graded in 1995). In both years all graders were "blind" to the identity of the students on all assessments, apart from tutorial participation grades, which could not be allocated without knowledge of the student in question.

Student Perceptions.

At the end of the 1995 and 1996 academic years, before the final examination, all supplementary attenders were asked to complete a questionnaire which gauged their perceptions of firstly, how much difficulty, if any, they felt they had experienced in studying DPS during the year, and secondly, how useful they considered the Supplementary Programme had been in ameliorating any difficulties. Students provided ratings on a four point scale (for example, (1) much difficulty, (2) some difficulty, (3) no difficulty, (4) unsure), and also provided open-ended comments to clarify their responses (Appendices XXXVIII and XXXIX).

Results.

English Language Proficiency Screening.

Farnill and Hayes (1996b) state that an AUSTEST score of 14 or below can be taken as an indication that difficulties with the English language may be present, but stress that this cut-off is elastic. The tester must determine whether this score is too high or low, depending upon circumstances and other available information, such as that collected during an interview with the testee (see Chapter III). At the University of Adelaide a score of 15 was taken as a marker of language difficulties for the 1995 and 1996 cohorts.

In 1995, 129 (92.1%) students completed the AUSTEST. Fifty students (38.8%) were identified by the test as experiencing difficulties with language. Scores ranged from

5 to 20 ($\bar{x}$=15.2, sd=3.4). In 1996, the AUSTEST was completed by 113 (95.0%) students, with 39 (34.5%) being identified as having English language problems. Scores for this group ranged from 4 to 19 ($\bar{x}$=15.5, sd=3.2).

Allocation to the Language Development Programme.

In 1995, 44 students (31.4% of first year enrolments) were required to attend the Language Development Programme. In 1996, 23 students, or 19.3% of the first year cohort attended the Language Development Programme.

DPS Assessments.

In 1995 and 1996, all summative assessment scores were negatively skewed, with few students failing outright on any single assessment measure. Five students (3.7%) failed to meet an overall satisfactory standard for DPS in 1995; in 1996 this number was 3 (2.5%).

English Language Proficiency Scores and DPS Assessments.

Spearman correlation coefficients for AUSTEST scores with the total end of year assessment score were low to moderate and significant in 1995 ($r=.31$, $p<0.001$) and 1996 ($r=.43$, $p<0.001$).

AUSTEST scores were categorized according to their position above or below the cut-off score of 15 and tested for their relationship with assessment scores. In 1995 and 1996 students identified as experiencing difficulties with the English language were more likely to perform less well than other students on both end of semester examinations (M-W tests, $p<0.005$) and the total end of year score for the subject (M-W test, $p<0.005$). A trend toward lower tutorial participation assessments was observed in 1995 (M-W test, $p<0.06$) but not in 1996. No significant differences were found between identified language difficulties and performance on the three practical reports in 1995. In 1996 students who scored below the cut-off of 15 on the AUSTEST were scored significantly lower on their survey and medical communication skills practical reports (M-W tests, $p<0.005$).

English Language Proficiency of Supplementary Programme Attenders.

Supplementary Programme attenders scored significantly lower on the AUSTEST than non-attenders in 1995 and 1996 (M-W test, $p<0.001$). In 1995 attenders scored an average of 11.9 on the AUSTEST (sd=3.3, range 5 to 17) compared with non-attenders ($\bar{x}$=17.2, sd=1.4, range 14 to 20). Similarly, in 1996 attenders performed more poorly on the AUSTEST than non-attenders ($\bar{x}$=12.6, sd=3.4, range 4 to 18 versus $\bar{x}$=16.7, sd=2.2, range 14 to 20).

Supplementary programme attenders versus non-attenders.

No significant differences were found for the 1995 cohort between the summative assessments on the three practical reports or tutorial participation for students who did, as compared with those who did not attend the Supplementary Programme. However, on the two semester examinations and final assessment score, supplementary attenders performed more poorly than non-attenders (M-W test, $p<0.001$). Of the five students who failed to reach a satisfactory standard for the requirements of the course overall, three had attended the programme.

In 1996, students who attended the Supplementary Programme performed less well than other students in both the survey and medical communication skills practical reports (M-W test, $p<0.005$), the first semester examination (M-W test, $p<0.001$) and the final assessment score (M-W test, $p<0.005$). No significant differences were observed for the clinical observations practical report, the second semester examination or tutorial participation assessment. The three students who failed to satisfy the requirements of the DPS course overall had been attenders of the programme.

The role of English language proficiency level (as measured by the AUSTEST) in the Supplementary Programme was examined by separate one-way analyses of covariance performed on the summative assessment measures discussed below. In 1995, the lower performance of supplementary attenders on the first semester examination ($F=20.48$ (1, 123), $p<0.001$), second semester examination ($F=8.67$ (1, 122), $p<0.005$) and final

assessment score (F=14.89 (1, 122), p<0.001) was attributable to students' English language proficiency. Similar results were found for the 1996 supplementary attenders on the survey practical report (F=12.14 (1,110), p<0.005), medical communication skills practical report (F=13.07 (1,109), p<0.001), the first semester examination (F=20.62 (1, 109), p<0.001) and the final assessment score (F=35.73 (1,109), p<0.001).

Supplementary Programme attenders, 1995 versus 1996.

A comparison of the academic performance of Supplementary Programme attenders on the three practical reports, the two end of semester examinations, tutorial participation assessment and the final assessment score in 1995 against 1996 showed no significant differences.

Non-English speaking background (NESB) Supplementary Programme attenders versus non-attenders.

The number of non-English speaking background students who did not attend the supplementary programme in 1995 was too small for meaningful statistical comparisons to be made. In 1996, 19 students from a non-English speaking background considered that they did not require supplementary assistance. A comparison of the NESB students who attended the programme with those who chose not to yielded no significant differences between the two groups in performance on the clinical observations practical report, the second semester examination or the tutorial participation assessment. However, students from NESB who did not attend the programme scored significantly higher on the survey practical report (M-W test, p<0.05), the first semester examination, the psychological development practical report and the final assessment score (M-W tests, p<0.005). These students were also less likely to be identified by the AUSTEST as experiencing difficulties with language (M-W test, p<0.005).

Supplementary and Language Development Programme Attendance and DPS Assessments

Tables LXXXIV and LXXXV indicate that most students in 1995 and 1996 who elected to participate in the Supplementary Programme also attended the Language

Development Programme. However, students in both years who had not been selected for language development did select themselves for the Supplementary Programme.

Table LXXXIV: Supplementary and Language Development Programme Attendance, 1995 (n=140).

		Supplementary Programme	
		yes	no
Language Development	yes	42 (30.0%)	2 (1.4%)
	no	12 (8.6%)	84 (60.0%)

Table LXXXV: Supplementary and Language Development Programme Attendance, 1996 (n=119).

		Supplementary Programme	
		yes	no
Language Development	yes	18 (15.1%)	5 (4.2%)
	no	15 (12.6%)	81 (68.1%)

A comparison of the Supplementary Programme attenders according to whether or not they had been selected for the Language Development Programme on DPS assessment scores showed no significant differences in 1995. In 1996, students in the Language Development Programme scored significantly lower on the second semester examination (M-W test, $p<0.05$) with a trend toward an overall poorer final assessment score for DPS (M-W test, $p<0.06$).

Student Perceptions.

Students' perceptions of the difficulty of DPS indicate that the majority of students attending the Supplementary Programme in 1995 and 1996 considered the course to be "very difficult" to "somewhat difficult" (Table LXXXVI). Generally speaking, students believed that the programme was "very useful" to "somewhat useful" in helping them overcome the challenges posed by the course (Table LXXXVII).

Table LXXXVI: Responses (%) to the questionnaire items "How much difficulty did you have in DPS in terms of. . ."

	Much	Some	None	Much	Some	None
Item	**1995**			**1996**		
reading and understanding the textbook?	33.3	56.4	10.3	7.4	77.8	14.8
understanding the lecture material?	19.4	75.0	5.6	22.7	77.3	-
taking lecture notes?	26.3	63.2	10.5	42.9	42.9	14.3
writing practice examination essays?	15.4	84.6	-	26.1	65.2	8.7
writing practical reports?	2.9	61.8	35.3	11.5	65.4	23.1
contributing verbally in tutorials?	6.1	45.5	48.5	16.7	29.2	54.2
overall difficulty?	13.9	80.6	5.6	7.7	88.5	3.8

Notes: All "unsure" responses have been coded as missing data.
Response rates: 1995 - 39 (72.2%). 1996 - 29 (87.9%)

Table LXXXVII: Responses (%) to the questionnaire items "How useful were the supplementary tutorials in helping you to . . ."

	Very	Some what	Not	Very	Some what	Not
Item	**1995**			**1996**		
take better lecture notes?	81.6	18.4	-	86.2	10.3	3.4
understand the lecture material?	89.5	10.5	-	93.1	6.9	-
write practice examination essays?	61.3	38.7	-	72.0	28.0	-
write the practical reports?	63.6	33.3	3.0	50.0	46.4	3.6
contribute verbally in tutorials?	40.0	48.0	12.0	36.4	50.0	13.6

Notes: All "unsure" responses have been coded as missing data.
Response rates: 1995 - 39 (72.2%). 1996 - 29 (87.9%)

The open-ended responses to the questionnaire were extensive and varied and have not been included in this Chapter, as their insight into the tertiary learning experiences of non-English speaking students are worthy of study in their own right. The transcribed responses can be found in Appendices XXXX and XXXXI.

Discussion

Many of the students studying medicine at the University of Adelaide from non-English speaking backgrounds experience difficulties in the medical course because of their English language skills. A number of authors have underscored the obligation of tertiary institutions to ensure support systems are in place for NESB students (Ballard, 1987; Munro, 1988; Burke, 1990; Burns, 1991; Kennedy, 1992; Isaac, 1993; Webb, 1993; McGowan and Cargill, 1997; Choi, 1997). In response to the need for teaching assistance the Faculty has implemented both the Language Development Programme, to address problems related directly to reading, writing, speaking and listening in English, and the Supplementary Programme, which focuses upon supporting students in their studies in the medical course. The teaching of academic study skills within the Supplementary Programme is consistent with the recommendation of Mullins, Quintrell and Hancock (1995), who surveyed both international and local students at three South Australian universities to determine their perceived strengths and weaknesses of the tertiary education offered to them by academic staff. Although the two programmes target many of the same students within the first year cohorts, their educational aims and purposes are based upon quite divergent philosophies. As its title implies, the Language Development Programme is developmental in its focus and encourages life-long language learning (McGowan, 1995a). The Supplementary Programme's primary aim, in the shorter term, is to ensure that self-identified "at risk" students succeed in the summative assessment criteria of a specific course, though learning strategies are taught which are intended for longer term benefits, in addition to an appreciation of the role of the behavioural sciences in medicine. It is important that medical students, regardless of language proficiency, are well educated in the psychosocial sciences (Bolman, 1995). Neumann (1847, in DiMatteo and DiNicola, 1982) said "medical science is intrinsically and essentially a social science, and as long as this is not recognized in practice we shall not be able to enjoy its benefits and shall have to be satisfied with an empty shell and a sham" (p 9). Students with language difficulties who do not receive appropriate interventions may be disadvantaged in their preparatory education about psychosocial issues which could be reflected in their later practice as a medical professional. Without a longitudinal research design and reliable and valid

outcome measures, including the definition of a "good doctor" (Huxham, Lipton, Hamilton and Chant, 1989), this can be merely speculation.

Supplementary attenders as a group did not score higher than other students in any of their summative assessments. However, on many measures there were no differences in performance between the two groups, and by the end of the academic year very few students had failed outright. Although these results cannot be categorically attributed to the effectiveness of the Supplementary Programme, since they are also due to the hard work of many other staff, and not in the least due to the efforts and perseverance of the students themselves, it can be posited that without the programme the disparity in performance may have been much greater and the fail rates far more substantial. It is not possible to test this without dividing the students into those with and those without language difficulties, who then are, or are not, given additional tuition. Such an experiment is of course unviable and unethical.

Isaacs (1989) in a review of the relationship between lecture note-taking and learning, has suggested that students may recall more and become more active in their approach to subject material where they have the skills to take lecture notes. The Supplementary Programme emphasized the teaching of lecture note-taking strategies and encouraged students to aim for accurate and complete notes. Whether this tuition was successful in both the short and long-term cannot be established from this study. Isaacs indicates that virtually no literature exists on the benefits or otherwise of note-taking in lectures within the medical education literature. This is an area for further research that requires exploration.

The finding that the underlying explanation for poorer performance on some assessment measures for the supplementary attenders could be attributed to students' existing language proficiency is a most important one. It implies that these students are considerably disadvantaged by comparison with their peers. Graham (1987) has discussed the situation where a student falls below a minimum level of sufficient English language

proficiency to negotiate a given course. Certainly, the range of AUSTEST scores in these cohorts would indicate that for some students this was the case. The fact that on some assessment measures students in the Language Development Programme were scoring significantly lower than others after almost a year in tertiary study and the Language Development Programme is a finding that requires further exploration. Farnill and Hayes (1996c) have demonstrated that students with language difficulties who study medicine do not improve significantly in their language skills simply through mainstream participation in the course. Future research efforts are necessary to examine the effectiveness of specific language tuition in improving proficiency for targeted students, and the relationship between language interventions and academic and clinical performance where a comprehensive language programme has been put in place. A longitudinal research design would be most instructive. Ninnes (1994), in a study of pass rates in the University of Adelaide medical course during the first three years showed that whilst in first year a negative effect was evident between non-English speaking language backgrounds and performance, by third year this trend had reversed, with students from English and East and Southeast Asian language backgrounds performing on par, and those of Southwest Asian language backgrounds performing the most poorly, in contrast with the remainder of non-English speaking background students, who scored highest. This finding is intriguing when considered in light of the fact that his student sample would not have had the advantage of the Language Development Programme, having completed first year before 1994. However, Ninnes used end of year assessment scores only, which may have masked important patterns. Similarly, he analysed the data according to the language students reported to speak at home, but took this to be the same as language background, which is not necessarily the case. Furthermore, a measure of English language proficiency was not available for his study. In addition, the statistical model (partial least squares path analysis) used by Ninnes (1994) accounted for only 11% of the variance in pass rates at the first year level, increasing to 33% at the third year for five variables (socio-economic status, language spoken at home, gender, age and performance in Matriculation). Therefore more than 89% of the variance cannot be explained by the language factor in first year, and over 67% is unexplained in third year. For these reasons Ninnes' results

should be interpreted with caution. In recognition of sometimes contradictory findings across studies, and problems in the interpretation of findings, Mullins, Quintrell and Hancock (1995) have stated that;

> "The relative academic success of students was discussed by Burke (1988) who found that academic performance of a small sample of students, who were new arrivals experiencing problems with social adjustment, was in fact better than the performance of local and international students who had completed their high school studies in Australia. This study, as well as preliminary data on comparative undergraduate progression rates at the University of Adelaide indicate that international students studying in Australia are doing well, highlighting the need for more readily available comparative data on the progression rates of local and international students" (p 203-204).

Correlations for end of year academic performance with the AUSTEST in both student cohorts are consistent with those reported by colleagues from the University of Sydney, Australia (Farnill & Hayes, 1996b), and are also similar to the correlations yielded when students at that university were screened for English language proficiency using the STAL (Hayes & Farnill, 1993a; 1993b). An interesting study might be an investigation into whether correlations between AUSTEST scores and academic performance weaken over time. It would also be instructive to consider whether the AUSTEST holds predictive validity for clinical performance measures. Comparisons between medical schools are interesting, but it must be remembered that whilst "it seems likely that there is a minimal level of English proficiency required before other factors assume more importance, (w)hat that minimal level is will almost certainly vary from institution to institution and, indeed, from program to program" (Graham, 1987, p 517). Thus, whether the medical schools have similar types of courses (for example, problem-based versus traditional) and methods of entry (undergraduate or postgraduate) will have implications for the validity of comparing the effects of language proficiency upon performance. However, generally speaking, at the time of this research and that of Hayes and Farnill (1993a; 1993b) and Farnill and Hayes (1996b) Adelaide and Sydney Universities did have comparable curricula.

There are several methodological limitations to this study, not the least of which is the assumption that the summative assessments of DPS are valid indicators of student

learning. In the medical education literature there has been for some time questions over the reliability and validity of both summative and formative assessment methods and their appropriateness under different situations (for instance, Bandaranayake, 1978; Newble & Cannon, 1994; Rolfe and McPherson, 1995). The influence of the type of assessment tool used to allocate summative scores for students must be considered, as this may result in differences in students' performance as a result of English language proficiency rather than reflecting true differences in ability or knowledge base. For instance, it is possible that students experiencing language difficulties are disadvantaged in assessment situations requiring lengthy essay answers or where vivas form part or the entire assessment. Dissanayake, Ali and Nayar (1990) have gone part of the way in investigating this, by examining the frequency and pattern of answer changes in multiple-choice question examinations as a function of the English language proficiency of Saudi Arabian students. Patterns of changes to answers were incorrect to correct, incorrect to incorrect and correct to incorrect. The authors concluded that language proficiency, as measured by the students' score in a first year subject of English, was not influential in the alteration of answers. Unfortunately Dissanayake, Ali and Nayar (1990) failed to report the relationship between English language proficiency and the incidence of total correct versus incorrect answers.

Huxham, Lipton and Hamilton (1975) conducted research investigating the performance on multiple choice as compared with essay examinations for 146 of the total cohort of 204 University of Queensland second year medical students enrolled in Physiology in 1972. Eight percent of the total group were from a non-English speaking background, but only half of these (4.0%) were included in the study. They concluded that students from non-English speaking backgrounds performed better on essay tests, hypothesising that perhaps whilst such students may lack the ability to understand the wording of a multiple choice statement which must increase the likelihood of an incorrect response, the assessor of an essay may be forgiving of errors in grammar and vocabulary providing the meaning in the answer is clear. It is possible that a similar bias towards

students who could not express themselves in "perfect" written English occurred with the DPS assessors in the present study.

That the interrater reliability across markers has not been controlled for in this paper is a methodological flaw; unfortunately complete data were not gathered in this area. A further methodological consideration is that between 1995 and 1996 there were some variations in the course and several different staff members were involved in teaching from one year to the other. Essentially though, the course content and structure remained constant and three of the five teachers remained the same. Nevertheless, staff changes may have impacted either upon student learning or the standards of marking, or a combination of these.

Perhaps students' questionnaire feedback is more informative than the marks that they achieved in evaluating the effectiveness of the programme. The questionnaire responses by the supplementary attenders on most items indicated that very few students felt the programme was not useful to them. The exception was the item related to the helpfulness of the programme in enhancing their ability to participate in tutorials. This is probably because around 50% of the students surveyed reported that they had not found contribution in tutorials to be particularly difficult for them, though it is also important to remember that the remaining students were challenged by this task, a finding consistent with other research into the experiences of foreign students studying in Australia (Samuelowicz, 1987; Barker, Child, Gallois, Jones & Callan, 1991; Choi, 1997).

A second index of the effectiveness of the programme may be attendance levels. Certainly no medical student would submit themselves to up to 85 hours of additional tutorial time per year on a voluntary basis if they felt that this investment of time was not going to benefit them in a tangible way, particularly in a medical course that encourages competition and rewards students who have been ranked highest in the class. Lending weight to this argument is that the pursuit of knowledge purely for intrinsic intellectual rewards is not necessarily a philosophy readily embraced by many of the students who

were the subjects of this paper. The attendance levels at the Supplementary Programme are consistent with Felix (1993, in Mullins, Quintrell and Hancock, 1995), who has suggested that additional contact outside of accredited course work is viewed more favourably by students if it involves lecturers from within the discipline.

The non-specific benefits of attendance in the Supplementary Programme cannot be measured, since ethical restraints do not allow the methodology necessary to establish whether these effects play a role here. Some of the non-specific benefits for students might include acknowledgement that they have special educational requirements, that they are allowed the opportunity to give their opinions without fear of ridicule or shame if their language is not equivalent to a native speaker's, the realisation that they are one of many students with the same problems, that an identified faculty staff member can be approached and act as a mentor if desired and that staff member's advice will pertain specifically to course content rather than abstract or non-specific material, which may further complicate misunderstandings.

It had been predicted that the 1995 Supplementary attenders would perform significantly better overall in DPS than those in 1996 because the latter group had limited supplementary assistance during first semester. Why this was not the case can only be conjecture, but it is possible that the benefit of any additional help outweighs the amount of help given, as measured by the number of hours. Again, without a methodology that includes a control group of NESB students with language difficulties who have no assistance whatsoever, this cannot be substantiated.

It is also necessary to realise that English speaking background students could receive additional help where necessary and were indeed encouraged to do so. This has not been controlled for in analyses since no records were available to determine the amount or type of assistance they were given. Some supplementary attenders did state the language they spoke at home was "English". Although the students may have identified this as their main language, it was not their first language. They may have been more

fluent in another language, but through circumstance or choice spoke English at home. Any student whose first language was English and who spoke English at home was not considered as being suitable to join the supplementary programme. This was because of restraints on resources, and also because previous experience demonstrated that where fluent English speakers were permitted to join the group, other students sometimes felt intimidated and humiliated, thus negating the fundamental aims of the programme.

Students who did not attend the Supplementary Programme were not asked to rate their perceptions of the difficulty of the course or the usefulness of teaching assistance through tutorials and the additional help, if any, that they actively sought when in need. To make more meaningful comparisons between groups of students this information should be collected for future cohorts. In future research it would also be feasible to compare the performance of the Supplementary attenders in DPS with results in other first year subjects, where no such programme is offered.

Medical educators need to beware of relying upon anecdotal evidence or their intuition that teaching interventions (or their absence) significantly influence students' academic performance without empirical evidence to support such claims. It is not always easy to know how best to define and measure outcomes, and frequently the data will not be "pure", but the intellectual energy expended in considering this and formulating solutions inevitably results in critical reflection about one's teaching methods and their true effectiveness (Ballard, 1987; Phillips, 1990, 1991; Biggs, 1996). Where they are found wanting, strategies for improvement can be planned and teaching methods evolved and adapted to suit pedagogical needs (Cleary, 1996). The resultant innovations may then themselves be held under scrutiny. To be realistic, though, persuading medical faculty staff to re-evaluate their teaching and implement new strategies and approaches is notoriously difficult (Craig and Bandaranayake, 1993).

The present study could be taken as an argument to support the position that incoming students' English language proficiency should be subjected to closer scrutiny or

more rigorous evaluations before their applications to medicine are accepted. Perhaps the development of language needs to occur before they enter the course rather than whilst they are undertaking it. Possibly the demands of both the course and improving language skills at the same time are unreasonable to ask of students, particularly those who are also establishing themselves in a foreign country and may be experiencing "culture shock" (Furnham & Bochner, 1986) and coping with approaches to learning and teaching that may be alien to their previous schooling, resulting in "study shock" (Burns, 1991).

Although this study did find that English language proficiency was moderately predictive of academic success there must be many non-language factors which contribute to academic success or failure (Light, Xu and Mossop, 1987) and thus account for the remaining variance. These factors need to be identified and explored because help can be provided by facilitating the positive elements and minimising those which inhibit progress. Some of the relevant factors in the enrolment and participation of non-English speaking background students in tertiary education identified by Isaac (1993) include: socio-economic background, birthplace, mother tongue, language spoken at home, age, gender, length of residence, motivation level, academic performance, location and quality of secondary schooling and previous tertiary qualifications. It may also be that language skills contribute to difficulties in academic and clinical performance that are also due to other factors such as preferred learning styles, the ability to use critical thinking, personality characteristics, adjustment to new cultural norms and to the unfamiliar expectations of Australian tertiary education (Ballard, 1987; Barker, Child, Gallois, Jones and Callan, 1991). Webb (1993) has identified seven areas in which students from a tertiary population could potentially encounter difficulties (p 7); English language proficiency; approaches to learning i.e. "surface" versus "deep"; insight into the "hidden curriculum" i.e. knowing how to be selective in deciding what should be covered in a course and what can be ignored (also identified by Phillips, 1990); familiarity with the milieu of higher education, including how tutorials and lectures function, an expectation of independent learning, assumed background knowledge; access to external supports such as family and friendship groups and financial resources; attitude to the "authority" of

lecturers; and knowledge of the conventions of academic discourse and discipline-specific practices, such as acceptable forms of essay writing and the appropriate use of language register.

McLaine and Stevenson (1993) reported that the best predictor of academic success, after selection procedures, is preparation for the skills required for tertiary learning, along with the motivation to succeed. Roessler, Lester, Butler, Rankin and Collins (1978) found that personality trait measures were useful in predicting achieved grades in the basic sciences for two cohorts of medical students. This was also demonstrated by Huxham, Lipton and Hamilton (1980) and Lipton, Huxham and Hamilton (1984) who considered personality test scores as predictors of academic and clinical achievement for medical students from the University of Queensland.

The medical course may be difficult for all students regardless of language background or proficiency, particularly the transition from the first three preclinical years of the medical course to the last three clinical years (Vernon-Roberts and Yeatman, 1994). Students with English language difficulties may struggle through the course, at least in their first year DPS, as evidenced in this study, taxing staff and financial resources due to their special needs. Nevertheless, it is imperative that they remain in medicine. This is because firstly, Australia is a multicultural society and doctors from diverse language and ethnic backgrounds are needed (Ahmad, Kernohan and Baker, 1991). Secondly, fee-paying students who return to their country of origin after graduating strengthen intellectual and research ties with our Asian and other international neighbours. For example, Malaysia aims to achieve a doctor-population ratio of 1:1500 and a ratio of one specialist for every 100 general practitioners by the year 2000 (Razali, 1996), and are unable to produce these numbers relying solely on their local medical schools. Australia is currently educating a number of Malaysian students who will return to their country to practice medicine. Thirdly, overseas students contribute substantially to our course by bringing with them different ways of viewing the world and enriching teaching sessions for all students and staff. Fourthly, the teaching expertise that develops to aid non-English

speaking background students inevitably enhances the quality of teaching for English speaking background students as well (Ballard, 1992; Morris and Hudson, 1995). Of note too are the multitude of personal and professional benefits for the students themselves, who come to live, study and work in a country other than their own (Gastel, 1995b). In addition, from 1997 the University of Adelaide has formally recognised the value of students as interpreters for patients who are not fluent in English, by providing training for both students and staff at two major South Australian public hospitals (Hailstone, 1997). Farnill, Todisco, Hayes and Bartlett (1997) have also reported on an innovative teaching programme to teach Australian undergraduate medical students interviewing skills with non-English speaking background people. This is all apart from the obvious financial benefits to the Australian economy (Sweetman, 1996; Klimidis, Minas, Stuart & Hayes, 1997) and the university community as a whole. Providing high calibre academic support to students from non-English speaking backgrounds has mutual benefits for all concerned.

Chapter X.

General Discussion and Conclusions.

The primary aims of this thesis were to investigate the English language proficiency of successive cohorts of undergraduate medical students at an Australian university (Aim I) and to establish the impact of that proficiency upon academic and clinical performance (Aim II). For the purpose of this thesis the tests employed to assess English language proficiency have also been used to define proficiency. It is realised that this is a circular argument that may not necessarily be accepted by other scholars. "Academic performance", as conceptualised in the thesis, refers to the summative outcome measures of an examination in medical communication skills and performance in a one-year behavioural science course. "Clinical performance" has been taken in this thesis to refer to interactions between students and patients, both "real" and standardized. As with the definition of "English language proficiency", it is acknowledged that these latter two definitions may be debated or be unacceptable to some readers.

In order to address the two primary aims of the thesis seven interrelated studies were presented which addressed the following areas.

Area I. Profiles were constructed of cohorts of University of Adelaide Medical School students with regards to both their language background and level of English language proficiency.

Area II. An investigation into whether any significant differences in students' academic and clinical performance could be attributed to the introduction of the Faculty's Language Development Programme was undertaken.

Area III. A methodology whereby English language skills and medical communication skills could be examined within a medical education framework was utilised.

Area IV. Students' ability to employ appropriate English language skills when interacting with patients was examined.

Area V. Studies involving the overt use of standardized patients within the context of a clinical encounter to assess the English language skills of undergraduate students with a range of proficiencies by means of a rating scale were designed and executed.

Area VI. Students' ability to convey clearly and accurately in writing a brief clinical encounter between themselves and a standardized patient was assessed.

Area VII. The predictive validity of measures of English language proficiency for academic performance in a specific subject within the curriculum was explored.

Area VIII. The role of a Supplementary Programme in the academic performance of students from non-English speaking backgrounds in a first year subject was considered.

Study I (Chapter III) addressed Aim I of the thesis. This Study reported a process which involved screening a 1994 first year cohort of students for English language proficiency with a standardized instrument. The scores derived from that instrument were then used to formulate a profile of English language proficiency for that particular student cohort and to determine the relationship between language proficiency and a number of demographic characteristics. The scores were also used, along with a structured interview, to allocate students for compulsory English language intervention. Study I addressed Area I of the thesis and also formed part of the baseline data necesary for consideration of Area II (whether academic and clinical performance might be influenced by participation in the Language Development Programme).

As a result of the findings of Study I several issues for future consideration were identified, including the need to collect complete data from both non-English speaking background and English speaking background students and those identified as having acceptable English language proficiency as well as those identified as experiencing difficulties with language. A further area for future research identified was whether the number of social supports and mental health have any bearing on students' academic and clinical performance. Many authors have considered the emotional impact of leaving one's own country to study elsewhere (Posen, 1968; Bochner and Wicks, 1972; Brislin, 1979; Argyle, Furnham and Graham, 1981; Furnham and Bochner, in Bochner, 1983;

Zwingmann and Gunn, 1983; Furnham and Bochner, 1986; Radford and Wong, 1986; Burns, 1991), but a search of the literature has failed to locate empirical research (as opposed to anecdotal evidence) that has considered the emotional impact that poor English language proficiency *per se* may have upon medical students. In their study of 2,536 international students Klineberg and Hull (1979) reported that depression was a problem for 25% of the sample, although the relationship between feelings of depression and proficiency in the language of the host country was not explored. Lucas, Lenstrup, Prinz, Williamson, Yip and Tipoe (1997) queried whether high levels of depression and stress reported in medical students at the University of Hong Kong could be attributed to English language problems, but this remains a largely unexplored question to date. Accardo, Haake and Whitman (1989) have written about learning-disabled medical students who experience academic difficulties, which could be viewed as an analogous situation to students struggling with language problems. Learning-disabled students include those with dyslexia (reading problems), dyscalculia (mathematics problems), dysgraphia (handwriting problems), problems with processing material in a logical, analytical manner (sequential learning) and problems with synthesising material into a meaningful whole (simultaneous learning) (Accardo, Haake and Whitman, 1989). Having to adjust to their handicap is both difficult and challenging (Accardo, Haake and Whitman, 1990; Policastro, 1990) and may result in "low self-esteem, poor self-concept, social isolation, withdrawal, anxiety and depression (as an outcome of) years of frustration and misunderstanding" (p 257). Quintrell and Westwood (1994) reported that first year undergraduate students who were paired with a local host student were more likely to describe their first year experience positively, to use student services offered by the university and to report gains in perceived language fluency. Quintrell and Westwood (1994) found no impact of hosting upon academic performance. Thus, an area of future research might be the exploration of the use of peer tutoring and support schemes for medical students. However, although emotional issues as a result of difficulties with language have not been explored in this thesis, the comments made by students that were reported in Studies I and VII indicate that they are salient and deserve further investigation as related to the medical student experience.

Study I also highlighted the fact that in evaluating the impact of English language proficiency upon academic and clinical performance in medical school, a longitudinal research design is preferable to a cross-sectional design, where this is a feasible option. Cross-sectional designs and comparisons between cohorts are valuable, but greater reliability and validity can be achieved by studying one or more complete cohorts of students over time.

In terms of contributing to knowledge, Study I was useful in providing the Faculty of Medicine at the University of Adelaide with an objective profile of the English language proficiency of a first year cohort of students. This information was used for a specific purpose (that is, the allocation of students to language classes), but additionally, could be used as comparative data regarding the incidence of likely English language difficulties for future cohorts, either at the University of Adelaide, or other institutions. Finally, the findings of Study I threw into question the validity of assuming that students' performance in medical communication skills examinations could be used as an index of English language proficiency.

Study II addressed Aim I and Areas I and II. In order to assess increments in English language proficiency as a result of the Language Development Programme it was necessary to compare a cohort of students who had been given access to the programme with a cohort who had not. Thus, in Study II the 1994 third year students were screened for English language proficiency with the same standardized instrument administered to the 1994 first year student cohort. Results of a previous English language test taken by the 1994 third years when they were in first year were also presented. Study II served the purpose of establishing baseline data for comparisons with the 1994 first year students.

Study II's secondary aim was to determine whether the results of the two tests taken by the 1994 third years (the Word Knowledge Test (WKT) when they were in first year and the Screening Test of Adolescent Language (STAL) in their third year) differed significantly. Significant increments in performance between the first and second test

would indicate that English language proficiency could increase over time without the need for specific language based interventions. Unfortunately, this secondary aim could not be validly achieved, due to a number of methodological flaws. As has been discussed in Chapter IV, a major problem was that complete data were unavailable for both cohorts for both tests. Thus, attrition rates may have impacted upon the findings, with students failing to proceed to third year also being the students who were experiencing difficulties with the demands made upon their English language skills. It should be stressed, however, that this is speculation, as no evidence was collected to support this. Also problematic in interpreting the results of this study is the fact that no information has been published regarding the concurrent validity of the STAL and WKT. It would appear that the two tests are measuring some of the same skills, (as evidenced by Table XXXIII presented in Chapter IV), but because of the sampling limitations just described, this conclusion must be viewed with caution. Furthermore, even the highest correlation of 0.67 leaves almost one-third of the variance unexplained. In any future research which seeks to determine whether English language proficiency improves over time, either with or without an intervention, two versions of the same standardized test should be administered, such as Forms E and F of the WKT. This study was necessary for inclusion in the thesis, but is limited as a wider contribution to knowledge, in the absence of concurrent validity data for the two language tests.

Study III addressed Aim II and Area III of the thesis, in that it sought to distinguish English language proficiency from medical communication skills as assessed in an examination. Since the examination was one component of the summative assessment of third year MBBS, it can be considered part of students' academic performance, and given that the examination took the form of an Observed Structured Clinical Interview (OSCI) utilising standardized patients, it may also be taken to indicate clinical performance.

Study III's findings showed that English language proficiency and medical communication skills should not be used as interchangeable terms. Unsatisfactory performance in the examination of medical communication could not be attributed to

language background or STAL results. However, ratings of spoken language were associated with performance in the examination, with students identified as having unsatisfactory fluency of speech more likely to fail. Methodologically, this study was sound in that the STAL and the medical communication skills examination were taken in fairly close temporal proximity, and also in that the examination was an actual Faculty requirement for summative assessment. On the other hand, one could query the extent to which behaviour in an examination may be extrapolated to actual behaviour in a clinical setting with patients. If it is argued that performance in this examination is not indicative of a student's ability to communicate with patients, the extent to which this study addressed Aim II could be questioned. However, since clinical performance is defined as student interaction with actual or standardized patients, Aim II was satisfied, for the purpose of this thesis.

Study III makes an important contribution to medical education because it clearly distinguishes between the terms "language proficiency" and "communication skills", which have often been used to refer to the same skill, as discussed in Chapter II. Thus, it cannot be assumed that students from a non-English speaking background will necessarily perform poorly in tests of medical communication, or that English speaking background students will easily acquire medical communication skills. However, students who have poor spoken fluency are more likely than students with clear spoken English to be rated lower by examiners in demonstrating medical communication skills. Thus, medical educators may wish to pay particular attention to students with poor spoken fluency, perhaps by allocating them to language interventions where these are available, to increase their ability to speak clearly and fluently. Students from non-English speaking backgrounds who believe that the learning of medical communication skills is partially contingent upon coming from an English speaking background can be reassured that this is not the case. Conversely, students from English speaking backgrounds who become complacent about learning medical communication skills, because they feel that they communicate satisfactorily every day, can be shown that medical communication is a specialized art and core clinical skill that must be mastered (Lipkin, 1991, in Cohen-Cole,

1991) apart from mastery of a specific language, which in itself is not sufficient. Researchers should endeavour not to confound language proficiency and communication skills when both designing studies and interpreting the results.

Study IV partially addressed Aim II of the thesis, by considering the impact of students' language upon clinical performance. This study was a retrospective review of clinicians' comments regarding students undertaking their clinical placements over years III, IV and V of the MBBS degree at the University of Adelaide. It was considered necessary for two reasons; first, to establish specific areas of concern related to language when students interacted with patients, as opposed to the generic concern that students had "language problems", and second, to identify suitable items for inclusion in a scale designed to assess students' English language proficiency when interacting with a standardized patient, as described in Study V. Study IV indicated that students' language difficulties did impact upon their interaction with patients, as evidenced by the comments made by clinical teaching staff. However, conclusions about the extent to which comments pertaining to language were associated with poor summative assessments on the Clinical Skills Tutor Assessment Sheet (Appendix XVII.I and XVII.II) could not be made, since numerical gradings for each student were not collected for this study. It could be argued that a weakness of this study was that so few clinicians commented about language on the assessment sheets. As discussed in Chapter VI, there are several explanations for why this might be so, including that clinicians are generally not trained to evaluate language skills, and therefore might be reluctant to express their opinions about individual students in this regard, for fear of recrimination. Thus, the comments that were made by clinical staff are all the more valuable. However, it is acknowledged that a far more rigourous research design would collect written comments about every student's language proficiency. Unfortunately this was impossible for this study, as the researcher was collecting archival material and had not had control over the way in which clinicians were asked to complete assessment sheets for each student.

As well as serving its purpose as the basis for the design of a scale to rate students' spoken language skills, Study IV was an extremely useful exercise, in that planning teaching strategies to assist students with English language difficulties may focus upon specific problematic areas as identified by clinical teachers. Indeed, on the basis of comments about students' problems with the use and comprehension of Australian colloquial language, a teaching project was designed and implemented at the University of Adelaide (Chur-Hansen and Barrett, 1996).

Study V addressed Aim I, partially addressed Aim II of the thesis and considered Areas IV and V, since students' spoken English language proficiency was rated in the context of a clinical interview with a standardized patient. This study's main purposes were twofold; first, as a pilot study to investigate the utility of an instrument to rate spoken language (the Language Rating Scale), and second, to collect data on that instrument for the 1994 third year cohort so that comparisons could be made between their performance and that of the 1994 first year cohort, who, as second years in 1995, had benefited from language intervention where appropriate. That is, this study served as the basis for comparison of the 1994 third year student cohort with the 1994 first year cohort (when these students had proceeded to second year in 1995). In this regard Study V served as a basis for addressing Area II in Study VI.

Extensive analyses were conducted on the ratings generated from the standardized patients and the three other raters (two English language specialists and the researcher), in order to explore the reliability and validity of the Language Rating Scale (LRS) as an instrument to assess spoken English language proficiency. Chapter VII presents a lengthy discussion of the conclusions drawn regarding the psychometric properties of the LRS. The LRS and the method by which it was administered were both found to have good face validity. Internal consistency, inter-item, content and construct validity were all good to excellent. Concurrent validity with standardized tests of language proficiency (the STAL and the WKT) were moderate; low to moderate correlations were found between the LRS and ratings of the fluency of spoken language made during an examination of medical

communication skills, discussed in Study III. These findings would indicate that the LRS is a useful tool for standardized patients to rate students' spoken language in a clinical context. However, there were a number of problems identified in this study which detracted from the LRS's utility. In particular, inter-rater reliabilities varied from high to low, depending upon the item and the sets of raters in question. It could be that one explanation for low inter-rater reliabilities concerned raters using their own idiosyncratic definitions of items on the scale, rather than adhering strictly to the criteria provided to them. Differing conceptualizations of what is meant by "fluent" language may similarly explain low concurrent validity between the LRS and ratings of language during the medical communication skills examination. This study, as a pilot, was very useful in identifying that the consistency and validity of the standardized patients' presentations was wanting; this too, could help explain poor inter-rater reliabilities, due to problems in generating speech from students to rate, and contributing to the likelihood of missing data.

Despite its weaknesses, this study constituted an important contribution to knowledge. It presented a new scale for the assessment of spoken language, based upon an extensive review of the relevant literature (presented in Chapter II) and local feedback made by clinical teachers (as described in Study IV). The methodology employed (the overt use of standardized patients in a clinical context) was found to be viable as a means by which to administer the scale. Whilst the study has been discussed here in terms of its role for this thesis, future researchers may wish to replicate the methodology (with or without the inclusion of a standardized rating scale of spoken language) in order to examine how students of differing language proficiencies interact with patients, allowing for the identification of those who may need additional assistance with their language skills.

Study VI addressed Aim I and partially addressed Aim II, in that clinical performance was evaluated. Areas II, IV, V and VI were also addressed by this study. This was a replication of Study V, employing a cohort of students who had been provided with access to the Language Development Programme (the 1994 first years, now in second

year in 1995), so that comparisons between this cohort and that of Study V (comprising students who had not had access to the Language Development Programme) could be made. This study sought to rectify methodological weaknesses that were identified in the pilot (Study V), primarily by the inclusion of a further two items on the Language Rating Scale, clearer definitions of each item and rating criteria and more intensive training of standardized patients.

Study VI extended Study V, in that a rating scale to evaluate written English language proficiency in a clinical context was developed and piloted. Comparisons were made between the two cohorts of students (the 1994 third years and 1995 second years) on this measure (the Written Language Rating Scale). Preliminary analyses were performed in order to assess the reliability and validity of the Written Language Rating Scale. Further exploration into the reliablity and validity of the Language Rating Scale was also undertaken. As found in Study V, the internal consistency and construct validity of the LRS was good to excellent and concurrent validity with the STAL was again found to be generally moderate. However, inter-item consistency and content validity differed depending upon which set of ratings were employed in analyses. Inter-rater reliabilities were low to moderate on nine of the eleven scale items for which this statistic could be calculated (the twelfth, non-verbal communication, could not be analysed). Because of the differences between raters, findings in relation to the relative performance of the two cohorts of students were somewhat contradictory. If the ratings made by standardized patients are taken as valid, it could be concluded that students who participated in the Language Development Programme performed significantly better on the LRS than students who did not. However, if the English Language Specialist's ratings are taken as the indicator, this conclusion cannot be made. Unfortunately, it is not possible from the research presented in this thesis to determine which set of ratings should be taken as the gold standard. A methodological weakness of Study V and Study VI was that a standardized test of spoken English other than the instrument designed for this thesis (the Language Rating Scale) was not administered. Therefore, the validity of different raters' evaluations cannot be determined. Any future research would have to ensure concurrent

validity between an established measure of spoken English and any newly designed measure of spoken English for more meaningful results to be gathered. Therefore, whilst the Language Rating Scale shows potential as an instrument to evaluate spoken language proficiency, it should not be used in future studies until its concurrent validity can be demonstrated to be sound.

The Written Language Rating Scale (WLRS) was found to have good face validity and good internal and inter-item consistency. The scale's content validity and construct validity were also found to be good. The WLRS showed higher concurrent validity with the WKT than with the STAL. Whilst this shows that the tests are tapping into some of the same skills, it would be more instructive in future research to determine the concurrent validity of the WLRS with another standardized test that assesses writing skills based on the reporting of a passage or interaction, as the WLRS aimed to do, rather than comparing it with tests that assess isolated words or sentences. Inter-rater reliablity for this scale could not be examined, as only one rater used the WLRS - this aspect of reliability should be established in any future research. Concurrent validity of the WLRS with the LRS depended upon which set of ratings were employed in analyses. Where the English Language Specialist's were used, this validity was higher than when the standardized patients' ratings were employed. The English Language Specialist was the rater of the WLRS in this study. As already has been discussed, it is not possible to know which of the two sets of ratings on the LRS are the more valid. Therefore, before the WLRS is used again, this methodological flaw must be rectified.

Bearing in mind the limitations of the WLRS, the ratings made by the English Language Specialist on the scale indicated that participation in the Language Development Programme generally improved writing skills, with students in the 1994 third year student cohort scoring lower than the 1995 second year student group on most items on the scale.

This study was useful in further developing the Language Rating Scale by the inclusion of new items, but the study failed to verify the validity of the scale. Study VI

explored the utility of a new scale to assess written language. Again, the validity of this scale was not ensured and further developments are required. The study was able to objectively compare two cohorts of students using the same instruments, in an attempt to evaluate the effectiveness of the Language Development Programme in improving language proficiency. Such methodology could certainly be used again, but only after the instruments involved in the assessment of English language proficiency have been stringently validated and standardized.

Study VII addressed Aim I of the thesis and considered Aim II in the light of academic performance. Areas II, VII and VIII were addressed. This study sought to investigate the predictive validity of English language proficiency for academic performance in a first year behavioural science course. The tests of English language proficiency designed for use in this thesis were not utilized in this study, in favour of a standardized test of written and aural language proficiency, the AUSTEST. The AUSTEST was also chosen for use with the two cohorts of students who were the subjects of this study, because as first year students, they had no opportunity to interact with either actual or standardized patients, the situation for which the Language Rating Scale had been intended for administration.

Study VII examined academic performance in a single first year subject as the outcome measure in preference to considering performance over the six years of the MBBS. Since most researchers have used the latter as an indicator of academic performance, as discussed in Chapter II, this study can be considered as contributing to the existing literature in this field. The study was also conducted in order to evaluate the influence of two interventions (the Language Development Programme and the Supplementary Programme) upon academic performance. In this regard, it is also adding to knowledge within the area of medical education.

This study demonstrated that English language proficiency as assessed by the AUSTEST was moderately predictive of academic performance. Future research needs to

isolate other factors that may be contributing to academic success or failure, so that educators can concentrate upon these in teaching strategies. The results concerning academic performance as a function of attendance at the Supplementary Programme showed that greater or lesser access to the programme did not significantly influence summative assessment scores. Students who attended the programme generally performed more poorly than other students on a number of assessments, or performed no differently. Since students who attended the Supplementary Programme also tended to be those allocated to the Language Development Programme, these two interventions were confounded and thus the utility of the Language Development Programme could not be reliably evaluated in this study.

The finding that Supplementary Programmme attenders did not perform at higher levels than other students, despite intensive additional assistance, could be interpreted in two ways. It could be argued that as students are so disadvantaged by their language capabilities that intervention programmes such as this are a waste of resources. However, another explanation lies in the possibility that without any additional intervention students in need of supplementary tuition would have performed far more poorly than was observed to be the case. Indeed, very few of these students failed to progress to the second year of the MBBS. Furthermore, student perceptions of the usefulness of the programme would support the argument that it is beneficial. As has been discussed in Chapter IX, to objectively determine the usefulness of interventions such as the Supplementary and Language Development Programmes would require unethical methodological designs, as discussed in Chapter IX.

Whilst this thesis explored a number of areas pertinent to English language proficiency and academic and clinical performance, and aimed to achieve a rigorous and sound methodology, this was not always possible. Ethical restraints meant that students could not be randomly allocated to interventions and logistics disallowed a true longitudinal research design following successive cohorts of students over the six year period of their medical degree and beyond. Comparisons between cohorts were made on

the basis of tests which might be argued to be questionable as assessments of "English language proficiency". Furthermore, it was not possible to conduct test-retest research, by administering the same test twice to one cohort. To have done so would have strengthened the methodology and thus the reliability and validity of some of the conclusions. Finally, Crystal (1976) has pointed out that any study on language and communication will betray the theoretical bias from which the researcher is grounded, depending upon his or her academic training. Thus, it should be borne in mind that as an academic psychologist working in a medical education setting, it is inevitable that my approach and interpretations, as well as that which I have chosen to include or omit from a review of the relevant literature, will be coloured by my own educational and vocational background.

It should be stressed that this research is not based upon the deficiency model (Jordan and Tharp, 1979), which postulates that certain cultural or social groups are inferior, and the author vigorously rejects such a model. The rationale behind this thesis comes from the need to objectively evaluate the role of English language proficiency as a prerequisite skill for successfully undertaking medical studies. The author agrees with Maher (1993) who has applauded the fact that medical education must now take place in a multicultural and multilingual world. He states;

> "Medical students are not a 'problem' because they come from a 'minority' or have accented speech or require language training. A more sophisticated view is to view their monolingual peers as linguistically disadvantaged" (1993, p 4).

Similarly, one can quote Felix (1993), who in referring to international students stated;

> "It is unfortunate that during the course of their studies abroad so much emphasis is placed on their deficiencies when they might be congratulated for taking on a task as sophisticated and courageous as theirs." (Felix, 1993, p 6 in Mullins, Quintrell and Hancock, 1995, p 203).

In this thesis, students were categorized depending upon the language they reported to speak at home on their enrolment form (Appendix IX). It should be recognized that some students who reported that they spoke English at home were sometimes identified

through language testing as experiencing difficulties with language. This may be a reflection of the reliability and validity of the language screening instruments, the accuracy of self-report, or both. Similarly, "country of birth" as assessed by this thesis was also taken from students' enrolment form data. Although a student may have been born in a country other than Australia, their length of residence was not taken into account. Thus students may have lived elsewhere for only a very short period of time, and at varying stages of their life.

Although there are methodological problems associated with employing demographic material generated from enrolment records, this was considered to be a more reliable source of data than other classification systems. For example McManus and Richards (1985) have noted that to ask people through questionnaires to state their race, creed and colour is difficult and may be unacceptable. Therefore, some researchers have classified subjects into European versus non-European according to surnames (see for example, McManus and Richards, 1985; Collier and Burke, 1986; McManus, Richards and Maitlis, 1989; Wakeford, Farooqi, Rashid and Southgate, 1992; McManus, Richards, Winder, Sproston and Styles, 1995; McManus, Richards, Winder and Sproston, 1996). As McManus and Richards (1985) have acknowledged, using names to determine ethnicity may not be entirely valid, although McManus, Maitlis and Richards (1990) concluded that surnames are valid and reliable indicators of ethnic origin. However, in Australia, people who call themselves "Australian" may have either European or non-European surnames. They may have a European surname through marriage or conversely, may have chosen to Anglicise a non-European name for personal reasons. Also, individuals with non-European surnames may come from families who have resided in Australia for several generations. For the purpose of this thesis, classification of students according to surname was not conducted, since alternative information was available, although it might be feasible to consider this classification system for future research involving medical students at Australian universities.

Throughout this thesis the term "non-English speaking background students" (NESB) has been employed. Just what this actually means is not necessarily straightforward. Holton and Salagaras (1988) define NESB as;

> "linguistic and cultural ancestry and current identity, and involves one or more factors like birthplace of father/mother and or self in the non-English speaking world, significant use of a language other than English within the household, experience of a first language other than English, and ethnic identification with non-English speaking ancestry. As such the term refers to both Australian citizens and permanent residents who come from non-English speaking backgrounds, and, to overseas students from non-English speaking backgrounds on temporary student visas." (p 4).

Where possible, in the analyses performed for this thesis the category of NESB has been refined into subcategories, such as the MATES students, permanent Australian residents, private fee-paying overseas students and so on. It is important to ensure that whilst students may be categorized in order to conduct meaningful statistical procedures, in discussing results one does not lose sight of the fact that the "NESBS" are not a homogeneous group (Isaac, 1993; McLaine and Stevenson, 1993). For example, students who are from non-English speaking families and who have been born and educated in Australia clearly have different experiences, strengths and weaknesses in language when compared with students who have recently arrived in this country and who have had very limited exposure to Australian English and the education system here (Tse, 1990). In turn, this latter group of students will also have advantages and disadvantages in dealing with the demands of language and tertiary study (Tse, 1990). It is the responsibility of medical educators to ensure that they do not stereotype all non-English speaking students as necessarily requiring the same support and teaching and learning strategies, for this is a misnomer. At a medical school where a substantial number of students are from a non-English speaking background and have varying proficiencies in English it is also the responsibility of teaching staff to make efforts to take this into account. These efforts may include the modification of individual teaching practices (Lucas, Lenstrup, Prinz, Williamson, Yip and Tipoe, 1997), or referring students at risk of failure or poor performance due to language, to support programmes where they exist, such as the Language Development and Supplementary Programmes described in this thesis. Where they are not available, these services could be implemented by concerned staff. On an

individual level, there is a growing literature and increasing academic activity including conferences and refereed journals which are devoted to the pursuit of pedagogical excellence, including the special needs of medical students and graduates who are not proficient in the language of instruction. Attending and contributing to conferences and keeping abreast of the literature will undoubtedly improve teaching practices and thus benefit students.

This thesis has largely omitted the issue of cultural background. Certainly the cultural background and identity of students may influence their academic and clinical performance and it has been argued that language and culture are inextricable (Isaac, 1993). However, whilst not debating this assertion, there is a voluminous body of literature on cross-cultural issues and education (see for example Bhawuk, 1990) and a focus on culture and medical education deserves more attention than has been possible here. Thus, for the purpose of the present research it was deemed acceptable to focus more upon language-related variables rather than any others, including cultural factors.

A methodological strength of this thesis is the fact that all students were tested for language skills and data were collected on all of them regarding their language background and the language spoken at home. In this regard it differs from other research that has chosen to focus upon non-English speaking background students or students from overseas, at the expense of local, English speaking background students. The exception here is the work of Farnill and his colleagues, who have studied whole intakes of medical students. Future research should concentrate upon the language abilities of all students in a given course, to establish the importance of the role of language background and language proficiency as opposed, for instance, to cultural knowledge or conceptual capabilities (Neumann, 1985; Isaac, 1993).

It should be noted that very little has been written about how to adjust curricula to suit the academic needs of foreign medical graduates (Cole-Kelly, 1994), and similarly, there is a dearth of writing about how to facilitate non-English speaking background

undergraduate medical students' learning. This could be construed as a shirking of the responsibility of tertiary institutions to ensure that courses are accessible to all students regardless of cultural or language background (Sanson, Augoustinos, Gridley, Kyrios, Reser and Turner, 1997). Morris and Hudson (1995), writing from the Universities of Wollongong and Griffith stated that; "Allowing for the fact that different types of English and different genres will be appropriate for different cohorts, all Australian students would benefit from better English language skills" (p 72) and that "...the traditional emphasis on 'correct' i.e. standard English may have to be waived in favour of a much greater emphasis on argumentation, even if the language of presentation is non-standard, albeit discriminating and effective" (p 73). This is debatable, because in the medical profession inaccurate language could potentially result in serious errors, so whilst this may be a true statement for undergraduate essay writing, it is not necessarily acceptable in case notes and letters of referral for example. In addition, Phillips (1990) has underscored the dangers of lowering standards to accommodate students who do not have the prerequisite skills or knowledge;

> "(A) major dilemma is centred on the economics of encouraging a high success rate in overseas student studies. The question seems to be whether academics should insist that overseas students meet the same standards they maintain and apply for local students. Even though this action might result in student failure and the loss of funds coming from overseas students. Or should there be a hidden curriculum for overseas students, whereby sufficient reductions of standards are made, so that such students pass their courses and consequently, overseas student funds are preserved? It is a hidden curriculum because it is realized that overseas countries insist on taking out the same qualifications that local students acquire. Should they suspect that standards have been adjusted, they might place their students elsewhere." (p 5).

Future research needs to investigate the fate of international students once they have returned to their country of origin. Given the data that have been collected on the students who have been the subjects of this thesis, it would be possible to follow them through in a longitudinal fashion. A question that requires immediate addressing involves whether the academic and clinical training provided at the University of Adelaide holds students in good stead with their own country's professional requirements, as well as catering for local students and their future patients (Hamilton, 1995). An important task for any medical school is to ensure that the objectives, goals and aims of teaching are realised (Vroeijenstijn, 1995). That is, are we teaching students the skills that they need to

be effective doctors regardless of the country in which they will work? In other words, can the skills students learn and develop during their time of study at the University of Adelaide be generalised across countries and across cultures? Do our students graduate with both the ability to perform specific tasks and work within a given cultural setting - which has been termed by Stevens (1995) as "technical" and "cultural" competence respectively? Methodological considerations are paramount in designing a research project which could investigate the outcome of students in professional terms on graduation in a valid and meaningful way (Klineberg and Hull, 1979; Rolfe and Powis, 1997), and would require amongst other things, both internal and external assessments (Vroeijenstijn, 1995) and reliable and valid outcome measures.

A factor that should be considered when evaluating whether or not we have adequately prepared our students for medical practice in their country of origin is the teaching philosophy of our medical educators. For example, a deficit model of teaching may be counterproductive and indicative of a tendency towards Australian parochialism (Hay, 1972). Professor Kerry Kennedy, then Dean of the School of Education at the University of Southern Queensland stated:

> "According to a [deficit] model, remedial action is seen to be necessary so that the backgrounds and values of overseas students can be replaced by a new set of values that have more currency in the Australian context. Such a remedy may well assist overseas students, but at what cost? How will they adapt when they return to their countries of origin? Can Australian institutions legitimately engage in such a subtle form of cultural imperialism even though it is designed to provide real assistance to students?" (1992, p 18-19).

Likewise, we need to consider students' personal adjustment upon returning to their country of origin, having spent at least six years studying medicine in Australia (and for some students, lamentably, up to nine years). Much discussion has revolved around the need for preparatory courses and bridging programmes to facilitate international students' entry into tertiary courses. However, there appears to be a dearth of corresponding work regarding how to assist students' exit from the host country and to reintroduce them into their own place of permanent residence (Brislin, 1979).

The vast majority of students who were the subjects of this thesis entered the University of Adelaide Medical School prior to selection procedures introduced in 1996 for the 1997 incoming students, which considered not only academic abilities but also personal qualities (Powis, Neame, Bristow and Murphy, 1988; Neame, Powis and Bristow, 1992; Aldous, Leeder, Price, Jervie Sefton and Teubner, 1997). An exception here are the students who entered through the MATE scheme; their selection process has been described in Chapter I. Thus, it may be that were the studies employing the Language Rating Scale, the Written Rating Scale and the Medical Communication Skills examination to be replicated with students entering under the new system, significant differences may be found. This may be one focus for future research evaluating the benefits of selecting students through interview and psychometric testing in place of matriculation scores for increasing levels of English language proficiency within cohorts. Whether or not this increased proficiency results in greater academic and clinical achievement would also need to be monitored.

An area which has not been considered in this thesis is racial discrimination. It is quite possible that some medical students from non-English speaking backgrounds may be treated differentially (and unfairly) by their peers, teachers or patients. In a comprehensive longitudinal study by Klineberg and Hull (1979) approximately one third of 2,536 overseas students studying various courses in one of eleven different countries (but not including Australia) reported having experienced discrimination. Harris and Jarrett (1990) have written about the potential costs and benefits of educating overseas students in Australia, and they warn;

> "It is likely . . . that racial prejudice amongst Australians will increase if there is an increase in concentration in those universities and disciplines where the overseas students already form a large proportion of the student group. . . . Despite the graffiti on some toilet walls - and it is not certain that these are all student authored - the evidence suggests that overseas students have not engendered attitudes of racial intolerance by Australian students. However, there may well be an upper limit to the acceptance of overseas students in certain areas." (p 68-69).

Morris and Hudson's (1995) recommendation to research Australian students' attitudes towards international students would thus appear timely. Certainly stereotypes

exist in Australia, as elsewhere, which are used to describe groups of students and foreign medical graduates at the expense of considering individual differences (Varki, 1992). For instance, Lowry and MacPherson (1988) needed to make the seemingly obvious point that "discriminating against all those who have foreign names or black faces is an inefficient way of excluding those with a poor command of English" (p 658).

Tan (1977) criticised the United States Educational Council for Foreign Medical Graduates (ECFMG) Examination and research comparing the clinical competence of doctors with foreign qualifications with their American counterparts on the grounds that attempts had been made to denigrate and unfairly exclude overseas trained doctors. It has similarly been alleged by the Human Rights and Equal Opportunity Commission that the Australian Medical Council (AMC) has been discriminatory in its examination of foreign medical graduates for registration to practice in Australia (Thong, 1990; Moss, 1992), a position that has been vigorously denied by Gerber (1991, 1992), Jones (1992) and Milton (1992). Gerber (1991) refuted the notion that testing foreign graduates' comprehension of spoken English constitutes racial discrimination, on the grounds that all candidates who have not qualified in Australia, including United States and Australian citizens must likewise be tested for language proficiency.

Notwithstanding arguments to the contrary, it is conceivable that students and graduates are probably exposed at some time to the prejudice and ethnocentrism of their colleagues (Cole-Kelly, 1994; Richards, 1994). In an article entitled "Letters to a Young Doctor" Rhodes (1983) provides the — hardly encouraging — advice to overseas doctors that medical professionals in Britain " . . . all find the customs of others unusual because we judge from different standpoints and upbringing (and) . . . we tend to be intolerant of others" (p 1136). The truth of this xenophobic statement was reinforced by Singh (1994) who wrote as an overseas registrar working in the United Kingdom:

> "It was hard for me to accept that the entire Western education that I had proudly received in India had resulted only in producing a clever wog. It was even harder to understand how intolerance, ignorance, and prejudice could occur in an affluent society when I had always believed that these in India reflected harsh socio-economic conditions" (p 1169).

Several researchers have empirically tested the possibility that applicants to medical school are treated differentially according to their "ethnicity", which is considered by McManus, Maitlis and Richards (1990) to be "whether [one's] physical appearance is similar to that of the majority of persons from particular parts of the world" (p 73). The dominant ethnic groups they have proposed are White European, Arab/Middle Eastern, Asian (Indian), African, Caribbean and Far Eastern (McManus, Maitlis and Richards, 1990).

One of the first systematic investigations of racial discrimination in the selection of students for London medical schools was conducted by Collier and Burke (1986). They called for a review of admission policies on discovering consistent discrimination on the basis of race. Esmail, Nelson, Primarolo and Toma (1995) investigated the acceptance rates of applicants to medical schools in the United Kingdom, by classifying candidates according to their ethnic group. The authors recognised that by considering only ethnicity, and no other demographic variables, their study was limited in the conclusions that could be drawn from it, but cautiously stated that certain medical schools were found to be preferentially selecting "white" students over others with the same A level scores.

Upon graduation, doctors from ethnic minorities were found by McKeigue, Richards and Richards (1990) and Esmail and Everington (1993) to be less likely to be shortlisted for interview for a graduate training post. Smith (1987) also wrote that overseas doctors tend to make multiple applications for employment before being successful, to be allocated positions in the least popular areas and specialties and to make slower progress in their careers as compared with British doctors of comparable qualifications.

McManus, Richards and Maitlis (1989) assessed the likelihood that ethnic background influenced acceptance into United Kingdom medical schools. They found that applicants from ethnic minorities were less successful in securing a place in medical school as compared with others, and that this difference could not be justified on the

grounds of poorer academic performance or non-academic factors. They contrasted this finding with the results of a previous study (McManus and Richards, 1985), where the lower acceptance rates of minority students could be explained in terms of their failure to satisfy non-academic criteria, such as contributions to the community. Smith (1987), however, queried the explanation given by McManus and Richards (1985) that "non-academic suitability" (Smith, 1987, p 329) might reasonably be used as a factor for excluding medical school candidates, bringing their conclusion that discrimination had not occurred into question. Likewise, Collier and Burke (1986) noted that non-academic criteria are often determined upon the idiosyncrasies of senior staff, and the requirements to fulfil these criteria and how they are weighted in the application process are not usually relayed to potential students in the prospectus.

In a prospective study McManus, Richards, Winder, Sproston and Styles (1995) demonstrated that after controlling for educational qualifications, applicants to medical schools in the United Kingdom who were from ethnic backgrounds were disadvantaged, supporting previous research (McManus, Richards and Maitlis, 1989). Furthermore, McManus, Richards, Winder, Sproston and Styles (1995) found that a European surname was a better predictor of success in securing a place in a medical school than was ethnic background *per se*.

Research has not only been concerned with applications to medical school, but has also focused on the academic performance of students from ethnic backgrounds. McManus, Richards, Winder and Sproston (1996) considered the possibility that students from ethnic minorities may perform less well than other United Kingdom medical students in undergraduate clinical examinations due to racial discrimination. Although they did find that ethnic minority students were more likely to fail multiple-choice questions, essays, vivas and clinical examinations, they believed that this could not be attributed to discrimination on the basis of race, as firstly, multiple-choice examinations were machine scored, and secondly, whilst ethnic students who were permanent residents in the United Kingdom performed more poorly in examinations, ethnic students who were temporary

residents for the duration of their studies performed better than Anglo-Saxon students. McManus, Richards, Winder and Sproston (1996) were unable to explain this finding, and made a call for further research. This thesis would suggest that perhaps language proficiency may have partially explained their finding, had it been taken into account in analyses.

Wakeford, Farooqi, Rashid and Southgate (1992) concluded that the United Kingdom membership examination for the Royal College of General Practitioners did not discriminate against ethnic doctors. They did find, however, that ethnic doctors trained outside of the United Kingdom performed poorly on the examination as compared with doctors trained in the United Kingdom, regardless of ethnicity. Wakeford, Farooqi, Rashid and Southgate (1992) indicated that doctors trained in the Indian subcontinent performed especially poorly, but took this as a reflection of inadequate initial training and a lack of preparation for working in the United Kingdom, rather than evidence of discriminatory practices by the College. They also speculated that these doctors may have experienced language difficulties, thus resulting in poorer examination performance, although they had no evidence to support this, other than the fact that the Indian group of doctors performed significantly worse than others on one of two oral examinations.

The possibility of racial discrimination existing at the University of Adelaide medical school has not been explored. Previous survey research has found that racism is a concern for up to 13% of the University of Adelaide overseas student community (Quintrell, 1991). On the basis of one or two survey questions it is difficult to ascertain whether students had in fact encountered racial discrimination and racist behaviour, or had misunderstood Australian cultural morés. Interviewing or in-depth surveying of overseas or non-English speaking background medical students and graduates about their experiences with their teachers, peers and patients would be ways to establish whether discriminatory activities occur and how frequently. However, as noted by Harris (in Dillner, 1995) this is a very sensitive area and by exposing these problems, if they exist, will beg the need for action (Lowry and MacPherson, 1988). How the problem might be

tackled once students have entered the course has not been widely discussed in the literature, though several authors have suggested ways of avoiding discrimination on applying for entry to medical school or graduate training (McKeigue, Richards and Richards, 1990; Esmail and Everington, 1993; McManus, Richards, Winder, Sproston and Styles, 1995). Dillner (1995), reporting on the admission by the University of Manchester Medical School that Asian male students had been treated differentially in clinical examinations, outlined some of the strategies to be implemented by Manchester. These included a prospective study to follow students by ethnicity through the course, structured examinations with objective grading criteria, a training course on gender and ethnic issues for medical teaching staff and an effort to ensure that patients participating in clinical finals examinations are representative of the cultural mix of the population. Esmail (in Dillner, 1995) also applauded Manchester's public approach to the problem and immediate response to clarify how and where discrimination might be occurring. McKenzie (1995) argued that careful monitoring of any strategies to eradicate discrimination should be systematically monitored, and concluded that a profession that cannot treat its own members fairly will be hard pressed to provide equitable health care for its patients.

Appendices.

Appendix I.I.

Vernon-Roberts, J. and Chur-Hansen, A. (1995)
Letter to the Editor. Communication skills of interns in New South Wales.
Medical Journal of Australia, **163**, 112.

Vernon-Roberts, J.M. & Chur-Hansen, A. (1995) Communication skills of interns in New South Wales.
The Medical Journal of Australia, v. 163, p. 112

NOTE:
This publication is included after page 362 in the print copy
of the thesis held in the University of Adelaide Library.

Appendix I.II.

Farnill, D., Hayes, S.C. and Chur-Hansen, A. (1995)
Interrater reliability of the scoring of the Screening Test of Adolescent Language.
Psychological Reports, **76**, 1027-1032.

Farnill, D., Hayes, S.C. & Chur-Hansen, A. (1995) Interrater reliability of the scoring of the screening test of adolescent language.
Psychological Reports, v. 76(3), pp. 1027-1032

NOTE:
This publication is included after page 363 in the print copy of the thesis held in the University of Adelaide Library.

It is also available online to authorised users at:

http://doi.org/10.2466/pr0.1995.76.3.1027

Appendix I.III.

Chur-Hansen, A. (1997)

Language background, proficiency in English, and selection for language development.

Medical Education, **31**, 312-319.

Chur-Hansen, A. (1997) Language background, proficiency in English and selection for language development.
Medical Education, v. 31(5), pp. 312-319

NOTE:
This publication is included after page 364 in the print copy
of the thesis held in the University of Adelaide Library.

Appendix I.IV.

Chur-Hansen, A., Vernon-Roberts, J. and Clark, S. (1997)
Language background, English language proficiency and medical communication skills of medical students.
Medical Education, **31**, 259-263.

Chur-Hansen, A., Vernon-Roberts, J. & Clark, S. (1997) Language background, English language proficiency and medical communication skills of medical students.
Medical Education, v. 31(4), pp. 259-263

NOTE:
This publication is included after page 365 in the print copy
of the thesis held in the University of Adelaide Library.

Appendix I.V.

Chur-Hansen, A. & Vernon-Roberts, J. (in press)
Clinical teachers' perceptions of medical students' English language proficiency.
Medical Education.

[redacted]

MEDICAL EDUCATION

The International Journal of Undergraduate,
Postgraduate and Continuing Medical Education

Blackwell Science Ltd
23 Ainslie Place
Edinburgh
EH3 6AJ
UK

Telephone (0131) 226 7232
Fax (0131) 226 3803
email: mededit@blacksci.co.uk

EDITORS
Professor H. J. Walton
Dr E. G. Buckley

Ref No: MEd/1997/000058

27 October 1997

Dr Anna Chur-Hansen
Dept of Psychiatry
University of Adelaide
South Australia 5005
Australia

Dear Dr Chur-Hansen,

Clinical teachers' perceptions of medical students' English Language proficiency

I am pleased to inform you that your revised paper has been accepted for publication. Proofs and an order form for offprints will be sent to you in due course.

Yours sincerely

Editor

Appendix I.VI.

Chur-Hansen, A. (in press)

Teaching support in the behavioural sciences for non-English speaking background medical undergraduates.

Teaching and Learning in Medicine.

Teaching and Learning in Medicine

An International Journal

Terrill A. Mast
Editor

Howard S. Barrows
Associate Editor

Editorial Board
John H. Shatzer, Chair
Stephen Abrahamson
M. Brownell Anderson
Dave Davis
Beth Dawson
Jack Ende
Brian C. Jolly
Dick Martenson
Nobutake Matsuo
John J. Norcini
S. Scott Obenshain
Emil Petrusa
Paul G. Ramsey
Richard J. Reznick
Peter L. Schwartz
Sharifah H. Shahabudin
Nu Viet Vu
Reed G. Williams
Christel Woodward
James O. Woolliscroft

Research Methodology
Jerry A. Colliver

International News
Don Clarke

Publishers
Lawrence Erlbaum Associates, Inc.

Southern Illinois University
School of Medicine
Dept. of Medical Education
P.O. Box 19230
Springfield, IL 62794-1217
(217) 782-7765
FAX: (217)524-0192
E-mail: tmast@siumed.edu

February 1, 1998

Dr. Anna Chur-Hansen
The University of Adelaide
Department of Psychiatry
Faculty of Medicine - Division of Health Sciences
GPO Box 498
Adelaide, South Australia 5005

Re: Manuscript No. 828
Teaching Support in The Behavioral Sciences for Non-English Speaking Background Medical Undergraduates

Dear Dr. Chur-Hansen:

We have received your manuscript for potential publication in Teaching and Learning in Medicine. Your manuscript has been reviewed by three well qualified reviewers, whose comments are enclosed.

Based on these three reviews, we would like to consider your manuscript for the Developments section of Teaching and Learning in Medicine. I would also make the suggestion that you reduce the size of the manuscript. The content is too long. Please consider these comments from your peers before revising for resubmission.

We require structured abstracts and enclose a copy of our guidelines. Your manuscript for the Development section should include a Background, Description, Evaluation and a Conclusion.

In making your revisions, please keep in mind that the reviewer comments reflect the opinions of your peers; you need not make every change suggested. However, please send a **detailed cover letter** explaining how the reviewers' (and our) comments have been addressed (or not addressed). We will need three copies of your revision.

Along with your revision, send your telephone and fax numbers, and your e-mail address.

We look forward to seeing your revision.

Sincerely,

Terrill A. Mast, Ph.D.
Editor

TAM:slb
Enclosure

Appendix II.I.

University of Adelaide Teaching Development Grant (1995)

"The assessment of spoken language in a clinical setting" with Dr Jane Vernon-Roberts.

For the sum of $3,500.00

THE UNIVERSITY OF ADELAIDE
Department of Psychiatry

Tuesday, 14 March 1995

The Dean
The Medical Faculty
University of Adelaide

Dear Professor Frewin,

The Language Development Committee is currently planning for a language examination at the end of second year. As you will recall, it was the intention of the curriculum committee that we examine every student in the year, not just those who have done the language course. This is a large undertaking, but we have devised an examination which we think will be very efficient.

We examine oral language by means of a viva. Students will interview a surrogate patient who works with a basic script to do with, say, indigestion. Their spoken language will be marked by a lecturer in ESL who is experienced in the assessment of English. It will also be evaluated separately by the surrogate patient (who will be trained to do this by Jane Vernon-Roberts and Anna Chur-Hansen, as part of Anna's longitudinal study, thereby providing data on the validity and reliability of the assessment process).

We will examine written language by taking two samples of the students' work. One will be a written account of the above interview; another will be a written answer, possibly from a DPS II examination. This written work will be assessed by the ESL lecturer.

We have also considered the introduction of more continuous forms of assessment so that the results do not depend solely on the end of year examination, but our plans in this area are not yet finalized.

Please find included the estimated expense of the end of year assessment. We have been careful to cost this accurately. We anticipate that we will require 6 surrogate patients and six ESL specialists for the examination, and therefore we have included a brief period of training to ensure inter-rater reliability.

We are pleased to report that the costing of the surrogate patients ($2,849.40) has been covered by a Teaching Development Grant won by Jane Vernon-Roberts and Anna Chur-Hansen. Therefore we are requesting Faculty support of $5,307.24 for the ESL specialists.

We feel confident that the Language Development Programme is already producing good results and making our faculty a leader in this field within Australia. We believe that a well organized examination at the end of second year—one that puts teeth into the language course by making students realize that we are very serious about it—is essential to the ongoing success of the programme.

Thank you for giving this budget estimate your careful consideration.

Yours sincerely,

Dr R J Barrett
Convenor
Language Development Committee

Dr Robert J Barrett
Senior Lecturer
University of Adelaide, Department of Psychiatry, Royal Adelaide Hospital, Adelaide, South Australia 5000
Tel: (08) 224 5141 Fax: (08) 232 3298

Appendix III.I.

Chur-Hansen, A. and Vernon-Roberts, J. (1996) Paper and Poster Sessions - "English language proficiency and performance in medical communication skills." Teaching Communication in Medicine Conference, St Catherine's College, University of Oxford, United Kingdom, July 24th-26th.

THE UNIVERSITY OF ADELAIDE

Language Background, English Language Proficiency and Medical Communication Skills in an Australian third year undergraduate medical student cohort.

Anna Chur-Hansen & Jane Vernon-Roberts.

Introduction

Medical communication skills are recognised as fundamental to effective health care. In reponse to this the University of Adelaide Faculty of Medicine has incorporated medical communication skills training into the curriculum.

The University of Adelaide, like several other Australian universities, teaches a substantial number of students from non-English speaking backgrounds who have a range of proficiencies in English.

There has been some question as to whether the medical communication skills of students from non English speaking backgrounds are poorer compared with the skills of English speaking background students. Rolfe and Pearson (1994) and the authors of several letters to the Editor of the Medical Journal of Australia (Glover 1995, Torda 1995, Young 1995) made the assumption that if a student is from a non English speaking background then it is more likely they will perform less well in examinations of medical communication skills, as compared with native English speakers. We noted that in the literature on this issue the terms "language" and "communication" were being used interchangeably, and that this reflected several basic assumptions about the relationships between language background, language proficiency and medical communication skills that had not been demonstrated with empirical evidence. We therefore designed a study to investigate the relationship between:

1. Language background
2. English language proficiency, and
3. Medical communication skills

The primary aims of the study were to determine:

1. whether coming from a non English speaking background is associated with lower proficiency in English, and
2. whether low proficiency in English adversely influences the demonstration of medical communication skills

Method

Subjects

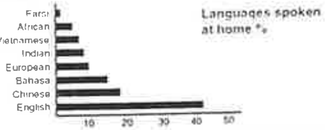
Subjects for this study were the University of Adelaide third year class of 1994, comprising 149 students, 54.4% males and 45.6% females.

Gender of Students

English was the language spoken at home by 40.9% of the students. The remaining 59.1% were from a non-English speaking background.

Language Background
ESB = English Speaking Background
NESB = Non English Speaking Background

Languages other than English spoken at home were Chinese, Bahasa Malaysia, Indian languages, Vietnamese, European languages (including Greek and Italian), an African dialect and Farsi.



Three quarters of the group were permanent Australian residents, with a quarter on temporary visas for the duration of their medical studies, meaning they would not be staying in Australia to practice medicine after graduation.

Resident Status of Students

Procedure

Assessment of Language Proficiency

One hundred and 42 of this group of students (95.3%) were screened for English language proficiency using a standardised instrument, the Screening Test of Adolescent Language (STAL) (Prather, Breecher, Stafford and Wallace, 1981). This is a 23 item test which can be administered to a large group of students at one time, by reading out the items, whilst they record their answers on paper. It takes about ten minutes to complete, and can be quickly and reliably scored using a handbook that comes with the test. There are four subtests:

1. Vocabulary (12 items), where the student must give an alternative word to fit a sentence. Example: "Gigantic. The room is gigantic. Give another word for gigantic."
2. Auditory memory span (3 items), where a sentence is dictated and must be recorded by the student word for word. Example: "The fire drill that we had last week turned out to be the real thing."
3. Processing language (5 items), by explaining what it is about a sentence that does not make sense and why it doesn't make sense. Example: "The sun was shining so brightly last week on Tuesday that I had to wear my sunglasses in the movie theatre."
4. Explaining the meaning of proverbs (3 items). Example: "Practice makes perfect."

Assessment of Medical Communication Skills

To assess the students in medical communication skills, our colleague, Dr Sheila Clark has designed an examination which takes the form of an Observed Standardised Clinical Interview (OSCI), administered after a 15 hour interpersonal skills training programme.

During the OSCI each student interviewed a role-playing patient, observed by a medical practitioner. At the end of the interview, which took about 10 minutes, the doctor and the role-playing patient used a five point scale from 1 (poor) to 5 (excellent) to rate the student on twelve aspects of the interview, related to the use of medical communication skills. These were:

1. introduction, including explaining the purpose of the interview and obtaining informed consent
2. exploration of the patient's knowledge of the problem,
3. exploration of the patient's concerns
4. summarising, checking and finishing the interview
5. rapport and attending,
6. listening,
7. appropriate use of questioning techniques
8. empathy skills,
9. personal and professional qualities, such as warmth, concern, sincerity and respect
10. the use of clear, unambiguous and fluent English
11. the use of appropriate language to the person and context, and
12. an overall feeling of empathy, reflected in the creation of an atmosphere of trust, support and sensitivity to the patient's emotional state

Results

Language Proficiency

Following Hayes and Farnill (1992) a benchmark of 19 out of a possible total of 23 on the STAL was used to identify students experiencing difficulties with the English language.

Results for this cohort of students were negatively skewed, with 70.4% scoring 20 or higher, identifying the remaining 29.6% (n=42) as experiencing difficulties with English language proficiency (mean=20.01, sd=3.15, range=7 to 23).

Medical Communication Skills

Forty students were assessed as a result of the OSCI examination as being unsatisfactory in their ability to demonstrate medical communication skills to their patients.

Ratings of spoken language on the scale from 1 (poor) to 5 (excellent) with a midpoint of 3 (adequate) resulted in the majority of students (85.9%, n=128) assessed as "adequate" to "excellent" in their use of clear, unambiguous and fluent English during the OSCI. The remaining students were rated as "poor" (3.4%, n=5) to "less than adequate" (10.7%, n=16).

Language background, English Language Proficiency and Medical Communication Skills

The following table lists the Spearman correlation coefficients for Medical Communication Skills demonstrated during the OSCI, language background, performance on the STAL, ratings of "clear, unambiguous and fluent English" and ratings of "appropriate language for the person and context". Language background was least often correlated with medical communication skills, with the two ratings of spoken language most often correlated.

Medical Communication Skills demonstrated during OSCI	Language Background	STAL	Clear, unambiguous and fluent English during OSCI	Appropriate language for person and context during OSCI
Introduction	[illegible]	[illegible]	[illegible]	[illegible]
Exploration of problem	[illegible]	[illegible]	[illegible]	[illegible]
Exploration of concerns	[illegible]	[illegible]	[illegible]	[illegible]
Summary, checking, finishing	[illegible]	[illegible]	[illegible]	[illegible]
Rapport and attending	[illegible]	[illegible]	[illegible]	[illegible]
Listening	[illegible]	[illegible]	[illegible]	[illegible]
Questioning	[illegible]	[illegible]	[illegible]	[illegible]
Empathy skills	[illegible]	[illegible]	[illegible]	[illegible]
Personal and professional qualities	[illegible]	[illegible]		[illegible]
Clear, unambiguous, fluent English	[illegible]	[illegible]	[illegible]	
Appropriate language	[illegible]	[illegible]	[illegible]	
Overall empathy	[illegible]	[illegible]	[illegible]	[illegible]
Total OSCI mark	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
STAL performance	[illegible]		[illegible]	[illegible]
Language background		[illegible]	[illegible]	[illegible]

* p < 0.05

To further investigate these correlations associations between the students' language background, English language proficiency as assessed by the STAL, the rating of "clear, unambiguous and fluent English" during the OSCI and the students' overall satisfactory/unsatisfactory performance on medical communication skills during the OSCI were tested employing the chi-square statistic.

There was a statistically significant association between whether a student came from an non-English speaking background or not and whether the student was identified by the screening test as experiencing difficulties with language. Students from non-English speaking background were most likely to be identified as experiencing language difficulties.

Language Background and Performance on the Screening Test of Adolescent Language

LANG/STAL	satisfactory	unsatisfactory	totals
esb	56	1	57
nesb	44	[illegible]	85
totals	100	42	142

$\chi^2=25.11$, df=1, p<0.001

There was no statistically significant association between language background and performance on the OSCI medical communications skills examination.

Language Background and Performance on the Medical Communications Skills Examination

LANG/OSCI	pass	fail	totals
esb	47	14	61
nesb	62	26	88
totals	109	40	149

not significant

Performance on the language screening test and the OSCI medical communication skills examination were also not associated.

Performance on the Screening Test of Adolescent Language and the Medical Communications Skills Examination

STAL/OSCI	pass	fail	totals
pass	79	21	100
fail	26	16	42
totals	105	37	142

not significant

The rating of clear, unambiguous and fluent spoken English made during the OSCI was found to be most often associated with other variables. It was significantly related to language background, whether the student was a permanent or temporary resident, performance on the language screening test, and performance on the medical communications skills examination.

Students identified as having a poor fluency of speech were more likely than other students to fail the OSCI medical communication skills examination.

Rating of Spoken Language and Performance on the Medical Communication Skills Examination

SPEECH/OSCI	pass	fail	totals
satisfactory	[illegible]	[illegible]	128
unsatisfactory	6	15	21
totals	[illegible]	[illegible]	149

$\chi^2=13.29$, df=1, p<0.001

Conclusions

In summary, we found that knowing whether a student was from non English speaking background or not was not predictive of their performance in a medical communication skills examination. Performance on the screening test of language proficiency and satisfactory performance on the medical communications skills examination were weakly, though significantly correlated. It was the rating of clear, unambiguous and fluent spoken language that was most predictive of performance in the medical communication skills examination.

This supports our hypothesis that language background and the ability to demonstrate medical communication skills cannot be assumed to be associated. Furthermore language proficiency as assessed by a standardised screening instrument was not as strongly correlated with performance in the medical communication skills OSCI as were ratings of spoken language.

On the basis of these findings we would argue that within the medical education arena it would be most appropriate to refer to "language" and "communication" as different skills in assessing students on examinations of "medical communication skills". It would therefore be inappropriate to make the assumption that a student weak in one of these skills will necessarily be weak in the other. That spoken language was a stronger predictor of performance in a medical communications skills examination than was language as assessed by a standardised screening test is important. Anna Chur-Hansen's research is currently examining more closely the aspects of spoken language that adversely affect the demonstration of medical communication skills.

In future research it would be valuable to investigate more closely the ability to demonstrate specific medical communication skills, such as empathy, or rapport, and assess their relationship to other predictive variables aside from language proficiency. The correlational data from this study between language related factors and performance in a medical communication skills examination were low to moderate. This would suggest that factors apart from linguistic play a greater role in medical students' ability to demonstrate medical communication skills in clinical settings. This hypothesis would support further our premise that it cannot be assumed that students from non-English speaking backgrounds, or those whose language proficiency is comparatively weaker than other students, will necessarily perform poorly on examinations of medical communication skills.

References

Glover, W.E. (1995) Letter to the Editor. Medical Journal of Australia, 162, 497

Hayes, S.C. and Farnill, D. (1992) A study of the concurrent validity of the Screening Test of Adolescent Language with recent immigrants. Psychological Reports, 71, 175-178

Prather, E.M., Breecher, S.V.A., Stafford, M.L. and Wallace, E.M. (1981) Screening Test of Adolescent Language (STAL). University of Washington Press, Seattle and London

Rolfe, I.E. and Pearson, D. (1994) Communication skills of interns in New South Wales. Medical Journal of Australia, 161, 667-670

Torda, T.A. (1995) Letter to the Editor. Medical Journal of Australia, 162, 497-498

Young, J.A. (1995) Letter to the Editor. Medical Journal of Australia, 162, 497

ANNA CHUR-HANSEN
Department of Psychiatry
University of Adelaide SA 5005
Fax: 61-8-232 3298
E-Mail: achurhan@medicine.adelaide.edu.au

Appendix III.II.

Chur-Hansen, A. (1997) Presentation and Poster Session - "Teaching Support for Non-English Speaking Background Medical Students." Australasian and New Zealand Association for Medical Education (ANZAME) Silver Anniversary Conference: Communication - Art and Science, Ormond College, University of Melbourne, Victoria, July 6th-9th.

Teaching support for non-English speaking background medical undergraduates in the behavioural sciences.

Anna Chur-Hansen, Department of Psychiatry, Royal Adelaide Hospital, University of Adelaide

A number of students who enrol in the Faculty of Medicine at the University of Adelaide, South Australia, are from non-English speaking backgrounds.

Currently, the Faculty screens students for English language proficiency after entry into the medical course, and it is apparent that some students are less well equipped than others with the English language skills required for tertiary study.

Learning in a first year course in the behavioural sciences presents a particular challenge to students experiencing difficulties with English because:

- teaching and learning strategies include essay and report writing, tutorial participation, oral presentations, debate and group discussions
- the critical evaluation of material is highly encouraged
- reading matter may demand a high level of English language proficiency
- the disciplines of psychology, sociology and anthropology are often based upon Western cultural values and assumptions
- a familiarity with Australian culture is assumed knowledge
- as first year university students the culture of tertiary study is unfamiliar to them.

When asked about how they perceive the study of behavioural science, typical comments from non-English speaking background students include:

"This subject is difficult for me because there's a lot of critical thinking to do. It's hard writing in English let alone thinking in English."

"I think all the lecture materials given basically are based on studies done in Western culture. Sometimes I find it hard to relate to my own culture and therefore it's hard to do a good critical thinking."

"This is a somewhat difficult subject for me because of language. I'm not really good at English, it is very hard for me to understand. ... However, I really like the subject because it is interesting and makes me think!!!"

The Teaching Programme

A programme of tutorials has been developed and offered since 1987 to assist non-English speaking background students in their study of behavioural science. The tutorials are subject-based rather than focusing on English language *per se*.

Attendance is voluntary and open to all students from non-English speaking backgrounds. Students may attend for up to four hours of tuition per week, or up to approximately 85 hours per year. There is no assessment or penalty for non-attendance. Because the tutorials are an adjunct to the behavioural science course, the programme has come to be known as the "Supplementary tutorials".

It is clear that English language proficiency is only one aspect that needs to be addressed for students who attend the Supplementary tutorials; cultural issues, expectations about learning and teaching, and isolation from other students are also pertinent issues that must be dealt with.

To address student needs, the programme includes:

- discussion of the content of lectures
- explanation of both everyday vocabulary and professional terminology
- placing course material into a cultural context
- asking and answering of questions
- workshops on essay writing, report writing, exam strategies and critical thinking
- discussions concerning how to interact with Australian peers
- activities to familiarise students with colloquial language used in the medical setting
- formative feedback on assignments to be submitted for summative assessment

In 1996, 33 (27.7%) of the first year class of 119 students chose to take advantage of the Supplementary tutorials. Their language backgrounds were:

Language	1996	Language	1996
Bahsa Malaysia	21 (63.6%)	Hindi	1 (3.0%)
Cantonese	2 (6.1%)	Mandarin	4 (12.1%)
English	3 (9.1%)	Vietnamese	1 (3.0%)
Farsi	1 (3.0%)	Total	33 (100%)

English Language Proficiency Screening

All enrolling students were screened for potential English language difficulties with the Australian Tertiary English Screening Test (AUSTEST), a twenty item instrument that has been validated with Australian undergraduate medical students.[1,2] The AUSTEST is composed of four subtests:

Subtest 1: Vocabulary (10 items), where an alternative word must be provided to the prompt word, which is presented in a sentence;
For example, another word for "antique" in the sentence "this is an antique chair".

Subtest 2: Dictation (3 items), where a sentence must be transcribed verbatim after having been read out only once;
For example, "The old house with red tiles was the one we once rented".

Subtest 3: Absurdities (4 items), where a nonsense sentence must be explained in terms of what does not make sense and why;
For example, "By a strange coincidence the mother and her child happened to share the same birthday, the 6th of June, 1953".

Subtest 4: Proverbs (3 items), where the meaning of the proverb must be given.
For example, "Strike while the iron is hot".

An AUSTEST score of 14 out of 20 or below can be taken as an indication that difficulties with the English language may be present, but this cut-off is elastic.[1,2] At the University of Adelaide a score of 15 was taken as a marker of language difficulties for the 1996 cohort.

Students who attended the Supplementary tutorial programme scored significantly lower on the AUSTEST than students who did not attend (M-W test, p<0.001):

	Attenders	Non Attenders
Number	33 (27.7%)	86 (72.3%)
Range	4 - 18	14 - 20
Mean	12.6	16.7
SD	3.4	2.2

Student Perceptions

At the end of the academic year all Supplementary tutorial programme attenders were asked to complete a questionnaire which gauged their perceptions of:

- how much difficulty, if any, they felt they had experienced in studying behavioural science during the year
- how useful they considered the programme had been in ameliorating any difficulties.

Responses (%) to the questionnaire items "How much difficulty did you have in terms of. . ."

Item	Much	Some	None
Reading & understanding the textbook?	7.4	77.8	14.8
Understanding the lecture material?	22.7	77.3	-
Taking lecture notes?	42.9	42.9	14.3
Writing practise examination essays?	26.1	65.2	8.7
Writing practical reports?	11.5	65.4	23.1
Contributing verbally in tutorials?	16.7	29.2	54.2
Overall difficulty?	7.7	88.5	3.8

Notes: All "unsure" responses have been coded as missing data.
Response rate: 29 (87.9%)

Responses (%) to the questionnaire items "How useful were the additional tutorials in helping you to . . ."

Item	Very	Some	Not
Take better lecture notes?	86.2	10.3	3.4
Understand the lecture material?	93.1	6.8	*
Write practise examination essays?	72.0	28.0	*
Write the practical reports?	50.0	46.4	3.6
Contribute verbally in tutorials?	36.4	50.0	13.6

Notes: All "unsure" responses have been coded as missing data.
Response rate: 29 (87.9%)

Typical comments from students who attended the Supplementary tutorials in response to open-ended questions regarding the usefulness of the programme included:

"(As a result of the Supplementary tutorial programme) my notes are clearer and more direct, especially in relating new knowledge with medicine."

"This programme is definitely useful. Not only are the examples given by the tutor clear, but she helps me to determine what are the important points (in the lecture)."

"You explain the meaning of words that I don't understand. I really need Supplementary tutorial"

Summative assessments for the behavioural science course were compared for students attending the Supplementary tutorial programme against those who did not. Attenders did not score higher than non-attenders on any summative measures, but for several assessments no significant differences were found in the performance of the two groups.

Conclusions

Non-English speaking background students' perceptions of a first year course in the behavioural sciences indicate that generally they found it to be moderately to very difficult. Their open-ended comments indicate that some of this difficulty can be attributed to their English language skills, and this is supported by the results of a standardised screening instrument designed to identify students with English language problems.

Students' perceptions of a teaching programme aiming to help them with their study of behavioural science indicate that they overwhelmingly found the assistance to be useful. Open-ended comments made by students were generally favourable. That many students were willing to commit up to 85 hours per year of their time to voluntary, non-assessable tutorials also attests to the value of the programme.

The Supplementary tutorials may have a number of "non-specific" benefits for non-English speaking background students. The importance of these for self-esteem and confidence should not be underestimated. They may include:

- formal acknowledgement of their special educational requirements
- the opportunity to voice opinions and ideas in a supportive environment
- an environment where the student is less likely to feel embarrassed if their language is not as that of a native speaker's
- the realisation that they are one of many students with similar difficulties
- an identified faculty staff member who can serve as a student mentor if needed
- the provision of concrete, specific explanations directly related to the course content to explain language-related or culturally-based problems with the course material

Attenders did not achieve higher grades than non-attenders on summative assessments. Whether these students might have performed significantly less well without the programme can only be speculation, although open ended comments suggest that it is the students' belief that they would be less likely to succeed with no additional support. Whether summative assessment methods are valid indicators of student learning is also a point for conjecture.

The evaluation of teaching interventions is not always simple. Outcome measures may not be easy to define in a valid or meaningful way. Ethical considerations may prohibit the use of control groups or the random allocation of students to classes. Nevertheless, the energy expended in critically reflecting upon pedagogical strategies is well spent where methods for improvement can be identified and subsequently put into practice. The Supplementary tutorial programme described in this poster is one such effort to reflect upon and improve medical education at the University of Adelaide.

References

1. Feinnil, D. & Hayes, S.C. (1996a) Screening Higher Education Students for English Language Problems: Development of the Australian Tertiary English Screening Test. *Higher Education Research and Development*, 15(1), 61-71.
2. Feinnil, D. & Hayes, S.C. (1996b) A User's Guide to the Australian Tertiary English Screening Test: AUSTEST. Behavioural Science in Medicine, University of Sydney, Glensdale Press.

Note: This poster is based upon a journal article submitted for publication to *Teaching and Learning in Medicine* in February 1997.

Appendix IV.I.

O'Hanlon, A., Winefield, H., Hejka, E. & Chur-Hansen, A. (1995) Initial responses of first-year medical students to problem-based learning in a behavioural science course: role of language background and course content.

Medical Education, **29**, 198-204.

O'Hanlon, A., Winefield, H., Hejka, E. & Chur-Hansen, A. (1995) Initial responses of first-year medical students to problem-based learning in a behavioural science course: role of language background and course content.
Medical Education, v. 29(3), pp. 198-204

NOTE:
This publication is included after page 371 in the print copy of the thesis held in the University of Adelaide Library.

Appendix IV.II.

Chur-Hansen, A. and Barrett, R.J. (1996)
Teaching colloquial Australian English to medical students from non-English speaking backgrounds.
Medical Education, **30**, 412-417.

Chur-Hansen, A. & Barrett, R.J. (1996) Teaching colloquial Australian English to medical students from non-English speaking backgrounds.
Medical Education, v. 30(6), pp. 412-417

NOTE:
This publication is included after page 372 in the print copy
of the thesis held in the University of Adelaide Library.

Appendix IV.III.

Barrett, R.J. & Chur-Hansen, A. (in press)

Getting off at Redfern: Gathering Colloquial "Aussie" English.

TESOL Journal.

TEACHERS OF ENGLISH TO SPEAKERS OF OTHER LANGUAGES, INC.

TESOL is an international professional organization whose mission is to strengthen the effective teaching and learning of English around the world while respecting individuals' language rights.

February 24, 1997

Rob Barrett
University of Adelaide
Department of Psychiatry
Royal Adelaide Hospital
Adelaide, South Australia
AUSTRALIA 5005

Dear Rob and Anna:

Chris Stryker, my co-editor finally had a chance to read your paper and loved it as much as I did. We are pleased to accept "Getting off at Redfern: Gathering Colloquial 'Aussie' English" for the Tips from the Classroom section of the *TESOL Journal.* We also have received the letter of permission from *Medical Education* . Thanks for taking care of that important detail. We will, of course, give them due acknowledgement.

I'm not able to give you a definite publication date for the article right now. At around the same time that we received your paper, we received a great paper for an ESL business class, and we're thinking about putting together a special issue on English for Specific Purposes. We also have a bit of a backlog of materials, so I should warn you that it could be six months to a year before your paper is published. Please let me know if that presents any problems. In the meantime, if you are able, would you please send a 3.5 inch disk, with the format and software clearly written on the label (e.g., IBM WordPerfect for Windows 6.1; or MAC MS Word 5) to: Marilyn Kupetz, TESOL Central Office, 1600 Cameron St., Suite 300, Alexandria, VA 22314. I also need a short bio for both of you -- no more than a line or two. You can add that to the disk and/or E-mail it to me.

Thanks so much for the work you put into condensing and rewriting the article to fit into our format. Teaching English in specific contexts, such as medicine, is an increasingly important area in TESOL, and we get so few submissions dealing with these issues. I'm pleased that you've given us such a great idea to publish.

Sincerely

Dorothy M. Taylor
Co-editor, "Tips from the Classroom"
19 Putnam St.
Buffalo, NY 14213
(716) 885-8831
E-mail: dmtaylor@buffnet.net

cc: Christian Faltis
Marilyn Kupetz

Appendix V.I.

Barrett, R. J. (1994)

Personal Communication (Letter).

Department of Psychiatry at the Royal Adelaide Hospital, University of Adelaide.

THE UNIVERSITY OF ADELAIDE

Department of Psychiatry

RJB:jkl

19 December, 1994

Ms Anna Chur-Hansen
Department of Psychiatry.

Dear Anna,

I am just reviewing the work of the English Language Development Committee for 1994, and I would like to personally thank you for the effort that you have put in this year. I think that it is crucial that when such an innovative programme is developed, it be researched and evaluated. I think we are indeed fortunate to have somebody who can apply such scientific rigour to this job. I look forward to working with you again next year.

Best wishes for the New Year.

Dr. R.J. Barrett
Convenor,
Language Development Committee

The University of Adelaide, South Australia 5005
Tel: (08) 303 4455 Fax: (08) 224 0464 Telex: UNIVAD AA 89141

Appendix V.II.

Farnill, D. & Hayes, S.C. (1996)

A User's Guide to the Australian Tertiary English Screening Test: AUSTEST.

Behavioural Sciences in Medicine, University Of Sydney, Glensdale Press.

Acknowledgement.

"We thank the Faculty of Medicine of the University of Sydney for supporting this work. We are also very grateful to Anna Chur-Hansen and other members of the University of Adelaide for providing information on reliability and validity based on their use of the *Austest* and to Sue Barrett our research assistant, who has been involved in this project from the beginning.

Appendix VI.

Required subjects for study, MBBS, University of Adelaide, 1992 - 1996.

1992

Anatomy I
Behavioural Science I
Biology I
Biomedical Statistics I
Chemistry I
Introductory Medicine I
Medical Physics I

Anatomy II
Biochemistry II
Genetics II
Human Physiology II
Community Medicine II

Clinical Science and Skills III
Human Physiology III
Microbiology and Immunology III
Pathology III
Pharmacology III
Social and Preventative Medicine III

Clinical Science IV
Clinical Skills IV
Psychiatry IV
Research Project

Clinical Science V
Clinical Skills V
Obstetrics and Gynaecology V
Paediatrics V

Applied Pathology and Forensic Medicine VI
Clinical Competence VI
Community Practice VI
Medicine and Surgery VI
Paediatrics VI
Psychiatry VI

1993

Anatomy I
Behavioural Science I
Biology I
Biomedical Statistics I
Chemistry I
Introductory Medicine I
Medical Physics I

Anatomy II
Biochemistry II
Genetics II
Human Physiology II
Community Medicine II

Clinical Science and Skills III
Human Physiology III
Microbiology and Immunology III
Pathology III
Pharmacology III
Social and Preventative Medicine III

Clinical Science IV
Clinical Skills IV
Psychiatry IV
Research Project

Clinical Science V
Clinical Skills V
Obstetrics and Gynaecology V
Paediatrics V

Applied Pathology and Forensic Medicine VI
Clinical Competence VI
Community Practice VI
Medicine VI
Surgery VI
Paediatrics VI
Psychiatry VI

1994

Cell and Molecular Biology I
Doctor, Patient and Society I
Human Structure and Function I
Introductory Medicine I

Anatomy II
Biochemistry II
Genetics II
Human Physiology II
Community Medicine II

Clinical Science and Skills III
Human Physiology III
Microbiology and Immunology III
Pathology III
Pharmacology III
Social and Preventative Medicine III

Clinical Science IV
Clinical Skills IV
Psychiatry IV
Research Project

Clinical Science V
Clinical Skills V
Obstetrics and Gynaecology V
Paediatrics V

Clinical Competence VI
Applied Pathology and Forensic Medicine VI
Medicine VI
Surgery VI
Community Practice VI
Paediatrics VI
Psychiatry VI

1995

Cell and Molecular Biology I
Doctor, Patient and Society I
Human Structure and Function I
Introductory Medicine I

Cell and Molecular Biology II
Doctor, Patient and Society II
Human Structure and Function II

Clinical Science and Skills III
Human Physiology III
Microbiology and Immunology III
Pathology III
Pharmacology III
Social and Preventative Medicine III

Clinical Science IV
Clinical Skills IV
Psychiatry IV
Research Project

Clinical Science V
Clinical Skills V
Obstetrics and Gynaecology V
Paediatrics V

Clinical Competence VI
Applied Pathology and Forensic Medicine VI
Medicine VI
Surgery VI
Community Practice VI
Paediatrics VI
Psychiatry VI

1996

Cell and Molecular Biology I
Doctor, Patient and Society I
Human Structure and Function I
Introductory Medicine I

Cell and Molecular Biology II
Doctor, Patient and Society II
Human Structure and Function II
Introductory Medicine II

Clinical Science and Skills III
Human Physiology III
Microbiology and Immunology III
Pathology III
Pharmacology III
Social and Preventative Medicine III

Clinical Science IV
Clinical Skills IV
Psychiatry IV
Research Project

Clinical Science V
Clinical Skills V
Obstetrics and Gynaecology V
Paediatrics V

Clinical Competence VI
Applied Pathology and Forensic Medicine VI
Medicine VI
Surgery VI
Community Practice VI
Paediatrics VI
Psychiatry VI

Appendix VII.I.

Language Development Committee's Written Language Skills Assessment Criteria.

WRITTEN LANGUAGE SKILLS

Test paper cover sheet

Student Number:

Marking Criteria

	UNSAT	SATIS	GOOD
1. TASK FULFILMENT (Has the question asked been satisfactorily answered?)			
2. COHESION AND ORGANISATION (Is the material being used to construct a well organised argument, are ideas developed and a conclusion reached?)			
3. SUBJECT SPECIFIC STYLE (Is the style and level of formality of language used appropriate for the task?)			
4. CONVENTIONS OF PRESENTATION (Is the answer clearly set out according to required criteria for layout, spelling and punctuation, referencing?)			
5. GRAMMATICAL CONTROL (Does the grammar aid or interfere with the meaning of what has been written?			

	UNSATISFACTORY	SATISFACTORY
6. Overall impression		

7. Recommendation

Appendix VII.II.

Language Development Committee's Written Language Assessment Feedback Sheet.

MEDICINE II: Written Language Assessment Feedback Sheet

STUDENT ID: ____________

DPS II ASSIGNMENT No 3
DATE. September 20. 1995

LLS Assessment No: GK________

severe problem<------->no problem

	5	4	3	2	1
A. Errors in grammar / vocab potentially interfere with meaning	❑	❑	❑	❑	❑

in. for example. the use of:

- verbs (tenses)
- prepositions (to..for..on..with..etc)
- definite / indefinite articles (the. a)
- incorrect (misleading) word choices

	5	4	3	2	1
B. Cohesion: logical development is hindered	❑	❑	❑	❑	❑

in. for example:

- omissions. or the use of
- pronouns (it..they etc)
- demonstratives (this..these..etc)
- conjunctions (although..since..who..etc)
- conjunctive adverbs (therefore..on the other hand..etc)
- unclear word choice ('lexical cohesion')

	5	4	3	2	1
C. Style is inappropriate for report	❑	❑	❑	❑	❑

Examples:

- too informal. conversational vocabulary
- lay expressions instead of medical ones
- faulty use of specialist terminology
- poor incorporation of source material / referencing

	5	4	3	2	1
D. Editing of report is inadequate	❑	❑	❑	❑	❑

Examples: minor errors in

- subject / verb agreement (e.g.final -s for 3rd person singular)
- past participles (e.g. final -d)
- spelling
- punctuation
- minor omissions

LANGUAGE ASSESSMENT:

SATISFACTORY	UNSATISFACTORY [submit DPS No 5 for language assessment]	RECOMMENDATION / COMMENTS
		❑ check language comments (copy attached) ❑ seek individual language development feedback ❑ attend vacation language sessions (Feb or April 96 - TBA)

Appendix VIII.I.

Language Development Committee's Oral Language Skills Assessment.

ORAL LANGUAGE SKILLS ASSESSMENT (OSCI)
DATE: November 21, 1995

Student Number:

ASSESSOR:

ASSESSMENT	severe problem				no problem
A. Ability to express self without being misunderstood	5	4	3	2	1

with particular reference to the ability to
- respond to patient's opening statements, move from point to point and to interview conclusion
- control vocab and grammar for successful communication

Where a rating of 4 or 5 is given, please comment briefly on the reason:

B. Ability to retrieve communication if patient misunderstands	5	4	3	2	1

in particular, the ability to
- recognise when patient has not understood
- re-phrase and communicate successfully
- use language which is appropriate for patient

Where a rating of 4 or 5 is given, please comment briefly on the reason:

C. Ability to clarify when unable to understand patient	5	4	3	2	1

in particular
- asking appropriate clarifying question(s), leading to
- success in understanding patient's meaning

Where a rating of 4 or 5 is given, please comment briefly on the reason:

D. Ability to speak intelligibly	5	4	3	2	1

in particular, does the student
- project voice appropriately
- modify enunciation / rhythm / intonation where needed

Where a rating of 4 or 5 is given, please comment briefly on the reason:

ORAL LANGUAGE ASSESSMENT:

UNSATISFACTORY | SATISFACTORY

UMcG ACUE 15/11/95
Med II Oral language assessment criteria

Appendix VIII.II.

Language Development Committee's Written Language Skills Assessment for the Language OSCI.

WRITTEN LANGUAGE SKILLS
WRITTEN ACCOUNT OF OSCI INTERVIEW
DATE: November 21, 1995

Student Number:

ASSESSOR:

A. OVERALL COMMUNICATION OF INFORMATION

severe problem <----------> no problem
5 4 3 2 1
❑ ❑ ❑ ❑ ❑

Assessment criterion: Does the account communicate clearly to another health care professional what took place?

Where a rating of 4 or 5 is given, please comment briefly on the reason:

B: ACCURACY IN THE CHOICE OF WORDS AND GRAMMATICAL CONSTRUCTIONS

severe problem <----------> no problem
5 4 3 2 1
❑ ❑ ❑ ❑ ❑

Assessment criterion: Is the health professional likely to misunderstand the events being recounted because of the following? (circle, highlight or make notes)

- incorrect (misleading) word choices
- unexplained specialist vocabulary
- omissions
- incorrect links (although..since..who..therefore..on the other hand..etc)
- any other (specify)...
- verbs (tenses)
- prepositions (to..for..on..with..etc)
- definite / indefinite articles (the, a)

C. PRESENTATION AND EDITING

severe problem <----------> no problem
5 4 3 2 1
❑ ❑ ❑ ❑ ❑

Assessment criterion: Are any of the following likely to cause difficulty or irritation to a reader of the account: legibility; lack of editing in punctuation, spelling, omissions (including omission of final -s or final -d)?

WRITTEN LANGUAGE ASSESSMENT:

UNSATISFACTORY | SATISFACTORY

Appendix IX.

University of Adelaide Student Enrolment Form.

PLEASE COMPLETE ALL SECTIONS AND NEATLY AMEND IN THE NON-SHADED AREAS ANY DETAILS WHICH ARE INCORRECT OR MISSING.

ENROLMENT FORM

PERSONAL DETAILS

STUDENT NUMBER | COURTESY TITLE: | CONTACT DEPT - CODE
SURNAME | - LOCATION
OTHER NAMES
SEX | OVERSEAS STUDENT PLEASE SPECIFY YOUR COUNTRY
DATE OF BIRTH
WORK PHONE | OVERSEAS TYPE (OFFICE USE ONLY)

SEMESTER ADDRESS (MARCH TO NOVEMBER)

POST CODE | SEMESTER PHONE

HOME ADDRESS (DECEMBER TO FEBRUARY)

POST CODE | HOME PHONE

STATISTICAL DETAILS

CONFIDENTIAL: The information you provide below is required by the Commonwealth Government. The University undertakes to provide the information to the Commonwealth Government as statistical data only and will not provide information on individual students.

Are you an Aboriginal or Torres Strait Islander? 2 No 1 Yes

Citizenship and residence status during this semester (tick one box):

1 Australian citizen
2 New Zealand citizen
2(a) If a N.Z. citizen, or Australian permanent resident, month & year of arrival ____ and, Month Year
2(b) Have you resided continuously in Australia since then
3 You (or a parent) have been granted permanent resident status and you are not an Australian or New Zealand citizen
4 You have a temporary resident visa, a visit visa or a student visa.
5 You are residing in Australia during this semester and are a diplomat or a dependant of a diplomat
6 You are residing overseas during this semester but are not an Australian or New Zealand citizen

In what country were you born? Australia / Overseas country ➡ Name of country

In what year did you first arrive in Australia? 19

Do you speak a language other than English at your permanent home residency?
01 No Yes ➡ Main non-English language spoken at your permanent home residence

Previous Studies: For each of the 8 items below, indicate whether you have either never commenced, commenced or completed the type of study.

answer all items (6) through (13):–

	Commenced/ completed	Last year of enrolment	Never commenced	Commenced but not completed	Completed	Last year of enrolment (if boxes 2 or 3 ticked)
7. Postgraduate course of any type			1	2	3	19
8. Bachelor degree course			1	2	3	19
9. Diploma or Associate Diploma course at an institution other than a TAFE college			1	2	3	19
10. Diploma or Associate Diploma course at a TAFE college			1	2	3	19
11. TAFE award course other than the above (do not count secondary education or hobby/recreation/leisure or personal enrichment courses)			1	2	3	19
12. Final year of secondary education at a high school, technical high school, secondary school or secondary college				1	2	19
13. Final year of secondary education at another institution				1	2	19
14. Some other qualification or certificate of attainment or competence which may be recognised for tertiary entrance				1	2	19

LIBRARY SERVICES

15. Do you give permission for your name to be disclosed to a reader who wishes urgently to consult a book that you have on loan from the Library?
YES NO

16. Do you give permission for your name and library barcode number to be passed to other University libraries in South Australia in the event that reciprocal borrowing privileges can be arranged?
YES NO

DISABILITIES QUESTION

18. Answering this question is optional. Please note this information will be treated with the utmost confidentiality. It is requested for the purpose of providing appropriate student support.

Do you have a disability that could affect your performance as a student? YES NO

Would you like to receive advice on support services, equipment and facilities which may assist you? YES NO

Please tick the following boxes if they apply to you.

RESTRICTED MOBILITY | HEARING IMPAIRMENT | VISION IMPAIRMENT
LEARNING | MEDICAL | OTHER

OFFICE USE ONLY
STATISTICAL DETAILS
PERSONAL DETAILS

EXTERNAL STUDENTS ONLY

17. Do you give permission for your name to be released to another external student?
YES NO

OFFICE USE ONLY
ADMISSION DETAILS

CATEGORY:

Appendix X.

Screening Test of Adolescent Language (STAL).

SCREENING TEST OF ADOLESCENT LANGUAGE

I. *VOCABULARY*

_____	1. Gigantic	The room is gigantic._____
_____	2. Kettle	The kettle is copper._____
_____	3. Unmarried	My teacher is unmarried._____
_____	4. Penalty	What is the penalty for breaking the rule?_____
_____	5. Duplicate	Can you duplicate this?_____
_____	6. Bright	The diamond was bright._____
_____	7. Plump	The child is plump._____
_____	8. Address	He made a political address._____
_____	9. Cluster	I saw a cluster of students._____
_____	10. Govern	She will govern._____
_____	11. Annoy	They annoy him._____
_____	12. Peaceful	It was a peaceful evening._____

TOTAL I_____

II. *AUDITORY MEMORY SPAN*

_____ 1. The fire drill that we had last week/turned out to be the real thing.

_____ 2. The school on the west side of town/has more new students than our own school.

_____ 3. Last night I went to a movie with my friend/at the theater that takes coupons.

TOTAL II_____

III. *LANGUAGE PROCESSING*

	what	why	
_____	☐	☐	1. The sun was shining so brightly last week on Tuesday that I had to wear my sunglasses in the movie theater.
_____	☐	☐	2. I went with my sister to the shoe store to buy a pair of combat boots to wear to the Junior Prom.
_____	☐	☐	3. After climbing up ten flights of stairs two steps at a time yesterday morning, the man finally reached the basement.
_____	☐	☐	4. The most recent set of identical twins born at the hospital was a girl and a boy.
_____	☐	☐	5. Last night after we adjusted the antenna, unplugged the television set, and changed the channel, the picture became much clearer.

TOTAL III_____

IV. *PROVERB EXPLANATION*

_____ 1. Practice makes perfect.

_____ 2. Actions speak louder than words.

_____ 3. Better late than never.

TOTAL IV_____

TOTAL TEST SCORE_____

Appendix XI.

STAL Interview, 1994.

STAL interview, 1994.

Student's name; _____

Interviewed by; _____

Date; / /94 Time;

1. What is your first language? _____

2. At what age did you begin learning English? _____

3. Have you formally studied English? (For example at school, or within a course such as ELICOS) yes no

   If yes, what was the course? _____

   In what medium of instruction was English taught? _____

4. On a rating scale of "very poor", "poor", "fair", "good" or "excellent", how would you rate your ability to use English for;

   - speaking? very poor poor fair good excellent
   - understanding spoken? very poor poor fair good excellent
   - reading? very poor poor fair good excellent
   - writing? very poor poor fair good excellent

5. Where were you born? _____

6. When did you arrive in Australia? Month _____ Year _____

7. Where did you complete your final year of secondary schooling?

   Adelaide Other (if so, where) _____

8. When did you complete the final year of secondary schooling?

   1993 Other (if so, when) _____

9. In which language did you study prior to coming to this university?

   English Other (if so, which) _____

10. Do you live with your family or friends or at college?

    _____

11. How long have you lived there? _____

12. What is the main language that you use here in Adelaide?

    English Other (if so, which) _____

13. What was the main language you used when living with your family?
(If you are living elsewhere whilst studying) _____

14. Do you have friends in Adelaide? yes no

15. If so, can you tell me about them? If not, can you elaborate as to why this is so, and how you feel about it.

- how many? _____
- how often do you see them?

  everyday once per week once per fortnight

  other _____
- are they "close" friends? yes no unsure

  Comments; _____

  _____
- do they speak English? yes no

  If yes, is their English; excellent good fair poor very poor
- when you are together, in what language do you mainly speak?

  English Other (specify) _____

  Comments; _____

  _____

16. Are any of these friends "Aussies"? yes no

17. How do you feel about getting to know your Australian classmates?

_____

_____

_____

18. Since you have been here (if a recent arrival), or in the last year (if not), have you had any difficulty with;

- understanding Australian speakers in general?

  a great deal somewhat not at all unsure

  Comment; _____

  _____
- making yourself understood?

  a great deal somewhat not at all unsure

Comment; _____

_____

- understanding Australian customs (that is, the way of life)?

  a great deal somewhat not at all unsure

  Comment; _____

  _____

19. Why did you choose to study Medicine?

_____

_____

_____

Thank you for answering these questions. Do you have anything that you would like to add about your language skills, or anything that you would like to ask us?

_____

_____

_____

_____

_____

Appendix XII.

Introductory talk about the STAL, delivered to 1994 first year medicine cohort.

ENGLISH EXAMINATION

INTRODUCTORY TALK

Delivered to 1994 first year medical students by Dr Rob Barrett, 25/2/94.

I want to explain the English Language Development Program to you this morning.

History of the development.

For some time now in this faculty, a strong emphasis has been placed on the importance of communication in medical practice.

We train doctors to go into the community and treat patients. A central element of accurate diagnosis and effective treatment is good communication between doctor and patient. If you cannot talk clearly with your patient, you cannot begin to function as a doctor.

We have become concerned in recent years that a number of students have been struggling through the pre-clinical years of the course then failing their clinical years through a lack of capacity to talk to their teachers, their fellow students, and then to the patients. These people are bright and intelligent, but they have insufficient command of English.

As a response to this we have developed a program to develop the English skills of these students.

You are privileged to be the first group at this medical school, and in fact we believe the first group in Australia, to benefit from this program.

We are fortunate to have Helen Mullins join the faculty. She is a specialist teacher in English and is experienced in teaching English to university students and professional people.

On Friday, there will be a test of English at 11 o'clock here in this lecture theatre.

It is compulsory. If you do not attend, you cannot proceed with your medical studies.

Those who, on the basis of this examination, do not show sufficient skills in English comprehension will join the English Language Development Program.

First, you will be interviewed by a member of the Faculty of Medicine. This interview will last between 5 and 10 minutes. We will ask some specific questions and also some general questions, in which you will be asked to talk in a more open-ended way.

The program itself will be for one and a half hours per week and will run over two years.

At the end of it there will be a written and oral examination in English. Students who do not pass this will not be able to continue on to third year.

I will be explaining the details of the examination on Friday but today let me just run through the format of the examination.

It is called the Screening Test of Adolescent Language, or STAL for short. It comprises 23 items which are divided into 4 subtests.

The test has been developed by speech experts who needed a test to screen high school students.

We use this particular test because we have evidence from Sydney University that those who have difficulties with it, subsequently run into problems in their clinical years.

There are a range of questions. Some of the questions are easy, others are difficult.

Subtest 1 examines vocabulary.

On Friday I will say a word and then I will use it in a sentence. You will write down another word that means the same thing, and which could fit into the sentence. You must only write one word.

I will give you an example of this today.

Carve.
Carve the turkey.
Write down a word for carve.

Acceptable answers are slice or cut.

Subtest 2 is a dictation exercise.

On Friday I am going to say a sentence and I want you to write it down exactly as I have said it. I will say it only once. You will have to listen carefully.

I will give you an example.

The new dentistry students walked quickly into the lecture theatre on Friday morning.

Subtest 3 checks how well you can recognise illogical sentences and explain why they don't make sense.

You might like to imagine now, that you are writing down the answer.

I will read you a sentence. I want you to listen and write down two things.
What does not make sense, and
Why it does not make sense.
You must listen carefully: I can read this only once.

Let me give you an example.

Yesterday morning I arrived at work very early because my car ran out of petrol and I could not find a service station to buy some more.

What does not make sense: Arriving early.
Why does it not make sense: Because running out of petrol makes you late.

Subtest 4 tests your understanding of idiomatic English by asking you to explain the meaning of a proverb.

Again, imagine that you are writing down the answer.

I am going to read out a proverb and I want you to write down what it means.

Let me give you an example.

Too many cooks spoil the broth.

An acceptable answer might be too many people involved in organising something can make it go wrong. Alternatively you could answer something like, sometimes it is better for one person to do something than a group of people.

CONCLUDING REMARKS

These were examples only.

The examination proper is at 11 am on Friday in this lecture theatre.
You must all be seated by 11.00 o'clock.
The doors will be closed at 11.05.
We cannot admit people who come late because it is an aural examination which requires silence.
Examination conditions will apply. That is, there will be no talking during the course of the exam. You cannot look at the answers written by other people.
Absolute silence will be required.
Anyone with hearing difficulties please sit near the front.
The examination is compulsory. Those who do not take it, cannot proceed with their medical course.
It is necessary to bring your student ID number with you. It is on your library card.
It is also necessary to bring a pen.

Appendix XIII.

Screening Test of Adolescent Language Answer Proforma.

EDUCATIONAL SKILLS SCREENING

STUDENT NO. _____

SURNAME _____ GIVEN NAMES _____

DATE OF BIRTH _____ AGE _____

SCREENING TEST

I WORDS

example _____

1. _____ 7. _____
2. _____ 8. _____
3. _____ 9. _____
4. _____ 10. _____
5. _____ 11. _____
6. _____ 12. _____

Total I _____

II DICTATION

1. _____

_____

2. _____

_____

3. _____

_____

Total II _____

III REASONS

1. What? _____

   Why? _____

2. What? _____

   Why? _____

3. What? _____

   Why? _____

4. What? _____

   Why? _____

5. What? _____

   Why? _____

Total III _____

IV EXPLANATIONS

1. _____

   _____

2. _____

   _____

3. _____

   _____

Total IV _____

TOTAL TEST SCORE _____

Appendix XIV.

Quotes from STAL Interview, 1994, Questions 17 and 18, with resultant coding categories.

Q17: "How do you feel about getting to know your Australian classmates?

Comments Coded by raters as "positive".

"I want to get to know them."

"I haven't actually gotten to close contact with them really. No, they're fine, friendly, really helpful. It will be interesting really, 'cause, well, ahh, since you have a different cultural background. It will be interesting."

"They are quite nice actually, they are very talkative, um, they are very approachable, umm, and basically we can start by saying which year, or where we are from, and there is no problem."

"Great. Learn about a new culture, their way of thinking, and feeling. So it's good. Quite easy [to make friends] at College, because everybody's an individual, so they're all open to friendships."

"I feel great. I get along with them very well."

"One main thing is that they are more open, speak loudly and their actions are big, compared to Malaysians, bit subdued. Sometimes [that makes it more difficult], but mostly I take it in my stride. I can cope with that."

Comments Coded by raters as "mixed".

"All of them are my coursemates. I don't see any major problem in mixing with them, 'cause they are quite friendly. I suppose it's just a matter if we want to start the conversation. They seem to be helpful."

"At first I find that it is hard for me to mix with them because they have the advantages in terms of language, can express their own feelings, but then, I think as we go along, we can make friends if we could start it and if we could just sometimes, I think if we have the abilities in terms of studies, we can be their friends. Yeah, it is possible. I think they just need some time for us to mix, to get to know us."

"Well, I like to know them, but since they tend to be among themself, and you know, I feel, like, awkward, just to go and say 'hi', and , and also because of the language barrier, I feel quite uneasy, you know, like awkward, just go and mix with them, and they speak fast, and sometimes I don't understand them, so yeah, and I sometimes shy too, yeah, I would like to make friends with them."

". . . I am happy because they can accept me, but sometimes I feel quite little and tiny to them because they are good, and especially in the language we are not that good, so we feel belittle ourselves. But try, because sometimes they are very helpful at times."

" I think the girls are quite approachable, but I'm not sure about the guys. But not all the girls are approachable. Sometimes try to approach them, but they ignore us."

Comments Coded by raters as "negative".

"At first I feel, very, how do you say it, inferior, because you see, most of the Aussies, they speak so fast, compared to me, I speak just word by word. At first I

feel that, very you know, inferior, but I try to, aahh, speak with them, just ignore that feeling. And then, I think, as time goes, it's getting better."

"I think I can get close to them, because some of them are very friendly, but sometimes I feel quite shy to approach them. So actually I don't know how to start the conversation."

"Sometimes quite difficult, because, like, aahh the, we don't really have that much in common, so you keep on talking the same topic over and over again, like every time you meet them. Yeah. And like 'cause sometimes it's quite hard because they have like, their own, their own clique, and it's quite hard to join them."

"Ok, but umm, the problem is maybe we have not many same interests, so after introduce myself, he or she introduce himself or herself, after chatting for a while then I have no other topic to talk about because different cultural understanding."

"Rather nervous, actually. In comparison with their English, mine is somewhat not so good. And they tend to speaks very fast. And sometime it's quite hard to get the actual meaning of what they're saying, when they're speaking so fast."

"Sometimes, umm, it's quite embarass, you know, it's sort of when I talk with them, or when I talk with some guy, they don't or he doesn't understand me, and they say 'pardon' and I say it again, and then he still say 'pardon', and then teasing, umm, and doesn't want to listen again, so I feel embarrassed and sometimes I feel humiliating. So, but my close friend, she's an Australian, she's very good, and because she's very patient, she try to understand me and, first time my expression not very good, so I change another expression again, so she try to understand me, I ask her which part I wrong, she will help me, 'You should say this, this , this.' 'Usually it's said in this way', or something. So, that means when I talk to my close friend I feel very comfortable, sorry comfortably. But I mean when I talk to other Australian, they very impatient, I feel very embarrassing. And sometimes when I have dinner or lunch or breakfast [at College] and I sit on same table and I'm just only Chinese, they all speak English and they use some slangs or something I just try to understand. I don't understand what they talking about, because maybe they talk very fast, sometimes they may use some slang, don't understand."

"Not that hard, but still can't really mix with them because I don't know their, I mean, everyday language. I mean slang and all that. I mean all their making jokes and all that."

Q18a: Since you have been here (if a recent arrival), or in the last year (if not), have you had any difficulty with understanding Australian speakers in general?"

"I haven't heard 'fair dinkum'. Instead of 'thank you' should be 'you're welcome, they say 'don't worries'."

"Maybe because of their slang and because they talk quite fast and it is quite hard for me to catch what they're trying to say."

"Yes, at first, but now I'm getting used to it, getting less."

"It's ok, but when you start using slangs, um, that's the time I have a bit of a problem, when they're talking in every day language. . . "

'I would say a great deal of understanding if I speak to someone from Adelaide, but from, I met a couple, I mean, from the country. I can understand most of what they say, but a few words of slang that I don't."

"Ummm, depends, if they sort of like use their Aussie accent, there's a few words that I couldn't understand."

"It's very difficult to understand the oldies, older people, they speak very quick, and then, then, they speak very quick, I don't know, they are very quick, hard to understand them."

"The old people. Maybe they talk quite fast."

Q18b: Since you have been here (if a recent arrival), or in the last year (if not), have you had any difficulty with making youself understood?

"Sometimes when we speak, it's like they couldn't hear us very well. It's not that they're deaf, but they tend to speak very slow, and it's not according with their accent. . . "

'Sometimes I do have to think about what I'm saying. 'Cause it's like converting from Vietnamese to English, yeah, so sometimes I do have problems conveying my ideas, but normally I don't."

"Not very often, but sometimes when I talk very fast, I tend to talk very fast sometimes, and couldn't get it, what I'm saying, I've got to repeat it. But that's not really a problem."

"I sometimes try to say something but people tend to understand about something else. Maybe I'm not expressing myself quite well. Maybe I don't have the vocabulary."

"It's quite hard to choose this rating. Because some of my Australian friends say that they can understand me, but I feel they don't understand me, they say 'hey?'".

". . . sometimes they can understand me easily, sometimes they cannot. Probably it is the idea, when it comes to expressing an idea, and it is very big, that seems to be a problem. Short ideas not a problem. . . . I had this difficulty of a presentation last year, in science project, supposed to read an article and summarise it in our own way, and that seemed to be a problem. I don't know what was the problem. . . "

"Some people can understand me, I speak clearly, they understand me. If we talk about fields that we learn at school then that's not very difficult for me, because I can do in English, but if they try to talk something else, about let's say, umm, I mean something my knowledge not very high, then I have difficulty."

"Sometimes, usually they can understand me, but just rare instances they can't understand me. . . Sometimes, like, they have to ask you again. I have to repeat."

Q18c: Since you have been here (if a recent arrival), or in the last year (if not), have you had any difficulty with understanding Australian customs (that is, the way of life)?"

"I mean, what they expect us to say and how, what they expect us to answer, when they ask us something. When we are a good friends of them, what they expect, what will they expect of us."

"Social circle . . . compared with my culture. In my culture - I'm not saying my culture is good - there is a boundaries between males and females, so, compared with Australia, where the people are very open-minded, so I just, I sometimes I can connect it with some other problems, like the collapsing of moral problems . . "

"It's the difference . . . yeah, they drink a lot actually, alcohol, and they talk when they group they don't usually talk about the work, but they talk about something

else, like sports, and I've been to quite a lot of parties, and that was one thing, I was quite surprised, the men in one corner and the women in the other. So yeah, and I found it quite interesting though, because everybody is friendly, you don't have to know anything about it."

"I'm not really used to your customs. Drinking. Haven't thought about it."

Appendix XV.

Word Knowledge Test (WKT) - Sample Questions.

DIRECTIONS This is a test of your knowledge of word meanings. Each word in capital letters is followed by five other words or phrases. You are to choose the word or phrase which has the same or most nearly the same meaning as the word in block letters, and mark your answer on the separate answer sheet. First do the practice examples.

PRACTICE EXAMPLES

P1 QUICK
- A soft
- B quiet
- C fast
- D early
- E sharp

P2 FALL
- A kill
- B sit
- C cry
- D drop
- E jump

P3 TALK
- A see
- B speak
- C smile
- D walk
- E take

P4 HIT
- A point
- B rise
- C strike
- D swing
- E lie

P5 POOR
- A not rich
- B foolish
- C dirty
- D old
- E wicked

P6 SHIP
- A water
- B boat
- C tree
- D stick
- E sail

P7 ABOVE
- A under
- B near
- C over
- D good
- E right

P8 PARCEL
- A string
- B post
- C book
- D package
- E store

P9 HAPPINESS
- A forgive
- B pipe
- C sadness
- D depression
- E joy

NOTE There are 72 questions to be answered in the test that follows.

You will have 10 minutes to do the test. Try to complete as many questions as possible. You may mark an answer even if you are not absolutely sure it is correct, but do not guess blindly.

Mark all your answers on the separate answer sheet. Do not make any marks in the test booklet.

DO NOT TURN THE PAGE UNTIL YOU ARE TOLD

Appendix XVI.

Medical Communications Skills Assessment Mark Sheet.

COMMUNICATIONS ASSESSMENT MARK SHEET

Student Name:
Student Number:
Time:

Roleplayer:
Examiners:
Date:

		POOR		ADEQUATE		EXCELLENT
I	THE TASKS					
	INTRODUCTIONS/explanation of purpose of interview/obtaining informed consent	1	2	3	4	5
	PATIENT'S UNDERSTANDING OF THE PROBLEMS (did the student adequately explore the patient's knowledge of the problem)	1	2	3	4	5
	PATIENT'S CONCERNS (did the student adequately explore the patient's concerns)	1	2	3	4	5
	SUMMARY/CHECKING/FINISHING (was the summary comprehensive of the patient's problems and concerns, did students check its accuracy and close appropriately)	1	2	3	4	5
II	LEVEL OF SKILLS					
	RAPPORT AND ATTENDING	1	2	3	4	5
	LISTENING (did the student really understand the patient and pick up appropriate cues)	1	2	3	4	5
	APPROPRIATE USE OF QUESTIONS (open, closed, clarification, checking & retrieval)	1	2	3	4	5
	EMPATHY SKILLS (responding, challenging, silence, tears)	1	2	3	4	5
	PERSONAL AND PROFESSIONAL QUALITIES (Did the student show warmth and concern, sincerity, respect and have a non judgemental attitude)	1	2	3	4	5
	USE OF CLEAR, UNAMBIGUOUS, FLUENT LANGUAGE	1	2	3	4	5
	USE OF APPROPRIATE LANGUAGE (to the person and context)	1	2	3	4	5
III	OVERALL FEELING OF EMPATHY (Did the student create an atmosphere of trust, identify the emotion expressed and give support to the patient by responding sensitively and appropriately)	1	2	3	4	5

FINAL ASSESSMENT

Satisfied Requirements ☐ Has not satisfied requirements ☐

Appendix XVII.I.

Clinical Skills Tutor Assessment Sheet, 1993 Third Year Students.

THE UNIVERSITY OF ADELAIDE : FACULTY OF MEDICINE

MEDICINE 3RD YEAR 1993

TUTOR ASSESSMENT - CLINICAL SKILLS
Royal Adelaide Hospital and The Queen Elizabeth Hospital

GRADED ASSESSMENT
The assessment should be an overall one based on the competence expected at third year level in history taking, physical examination and case presentation. It also includes attitude, knowledge, enthusiasm and attendance.

Marking Scale:	Distinction	75
	Credit	65 - 74
	Pass	50 - 64
	Fail	<50

If in your judgement a student's performance warrants a clear fail, please make this apparent with a mark of 40 or less.

STUDENT	
ASSESSMENT	
COMMENT:	

QUALITATIVE FEEDBACK
Please indicate below the level of English Language Proficiency of this student if it is applicable. Language development will be offered to those students requiring it.

NEEDS ATTENTION	BORDERLINE	ADEQUATE
COMMENT:		

Please complete this form and return it in the envelope provided to Dr J Vernon-Roberts, Clinical Studies Office, Level 7, Services & Teaching, Royal Adelaide Hospital.

Assessor: ..... Signature: .....

Ward: Date: .....

Thank you for making this valuable contribution

Appendix XVII.II.

Clinical Skills Tutor Assessment Sheet, 1993 Fourth Year Students.

THE UNIVERSITY OF ADELAIDE : FACULTY OF MEDICINE

MEDICINE 4TH YEAR : ASSESSMENT/ OBSERVED

CLERKSHIP 2- CLINICAL TERM ( ) - 1993

( ) ATTACHMENT ROYAL ADELAIDE HOSPITAL.

Please complete this student assessment form and return in the envelope provided to Dr Jane Vernon-Roberts, Clinical Studies Office, Level 7, Services and Teaching Block, Royal Adelaide Hospital as soon as possible after the end of the attachment.

ASSESSMENT should be an overall one, including ward involvement and knowledge of patients.

OBSERVED CASE Please include marks for history taking, physical examination, rapport with patient, synthesis of signs and symptoms and presentation of case.

Marking scale:	Distinction	75.
	Credit	65-74.
	Pass	50-64.
	Fail	<50.

If in your judgement a student's performance warrants a clear fail, make this apparent with a mark of 40 or less. A mark of 49 is borderline, and the student usually ends up with a pass mark. Because marks have to be scaled at the end of each year to create uniformity, individuals will not be given a precise mark until that time.

At the end of a six week attachment a student should be told whether they have passed of failed.

STUDENT	ASSESSMENT MARK	OBSERVED CASE
COMMENT		

ASSESSOR: SIGNATURE: .....

WARD: DATE: .....

Thank you for making this valuable contribution

JVR;ejmFeb93Stud Assess

Appendix XVIII.

Clinician comments on language and resultant coding categories (in parenthesis).

1. English skills are poor
(Specific problem not mentioned)

2. Needs more confidence and to improve language skills
(Specific problem not mentioned)

3. Needs much more practise with idiomatic colloquial English in conjunction with patient contact. I think the University needs to address this issue; lack of familiarity with colloquial English is preventing this student from getting adequate patient exposure
(Colloquial language/Patient difficulties)

4. Enthusiastic worker. Minor language difficulty
(Specific problem not mentioned)

5. Language will be a problem as tends to speak very fast when under pressure. Needs more clinical practice but is improving in knowledge and case presentation
(Rate of speech)

6. Attendance was v. good. Interested in daily ward activities. Needs to improve English especially when presenting cases
(Language hampering relay of knowledge)

7. Very keen and hard worker. His English ability holds him back but he's aware of this and making headway
(Specific problem not mentioned)

8. Keen student and regular attender. Fair history only but satisfactory observed examination. Discussion-pass only. Minor problems with English. Just needs to keep working
(Specific problem not mentioned)

9. Some lack of confidence possibly due to poor English. Seems to have good knowledge
(Language hampering relay of knowledge)

10. A reasonable performance. A little quiet and retiring. Language adequate but a little hard to understand at times. Very regular, punctual and attentive at tutorials and ward rounds
(Comprehensibility)

11. A conscientious student who is still having some language problems. Needs more talking experience
(Specific problem not mentioned)

12. Considerable improvement during the period (on the ward). Good potential particularly if English speech improves
(Specific problem not mentioned)

13. Spoken English is a problem. I have suggested elocution lessons
(Comprehensibility/Pronunciation)

14. This man's knowledge appears to be high but he has a problem with verbal expression despite having had his secondary education in Adelaide
(Language hampering relay of knowledge)

15. Very uncertain (as to what to comment). ? Difficulties with language and interpersonal skills. We suggest counselling
(Interpersonal skills/Specific problem not mentioned)

16. Good history; coped well with language barrier; sound examination; but interpretation somewhat lacking
(Specific problem not mentioned)

17. She was an attentive student who came regularly to tutorials. However she was quiet and retiring and said little. Her English is adequate but not perfect (she realises this). Her performance in clinical today was mediocre but I judged her to just pass.
I explained all the above to her afterwards urging her to be more forward in her clinical work and telling her that today's performance was only just adequate. On asking, she told me that she lives with Malaysians and speaks Malaysian 80% of the day. I told her to try and speak English all the time and to mix as much as possible with native English speaking people
(Personality factors/Language hampering relay of knowledge)

18. Main problem is poor English - means histories take x3 as long and makes his learning slow. Good technique with examination (physical). Conscientious
(Specific problem not mentioned)

19. Some language difficulties need attention
(Specific problem not mentioned)

20. Some language problems and reluctance to project knowledge; a very capable student
(Personality factors/Language hampering relay of knowledge)

21. Needs more practice with presentations, but English reasonably good
(Specific problem not mentioned)

22. Student has poor English communication skills and this is a problem with patients and clinical staff. Obviously intelligent - should she take English communication lessons? Needs encouragement
(Patient difficulties/Staff difficulties/Specific problem not mentioned)

23. The worst student ever. He attended irregularly and with no apology/explanation. When present he displayed communication difficulty, a marked lack of clinical skill/competence and an indifference to being corrected. In my opinion he does not have what it takes to graduate from Uni Adelaide Medical School and should be counselled accordingly. He disappeared whilst doing an observed case
(Specific problem not mentioned)

24. Shyness seems to be more of a problem than language but difficult to judge
(Personality factors)

25. Language borderline - adequate. Somewhat slow in delivery which makes it difficult conversing with patients. Obviously difficulty with vernacular, which will improve with exposure
(Colloquial language/Rate of speech/Patient difficulties)

26. Borderline language. Slow conversation. Difficulty with common phraseology
(Colloquial language/Rate of speech/Register difficulties)

27. Adequate language. Some minor difficulty
(Specific problem not mentioned)

28. The quietest member of the group. Says little even on direct questioning but seems to understand what is going on. Her language is adequate - I think it is
(Personality factors)

29. Language is only borderline - speech staccato and a bit abrupt. Will probably improve
(Rate of speech/Specific problem not mentioned)

30. Language sometimes a problem but should be ironed out with practice
(Specific problem not mentioned)

31. Some difficulties with language.
(Specific problem not mentioned)

32. His ability to take histories appears to be hampered by his command of English and possibly by cultural factors. I think he could overcome this with help.
(Cultural difficulties/Specific problem not mentioned/Patient difficulties)

33. Good performance, tends to be reserved. ??unsure of language vs medical terminology
(Personality factors/Register difficulties)

34. (He) has the potential to be very good but his "average" language skills are holding him back. His English is adequate for pass grade but for distinction needs work
(Specific problem not mentioned)

35. Early in year some difficulty either because of language or shyness but this improved greatly
(Personality factors/Specific problem not mentioned)

36. She appears to have a significant language and possibly cultural block to communication with patents which severely inhibits her clinical ability
(Patient difficulties/Cultural difficulties/Language hampering relay of knowledge)

37. Borderline (pass). Some minor problems communicating with patients compounded by a knowledge base clearly weaker than the other students. Spontaneous speech borderline/adequate. Sometimes problems with colloquial language
(Colloquial language/Patient difficulties/Conversation difficulties)

38. Borderline/adequate language. Some problems with some colloquial use of English
(Colloquial language/Specific problem not mentioned)

39. Major communication problems with language. Reasonable theoretical knowledge/examination technique
(Specific problem not mentioned)

40. Some problems with language
(Specific problem not mentioned)

41. Quiet, some trouble with English
(Personality factors/Specific problem not mentioned)

42. Has worked very hard to improve an initially very poor performance. Her language skills are improving
(Specific problem not mentioned)

43. Good basic knowledge and skills held back in some cases by language
(Specific problem not mentioned)

44. Poor communication skills. Poor attitude
(Specific problem not mentioned)

45. Major problems with English in patient communication
(Patient difficulties/Specific problem not mentioned)

46. Good progress, hard working, needs more English practice
(Specific problem not mentioned)

47. Some difficulty with English
(Specific problem not mentioned)

48. A quiet student, has reasonable knowledge <u>but</u> language background is a problem with presentations. Has improved
(Language hampering relay of knowledge)

49. English language quite good, but slow. A reasonable performance but rather lax in knowledge of patients. History, examination average in an easy case. Synthesis stumbling
(Rate of speech)

50. Communication skills need a bit of work
(Specific problem not mentioned)

51. Taking the language barrier into consideration he has performed well and should continue to do so
(Specific problem not mentioned)

52. Keen, persistent, punctual, attentive student. Her English is satisfactory but not ideal. I suggested she mix more with native Australians to improve fluency. Her general attitude and endeavour are good
(Specific problem not mentioned)

53. Communication skills need a bit of work
(Specific problem not mentioned)

54. Communication skills need a bit of work
(Specific problem not mentioned)

55. Good knowledge and enthusiastic commitment to ward activities. Some minor difficulties with nuances of English language
(Specific problem not mentioned)

56. Struggling with her conversational English. Knowledge better than she can express. A very keen student. She attends all clinic activities and tries very hard
(Language hampering relay of knowledge/Conversation difficulties)

Appendix XIX.

Language Rating Scale, 1994 pilot version.

LANGUAGE SKILLS. RATING SHEET.

Affix label here.

ROLEPLAYER: .....

TIME: ..... DATE: .....

	Poor		Adequate		Excellent
Use of correct tense	1	2	3	4	5
Use of appropriate register	1	2	3	4	5
Comprehensibility of speech due to accent	1	2	3	4	5
Appropriate rate of speech	1	2	3	4	5
Appropriate use of non-verbal communication	1	2	3	4	5
Response to requests, apologies, thanks	1	2	3	4	5
Understanding of informal language	1	2	3	4	5
Clarification where comprehension lacking	1	2	3	4	5
Fluency of speech	1	2	3	4	5
Overall impression of language proficiency	1	2	3	4	5

Comments:

Appendix XX.

Standardized patient script, 1994.

ENGLISH LANGUAGE PROFICIENCY TESTING - MED 3 PILOT PROJECT

INDIGESTION - CHRONIC UPPER ABDOMINAL PAIN

Play yourself or someone else you know well.

USE THIS MEDICAL SCRIPT

Patient Profile: Age, Marital Status, Occupation - You Choose

Presenting complaint and history of presenting complaint:

You have suffered with **PAIN** for 2 months.
The pain is in the upper abdomen and occurs several times a day. You think it is worse when you are hungry.

The pain comes on gradually and is constant in nature. It doesn't go away until you take "Quickeze". It is a discomfort to begin with and then becomes a gnawing burning sensation.

You describe it as indigestion and it is usually mild and does not bother you unduly. However, more recently it has begun to wake you at night and you have been taking lots of Mylanta and milk to relieve the symptoms.

On two occasions the pain has been severe enough to ring the doctor and that worries you. You think it may be an ulcer and you have had a friend with one that burst and necessitated serious surgery. You would like to avoid surgery.

Past Medical History:	Occasional diarrhoea, doesn't necessitate a trip to the Doctor.
Family History:	None of significance
Medication:	Panadol for headache
Lifestyle factors:	Cigarettes - 10 per day for years Alcohol - Wine with food for years Diet - Eats everything Exercise - None because not interested
Allergies:	None

YOU ARE LISTENING TO LANGUAGE ONLY

Student Instructions:
Hello, I have a problem with pain here (put hand on upper abdomen). I think I may have an ulcer.

Ensure use of the following informal terms:

1. Take a sickie
2. Feeling crook
3. Feeling under the weather

PTO

INTERVIEW FORMAT

The maximum time allowed is 10 minutes.

Students may leave sooner if the examiner decides that the interview is complete.

A BELL WILL BE RUNG AT THE TEN MINUTE MARK

Students will be shown into the interview room.

The "patient" will be seated inside that room.

The students will then proceed to establish and discuss the presenting problem.

NO MEDICAL HISTORY TAKING IS REQUIRED

There will be no discussion between the student and the examiner once the examination is completed.

The student is then required to record, in 5 minutes, a written account of the oral interview.

THE STUDENT WILL LEAVE THE EXAMINATION WHEN BOTH ORAL AND WRITTEN TESTS ARE COMPLETED.

Appendix XXI.

Student Information Sheet regarding English language proficiency testing, 1994.

THE UNIVERSITY OF ADELAIDE

Faculty of Medicine
Clinical Studies Office, Level 7, Services & Teaching, RAH

9th May 1994

ENGLISH LANGUAGE PROFICIENCY TESTING
MED 3 PILOT PROJECT 1994

The Faculty of Medicine has found that many students who do not have proficiency in the use of English language experience serious difficulties in communicating with patients when undertaking the hospital attachments in the third and subsequent years of the medical course.

Furthermore, the Faculty is proposing to introduce a major clinical communication skills assessment, including English language proficiency, at the end of the second year and a pass in this will be necessary for a student to be able to proceed to the third year of the medical course.

To ensure fairness and the validity of the instrument chosen to assess language proficiency at the end of second year, the language skills results of the third year medical students will be compared with their communication skills results at the end of Semester 1 and those of Clinical Science and Skills at the end of Semester 2. If our instrument is valid we would expect a clear correlation between language and clinical skills.

ATTENDANCE WILL BE COMPULSORY AT EACH OF THE FOLLOWING EVENTS

1. STAL TEST (English Language Proficiency)
   TUESDAY May 17th, 2pm in the ROBSON THEATRE, RAH

2. LANGUAGE SKILLS TEST (10 Minute ORAL, 5 Minute WRITTEN)
   TUESDAY June 21st, following the Communication Skills Examination

The results of both these tests will be confidential and available only to the Dean of the Faculty and the members of the Language Development Committee. Improvement strategies will be discussed with those students who fail to meet the benchmark standard.

INFORMATION ON THE STAL TEST WILL BE FOUND OVERLEAF

Mr Peter Devitt
Co-ordinating Examination
Third Year

Professor Derek Frewin
Dean
Faculty of Medicine

Information on the STAL test of English Language Proficiency

The English language proficiency test that you will take is called the STAL, which is short for Screening Test for Adolescent Language. It is a test that originated in the USA, and has been modified by the Medical Faculty at the University of Sydney for use in Australia.

The STAL comprises 23 items which are divided into four subtests:

Test 1 examines vocabulary. You will hear a word and will need to think of a synonym for that word. There are 12 different words, all in everyday use.

Test 2 is a dictation exercise. You will hear 3 sentences and write them down.

Test 3 checks how well you can recognise illogical sentences and explain why they don't make sense.

Test 4 deals with the understanding and explanation of idiomatic English.

There is nothing you can do to prepare for the test, as it is testing the language skills that you have developed over your lifetime.

The STAL will be given to you in a group. The test takes about 15 minutes in total. The tester will stand at the front of the class, and will read the items to you. You will be given an answer sheet on which to write. You need only to bring your own pen. If you have any hearing impairment you should sit at the front of the class, as the test depends upon you being able to hear clear what is read.

Appendix XXII.

Introductory talk about the STAL, delivered to the 1994 third year medicine cohort.

ENGLISH EXAMINATION

INTRODUCTORY TALK

Delivered to 1994 third year medical students by Dr Rob Barrett, 17/5/94.

Thank you for your compulsory attendance.

For some time now in this faculty, a strong emphasis has been placed on the importance of communication in medical practice.

We train doctors to go into the community and treat patients. A central element of accurate diagnosis and effective treatment is good communication between doctor and patient. If you cannot talk clearly with your patient, you cannot begin to function as a doctor.

We have become concerned in recent years that a number of students have been struggling through the pre-clinical years of the course then failing their clinical years through a lack of capacity to talk to their teachers, their fellow students, and then to the patients. These people are bright and intelligent, but they have insufficient command of English.

On June 21st we are introducing an English examination.

Today we are having a trial run of a screening test of English. This test is designed to identify people who may have problems with the English language. Ultimately we will be using it for students when they first come in to medicine, to pick up any problems they may be having right from the beginning.

We will be doing a statistical comparison of your results today with your results on June 21st.

It is important to say that your result today will not be used for your formal assessment.

I should also say that attendance at the June 21st examination is compulsory, but because this is the first time that we have conducted the examination, no-one will fail.

If, as a consequence of the examination, we identify any people with language problems, we will let you know and offer appropriate language development.

Appendix XXIII.

Proforma for students' written account of the interview with a standardized patient, 1994 third year medicine cohort.

JUNE 21ST 1994

MED 3

Affix label here

You have 5 minutes to record the interview you have just completed. Please write an account of this interview so that another health care professional would understand what took place and the decisions and conclusions drawn by yourself and the patient. Do not use medical jargon.

Appendix XXIV.

Language Rating Scale Pilot Study (Chapter VII).
Individual standardized patient's ratings.

Ratings of Standardized Patient I.

n=25	1	2	3	4	5	missing	$\bar{x}$	sd
tense	-	-	1 (4.0)	2 (8.0)	22 (88.0)	-	4.84	0.47
register	-	1 (4.0)	-	1 (4.0)	23 (92.0)	-	4.84	0.62
accent	-	1 (4.0)	3 (12.0)	2 (8.0)	19 (76.0)	-	4.56	0.87
rate	-	-	1 (4.0)	5 (20.0)	19 (76.0)	-	4.72	0.54
nvc	-	3 (12.0)	2 (8.0)	6 (24.0)	14 (56.0)	-	4.24	1.05
speech acts	-	-	1 (4.0)	5 (20.0)	19 (76.0)	-	4.72	0.54
informal	1 (4.0)	3 (12.0)	6 (24.0)	2 (8.0)	13 (52.0)	-	3.92	1.29
clarification	1 (4.0)	3 (12.0)	1 (4.0)	1 (4.0)	18 (72.0)	1 (4.0)	4.33	1.27
fluency	-	-	3 (12.0)	4 (16.0)	18 (72.0)	-	4.60	0.71
overall	-	1 (4.0)	3 (12.0)	7 (28.0)	14 (56.0)	-	4.36	0.86

Ratings of Standardized Patient II.

n=24	1	2	3	4	5	missing	$\bar{x}$	sd
tense	-	-	3 (12.5)	5 (20.8)	16 (66.7)	-	4.54	0.72
register	-	2 (8.3)	2 (8.3)	11 (45.8)	9 (37.5)	-	4.13	0.90
accent	-	-	1 (4.2)	5 (20.8)	18 (75.0)	-	4.71	0.55
rate	-	1 (4.2)	7 (29.2)	4 (16.7)	12 (50.0)	-	4.13	0.99
nvc	-	-	12 (50.0)	-	12 (50.0)	-	4.00	1.02
speech acts	-	-	3 (12.5)	6 (25.0)	13 (54.2)	2 (8.3)	4.46	0.74
informal	-	-	1 (4.2)	7 (29.2)	16 (66.7)	-	4.63	0.58
clarification	-	1 (4.2)	1 (4.2)	5 (20.8)	13 (54.2)	4 (16.7)	4.50	0.83
fluency	-	-	5 (20.8)	6 (25.0)	13 (54.2)	-	4.33	0.82
overall	-	-	4 (16.7)	9 (37.5)	11 (45.8)	-	4.29	0.75

Ratings of Standardized Patient III.

n=25	1	2	3	4	5	missing	$\bar{x}$	sd
tense	1 (4.0)	3 (12.0)	2 (8.0)	3 (12.0)	16 (64.0)	-	4.20	1.26
register	-	1 (4.0)	7 (28.0)	11 (44.0)	6 (24.0)	-	3.88	0.83
accent	-	1 (4.0)	6 (24.0)	5 (20.0)	-	13 (52.0)	3.33	0.65
rate	-	2 (8.0)	6 (24.0)	3 (12.0)	13 (52.0)	1 (4.0)	4.13	1.08
nvc	-	3 (12.0)	3 (12.0)	14 (56.0)	5 (20.0)	-	3.84	0.90
speech acts	-	-	2 (8.0)	11 (44.0)	12 (48.0)	-	4.40	0.65
informal	2 (8.0)	5 (20.0)	4 (16.0)	5 (20.0)	9 (36.0)	-	3.56	1.39
clarification	-	4 (16.0)	2 (8.0)	6 (24.0)	3 (12.0)	10 (40.0)	3.53	1.13
fluency	-	1 (4.0)	7 (28.0)	3 (12.0)	14 (56.0)	-	4.20	1.00
overall	-	1 (4.0)	7 (28.0)	6 (24.0)	11 (44.0)	-	4.08	0.95

Ratings of Standardized Patient IV.

n=25	1	2	3	4	5	missing	$\bar{x}$	sd
tense	-	2 (8.0)	4 (16.0)	4 (16.0)	15 (60.0)	-	4.28	1.02
register	-	1 (4.0)	4 (16.0)	6 (24.0)	14 (56.0)	-	4.32	0.90
accent	-	1 (4.0)	4 (16.0)	3 (12.0)	17 (68.0)	-	4.44	0.92
rate	-	-	4 (16.0)	5 (20.0)	16 (64.0)	-	4.48	0.77
nvc	-	-	3 (12.0)	6 (24.0)	15 (60.0)	1 (4.0)	4.5	0.72
speech acts	-	-	2 (8.0)	3 (12.0)	15 (60.0)	5 (20.0)	4.65	0.67
informal	-	2 (8.0)	5 (20.0)	5 (20.0)	11 (44.0)	2 (8.0)	4.09	1.04
clarification	-	-	5 (20.0)	-	13 (52.0)	7 (28.0)	4.44	0.92
fluency	-	3 (12.0)	4 (16.0)	5 (20.0)	13 (52.0)	-	4.12	1.09
overall	-	-	5 (20.0)	6 (24.0)	14 (56.0)	-	4.36	0.81

Ratings of Standardized Patient V.

n=25	1	2	3	4	5	missing	$\bar{x}$	sd
tense	-	-	7 (28.0)	14 (56.0)	4 (16.0)	-	3.88	0.67
register	-	-	9 (36.0)	12 (48.0)	4 (16.0)	-	3.80	0.71
accent	-	-	7 (28.0)	14 (56.0)	4 (16.0)	-	3.88	0.67
rate	-	-	7 (28.0)	14 (56.0)	4 (16.0)	-	3.88	0.67
nvc	-	-	9 (36.0)	12 (48.0)	4 (16.0)	-	3.80	0.71
speech acts	-	-	8 (32.0)	14 (56.0)	2 (8.0)	1 (4.0)	3.75	0.61
informal	-	-	13 (52.0)	7 (28.0)	3 (12.0)	2 (8.0)	3.57	0.73
clarification	-	-	2 (8.0)	9 (36.0)	2 (8.0)	12 (48.0)	4.00	0.58
fluency	-	-	11 (44.0)	10 (40.0)	4 (16.0)	-	3.72	0.74
overall	-	-	10 (40.0)	11 (44.0)	4 (16.0)	-	3.76	0.72

Ratings of Standardized Patient VI.

n=25	1	2	3	4	5	missing	$\bar{x}$	sd
tense	1 (4.0)	2 (8.0)	6 (24.0)	4 (16.0)	12 (48.0)	-	3.96	1.21
register	1 (4.0)	2 (8.0)	7 (28.0)	8 (32.0)	7 (28.0)	-	3.71	1.10
accent	3 (12.0)	1 (4.0)	8 (32.0)	1 (4.0)	12 (48.0)	-	3.72	1.43
rate	1 (4.0)	4 (16.0)	6 (24.0)	3 (12.0)	11 (44.0)	-	3.76	1.30
nvc	-	1 (4.0)	7 (28.0)	8 (32.0)	7 (28.0)	2 (8.0)	3.91	0.90
speech acts	-	-	2 (8.0)	12 (48.0)	10 (40.0)	1 (4.0)	4.33	0.64
informal	2 (8.0)	8 (32.0)	6 (24.0)	-	9 (36.0)	-	3.24	1.45
clarification	1 (4.0)	3 (12.0)	2 (8.0)	8 (32.0)	5 (20.0)	6 (24.0)	3.68	1.20
fluency	1 (4.0)	4 (16.0)	7 (28.0)	3 (12.0)	10 (40.0)	-	3.68	1.28
overall	1 (4.0)	2 (8.0)	8 (32.0)	3 (12.0)	11 (44.0)	-	3.84	1.21

Appendix XXV.

Open-ended comments on the Language Rating Scale 1994 third year cohort pilot study. and resultant coding categories [in parenthesis].

Key to coding categories.
1=difficulty arriving upon a rating
2.=difficulty keeping the student talking
3=nervousness/anxiety of the student
4=failure of student to clarify
5=non-verbal communication
6=use of jargon

Standardized patient I.

1. Excellent.
2. V good.
3. He sounded incomprehensible at first, but with a bit of time, he was ok. A lot of room for improvement however.
4. Simply excellent.
5. I would give her in fact a 3.5 (for overall proficiency) not a 3. [1]
6. She was perfectly spoken although she clearly did not understand colloquial terms like "taking a sickie", didn't understand "Quickeze" and at times, failed to ask for clarification. [4]
7. More like a 3.5 (for overall proficiency). [1]
8. Only problem with her speech patterns were the use of jargon (medical jargon) such as stools, abdominal etc. [6]
9. Not worth a 5 (for overall proficiency) but close to it i.e. 4.5. [1]

Standardized patient II.

1. Slow in responding. Inappropriate non-verbals - looking at my chest and not at me. No jargon. Not really with it - i.e. not really interested. [5]
2. Problems. Jargonistic and sat inappropriately with legs crossed and swinging and hands locked over knees. Strange! [5, 6]
3. Seemed to be slow in her thinking. Didn't summarise but seemed to know what was going on. [4]
4. Not very efficient, far too brief. Disinterested in asking too many questions. [2]
5. Too much jargon. [6]
6. Very good. Inappropriate introduction, but otherwise excellent.
7. Very good, excellent empathy and non-verbals.
   Not at all phased by this task. [5]
8. Very good. Good non-verbals - sat closely. [5]
9. Got off the track a bit - had to be brought back to the presenting complaint.
10. Excellent.
11. A bit slow in her thinking. Brief interview. [2]

12. Very good really - used some jargon and moved away from me at one stage but corrected himself quickly. Got caught in the middle of the interview but picked up again well. [5, 6]

13. Excellent - a bit nervous though. [3]

14. Bit hesitant as unsure of task but then relaxed and got on with it. [3]

15. Very good. Nervous, so tended to stare at me a bit,
but otherwise, competent. [3]

16. Very impressive - good summary.

17. Excellent!

18. Jargonistic at the end but otherwise good.
Inappropriate crossing of legs. [5, 6]

19. Spoke very quickly and hard to understand at times because of accent. Moved chair away from me at beginning of interview. [5]

20. Very good empathic and non-verbal skills. [5]

21. Not really aware of his role in this role-play.

22. Not very fluent - worried, anxious - her hair kept falling in her face and mouth! [3]

23. Hesitant - too many times. I think this student had a problem with this role play - like trouble believing and doing it.

Standardized patient III.

1. Not really sure about this one. Pleasant, good manner. [1]

2. Student reasonably comfortable with speech, a few tense mishaps and was still not sure what "Quickeze" were. Seemed to understand "take a sickie". Didn't ask about medication, lifestyle factors etc.

3. Tended to avoid eye contact somewhat. [5]

4. No problems here.

5. Very pleasant manner.

6. Very difficult to get to speak. Somewhat clipped conversation, didn't understand "sickie" or "under the weather". No real non-verbals, sat with some avoiding of eye contact. [2, 5]

7. Pleasant manner, good attitude. Didn't know what "Quickeze" was.

8. Avoided some eye contact. Not sure about "Mylanta". [5]

9. Reticent, didn't talk a lot, had to be encouraged. Reasonable language skills. Needs to overcome some shyness. [2]

10. Good manner, relaxed. Didn't ask about medication, lifestyle factors etc.

11. Didn't know what "under the weather" meant. No problems here except that. Good manner.

12. Didn't understand "under the weather". "Mylanta" - reasonable grasp.

13. Uses some professional jargon but generally fine. [6]

14. Speaks a little fast and clips words occasionally. Understood colloquialisms.

15. Used medical jargon and has a habit of muttering outloud (obviously going through the list of what to ask the patient). [6]

16. Another amateur counsellor! A bit flippant. (Not too good at history taking. Some jargon but language fine. No introduction, thanks etc.). [6]

17. This student thought I was supposed to interview him! Had no idea what Mylanta was. Manner was good.

18. Very casual (glad I wasn't really sick). Tended to avoid eye contact, leaning back in chair, used medical jargon. [5, 6]

19. Didn't know what "under the weather" meant. Didn't take a really good history, would have liked him to talk more. [2]

20. No real problems except didn't know what Quickeze or "under the weather" meant. Some jargon. [6]

21. Very uptight, not sure why he was here. Interviewing skills poor due to this anxiety but language fine. [3]

22. This one needed to talk more but couldn't get him to. Reasonable. [2]

23. A little abrupt but fair.

Standardized patient IV.

1. He would ask, but it was clear he still didn't understand. He wouldn't ask for further clarification. [4]

2. Obviously nervous but explained things like when I said I'd had a sickie, she said, "so it's caused you to take time off work". [3]

3. I think because he was nervous that he asked me the same questions over and over again. [3]

4. Lovely smile, encouraged you to open up. [5]

5. Very quiet. Didn't really ask what it meant, but did repeat "under the weather". Too many pauses thinking about what to say. [2]

6. He gave no indication that he didn't understand anything. Was quite good at asking questions.

7. Excellent.

8. Excellent, despite mild accent explained things very well. Thought I said "unwell" not "under the weather".

9. Excellent.

10. Asked about "under the weather".

11. A bit quiet. A feel that she needed more clarification but was too shy. [2, 4]

12. Used too much jargon. [6]

13. Asked when I used "under the weather" for the second time.

14. He didn't really say much, but he did clarify some of my expressions.[2]

15. She made me feel as though I should talk because she didn't say much. [2]

16. Excellent.

17. Used a bit of jargon. [6]
18. She seemed disinterested. Sat back, spoke a bit fast.
19. I am not sure if she understood everything but didn't give it away if she didn't understand. Had to ask her to repeat a couple of words due to her accent e.g. vomit. [4]
20. Used the word "stool". [6]
21. She repeated a few things I said to establish better comprehension.
22. Seemed very interested.

Standardized patient V.

1. No opportunity to check clarification, understanding of informal language (my fault). [4]
2. No opportunity to assess understanding of informal language. My mistake. [4]
3. Did not observe clarification where comprehension lacking. [4]
4. Used radiate (jargon), stools. [6]
5. Used "radiated" (jargon) once but rest of the time clear. Clarified what I meant by "sickie". [6]
6. Jargon - localized, epigastric. [6]

Standardized patient VI.

1. Excellent communication skills.
2. Did have a noticeable accent.
3. Native speaker but not at all fluent in speech - possibly due to nerves? [3]
4. Did not understand informal terms "under the weather" but asked for clarification.
5. Did seem to sit back a little but generally seemed concerned. [5]
6. Some words were used in the incorrect tense.
7. I am unsure whether he understood Quickeze.
8. Did not understand about Quickeze but did not clarify its meaning. [4]
9. I think this student did understand Quickeze and Mylanta but clarified.
10. His eyes darted back and forth quickly, which was a little off-putting. He said to me afterwards that he did not understand "Mylanta" but he did not clarify this during the interview. [4, 5]
11. Used quite a few medical jargon words. [6]
12. Some words were in the incorrect place in speech.

Appendix XXVI.

Spearman *rho* correlation coefficients, Language Rating Scale Pilot Study (Chapter VII).

Standardized Patient I.

register	NS								
accent	0.73	NS							
rate	0.68	NS	0.82						
nvc	0.51	NS	0.64	0.71					
speech acts	0.62	NS	0.58	0.79	0.73				
informal	0.42	NS	0.67	0.61	0.61	0.56			
clarify	0.33	NS	NS	NS	NS	NS	0.60		
fluency	0.71	NS	0.77	0.92	0.76	0.73	0.66	0.49	
overall	0.61	NS	0.77	0.75	0.76	0.59	0.85	0.64	0.80
	tense	**register**	**accent**	**rate**	**nvc**	**speech a**	**informal**	**clarify**	**fluency**

All values significant at $p<0.05$.

Standardized Patient II.

register	0.46								
accent	0.79	0.41							
rate	0.75	0.44	0.59						
nvc	0.69	0.54	0.57	0.94					
speech acts	0.85	0.58	0.75	0.84	0.88				
informal	0.76	NS	0.75	0.49	0.47	0.62			
clarify	0.50	NS	NS	0.57	0.53	0.45	0.60		
fluency	0.79	0.49	0.63	0.93	0.89	0.77	0.57	0.73	
overall	0.79	0.46	0.62	0.90	0.89	0.85	0.73	0.77	0.89
	tense	**register**	**accent**	**rate**	**nvc**	**speech a**	**informal**	**clarify**	**fluency**

All values significant at $p<0.05$.

Standardized Patient III.

register	0.71								
accent	0.69	NS							
rate	0.90	0.64	0.75						
nvc	NS	NS	NS	0.50					
speech acts	0.42	0.51	NS	0.54	0.55				
informal	0.66	0.61	0.85	0.71	NS	0.60			
clarify	0.66	NS	NS	0.54	NS	NS	0.57		
fluency	0.86	0.69	0.86	0.88	0.42	0.64	0.83	0.64	
overall	0.80	0.71	0.90	0.85	0.52	0.71	0.86	0.60	0.92
	tense	**register**	**accent**	**rate**	**nvc**	**speech a**	**informal**	**clarify**	**fluency**

All values significant at $p<0.05$.

Standardized Patient IV.

register	0.71								
accent	0.91	0.68							
rate	0.83	0.63	0.92						
nvc	NS	0.62	NS	0.51					
speech acts	0.65	0.60	0.54	NS	NS				
informal	0.80	0.77	0.71	0.73	0.75	0.73			
clarify	0.63	0.72	0.58	0.58	0.79	0.61	0.74		
fluency	0.69	0.86	0.65	0.55	0.46	0.52	0.72	0.79	
overall	0.85	0.93	0.87	0.79	0.56	0.63	0.80	0.72	0.88
	tense	**register**	**accent**	**rate**	**nvc**	**speech a**	**informal**	**clarify**	**fluency**

All values significant at $p<0.05$.

Standardized Patient V.

register	0.60								
accent	0.89	0.70							
rate	0.89	0.60	0.89						
nvc	0.70	0.60	0.60	0.70					
speech acts	0.45	0.53	0.52	0.58	0.82				
informal	0.61	0.47	0.63	0.61	0.63	0.49			
clarify	0.75	0.54	0.75	0.75	0.77	0.82	0.66		
fluency	0.83	0.71	0.83	0.83	0.81	0.72	0.62	0.59	
overall	0.87	0.75	0.87	0.87	0.75	0.65	0.58	0.66	0.95
	tense	**register**	**accent**	**rate**	**nvc**	**speech a**	**informal**	**clarify**	**fluency**

All values significant at $p<0.05$.

Standardized Patient VI.

register	0.75								
accent	0.84	0.64							
rate	0.90	0.67	0.84						
nvc	0.66	NS	0.54	0.73					
speech acts	0.68	0.66	0.68	0.77	0.64				
informal	0.90	0.70	0.86	0.87	0.60	0.72			
clarify	0.49	NS	0.57	0.53	NS	0.56	NS		
fluency	0.91	0.70	0.86	0.93	0.60	0.62	0.88	0.46	
overall	0.94	0.70	0.95	0.94	0.64	0.73	0.93	0.53	0.94
	tense	**register**	**accent**	**rate**	**nvc**	**speech a**	**informal**	**clarify**	**fluency**

All values significant at $p<0.05$.

Appendix XXVII.

Language Rating Scale, revised version, 1995.

LANGUAGE SKILLS RATING SHEET.
Med II, November 21st, 1995.

Affix label here.

ROLEPLAYER: .....

TIME: ..... DATE: .....

	Poor		Adequate		Excellent
Use of correct tense	1	2	3	4	5
Use of appropriate register	1	2	3	4	5
Comprehensibility of speech due to accent	1	2	3	4	5
Appropriate rate of speech	1	2	3	4	5
Appropriate use of non-verbal communication	1	2	3	4	5
Response to requests, apologies, thanks	1	2	3	4	5
Understanding of informal language	1	2	3	4	5
Use of informal language	1	2	3	4	5
Clarification where comprehension lacking	1	2	3	4	5
Audibility of speech	1	2	3	4	5
Fluency of speech	1	2	3	4	5
Overall impression of language proficiency	1	2	3	4	5

Comments:

Appendix XXVIII.

Written Language Skills Rating Sheet.

WRITTEN LANGUAGE SKILLS RATING SHEET

STUDENT: .....

RATER: **MARGARET GUNN.**

	Poor		Adequate		Excellent
Appropriate content	1	2	3	4	5
Use of appropriate register (jargon)	1	2	3	4	5
Use of appropriate register (value-judgements)	1	2	3	4	5
Use of appropriate vocabulary	1	2	3	4	5
Appropriate use of tense	1	2	3	4	5
Appropriate use of articles,pronouns,prepositions	1	2	3	4	5
Appropriate use of spelling,punctuation,capitals	1	2	3	4	5
Legibility of handwriting	1	2	3	4	5
Appropriate use of conventions	1	2	3	4	5
Fluency of written expression	1	2	3	4	5
Overall impression of written proficiency	1	2	3	4	5

Comments:

Appendix XXIX.

Proforma for students' written account of the interview with a standardized patient, 1995 second year medicine cohort.

NOVEMBER 21, 1995
MEDICINE II

Affix label here

You have 10 minutes to record the interview you have just completed. Write an account of this interview so that another health care professional would understand what took place, and the decisions and conclusions drawn by yourself and the patient. You should write in complete sentences, and in paragraphs. Your writing must be legible. Please avoid the use of specialised medical terminology. If it is used, ensure that it is explained.

Appendix XXX.

Standardized patient script, 1995.

ENGLISH LANGUAGE PROFICIENCY TESTING - MED 2 1995

INDIGESTION - CHRONIC UPPER ABDOMINAL PAIN

Play yourself or someone else you know well.

USE THIS MEDICAL SCRIPT

Patient Profile: Age, Marital Status, Occupation - You Choose

Presenting complaint and history of presenting complaint:

You have suffered with **PAIN** for 2 months.
The pain is in the upper abdomen and occurs several times a day. You think it is worse when you are hungry.

The pain comes on gradually and is constant in nature. It doesn't go away until you take "Quickeze" during the day. It is a discomfort to begin with and then becomes a gnawing burning sensation.

You describe it as indigestion and it is usually mild and does not bother you unduly. However, more recently it has begun to wake you at night and you have been taking a tablespoon of Mylanta and milk to relieve the symptoms during the night.

On two occasions the pain has been severe enough to ring the doctor and that worries you. You think it may be an ulcer and you have had a friend with one that burst and necessitated serious surgery. You would like to avoid surgery.

Past Medical History:	Occasional diarrhoea, doesn't necessitate a trip to the Doctor.
Family History:	None of significance
Medication:	Panadol for headache
Lifestyle factors:	Cigarettes - 10 per day for years Alcohol - Wine with food for years Diet - Eats everything Exercise - None because not interested
Allergies:	None

YOU ARE LISTENING TO LANGUAGE ONLY

Student Instructions:
Hello, I have a problem with pain here (put hand on upper abdomen). I think I may have an ulcer.

Ensure use of the following informal terms:

1. Take a sickie
2. Feeling crook
3. Feeling under the weather

PTO

INTERVIEW FORMAT AND PATIENT INSTRUCTIONS

Before each student enters the room, ensure that the tape-recorder is functioning. Clearly state the next interviewee's name onto the tape.

The student will knock at the door, and enter your room.

Introduce yourself as the standardised patient, and the ESL teacher.

Take TWO identification labels from the student, one for your rating scale, and one for the ESL teacher's rating scale.

Tell the student that he or she is to conduct an interview with YOU, and will not be addressing the ESL teacher.

Remind the student that you are aware that they are in second year, and that you will not expect a diagnosis, or the demonstration of medical knowledge during the interview.

Tell the student that you are going to tape-record the interview, turn the tape on, and inform the student that the examination has commenced.

Begin with "I have a problem with pain here (put hand on upper abdomen). I think I may have an ulcer."

The maximum time allowed is 10 minutes.

Students may leave sooner if the standardised patient decides that the interview is complete.

A BELL WILL BE RUNG AT THE TEN MINUTE MARK - THE INTERVIEW MUST END AT THIS POINT IF IT HAS NOT ALREADY CONCLUDED

There will be no discussion between the student and the "patient" or the ESL teacher once the examination is completed.

The student will leave the interview room, and is then required to record, in 10 minutes, a written account of the oral interview.

THE STUDENT WILL LEAVE THE EXAMINATION WHEN BOTH ORAL AND WRITTEN TESTS ARE COMPLETED.

Appendix XXXI.

Language Rating Scale Training Criteria, 1995.

LANGUAGE SKILLS RATING SCALE.
STANDARDISED PATIENT TRAINING SESSION.
November 14th, 1995.

Use of appropriate tense.

"Tense" is the relationship between the form of the verb and the time of the action or state that it describes.

Does the student employ the correct tense during the interview, in terms of relaying the intended meaning?

Rating
1. inappropriate use of tense creates considerable misunderstanding
2. inappropriate use of tense creates some misunderstanding
3. inappropriate use of tense creates unimportant misunderstanding
4. inappropriate use of tense creates no misunderstanding
5. no inappropriate use of tense

Use of appropriate register

"Register" is the style of speech that is used by a particular group of people, with its own special words or phrases, and sometimes with its own grammar, as in legal registers. Usually people who share a common register come from the same occupation, or have the same interests. Therefore doctors have their own register, and they use this when interacting on a professional level. It is inappropriate for a doctor to use a medical register when consulting with a patient. For example, the use of medical terminology ("jargon") without clear explanation of the meaning, is inappropriate.

Does the student use the appropriate register throughout the interview?

Rating
1. student often relies on specialist register to relay meaning
2. sometimes relies on specialist register to relay meaning
3. rarely relies on specialist register to relay meaning
4. sensitive to register, can modify speech to suit patient, minor errors
5. sensitive to register, can modify speech to suit patient, no errors

Comprehensibility of speech due to accent.

"Accent" is defined as a way of speaking which tells the listener something about the speaker's background, such as the country from which they come, their social class, and whether the speaker is a native speaker of the language. Sometimes accent can interfere with the listener's understanding of the speaker's words.

Does the student's accent influence the understanding of the meaning of the speech in any negative manner?

Rating
1. frequently pronunciation of words is poor, and speech difficult to understand
2. sometimes pronunciation of words is poor, and speech difficult to understand
3. accent is present, but does not significantly affect patient's comprehension
4. accent present, does not interfere with patient's comprehension
5. accent, if present, in no way hampers patient comprehension

Appropriate rate of speech

"Rate" refers to the rate or speed at which the person is speaking. Rate of speech can be objectively measured by counting the number of syllables per minute. However, one can have a subjective impression of the rate of speech as seeming more rapid or slower than one feels is usual.

What is the student's rate of speech? Does it interfere with the listener's understanding of the speaker?

Rating
1. speech is notably too fast or too slow, interferes with listener's comprehension
2. rate of speech sometimes interferes with listener's comprehension
3. rate of speech rarely interferes with comprehension
4. rate of speech is not of a native speakers, but does not affect comprehension
5. rate of speech is as a native speaker

Appropriate use of non-verbal communication

"Non-verbal communication" (nvc) includes facial expressions, head and eye movements, body and hand gestures, that support, emphasise or alter the meaning of the spoken word.

Are the student's non-verbal behaviours in accord with what he or she is saying?

Rating
1. nvc significantly inappropriate to interaction eg laughing, poor eye contact
2. nvc sometimes inappropriate
3. nvc rarely inappropriate
4. nvc almost always congruent with interaction
5. nvc always congruent with interaction

Response to requests, apologies, thanks

Requests, apologies, thanks and invitations are called "speech acts".

Does the student respond appropriately to the patient's speech acts?

Rating
1. poor grasp of social conventions, interferes with interaction
2. some difficulty with speech acts, interferes with interaction
3. some difficulty with speech acts, but does not hamper interaction
4. rarely inappropriate in response to speech acts
5. always responds appropriately to speech acts

Understanding of informal language

Includes figures of speech, colloquialisms, slang. In a "figure of speech" most commonly the speaker uses a simile or metaphor for special effect, but does not intend the literal meaning to be taken by the listener. "Colloquialisms" and "slang" are informal forms of speech, and include the examples you will use in the standardised interview script, such as "crook", "take a sickie" and "under the weather".

Does the student show by his or her response that they understand the use of these kinds of speech by the patient?

Rating
1. clearly does not comprehend informal language used by the patient
2. does not comprehend some informal language
3. does not always comprehend, but this does not significantly affect interaction
4. usually does understand informal language
5. understanding of informal speech is as native speaker's

Use of informal language

Does the student demonstrate familiarity with informal language by using it in their own speech?

Rating
1. never uses informal speech
2. attempts at informal speech are awkward, or incorrect
3. attempts may not be correct, but do not hamper interaction
4. may not be completely at ease with informal speech, but is appropriate
5. is clearly comfortable with informal speech and uses it appropriately

Clarification where comprehension lacking

"Clarification" refers to the skill of ensuring that one understands what the speaker has said, where there is doubt. For example, where the patient uses unfamiliar informal speech ("under the weather"), or the name of a drug ("Mylanta"), the student should clarify this by rephrasing, or asking the patient to explain.

Does the student clarify what they have not fully understood?

Rating
1. clearly misunderstands, and does not ever clarify
2. rarely clarifies, obviously has misunderstood some information
3. sometimes clarifies information which is not understood
4. always clarifies that which is not understood
5. may not clarify, but this is because the student has clearly understood

Audibility of speech

"Audibility" refers to the projection of speech.

Is the student's speech inappropriately loud or quiet?

Rating
1. speech is inappropriately loud/soft, interferes with interaction
2. audibility is inappropriate, sometimes interferes with interaction
3. audibility is inappropriate, but does not interfere with interaction
4. audibility is usually appropriate, does not influence interaction
5. student's projection is like that of a native speaker's

Fluency of speech

"Fluency" refers to features of speech that make it seem natural to the listener, including the use of pauses, rhythms, intonations, stress on syllables and words, the rate of speaking, and the use of interjections and interruptions. Thus fluency is a global indicator of the quality of speech. It can be described as a level of proficiency in language which includes the ability to produce spoken language with ease; to speak with a good, but not necessarily perfect command of intonation, vocabulary and grammar; to communicate ideas effectively; and to produce continuous speech without causing comprehension difficulties or a breakdown of communication.

Does the student appear to be speaking in English without effort, or does the student appear to struggle? Does the student have difficulty finding the right words, and in putting them together into a sentence? Are there difficulties in comprehension for either party? Does communication "break down" at any point, where "break down" is defined as a track of thought being abandoned due to lack of understanding, the student asking a question/giving an answer that is irrelevant to the patient's initial response.

Rating
1. fluency is poor, uneven, limited vocabulary means time spent searching for right words, words are not in correct place in sentence, patient must repeat self to be understood
2. hesitations as student searches for right words and grammatical constructions, patient may need to repeat self to be understood, student shows more difficulty in longer or more complex sentences than simple, briefer ones
3. appears comfortable with speaking, rarely needs to search for words, and can fill in gaps with other speech whilst searching for words where necessary, fluency is good
4. can use language fluently and accurately, using wide vocabulary
5. uses language as would a native speaker, with complete fluency, accuracy and vocabulary

Overall impression of language proficiency

All things considered, where would you rate this student on the scale from 1 to 5, based on their performance during the interview?

Do you think this student has a level of language sufficient to negotiate an interview with a patient in a medical interview setting.

Rating
1. language is extremely poor, unacceptable standard for interaction with patients
2. language is poor, and is unacceptable for patient interactions
3. language is acceptable for patient interaction, but requires improvement
4. completely acceptable standard of language for patient interactions
5. faultless performance, completely acceptable standard of language

Comments

The items on this scale take into account the **frequency** of a response, and also the **impact** that a response has upon the interaction. You may need to comment upon why you have rated a student as "poor" (1), if for example, they used tense incorrectly only once, but you feel that this error was an important one that significantly impacted upon the course of the interview.

Sometimes you will be unable to rate an item - explain why this was so here.

You may also find that you want to say something about the student's language that has not been included in the rating scale, and this should be noted here.

Your open-ended comments are extremely valuable. If there is not enough space on the front of the rating sheet, please feel free to write on the back, which is blank.

ACH 12/11/95

Appendix XXXII.

Student Instructions on the day of the Language OSCI, 1995.

Language OSCI Examination, November 21st, 1995.
Information for students on the day.
Spoken by Jane Vernon-Roberts.

You are about to enter the examination for language skills.
Anna will show you to your room.
When you enter the room, you will find two people in there.
One is a standardised patient.
One is an English as a second language teacher.
The standardised patient will introduce herself and the language teacher.
You should give one of your name labels to the standardised patient and one to the language teacher.

The examination will then commence.
You will be interviewing the standardised patient for up to 10 minutes. The interview will be tape-recorded.
The language teacher is preset as an observer - you should not direct your interview to them.
The language teacher is present to assess your spoken language only.
You will not be assessed for your medical knowledge.
This is not the "Communication Skills" OSCI exam. You will not be assessed on the various skills, like empathising, introducing yourself and so on, that you have been learning this year, although of course it will be appropriate to use these skills.
Remember, in this exam you will be assessed on your capacity to communicate using the English language.

After the interview you should wait outside the room, on the chair.
Do not talk.
Anna will direct you to a second room, and on hearing a bell, you should begin writing an account of the interview you have just completed. You have 10 minutes in which to do this.
Instructions for the written account are on the examination paper.
Make sure you have affixed your name label to your written account.
A bell will sound. You should remain seated until Ms Francis has collected your papers.
Ms Francis will then direct you to the exit. Please leave quietly.

ACH 1/11/95.

Appendix XXXIII.

Language Rating Scale Study (Chapter VIII).
Individual standardized patient's ratings.

Ratings of Standardized Patient I.

n=22	1	2	3	4	5	missing	$\bar{x}$	sd
tense	-	2 (9.1)	2 (9.1)	6 (27.3)	12 (54.5)	-	4.27	0.99
register	-	-	2 (9.1)	8 (36.4)	12 (54.5)	-	4.46	0.67
accent	-	-	1 (4.5)	5 (22.7)	16 (72.7)	-	4.68	0.57
rate	-	-	1 (4.5)	5 (22.7)	16 (72.7)	-	4.68	0.57
nvc	-	-	1 (4.5)	4 (18.2)	17 (77.3)	-	4.73	0.55
speech acts	-	-	-	2 (9.1)	20 (90.9)	-	4.91	0.29
informal 1	-	1 (4.5)	1 (4.5)	5 (22.7)	15 (68.2)	-	4.55	0.80
informal 2	-	1 (4.5)	1 (4.5)	1 (4.5)	10 (45.5)	9 (40.9)	4.54	0.97
clarification	-	-	1 (4.5)	6 (27.3)	15 (68.2)	-	4.64	0.58
audibility	-	1 (4.5)	1 (4.5)	3 (13.6)	17 (77.3)	-	4.64	0.79
fluency	-	-	4 (18.2)	7 (31.8)	11 (50.0)	-	4.32	0.78
overall	-	-	2 (9.1)	5 (22.7)	13 (59.1)	2 (9.1)	4.55	0.69

Ratings of Standardized Patient II.

n=22	1	2	3	4	5	missing	$\bar{x}$	sd
tense	-	1 (4.5)	2 (9.1)	3 (13.6)	16 (72.7)	-	4.55	0.86
register	-	-	3 (13.6)	3 (13.6)	16 (72.7)	-	4.59	0.73
accent	-	-	1 (4.5)	4 (18.2)	17 (77.3)	-	4.73	0.55
rate	-	-	-	-	22 (100)	-	5.00	0.00
nvc	-	-	-	2 (9.1)	20 (90.9)	-	4.91	0.29
speech acts	-	1 (4.5)	-	-	20 (90.9)	1 (4.5)	4.86	0.66
informal 1	-	1 (4.5)	1 (4.5)	2 (9.1)	18 (81.8)	-	4.68	0.78
informal 2	-	-	-	3 (13.6)	18 (81.8)	1 (4.5)	4.86	0.36
clarification	-	1 (4.5)	2 (9.1)	4 (18.2)	15 (68.2)	-	4.50	0.86
audibility	-	-	-	3 (13.6)	19 (86.4)	-	4.86	0.35
fluency	-	1 (4.5)	3 (13.6)	1 (4.5)	17 (77.3)	-	4.55	0.91
overall	-	-	4 (18.2)	2 (9.1)	16 (72.7)	-	4.55	0.80

Ratings of Standardized Patient III.

n=18	1	2	3	4	5	missing	$\bar{x}$	sd
tense	-	-	1 (5.6)	4 (22.2)	13 (72.2)	-	4.67	0.59
register	-	-	1 (5.6)	10 (55.6)	7 (38.9)	-	4.33	0.59
accent	-	-	2 (11.1)	5 (27.8)	11 (61.1)	-	4.50	0.71
rate	-	-	4 (22.2)	2 (11.1)	12 (66.7)	-	4.44	0.86
nvc	-	-	3 (16.7)	3 (16.7)	12 (66.7)	-	4.50	0.79
speech acts	-	-	1 (5.6)	2 (11.1)	15 (83.3)	-	4.78	0.55
informal 1	-	1 (5.6)	7 (38.9)	-	9 (50.0)	1 (5.6)	4.00	1.12
informal 2	-	-	-	1 (5.6)	15 (83.3)	2 (11.1)	4.94	0.25
clarification	-	1 (5.6)	3 (16.7)	8 (44.4)	6 (33.3)	-	4.06	0.87
audibility	-	-	-	6 (33.3)	12 (66.7)	-	4.67	0.49
fluency	-	1 (5.6)	4 (22.2)	12 (66.7)	-	1 (5.6)	3.65	0.61
overall	-	-	5 (27.8)	13 (72.2)	-	-	3.72	0.46

Ratings of Standardized Patient IV.

n=22	1	2	3	4	5	missing	$\bar{x}$	sd
tense	-	-	2 (9.1)	2 (9.1)	18 (81.8)	-	4.73	0.63
register	-	-	2 (9.1)	9 (40.9)	11 (50.0)	-	4.41	0.67
accent	-	-	1 (4.5)	1 (4.5)	20 (90.9)	-	4.86	0.47
rate	-	-	1 (4.5)	8 (36.4)	13 (59.1)	-	4.55	0.60
nvc	-	1 (4.5)	1 (4.5)	4 (18.2)	16 (72.7)	-	4.59	0.80
speech acts	1 (4.5)	-	1 (4.5)	1 (4.5)	19 (86.4)	-	4.68	0.95
informal 1	-	1 (4.5)	2 (9.1)	1 (4.5)	17 (77.3)	1 (4.5)	4.62	0.87
informal 2	-	-	-	2 (9.1)	3 (13.6)	17 (77.3)	4.60	0.55
clarification	-	-	-	8 (36.4)	12 (54.5)	2 (9.1)	4.60	0.50
audibility	-	-	-	3 (13.6)	19 (86.4)	-	4.86	0.35
fluency	-	3 (13.6)	3 (13.6)	2 (9.1)	14 (63.6)	-	4.23	1.15
overall	-	-	4 (18.2)	5 (22.7)	13 (59.1)	-	4.41	0.80

Ratings of Standardized Patient V.

n=21	1	2	3	4	5	missing	$\bar{x}$	sd
tense	-	-	-	5 (23.8)	16 (76.2)	-	4.76	0.44
register	-	-	-	3 (14.3)	18 (85.7)	-	4.86	0.36
accent	-	-	5 (23.8)	9 (42.9)	7 (33.3)	-	4.10	0.77
rate	-	-	3 (14.3)	7 (33.3)	11 (52.4)	-	4.38	0.74
nvc	-	-	1 (4.8)	4 (19.0)	16 (76.3)	-	4.71	0.56
speech acts	-	1 (4.8)	-	10 (47.6)	10 (47.6)	-	4.38	0.74
informal 1	-	-	5 (23.8)	7 (33.3)	7 (33.3)	2 (9.5)	4.11	0.81
informal 2	-	-	-	8 (38.1)	3 (14.3)	10 (47.6)	4.27	0.47
clarification	-	-	-	12 (57.1)	9 (42.9)	-	4.43	0.51
audibility	-	-	3 (14.3)	4 (19.0)	14 (66.7)	-	4.52	0.75
fluency	-	2 (9.5)	3 (14.3)	9 (42.9)	7 (33.3)	-	4.00	0.95
overall	-	-	2 (9.5)	15 (71.4)	4 (19.0)	-	4.10	0.54

Ratings of Standardized Patient VI.

n=22	1	2	3	4	5	missing	$\bar{x}$	sd
tense	-	-	1 (4.5)	2 (9.1)	19 (86.4)	-	4.82	0.50
register	-	-	4 (18.2)	13 (59.1)	5 (22.7)	-	4.05	0.65
accent	-	-	4 (18.2)	1 (4.5)	17 (77.3)	-	4.59	0.80
rate	-	-	2 (9.1)	4 (19.2)	16 (72.7)	-	4.64	0.66
nvc	-	1 (4.5)	-	3 (13.6)	18 (81.8)	-	4.73	0.70
speech acts	-	-	-	1 (4.5)	21 (95.5)	-	4.96	0.21
informal 1	-	1 (4.5)	2 (9.1)	5 (22.7)	12 (54.5)	2 (9.1)	4.40	0.88
informal 2	1 (4.5)	-	2 (9.1)	12 (54.5)	5 (22.7)	2 (9.1)	4.00	0.92
clarification	1 (4.5)	-	4 (18.2)	8 (36.4)	9 (40.9)	-	4.09	1.02
audibility	-	1 (4.5)	-	2 (9.1)	19 (86.4)	-	4.77	0.69
fluency	-	1 (4.5)	4 (18.2)	7 (31.8)	10 (45.5)	-	4.18	0.91
overall	-	-	3 (13.6)	11 (50.0)	8 (36.4)	-	4.23	0.69

Appendix XXXIV.

Open-ended comments on the Language Rating Scale 1995 second year cohort study and resultant coding categories [in parenthesis].

Key to coding categories.
1=use of jargon
2=tense
3=non-verbal communication
4=audibility
5=comprehensibility
6=colloquial language - positive comments
7=colloquial language - negative comments
8=accent
9=vocabulary
10=nervousness
11=fluency
12=failure to clarify
13=rate
14=pronunciation
15=difficulty keeping the student talking
16=difficulty arriving upon a rating

Standardized Patient I.

1. Used some medical jargon but corrected herself. Most difficulty encountered understanding the student due to problems with the use of inappropriate tense which caused her to stumble and search for words. This was balanced however with good use of non-verbal skills. [1, 2, 3]

2. Said "abdomen" and "altercation". Although he used the term "altercation" he was correct in assuming I could understand this particularly when told that I was an English teacher. He was rather soft and tended to mumble at times but overall was very easy to comprehend. [1, 4, 5]

3. Used may examples of informal language. No problems, easily understood. Much appropriate use of informal language. [6, 5]

4. Could not assess use of informal language. Queried "under the weather" but I feel she understood anyway. Although this student has an accent she has very fluent speech with no hesitations and she uses a wide vocabulary. I believe she understood all the informal speech however she wished to clarify one phrase, as above. She also understood "pap smear" and "pill" which are medical terms but in common usage. [16, 7, 8, 9]

5. Smiled and laughed appropriately - good eye contact. Used jargon - asked if I had brought up any sputum and if I had blood or mucus in my stools. She was able to explain those terms however, when they were queried. Otherwise fluent use of language. [3, 1, 11]

6. Non-verbal communication hampered by nervousness (looking at feet). Fluency of speech may have been affected by nervousness. Did not clarify informal language but clearly understood. Rather hesitant but probably related to nervousness than to fluency. [3, 10, 12]

7. "Aussie" accent tends to interfere with comprehensibility of speech. Sometimes speech a little fast. Has some difficulty finding appropriate word - has a reasonable but not a full vocabulary. Rather sloppy pronunciation but does not really interfere with comprehension. [8, 13, 9, 14, 5]

8. Register - said what is your "job stability". Smiled and laughed appropriately. Did not ask for clarification with use of slang terms and I feel she understood these. Overall fluent and clear. [1, 3, 12, 11]

9. Non-verbal behaviour and fluency of speech influenced by nervousness. This student was quite nervous and so reduced fluency but overall was very easy to understand and obviously understood all that was said to her. [3, 11, 10, 5]

10. Register - once said "does the pain radiate". Used informal language "is that the worst case scenario". No problems, very easy to comprehend. [1, 6, 5]

11. Used informal language - "the terrible twos", "kill two birds with one stone". Clear and fluent with extensive use of informal speech including use of well known proverbs. [6, 11]

12. Said "dietary restriction". Excellent use of non-verbal communication. Understood "take a sickie" but not "feeling crook". Needed an explanation of "gnawing". Although this student has quite a strong accent and sometimes uses the wrong tense she is quite fluent and very easy to understand. [7, 1, 3, 8, 2, 5, 11]

13. Said "inherited trait", spoke a little fast, used informals "do you drink?", asked "what does under the weather mean?". This student was reasonably easy to understand although he spoke quite softly (he has a rather hoarse voice). He would occasionally appear to be searching for words. [1, 13, 7, 4, 5]

14. Speech too fast - ? due to nerves in part. Some inappropriate laughter. Clarified "Mylanta". This student was very nervous and this may have made him speak so rapidly - he stumbled over his own words quite frequently. He did not ask for clarification of informal speech but I feel he did not fully understand all that I said to him. [13, 3, 10, 12]

15. Occasionally laughed inappropriately. I feel she needed to clarify more often, as I don't think she understood. Although it is difficult to assess her understanding of informal language I do not believe this student understood terms such as "under the weather", "crook" or "taking a sickie". Had difficulty thinking of words. E.g. "Your boss will be disappointed, no?" and I said "angry" and she said "yes", such that speech was hesitating at times. [3, 12, 16, 11, 7]

16. Used "abdomen" and "antacid". Non-verbals affected by nervousness. Some problems with tense. Fluent with little hesitation - some nervousness apparent. Did not clarify the use of slang terms but appeared to understand these terms. Tended to loose eye contact and kept looking at my abdominal area - where the pain was supposed to be - rather off-putting. [1, 3, 10, 2, 11, 6, 12]

17. Although this student is softly spoken and was a little nervous, her language skills were excellent. [4, 10]

18. Spoke too fast. Did not understand when I said "My husband thinks I am a wooz". Used informal language - "It has been a hassle for you". Speech too soft, quite difficult to hear her at times. Fluent speech but understanding hampered by reduced audibility. Understood most informal speech and was able to use herself. [13, 6, 4, 11]

19. Understood all informal language including fear of "coming under the knife". No difficulties in any areas. Understood all terms including slang - responded appropriately. [6]

20. Used jargon - "have you had any medical procedures". Used informal language - "you are afraid to be bludging from work". A little soft in speech, sometimes searched for words. This student was a little nervous and had some difficulty finding things to say but in terms of language skills was excellent. [1, 6, 4, 10]

21. Very good eye contact. Used informals - "for sure", "so he's a real sweetie". Extremely fluent - no difficulties comprehending at all. [3, 6, 11]

22. Used jargon - "family history", "character of the pain". Rate of speech too fast due to nervousness. Rather too formal in speech acts but was very nervous. Audibility too loud but due to nervousness. Language proficiency good but marred by initial nervousness which tended to make his presentation overly formal with the use of specialist register - improved as he became more comfortable. [1, 13, 10, 4]

Standardized Patient II.

1. A 3 for overall impression of language proficiency because some subsequent improvement may be required. I became used to the accent very quickly so after a time there was no difficulty.[8]

2. Clarified "run down". asked about "under the weather". Understood Quickeze and Mylanta. No problems whatsoever. [7]

3. Response to requests, apologies, thanks - I said "we'll leave it there" and he said "thanks".

4. He said "that makes you feel misery". Misery - not miserable. I'm really tossing up giving overall impression of language proficiency a 4 - but I feel as though it does "require improvement" but not a great deal. [16]

5. Audibility - a bit quiet. [4]

6. No problems!

7. No problems with language but he laughed when I said my parents died. [3]

8. Non-verbals a bit too enthusiastic, eyes widened etc. When I asked her a few things she replied "Mmm, Mmm". [3]

9. His leg was swaying a lot but I put this down to nerves. [3, 10]

10. Uses some jargon although appropriately. [1]

11. Explained appropriately meaning of "secretion". [1]

12. He paraphrased - he clearly knew what I meant. He said "You don't feel well about taking a sickie" - instead of "good". [6]

13. Demonstrated sleeping action with hands. Inappropriate register - reflux, oesophagus, gastric reflux, acidic . . . [3, 1]

14. Used the word faeces. [1]

15. Kicked her leg about a bit. [3]

16. Explained "sickies" by asking if that's the only time I've had off. [6]

17. Excellent non-verbals, moved forward. Speech a bit soft. Sometimes forgot "an" e.g. "you think is ulcer". [3, 4]

18. Used the word "stool". [1]

19. Some problems with tense but very minor. Often paused searching for words but I think this was more because he didn't know what to say rather than how to say it. [2, 11]

Standardized Patient III.

1. Giggled inappropriately due to nerves. Did not rate use of informal language - I don't think she is comfortable with it but not sure. Hesitations in speech (possibly worsened by feeling nervous). [3, 10, 7, 11]

2. Speaks fast at times. Used informal language: before the interview began used the term "no problems". Clarified sometimes, second time I used the term Mylanta, not initially. [13, 6]

3. Used "localized". [1]

4. Used informal language - "stick a tube down your throat". Nervousness affected his fluency, but not his comprehensibility. He did not appear comfortable with speaking. [6, 11, 5]

5. Used "antacids" but in context. Nervousness affected fluency but not a language problem. [1, 10]

6. Giggled inappropriately at one point. Clarified appropriately. Used terms "Mylanta" and "Quickeze". Confused "gnawing" with "annoying", but this did not interfere significantly with the interaction. [3]

7. Did not clarify with "sickie", "feeling crook" or "under the weather" although he appear to understand these terms. Did understand Mylanta and Quickeze. [7]

8. Used informal language - "that's cool". [6]

Standardized Patient IV.

1. nvc - kept unfolding his hands - irritating. One thing this student did was to say "for sure" after almost everything I said to begin with. He did stop this later when he became more familiar with his surroundings and task. May just have been a nervous thing - but God it was annoying. [3, 10]

2. A bit fast in speech, but ok. nvc excellent - appropriate leaning forward with concern. [13, 3]

3. Rate of speech - a bit fast(ish). Nervous - but good language skills. Used jargon - exacerbate. [13, 10, 1]

4. Cannot rate use of informal speech - didn't use any. Rate of speech - too fast. nvc - leaning forward and good eye contact. [16, 7, 13, 3]

5. Spoke quickly at times. Didn't use informal language. Excellent except for rate of speech from time to time. [13, 7]

6. Rate fast from time to time. Didn't use informal language. He was excellent. his facial expressions were so intense when talking to me - but good. [13, 7, 3]

7. Used jargon - gastrointestinal system, stools. An example of poor communication skills but good language skills. Awkward guy. [1]

8. Very easy to talk with. Pleasant, made all the right non-verbals and she kept the conversation going at times and so I found it easy and comfortable. [3]

9. Once I had to ask him to repeat himself because he spoke too fast. Excellent - well spoken most of the time. Good non-verbals. Interested and constantly enquiring. No trouble searching for words. He'll be a good doctor! [13, 3, 11]

10. A bit awkward in beginning with shacking hands and sitting down. Used "crook" after I said it. This student was very good with nvc and asking the right questions. He didn't mind a chat. [3, 6]

11. Excellent - pleasant, easy to talk to.

12. Did not use informal language and would be uncomfortable with it. Softly spoken. Appears comfortable with speech but not fluent. She was very good at some things but her tense, audibility and fluency were her main problems. [7, 4, 11, 2]

13. Excellent - but very nervous. Got into it as time wore on. [10]

14. Used "gastric secretions". Spoke a bit fast. Wringing hands - nervous. Nervous but good. He also had a cough - that did interfere with the interview. [1, 10]

15. I didn't know whether he understood my informal language or not. It was difficult to tell. [16, 7]

16. This girl was very nervous and not very talkative. She had a mental block in the middle of the whole thing. Non-verbal communication - not very good. Eye contact poor. Very stiff and formal interview. Typical example of good language skills but poor communication skills. [10, 15, 3]

17. Rate of speech fast(ish) but ok to understand. NVC - excellent. Audibility - bit soft to begin with. Very good - easy to talk to. [13, 3, 4]

18. Some hesitations and searching for the right words. Tried hard though and I think she did a good job in getting her meaning through. [11]

19. Fluency not that great - could be nerves. NVC - too starry eyed; large starry eyes - a bit off-putting but she maintained eye contact. [11, 10, 3]

20. NVC - head nodding, laughter, head shaking -oh - it was so bad. [3]

21. NVC - clicking his fingers, very annoying and distracting. Eye contact poor as well. Very nervous and a bit disoriented. [3, 10]

Standardized Patient V.

1. An excellent communicator. One slip but it was my fault. I was not very clear - she thought I meant cook not crook.

2. Not sure if his rate was a little fast and if he actually used any informal language. He appeared to understand the colloquials but did not respond to them. [13, 7]

3. Nervous and slightly stiff. Not sure if he used any informal language. This student has braces (on his teeth) - may have a slight lisp to his speech, that is why I gave him a 4. [10, 7]

4. This student, due to his nervousness, had a problem in trying to keep the conversation going towards the end, began to repeat himself so I terminated the session. However the agenda had been completely covered - I just felt that going on further would probably increase his nervousness and therefore mistakes. [10]

5. This student had a slight problem with fluency and hesitated during one segment of the conversation. I have given her a 2 because that was the only slot I could fit her into. Otherwise she was quite good, fairly clear and concise. [11]

6. Slight accent but tries to speak clearly. Not sure if she understood the informal, could not tell from her non-verbal cues; her response to each colloquial and to the medication names seemed appropriate. [8, 3, 6, 7]

7. Made no comment about the informal language. Voice a bit soft. Not sure if he understood informal language. could not gauge from his expression. He seemed nervous and was really trying very hard to think of something to say. [7, 10]

8. Non-verbals a little stiff and formal. Very good, he tried hard to keep the conversation running smoothly. Slight lisp made it difficult to understand or it just may have been his accent. [3, 8]

9. This student had the session under her control, as you will note from the tape. She was quite fluent her only problem was a slight lisp (did not interfere with comprehension). Not sure if she understood informal language or whether she missed the first two cues. [7]

10. Spoke a little fast. A little quiet. Tries hard to be fluent. Some problem with audibility and accent otherwise a good communicator. [13, 4, 8, 11]

11. Very good non-verbal cues alerted me to her understanding of informal language. Slight lisp - I had to listen carefully. [3, 6]

12. Did not use informal language - not sure if it's because he is not familiar with it. Very clearly articulated to overcome his accent. Very good language skills and quite fluent and relaxed. [7, 8, 11]

13. Uses plurals when unnecessary. Student clarified because she did not understand any of the informal speech. Spoke a little too softly. More fluent than previous students even though very nervous. [7, 4, 11, 10]

14. The language flowed very easily. This student has a good command of English and is quite confident in her speech. Very good. [11]

15. An excellent communicator. Good command of the English language. Could not fault her on anything. [11]

16. Very good language skills. I slipped up more than the student. [11]

17. Problem answering questions. Student has difficulty in communicating anything more than a sentence or two. Found this interview hard going, he seemed to be unable to do anything but ask questions and reflect. Perhaps it was my fault - not enough open questions? [11]

18. Problems with accent although student tries hard. This student tried very hard to cover up her accent - she did very well and di a fair bit of talking - seemed comfortable and confident in her role. Only a couple of very slight pauses in searching for the appropriate word. [8, 11]

19. Did not acknowledge use of informal language each time but it was clear she understood. Overall a clear communicator apart from one use of jargon - "precipitate". [6, 1]

20. Knee bounces when talking, stopped when listening. Not sure if rate of speech or accent were the most outstanding problems this student had - probably both. Needs to overcome his knee bounces when nervous or it may just be a bad habit of his. [3, 13, 8, 10]

21. Two occasions used slight medical jargon to summarize patient problem. Spoke a little fast due to nervousness. I believe the student to be highly nervous but covered up well. This nervousness perhaps influenced his rate of speech - at times a little fast. [1, 13, 10]

22. Perhaps his nerves stopped him from maintaining the flow of conversation, also he spoke a bit softly - not sure if he understood all the informal language but did nod. Spoke a bit fast - may be due to nervousness. [10, 4, 13, 7]

Standardized Patient VI.

1. Slouched, did not face patient directly, giggled a little - seemed to think the exercise was amusing. Did not rephrase or question the phrases. It is most likely he understood, but difficult to assess. [3, 12]

2. I gave him a 4 for his understanding of informal language because I felt he understood, but did not rephrase the phrases. [6, 12]

3. Did not use any informal language. [7]

4. I could not rate this student on understanding of informal language. When the informal phrases were used shewould give a slight frown, but did not use them, or try to clarify these phrases. Difficult to assess use of informal language, she does not use very informal language. She did not clarify phrases, (but) this student summed up the interview very well and accurately. She did not use the (colloquial language) however, and I can't judge whether or not she understood them. [16, 7, 12]

5. She understood informal speech, although not always immediately. Did not use the colloquial phrases or other informal speech phrases, but it would be difficult to say if using speech as informal as the given phrases would make her uncomfortable. [6]

6. Perhaps a little fast. She sometimes repeated or clarified the informal speech, but I believe she was doing this to show me she understood. Searched for a word on one occasion but I believe she was doing this for the patient's benefit; not because she didn't understand. Difficult to measure the difference between a 4 and a 5 for this. [16, 13, 6]

7. Referred to antacids - which I hadn't used at all - a specialist term which indicated he knew what Quickeze and Mylanta were; but probably inappropriate use of specialist register. I suspect he doesn't use informal speech much. I gave him a 4 for overall proficiency because of his use of specialist register. [1]

8. Used one or two words that might be difficult to understand in a conversation with a patient (eg debillitating). Was comfortable with my use of informal language - but did not rephrase my phrases. Asked for clarification, but did not use an equivalent. I gave her a 4 for overall proficiency because I felt she tended to use the medical register, rather than modify her speech to the register of the patient. [1, 6]

9. Used appropriate register overall - made a comment about symptoms paralleling friend's symptoms - which might confuse some patients; which is why I gave him a 4 for register. Spoke too quickly - but related to a large extent to nervousness. Difficult to assess his understanding of informal language, because he didn't really clarify the statements. My feeling is that he understood. He did clarify some things; because he was indicating he understood. [16, 13, 1, 10, 6, 7]

10. Rate of speech a little fast. [13]

11. Did not use informal language. Very softly spoken. [4, 7]

12. Speech a little fast at times. [13]

13. A bit difficult to assess understanding of informal language. He certainly did not appear baffled by the phrases and did not need to clarify them. My feeling is that he understood. Again difficult to assess use of informal language. He did not appear uncomfortable in his speech at all. [16, 7, 6]

14. Speech perhaps a little fast due to nervousness. Difficult to say if he understood all of the informal phrases. Did not use the phrases but speech included informal language - an explanation about acid. Perhaps a little quiet. [16, 13, 10, 4, 6]

15. Use of language appropriate, but did not go out of his way to use the informal phrases or expand upon them. [7]

16. This interview was interrupted by the tape ending in the middle.

17. Did not use clarification but I believe she understood. [12]

18. A very good interview - from patient's point of view.

Appendix XXXV.

Spearman *rho* correlation coefficients, Language Rating Scale Study (Chapter VIII).

Standardized Patient I.

	tense	register	accent	rate	nvc	speech acts	informal	informal under.	clarify	audible	fluency
register	NS										
accent	0.72	NS									
rate	NS	NS	0.57								
nvc	0.57	NS	NS	NS							
speech acts	NS	NS	NS	0.58	NS						
informal	0.50	NS	0.53	0.44	NS	NS					
informal under.	0.67	NS	0.85	0.81	NS	NS	0.99				
clarify	0.47	NS	0.50	0.43	NS	NS	0.99	0.99			
audible	NS	NS	NS	NS	NS	NS	NS	NS	NS		
fluency	0.89	NS	0.76	NS	0.52	NS	0.52	0.60	0.50	NS	
overall	0.77	NS	0.83	0.62	NS	NS	0.83	0.99	0.81	NS	0.73

All values significant at $p<0.05$

Standardized Patient II.

register	NS										
accent	0.70	NS									
rate	-	-	-								
nvc	NS	NS	NS	-							
speech acts	0.49	NS	NS	-	0.69						
informal	0.57	NS	0.88	-	NS	0.54					
informal under.	0.78	NS	0.85	-	NS	-	0.63				
clarify	0.91	NS	0.65	-	NS	0.46	0.53	0.72			
audible	NS	NS	0.46	-	NS	NS	0.48	NS	NS		
fluency	0.75	NS	0.99	-	NS	0.50	0.87	0.89	0.69	NS	
overall	0.83	NS	0.92	-	NS	NS	0.79	0.81	0.77	NS	0.92
	tense	**register**	**accent**	**rate**	**nvc**	**speech acts**	**informal**	**informal under.**	**clarify**	**audible**	**fluency**

All values significant at $p<0.05$

Standardized Patient III.

register	NS										
accent	0.82	NS									
rate	0.73	NS	0.73								
nvc	NS	0.57	NS	NS							
speech acts	NS	NS	NS	NS	NS						
informal	NS	NS	0.56	0.50	NS	NS					
informal under.	NS	NS	NS	NS	NS	NS	NS				
clarify	0.56	NS	0.60	0.58	NS	NS	0.76	NS			
audible	NS	NS	0.47	0.52	NS	NS	NS	-	NS		
fluency	0.51	NS	0.61	0.64	NS	NS	NS	-	NS	0.99	
overall	0.73	NS	0.81	0.72	NS	NS	0.60	NS	0.55	0.61	0.73
	tense	**register**	**accent**	**rate**	**nvc**	**speech acts**	**informal**	**informal under.**	**clarify**	**audible**	**fluency**

All values significant at $p<0.05$.

Standardized Patient IV.

	tense	register	accent	rate	nvc	speech acts	informal	informal under.	clarify	audible	fluency
register	NS										
accent	0.74	NS									
rate	NS	NS	NS								
nvc	NS	NS	NS	NS							
speech acts	NS	NS	NS	NS	NS						
informal	0.56	NS	0.54	NS	NS	NS					
informal under.	NS	NS	NS	NS	NS	NS	0.97				
clarify	NS	NS	NS	-0.46	NS	NS	0.61	NS			
audible	NS	NS	NS	NS	NS	NS	0.44	NS	NS		
fluency	0.42	NS	0.46	NS	NS	NS	0.60	0.97	NS	NS	
overall	NS	0.53	0.51	NS	NS	NS	0.57	0.97	0.46	NS	0.93

All values significant at $p<0.05$

Standardized Patient V.

	tense	register	accent	rate	nvc	speech acts	informal	informal under.	clarify	audible	fluency
register	NS										
accent	NS	NS									
rate	NS	NS	NS								
nvc	NS	NS	NS	NS							
speech acts	NS	NS	NS	NS	NS						
informal	NS	NS	0.53	NS	NS	NS					
informal under.	NS	NS	NS	NS	NS	0.67	0.63				
clarify	0.48	NS	NS	NS	NS	NS	NS	0.67			
audible	NS	NS	NS	NS	NS	NS	NS	NS	NS		
fluency	NS	0.48	0.50	NS	NS	NS	NS	NS	NS	NS	
overall	NS	NS	0.50	0.54	NS	NS	0.47	NS	NS	NS	0.71

All values significant at $p<0.05$

Standardized Patient VI.

register	NS										
accent	0.77	NS									
rate	NS	NS	NS								
nvc	NS	NS	NS	NS							
speech acts	NS	NS	NS	NS	0.54						
informal	0.61	NS	0.64	NS	NS	NS					
informal under.	NS	0.59	NS	NS	NS	NS	0.60				
clarify	NS	NS	0.52	NS	NS	NS	0.66	NS			
audible	NS	NS	NS	NS	NS	NS	NS	NS	NS		
fluency	0.59	NS	0.70	0.58	NS	NS	0.77	NS	0.73	NS	
overall	0.65	0.56	0.50	NS	NS	NS	0.75	0.58	0.61	NS	0.64
	tense	**register**	**accent**	**rate**	**nvc**	**speech acts**	**informal**	**informal under**	**clarify**	**audible**	**fluency**

All values significant at $p<0.05$

Appendix XXXVI.

Newspaper article from "The Advertiser", Adelaide, November 29, 1997.

NOTE:
This appendix is included on page 447 of the print copy
of the thesis held in the University of Adelaide Library.

Appendix XXXVII.

The Australian Tertiary English Screening Test (AUSTEST).

A User's Guide to the Austest

TEST ADMINISTRATION

After ensuring suitable seating arrangements and the availability of answer sheets and writing implements the administrator should explain the purpose of the test as follows.

"This test has 20 items in four sections and there will be ample time for you to write the answer to each question in the numbered spaces provided. Please be patient if an item is easy and you have time to spare because others may need more time."

SECTION ONE - WORDS.

"I will read out a word and then I will use the word in a sentence. I would like you to write another single word that means the same thing, a word which will fit into the sentence in the same way.
For example the word is 'CUT' and the sentence is 'CUT THE CAKE'. What is another word for 'CUT' in that sentence? Write it in the space marked example. Pause for 15 seconds. If you wrote "slice" or "carve", or a very similar word, that would be correct because they are single words that mean the same thing and would fit perfectly into the sentence. Remember your answer must be just one word each time."

Insert the following list of 10 words in the "frame" and pause 15 seconds after reading each sentence.

"Number ** the word is **** and the sentence is *** *** ***".**

ONE	CONSUMED	THE FIRE CONSUMED THE BUILDING.
TWO	FOSTER	WE WILL FOSTER GOOD RELATIONSHIPS.
THREE	ANTIQUE	THIS IS AN ANTIQUE CHAIR.
FOUR	FRIGHTENED	DOGS ARE FRIGHTENED OF THUNDER.
FIVE	ENORMOUS	THE MOUNTAIN IS ENORMOUS.
SIX	MISERABLE	HE WORE A MISERABLE EXPRESSION.
SEVEN	SYMBOL	THIS RING IS A SYMBOL OF LOVE.
EIGHT	SUSPENDED	THE LIGHT WAS SUSPENDED FROM THE CEILING.
NINE	TEMPEST	A TEMPEST STRUCK THE CITY.
TEN	TANTRUM	THE CHILD HAD A TANTRUM.

SECTION TWO - DICTATION

"I will read out a sentence and I would like you to write it down exactly as I say it. Listen carefully because I can only say it once."

(pause for 45 seconds after each sentence)

1. THE OLD HOUSE WITH THE RED TILES / WAS THE ONE WE ONCE RENTED.

2. THE FILM SHOWING IN THE NEXT THEATRE / ATTRACTS MORE YOUNG THAN OLD.

3. LAST WEEK I HAD LUNCH WITH MY PARENTS / AT THE CAFE THAT ACCEPTS ALL MAJOR CREDIT CARDS.

SECTION THREE - ABSURDITIES.

"I will read out a sentence and I would like you to listen and then write down two things. What does not make sense, and why it doesn't make sense."

"Remember, write down what it is that doesn't make sense and why."

"I can read it only once, and so listen carefully".

(pause for 75 seconds after each sentence)

1. THE SHAREHOLDERS WERE DELIGHTED THAT THEIR DIVIDEND HAD BEEN REDUCED BECAUSE OF THE BIG INCREASE IN PROFITS.

2. BY A STRANGE COINCIDENCE THE MOTHER AND HER CHILD HAPPENED TO SHARE THE SAME BIRTHDAY, THE 6TH OF JUNE 1953.

3. IT HAD RAINED HEAVILY ALL WEEK AND SO THE GARDENER THOROUGHLY WATERED ALL THE OUTDOOR PLANTS.

4. THE COFFEE WAS DELICIOUS AND INEXPENSIVE AND SO THE MAN ORDERED TEA TO BE SERVED TO HIS FAMILY INSTEAD.

SECTION FOUR - PROVERBS.

"I'll read out a sentence, and I'd like you to explain in your own words what it means".

(pause for 45 seconds after each sentence).

1. STRIKE WHILE THE IRON IS HOT.
2. ALL THAT GLITTERS IS NOT GOLD.
3. TOO MANY COOKS SPOIL THE BROTH.

"Thank you, please remember to put your name and number on your answer sheet...etc....etc" as is appropriate to manage the session.

Appendix XXXVIII.

Students with English as a Second Language.
Supplementary Tutorials, 1995 - Evaluation Questionnaire Semester II.

The year is nearly at an end. This is your last chance to give me feedback on the Supplementary tutorials.

Please be assured that your responses are **strictly confidential**, and are between you and me only. I will not divulge your personal responses to anyone else for any reason. I will use the information you give me to improve my teaching, and hopefully, your learning. Any reports that are produced from what you tell me will describe you as a group of students, but never as identified individuals.

If this questionnaire does not cover an issue that you would like to discuss, or if you need more space for comments, you may use the space below. If you want to discuss any matters with me, either related to this questionnaire or otherwise, remember that you are always welcome to visit me in my office to do so.

Also remember that over the next 5 years my door is always open for you. If you want to come and see me for whatever reason, even just to say "hello", **you are welcome!**

Good luck for your exams and for the future.

Name .....

1. On average, how many supplementary tutorials did you attend per week in Semester II?
   One Two Three More than three

2. How much difficulty did you have in Semester II DPS in terms of;
   i) reading and understanding the textbook?
   Much difficulty Some difficulty No difficulty Unsure
   Comment . . .

   ii) understanding the lecture material **before** the supplementary tutorials?
   Much difficulty Some difficulty No difficulty Unsure
   Comment . . .

   iii) taking lecture notes **without** assistance from the supplementary tutorials?
   Much difficulty Some difficulty No difficulty Unsure
   Comment . . .

   iv) writing practise examination essays?
   Much difficulty Some difficulty No difficulty Unsure
   Comment . . .

   v) writing the practical reports?
   Much difficulty Some difficulty No difficulty Unsure
   Comment . . .

vi) contributing verbally (asking and answering questions) in "regular" tutorials?

Much difficulty Some difficulty No difficulty Unsure

Comment . . .

3. Overall, do you feel that DPS is a difficult subject for you?

Yes, very difficult Somewhat difficult Not at all difficult Unsure

4. If you answered "yes, very" or "somewhat" difficult to Question 3, please explain why you think DPS is difficult for you.

5. How useful were the supplementary tutorials in Semester II;

(i) in helping you to take better lecture notes?

Very useful Somewhat useful Not useful Unsure

Comment . . .

(ii) in helping you to understand the lecture material?

Very useful Somewhat useful Not useful Unsure

Comment . . .

(iii) in helping you to write practise examination essays?

Very useful Somewhat useful Not useful Unsure

Comment . . .

(iv) in helping you to write the practical reports?

Very useful Somewhat useful Not useful Unsure

Comment . . .

(v) in helping you contribute verbally (asking and answering questions) in tutorials?

Very useful Somewhat useful Not useful Unsure

Comment . . .

6. Are there any ways in which the supplementary tutorials could be improved?

yes no unsure

7. If you answered "yes" to Question 6, please give your suggestions for ways in which the tutorials could be improved.

8. Have you found the Glossary of terms useful throughout this year?

Yes No Unsure

Comment . . .

9. Have you found the Colloquialisms handbook useful?

Yes No Unsure

Comment . . .

10. Have you added to the Colloquialisms handbook throughout the year?

Yes No

Comment . . .

Thank you for your co-operation in completing this questionnaire.

Appendix XXXIX.

Students with English as a Second Language.
Supplementary Tutorials, 1996 - Evaluation Questionnaire Semester II.

The year is nearly at an end. This is your chance to give me feedback on the Supplementary tutorials that I have conducted over second semester.

Please be assured that your responses are **strictly confidential**, and are between you and me only. I will not divulge your personal responses to anyone else for any reason. I will use the information you give me to improve my teaching, and hopefully, your learning. Any reports that are produced from what you tell me will describe you as a group of students, but never as identified individuals.

If this questionnaire does not cover an issue that you would like to discuss, or if you need more space for comments, you may use the space below. If you want to discuss any matters with me, either related to this questionnaire or otherwise, remember that you are always welcome to visit me in my office to do so.

Also remember that over the next 5 years my door is always open for you. If you want to come and see me for whatever reason, even just to say "hello", **you are welcome!**

Good luck for your exams and for the future.

Name .....

1. On average, how many supplementary tutorials did you attend per week in Semester II?
   One Two Three More than three

2. How much difficulty did you have in Semester II DPS in terms of;
   i) reading and understanding the textbook?
   Much difficulty Some difficulty No difficulty Unsure
   Comment . . .

   ii) understanding the lecture material **before** the supplementary tutorials?
   Much difficulty Some difficulty No difficulty Unsure
   Comment . . .

   iii) taking lecture notes **without** assistance from the supplementary tutorials?
   Much difficulty Some difficulty No difficulty Unsure
   Comment . . .

   iv) writing practise examination essays?
   Much difficulty Some difficulty No difficulty Unsure
   Comment . . .

   v) writing the practical reports?
   Much difficulty Some difficulty No difficulty Unsure
   Comment . . .

vi) contributing verbally (asking and answering questions) in "regular" tutorials?

Much difficulty Some difficulty No difficulty Unsure

Comment . . .

3. Overall, do you feel that DPS is a difficult subject for you?

Yes, very difficult Somewhat difficult Not at all difficult Unsure

4. If you answered "yes, very" or "somewhat" difficult to Question 3, please explain why you think DPS is difficult for you.

5. How useful were the supplementary tutorials in Semester II;

(i) in helping you to take better lecture notes?

Very useful Somewhat useful Not useful Unsure

Comment . . .

(ii) in helping you to understand the lecture material?

Very useful Somewhat useful Not useful Unsure

Comment . . .

(iii) in helping you to write practise examination essays?

Very useful Somewhat useful Not useful Unsure

Comment . . .

(iv) in helping you to write the practical reports?

Very useful Somewhat useful Not useful Unsure

Comment . . .

(v) in helping you contribute verbally (asking and answering questions) in tutorials?

Very useful Somewhat useful Not useful Unsure

Comment . . .

6. Are there any ways in which the supplementary tutorials could be improved?

yes no unsure

7. If you answered "yes" to Question 6, please give your suggestions for ways in which the tutorials could be improved.

8. Have you found the Glossary of terms useful throughout this year?

Yes No Unsure

Comment . . .

9. If you have any further comments or suggestions please list them here:

Thank you for your co-operation in completing this questionnaire.

Appendix XXXX.

Students with English as a Second Language.
Supplementary Tutorials, 1995 - Evaluation Questionnaire Semester II.
Open-ended reponses.

1. On average, how many supplementary tutorials did you attend per week in Semester I?

< One One Two Three More than three

2. How much difficulty did you have in Semester I DPS in terms of;

i) reading and understanding the textbook?

Much difficulty Some difficulty No difficulty Unsure

Sentences rather complicated.
Lots of the material is very difficult to understand. I've to get to the supp material before I can understand the textbook.
So much "big" words, I have to use the dictionary a lot. By doing so I lost interest in continue reading.
Some of the facts, I can't relate them to real life (can't apply).
So far, whatever, textbook material I have read made sense to me and was fairly easy to read.
I can understand most of the content but the main point in the book is not clear. The textbook is also boring probably because the way the material is presented is not very clear.
I'm not really understand the psychology, the words that written in the textbook is quite hard to understand. I often have to look in the dictionary. However, this is a very good experience for me.
Contents of book are a bit harder.
The language.
Written very wordy.
Long sentences with bombastic words.
Especially the front chapters - all chapters which is taught by X.
My vocab knowledge is incompatible with the words/terms in the textbook.
Lots of jargons and complicated sentences.
Found text book difficult to understand.
Quite hard to understand if just read the book without going to lecture/supps.
The lecture and supp tute give more info than the book. The red book sometimes were presented with diff seq than the lect/supp.

ii) understanding the lecture material before the supplementary tutorials?

Much difficulty Some difficulty No difficulty Unsure

The lecture was not well organised and used a lot of jargons and difficult words.
I tried to understand it but most the lecture notes was just jumping around.
Sometimes it was too boring.
It really depends on the lecture material. Sometimes it was quite okay and sometimes it was quite hard to understand. I guess supp tute made it easier when more simpler examples were given.
Depends on the lecture; some of the lectures were new and not from the textbook. Therefore difficulties may arise.
Depends on the topic.
I think I can understand the lecture allright if I did concentrate, but my problem is that I do lots of daydreaming during lecture. Therefore I did not follow what Dr X talking about.
Due to the lectures being relatively unstructured, it was diff to make sense of what the main points of the lecture were.

Sometimes I have problem to understand the lecture when the lecturer use word that I don't understand.

It is very-very hard, but sometimes it depend on what topic are we discuss about. Overall, I'm not really understand, since we deal with psychology stuff this semester.

Jumps around too much, not sure what is important.

Lecturer tend to jump all over the place. No clear introductions.

Some of the lectures was "disjointed" and were hard to understand.

Especially guest lectures.

Depends upon a topic covered.

Don't get what lecturer said in the lecture.

iii) <u>taking lecture notes without assistance from the supplementary tutorials?</u>

Much difficulty **Some difficulty** **No difficulty** **Unsure**

I really had to pay full attention to the lecture notes if I really wanted to take them down. If not I would miss some information.

Supp clarify my understanding.

Compared to Dr X's lectures, Dr X's lectures were somewhat unstructured. Writing notes required mental structuring/organization of notes. Thus, took time and I may have left out important points.

The lecture is very fast without stressing the important point. Most of the times, I can't generate my own sentences to sumarize the example.

The lecturer was not really organize so it is very hard to take note straight from the lecture.

Lecturer goes too fast.

Supp tutes help me to rearrange my notes.

Found it difficult to pick up the facts of the lecture.

Not clear of what being said in the lecture.

Because I didn't read beforehand the book, therefore hard to write the lect note.

iv) <u>writing practice examination essays?</u>

Much difficulty **Some difficulty** **No difficulty** **Unsure**

I was not sure exactly what was required.

Haven't write one yet.

I've the idea sometimes but it has to be proved by you first.

No comment at the moment because haven't try any past year questions.

I haven't started yet.

I have not done many yet.

I have not written out any of the exam q's in prose form. For all of them I have put down points which I think are appropriate.

Not really difficult because I can ask my tutor and also my senior, however ACH always the best!!!

Can't understand some questions.

Found it difficult to integrate all aspects of the course to write a good essay. Not always sure what they want.

I had a lot of problems at the beginning of this school year, but now I am getting what DPS essays request.

v) <u>writing the practical reports?</u>

Much difficulty **Some difficulty** **No difficulty** **Unsure**

It's not difficult with exam essay becoz we've revised it and I've you to discuss with, but sometimes I've no idea how to plan the question.

Luckily, with the help of a few people, I manage.

Frankly speaking, I prefer to write reports using lots of subheadings. However, I didn't think I had any difficulties in writing the practical reports, although to get them started was quite hard.

I didn't do any reports this year. I did not encounter any major difficulties last year.

Discussion clarify understanding.

I think I could integrate the course work and understand the aims of the reports.

Not really difficult. I just feel a little bit confuse when I want to choose the title and also to choose the relevant data for the topic. Overall, it is not really difficult.

Don't have to do it.

(No difficulty) with the help of Anna!

Found it good to be able to go through the rough copy for any fundamental problems.

I've already done it last year.

vi) **contributing verbally (asking and answering questions) in "regular" tutorials?**
Much difficulty Some difficulty No difficulty Unsure

Except that sometimes I feel hesitate to give out opinions because the way I perceive things are differ compared to the others (esp Aussies).

I always try to speak during the tutorial.

Sometimes, too many people/tute.

Well, I have to say that I am sort of a quite person by nature. That doesn't mean that I had much difficulty in contributing verbally in tutorials. It depends on the session. If, let say, in small group discussion (in tutorials) of course I'd speak something.

I got nothing to ask but then hard to get into class discussion as they're talk really fast. Hard to get in.

I often not really understand the lecturer so I used the supplementary tutorials to complete my notes and at the same time to understand the lecture. So, at that time I am not really prepared to ask any questions probably because I'm not really sure about what am I going to ask.

(some difficulty) but I feel I've improved compare to 1st semester.

Difficult to jot in.

Found it difficult to add my point of view when others were discussing a point.

When you don't know what to talk about.

I am very embarrassed due to English to speak up in front of others.

Tutorial group is a large group. Uncomfortable to talk.

3. Overall, do you feel that DPS is a difficult subject for you?
Yes, very difficult Somewhat difficult Not at all difficult Unsure

4. If you answered "yes, very" or "somewhat" difficult to Question 3, please explain why you think DPS is difficult for you.

Have terminology which is not applicable in 'medicine world'. (Not all eg schizo, depress, hypertension are used) - but other terms like operant conditioning.

Remembering things.

Since I got much work to do throughout this sem (eg reports etc), I don't have much time to make a revision. Perhaps if I studied constantly it is not as hard as I thought.

Because we need to think critically and we cannot really "see" it.

It's an interesting subject for me, but to prepare for the exam, it makes life miserable. Material in second semester it's more difficult than first semester.

To get informations from the red book is very difficult.

Understanding the red book. In fact, it took +/- two hours/chapter, sometimes more.

Sometimes, I do not have idea to integrate and write the essays. Sometimes it is easy to integrate whole topics together.

Because there are so much to understand and to remember and most of them are words, or should I say new words which I don't use in daily life.

The textbook is not readable. The lecturer is the worst in the year. But luckily subtut tutor is good.

You have to use your general knowledge. You have to write essays for the exam.

It involve writing essays and there is no correct answers so we have to learn a lot of pro and cons.

Maybe, I'm not really good at writing essays. Hard for me to start an essay but once I got the idea I think I can manage.

Because it's individual stuffs like emotions etc. That = very complex.

I think DPS is difficult mainly because my English is not good. Sometimes I found it very difficult to put all the point in the essay as the one you show me in the 1st semester. I think that, in DPS you get more marks if you good at how to put the points.

It is (somewhat difficult) because of language, I'm not really good at English, it is very hard for me to understand. I often take a long time to understand the lecture. The lecture is <u>not</u> really organized. It is quite hard to understand. However, I <u>really</u> like DPS, because this subject is very interesting and it make me think!!! actively. Since I learnt DPS, I start to understand - what does it means - by "human behaviour".

The questions in exams are not directly from the books or the lecture or anywhere. I have to intergrate the materials from those resources and have difficulty to priorities facts in answering the exam questions.

It's not the 'just memorising' subject like I'm used to. It requires critical thinking and integration of different lectures.

Somewhat difficult; some section in the textbook is difficult to understand (esp the one we have to read in sem II).

It is not very difficult but since doing psychology for the first time, it is difficult in the start to understand the concepts and their relevance. Later in the course there is a better understanding.

Can't really understand the lectures without going to the supp tute. The subjects prob. more complex than in previous sem.

As it needs a lot of intergration from various topics in DPS. It's hard to answer DPS questions (Sem II). I find it quite hard to understand the lecture material presented in second semester.

Language difficulties ie the meaning of the lecture is found sometimes vague. Lots of readings esp the red book is somewhat difficult to understand.

The materials themselves were quite difficult. Besides the text book - also difficult to understand.

The text book is my main problem. Sentences are just too complicated (sometimes) and lots of big words are used in the text book which slow down my learning.

Lecture material sometimes hard to be integrated.

DPS load of homework is too much, especially some the tutorials.

Probably because the only referrence book that contain the lecture material is hard to understand especially from Prf X's part.

Too much to read on complicated facts. Overloading of knowledge at once.

The work this semester seems to have gotten a lot harder. Perhaps with more expected eg more reports, tutes to do etc. Some concepts were easy to grasp, others not so. The amount of information to learn, also seems a little much.

Too much at one time. Reference book (the Red Book) quite difficult to understand.

I am familiarized with some aspects of psychology (my previous overseas medical course) but these aspects (particularly american studies and approach are new for me completely).

Because, the things that we learn in DPS usually is something that can't be seen. No real mechanism but integration of many other factors. Not as straight forward as learning about diseases.

Very boring to study. Not much interesting stuff. Just need to memorize all the facts and think about it only.

5. How useful were the supplementary tutorials in Semester I;

(i) <u>in helping you to take better lecture notes?</u>

Very useful Somewhat useful Not useful Unsure

Anna is a good lecturer.

For me the supp tute helped me to understand more of the lectures. This is because the way the lecturers gave the lecture and the way you explain the lecture were not the same. Of course taking down your notes was much more easier. But that doesn't mean I could take better lecture notes.

Helps in organizing the lecture materials better. Therefore have bett cognitive manipulation for the materials.

Sometimes I don't know what we are supposed to get out of the lecture.

It was very much appreciated this semester than last semester, due to Dr X's nature of lecturing.

It is very easy for me to take notes from you because you put the lecture notes in a point form.

I really need supplementary tutorial. However, I always can't comes to supplem. tut because of my schedule is too /, especially on Monday and Thursday because my HSF tut always changing.

Better structured notes.

Identified the important facts that I should have wrote down.

Especially things that do not covered in the Red Book eg guest lectures.

Anna, you teached me to connect things within material and use them as the facts to support statements in any moment.

Lecture and the book is very hard to understand and used vocab that is unfamiliar.

(ii) <u>in helping you to understand the lecture material?</u>
Very useful Somewhat useful Not useful Unsure

Not only helping me to understand the present notes but also the previous notes. Therefore I can integrate information much better.

You've done a wonderful job, especially bring up "supposedly" (!!) past knowledge and integrating them into the present lectures for clarity, comprehension and application of knowledge.

Sometimes you explaine the meaning of words that I don't understand.

Very useful Anna, you always make thing easier.

The information given in supp tutes were organised to that they join and make sense.

More clear approach to the topic.

(iii) <u>in helping you to write practise examination essays?</u>
Very useful Somewhat useful Not useful Unsure

I haven't practice writing yet.

Helps give more insight into how to tackle the questions better.

I have not tried to write practice examination essays yet.

Tutorial notes often clearer and easier to understand.

Had more clear understanding on the questions.

Identified what needs to be put into exam answers, but perhaps just needs more practice.

Did not tried it before (for sem II).

(iv) <u>in helping you to write the practical reports?</u>
Very useful Somewhat useful Not useful Unsure

Better if the discussion on the practical reports were done early; not when only one week left to submit the report.

Anna guides us.

Gave me the idea of how to write the reports (ie the right format etc).

If I were to do practical reports I can say that the supps will be very useful.

Can get better and clearer understanding of what's required.

However, sometimes, my tutor Anne and you give different format on how I should write my report. Anyway, both of you are very helpful.

Referred a lot on previous students' reports to get general idea.

Identified areas that are important in an essay/report.

(v) **in helping you contribute verbally (asking and answering questions) in tutorials?**

Very useful Somewhat useful Not useful Unsure

It doesn't work.
My fault. I did not participate the way I should ie use all the chance I got.
After got the clarification, I am more confident to talk about it.
I think previously I contributed a fair amount to tutorials, and if I am contributing more now I don't think it's because of the supp tutes but because of familiarisation with my tute group.
Still a little nervous about butting in to make a comment.
To talk in tutorial depends on the individual.

6. Are there any ways in which the supplementary tutorials could be improved?
yes no unsure

7. If you answered "yes" to Question 6, please give your suggestions for ways in which the tutorials could be improved.

Have a session which encourage students to participate eg asking questions.
Maybe instead of repeating the whole lecture we can just ask questions about what we missed - to save time.
Too many people in one tutorial sometimes is quite distractful and make me uncomfortable to contribute verbally.
Maybe, at the end of each month, there is a revision class so that it won't happen likes what happen now. Sometimes we've no idea about the past stuff.
Students involvement should be encouraged esp. in expressing their ideas.
Should do practice exam questions throughout the semester/year.
Give homework for students (esp those with major difficulties and have those works mark).
Smaller group. ask more questions.
Do next time just like what you did last time. Ask each person the lecture content.
I think the idea of having the tutes in smaller groups is pretty good, as then most people (or all) can contribute and effective discussion can be easier to structure.
I always use supplementary tutorial to complete my note, and sometimes I'm not really I understand, and when I want to ask question, I feel not really confident, because, I'm not really understand. So, I really hope, you can spend 10-20 minutes, a short discussion about the topic for my junior.
Every topic that had been covered, practiced at least 1 DPS essay questions - make it a compulsory for student to pass up the essay.
Do it under the shades (something outdoor and greeny) eg Bot garden.
Anna's recent approach, asking questions in supp tute was a v. good idea.
Fix the number of students at one class and fix their time. Less students will encourage more participation.
The time is certainly hard for me to attend because there have been a lot of changes in my timetable where PBL and PBRP are shifted to the slot where supp tute was. Sometimes the only time available is during the lunch hour, which i often spend to "sleep" somewhere. I don't complain to you becoz no need to change the whole time for tute just becoz of one person.

8. Have you found the glossary of terms useful throughout the year?
yes no unsure

Sometimes to make sure of some meanings.
Especially when writing the essays.
In order to find the definition and makes me more understand about the material of DPS.
Haven't really look at it.

I tried to use it while doing the practical reports although there was a time when I couldn't find the desired words in it.

Esp. while doing practical essays.

The glossary gives more better definitions on some terms than in the textbooks.

Not enough terms included.

I use it a lots when writing up report - the definitions are good.

Easy to write report.

Helpful in writing reports, reading of the text and exam revision.

Save my time, to find out the defination of the term that I have learnt in ?

Mostly in essay writting and report writting.

It's quite difficult to understand it.

Easy to find the words mention in the lect eg illusion and hallucination.

9. Have you found the Colloquialisms handbook useful?
yes no unsure

It is useful but I don't have much time to read it as we got lots of reports to do throughout the year. Perhaps, I will read it during the holiday.

Yes, it's a useful book, but, I think we'll feel it's more useful when we're dealing with patients.

Haven't really look at it.

I didn't really open the handbook after the last time we discussed about it in the supp tute (discussion group).

The thing is, I used some of the sentences once awhile but not all the times.

Haven't read through yet.

Not yet

Esp call turkey: during PBL session.

Personally I think it is useful and is a good idea, however, I have not looked it up to define any colloquials as such, apart from perhaps once or twice.

I learnt so many new words. I can use it during my interaction with my friend and also for my examination.

Not for exams, but it takes time for me to become familiar with the usage.

I used it when watching TV prog (Aussie). I hate Aust accent!!!

10. Have you added to the Colloquialisms handbook throughout the year?
yes no unsure

I learn from my guardian family (they're Australian).

It's a very useful book.

No intention - too bad.

I haven't encountered many slangs in everyday communications ie with friends, in PBL, in tute etc.

Haven't done so.

Sometimes.

My English teacher uses colloquials.

I may have added new words to my vocabulary, but being unfamiliar with the handbook I cannot say whether I have in fact added to it.

(Yes), but not really much, just 3-4, because, I don't have time. Probably, I can do it during this holiday.

Did not refer much.

Further comments.

Second semester DPS - a bit hard sometimes, the supp tute notes were too long. Red book is too hard to be understood. Lucky we have Anna . . .thanx very much. Love you!

Thank you Anna. You've been such a very helpful and nice tutor.

Thank you very much, I'm sorry I sometimes miss your supplementary tutorials because I always got sick.

Anna, thank you very much for your assistance throughout the year! It was a great pleasure to have you. See you!

Appendix XXXXI.

Students with English as a Second Language.
Supplementary Tutorials, 1996 - Evaluation Questionnaire Semester II.
Open-ended reponses.

1. On average, how many supplementary tutorials did you attend per week in Semester II?

One Two Three More than three

2. How much difficulty did you have in Semester II DPS in terms of;

i) reading and understanding the textbook?

Much difficulty Some difficulty No difficulty Unsure

Language, padding, confused sentences, high vocab use.
But with the lectures and supp tutorials I could overcome the problem.
Some topics are OK but some are quite difficult to understand.
Difficult to understand some words, therefore have difficulties in understanding the whole sentence, let alone trying to get to the point.
A bit difficult in understand X's topic.
It's not difficult to read but to understand about certain topics eg studies by Milgram and Asch, the explanations are quite vague.
However, there is certain parts in the textbook really need further explanation in order to be understood.
Not direct and not "student-friendly".
Lots of stuff in there. I don't really know what we need to know for the exams.
Slight difficulty sometimes due to confusing writing style.
Dr X stuffs!!! It's too (the english) is too high standard and not direct.
X's part is somewhat vague but sub tute really helped put it in context.

ii) understanding the lecture material before the supplementary tutorials?

Much difficulty Some difficulty No difficulty Unsure

Confused organisation and presentation of lectures. Lectures also fast.
Some lecture.
Dr X's lectures were quite OK to follow, but Dr X's gave me "much difficulty".
This may be due to the lecturer jumping from one point to another making it difficult to see the coherency.
Depends, if I managed to focus my full attention, my understanding is OK before supplementary tute.
Some of the lectures I did not understand much but most of it, I had some idea.
Depends on lecturer and the material that was being discussed.
Some of the lecture materials are fairly easy to understand. But sometimes, the points are not sequential. Therefore it's quite hard to determine which are the main points.
Especially lectures given by Dr X, the subtopics she lectured were not clear as she seldom clarify initially what she was about to lecture.
The level of difficulty still depends on how the lectures material was being presented because if the lecturer tend to jump from one point of discussion to the other, then it would be difficult to follow.
Difficult to grasp the subject content. Boring. Lecture not direct or clear.
X lectures alright. X and X - not much structure in lectures, hard to determine what's important.
It depends on which lecturer gave the lecture.
I don't usually read beforehand.
(much difficulty) X's!

iii) **taking lecture notes without assistance from the supplementary tutorials?**
Much difficulty **Some difficulty** **No difficulty** **Unsure**

Supps provide easyness to catch up with what's being catch-up with (sorry, terrible English use).
Need assistance for Dr X's lectures!!
Not sure of what is important and what should be left out.
That is my most main, cardinal problem.
As said earlier, the main points said during the lectures are not in order.
Taking lectures was sometime not too difficult and easy to understand. But when I came for supps I could gain extra understanding.
It is quite difficult in picking out necessary points from certain lectures cause some of the lectures were not properly structured.
Lecture not well organise and not direct to the point. Boring.
Miss out points (especially when the lecturers gave unorganized lecture).
Lectures are sometimes a bit fast and I tend to leave out some of the important points. Some are not that systematic. Having supp tutes - my notes are well arranged.
X - alright. X and X - lectures not much structure, therefore not much structure in notes.
It depends on which lecturer gave the lecture.
Lecture contents too scattered, no structure.
(much difficulty) esp X's!

iv) **writing practice examination essays?**
Much difficulty **Some difficulty** **No difficulty** **Unsure**

But right now it's quite okay as I've been told/explained how to structure my answer.
Find it hard to know which are the important points for the essays.
I haven't done any (actually writing them down) but I have discussed it verbally with friends.
Very helpful, could do with more.
I haven't attend any of these sessions.
One of the major reason its difficult is that I'm not prepared. Another, is I sometimes find it hard to construct brief and succinct sentences without babbling around.
The questions suck! Lecture is not exam oriented ie content is so diffuse that it is difficult to know how to apply the knowledge properly.
Difficult to develop critical thinking and use precise explanation.
I don't tend to take it seriously because I know that it's not a real exam.
If I know the material for the essay I can write it quite easily. If i don't know then I have some obvious problems. Also, I don't think I write fast enough as I seem to run out of time towards the end.
No time to learn for it.
I have difficulty in focusing the content that I should write therefore I end up by putting everything in the essay.
Because have to integrate everything so sometimes it's all jumbled up and unclear.

v) **writing the practical reports?**
Much difficulty **Some difficulty** **No difficulty** **Unsure**

After some time, I learnt what the report really wants and I also got the idea of how to do critical thinking. It was fun!
The clinical observation is a bit challenging.
Sentence structures and incorporating ideas.
(no difficulty) But the markers of the reports didn't think so. I don't care about their opinion at all.
It's quite hard to write a report especially when it requires to make an integration with the lecture materials and the textbook.

Again, constructing succinct sentences was difficult. being critical was not very easy sometimes.

Still believe that for the clinical observation essay, the question should be given before the observations were done.

The fear of not writing what the tutors want really worries me. Should provide model answers for similar questions, and advice on regular pitfalls.

To be precise in explanation. To omit the irrelevant findings.

I need to improve on my expression and I think I've tried my very best for the clinical case prac reports.

Have a look at past years' report and use their structure to put your info in.

Not much difficulty but unable to get high marks, my reports are often too long.

I still need some help like proof-reading and how to express my content.

Time management problem!!

vi) **contributing verbally (asking and answering questions) in "regular" tutorials?**
Much difficulty **Some difficulty** **No difficult** **Unsure**

I guess that I am a bit inhibited to talk. Most probably due to my personality.

My problem is talking too much.

Contribute ideas but seemed to be irrelevant to the topics and being objected by others.

I only like to ask when I really don't understand and prefer to answer questions only when being appointed.

Don't know what to say and can't be bothered.

I still have low self-confident (but I'm trying).

3. Overall, do you feel that DPS is a difficult subject for you?
Yes, very difficult **Somewhat difficult** **Not at all difficult** **Unsure**

4. If you answered "yes, very" or "somewhat" difficult to Question 3, please explain why you think DPS is difficult for you.

Textbook uses quite high English, confusing and long sentences. Lectures are mostly too fast and s'times unorganized. Difficulty in providing critical evaluation in exam answers, but nevertheless, most of the subject require common sense just difficulty in critical thinking.

It needs critical thinking and therefore full understanding of the topics.

Because firstly there's a lot of critical thinking to do. It's hard writing in English let alone thinking in English. It is also a bit difficult trying to gather what the lecturer really means and what they're really saying, the main points.

Because we have to integrate all the topics in answering question to get a good marks. A bit difficult for me to express myself in a good English.

Sometimes I don't really understand the topic and a little bit confusing.

The incorporation of ideas into essay and linking topics.

It's just the nature of the subject that there is no correct answer for the questions and I have problem in understanding what it is that you guys really want.

Too much emphasis on learning the red book; sometimes it gets a bit boring.

Because there're so many things to be remembered. And, the way it should be answered in exams make it even worse.

Because for me, I think all the lecture materials given basically are based on studies done in Western culture. Sometimes I find it hard to relate to my own culture and therefore it's hard to do a good critical thinking to answer the exam questions.

Major problem is to become a critical thinker.

It is a very language based subject. DPS is also very subjective.

I find it difficult in expressing ideas in English that can be understood by others (especially in exams)

I think that DPS is more towards English language usage. Therefore it's a disadvantage for students who come from non English language background.

The marking scheme is too strict and demanding. The boring lectures (Dr X is an exception here) really kill my interest in psychology. The comprehensibility of the red book is at the extreme low. The interrater reliability among examiners is unsatisfactory.

Need to remember and understand the concepts which are subjectives (ie perception, personality etc). The contents are too much.

I found that the guess lecturer stuffs were very hard to understand and to find out the source of the lecture materials.

Too many materials to cover, understand, remember and at the same time have to complete the assignment (report).

Lots of info to rote learn. 4 essays in 3 hr = very sore hand at end of exam.

Unsure about how to use critical thinking.

Everything is in English. All info abstract.

The quantity of what we have to learn is too much, I find it very difficult to remember any thing. Essay writing skill is yet to be improved, especially in an exam situation where time is limited.

It's abstract. No fixed way to answer exam question (have to integrate everything!).

5. How useful were the supplementary tutorials in Semester II;

(i) in helping you to take better lecture notes?

Very useful Somewhat useful Not useful Unsure

Because most lectures are too fast, not well organized (sometimes). Therefore Anna gives pts that is missed along lectures.

It helps me in better understanding about the lecture materials.

From the supp tute, the tutor gave the main points from the lecture and it taught me how to get the main points from the lecture without having to write down everything the lecturer says.

Helps filter unnecessary topics or discussions.

That's not why I came to the supp tutorials.

Excellent summary of the lecture.

It helps me to put all my lecture notes in order and have a better understanding about all the important points mentioned in the lecture.

Usually, coming to supps, added more notes to my notes.

Now know how to pick out the important points in the lecture.

The notes are clearer and more direct, especially in relating new knowledge with medicine.

Identify the relevant points.

More structured. Highlight important points in lectures. Critical thinking pts also helped.

Sup tuts help me to organise my lecture notes into something that actually mean something; otherwise, it's just a series of unrelated sentences and paragraphs.

I'm more clear and understand the lecture.

Sometimes tutor have somewhat different opinion from lecturer so sometimes end up with 2 sets of lecture material!!

(ii) in helping you to understand the lecture material?

Very useful Somewhat useful Not useful Unsure

You made all the lectures clearer than when it was given by the lecturer.

Group discussions especially.

This was the main reason for attending.

It has been definitely useful. Not only the examples given by the tutor (Anna) are clearer but she also helps me to determine what are the important point.

It help in clarifying the content of the lecture material so that it would make some sense.

It makes the contents of the lecture more understandable.

Better explanations given.

How to be critical in using the material.

Going over the material in sup tutes help me to understand and remember material. Also, definitions of some terms helped to clarify what was said during lectures.

Good organisation of info from the lectures makes understanding a lot easier.

(iii) **in helping you to write practice examination essays?**
Very useful Somewhat useful Not useful Unsure

Provide practise for critical evaluation.

I haven't written any yet.

Critical thinking session.

I got to know about "critical thinking" which you emphasize so much.

I sort of know what the standard is expected of me now.

I haven't attend this session yet.

Because in supp tute, not only I'll manage to understand and remember my lecture notes but I also learn how to make critical thinking that is an essential requirement for the exam.

Good, but more time should be given.

Somewhat clear feedback. I don't really receive some suggestions about the alternatives to provide better answers.

I think after we finish certain topic, we should have examination practise, instead of having lots of topics in one essay practice. maybe after we finish all the syllabus, we can have practise for the overall topics.

Critical thinking aspects to include in essays very useful. +structured notes make revision easier and faster therefore know more therefore better essays.

Haven't done enough of it. I've only done 1 essay so far.

It help me to know what is important or required for my essays esp critical thinking.

Didn't attend some!

(iv) **in helping you to write the practical reports?**
Very useful Somewhat useful Not useful Unsure

Actually somewhat useful, but I just wanna put very useful. It sounds better that way.

The supp tute for writing the reports was useful in getting me started with the reports.

As stated, my main problem is integrating ideas with relevant issues.

It was very good for me that you had a look at my last practical before I handed it, may be I should have done the same thing for the other one as well.

Becoz the examples given in the tute are easier to understand and thereforeeasier to relate it with the practical reports.

During supps, practicing critical thinking helped me to become critical.

Through the tutorial, lecture material and techniques in writing the report were throughly discussed. Thus, increas comprehension and it will be easier to put out the ideas.

Important notes were given but should inform students about the possible pitfalls too.

Whether or not the text/work was in the right track.

Structure was helpful but we also went thru structure in normal tute and can get get structure from past reports.

Sup tuts rarely touches on practical reports, nornal tuts are more useful in this aspect.

I get to know how to write those essays earlier and have someone to ask/turn to, if I have problem/not clear.

(v) **in helping you contribute verbally (asking and answering questions) in tutorials?**
Very useful Somewhat useful Not useful Unsure

Coz it's different between supps and tutes. Supps don't have big mouthed, talkative whites. Therefore harder to gain confidence to talk.

Increased my self confidence - helped my participation in PBL sessions as well!

Tutorials are most of the time boring. X (tutor) always tells jokes just to get the other's attention and trying to be cool. So not much critical thinking could be learned (from my point of view).

Because it functions as a semi-discussion group.

Sometimes the feedbacks about our comments were irritable and the comments seemed to be humiliating and turned me down.

I don't ask many questions in tutes as I don't have any questions.

Sup tuts and normal tuts are totally different, the material covered in each tut is also different.

I gain more confidence.

6. Are there any ways in which the supplementary tutorials could be improved?
yes no unsure

7. If you answered "yes" to Question 6, please give your suggestions for ways in which the tutorials could be improved.

Perhaps give us more chance to try answering question for the exam.

I think that it would not be a bad idea if the students had more chance of having a group discussion rather than repeat of lecture material. I really don't know if this would be practical or not?

Practice past year questions once a week.

I don't know. As for me it is excellent already. But of course, every thing can be improved.

It will be more helpful if the supplementary tutorials are carried out in a smaller group so that everyone has the chance to voice out their confusion or difficulties in understanding the lecture materials.

It is just right.

Try to link new info/data provided to past exam questions. You have tried that sometimes, but more emphasis should be put on this.

Provide clear alternatives when the points given by the students are not really correct, but the way this is done is in the positive/supportive atmosphere.

There's not much more you can do.

It is already good enough.

8. Have you found the Glossary of terms useful throughout this year?
Yes No Unsure

Very useful for report writing, some usefulness for textbook.

But there are some other words that were not in there and couldn't find it in the english dictionary either.

Actually only in the 2nd semester. The supp tutes directed my study towards the glossary.

It is very good and I found it very useful both for understanding the red book and also in my practical papers.

It was bloody good. Thanks Anna.

I think you might want to add few more words in it. Some of the words (that are no included) might seem easy but s'time are not familiar to us.

Honestly, I never used the Glossary of Terms because I personally think that the supplementary tute have answered all my queries about the lecture materials or jargons that I do not understand.

It provide a rather standardized definition of the term in the subject.

It helps no doubt.

To give definitions in the report. To understand the psychiatric jargons.
I only used it a few times when writing the communication skills report.
It's very useful in helping me writing my essay and reading the textbook.
Sometimes!

9. If you have any further comments or suggestions please list them here:

I think that you should continue with this program because I think thsat it's very helpful to me. Thank you very much!! and Forever grateful to you!!

Congratulations Anna, on your promotion!!

I'm grateful because you're willingly to have the supplementary tutorials for all of us even though sometimes we're pain in the "ass"! Maybe you can get a day, where the past years Q can be discuss but providing the students have gone through/have a look at the Q's. Lastly, thanx a lot! If in Malaysia, people will say this to you: May God bless you Anna!

Good job Anna?

It would have been very good if the supp tutes were present in the 1st semester as well, so that, we could get used to the thinking style demanded by DPS I.

I believe that the tutorials were helpfull and I would like to thank you for doing a very good job (I am not trying to be nice, you know me, this is not in my nature!!). I certainly feel more confident and ready for exams and hope to perform better in the exams. I really admire your close relationship with the Malaysian students and it is good for them to have someone like you because they are away from home in a rather strange environment. I hope that you keep up your good work next year. Thanks a lot for everything.

The DPS supplementary tutorials were perfect except you tended to concentrate too much on negative first impressions. Hmmm.

I think, more discussion with critical thinking practice should be done where we could get into groups and contribute. (But this may sometimes leave out those who are more quiet).

Thank you for your efford for doing the supp tute. It's very valuable and beneficial. Your effort is really appreciated!!

I would go crazy if I try to list them down here because it would take ages for me to complete it. My advice is that the subject, the teaching style and the marking scheme should go through a total revamp. My grading for this subject is no higher than the lowest possible grade given.

Don't shout in the class (tutor). It is good for students to ask questions during the class, but sometimes that made the class went over time and we/some of us would be late for the next class (ie the other obligatory tutorials). If possible, be punctual at the beginning and the ending of the class. The elaboration about the lecture materials were well presented. This effort should be improved and maintained for the future.

I suggest that you allocate fix time for the tutorial. Perhaps you could discuss it with the people who organize the timetable. For example, it is not very appropriate to have tutorial at 5 p.m. coz it's sort of late. I think you should start the Supplementary tutorials in first semester, because for me, I think I'm a little bit lost/do not understand lectures and having difficulty in taking lecture notes (especially for Dr X lecture) in first semester.

The sup tutes helped me a great deal this semester esp in lecture notes and critical thinking for exam essays. Keep up the good work!

I think it'll help me to do well in exams. Thank you sincerely!!!

References

Abdulrazzaq, Y. M. and K. I. Qayed (1993). "Could final year school grades suffice as a predictor for future performance?" Medical Teacher 15(2/3): 243-251.

Abrahamowicz, M., R. M. Tamblyn, J. O. Ramsay, D. K. Klass and M. L. Kopelow (1990). "Detecting and Correcting for Rater-induced Differences in Standardized Patient Tests of Clinical Competence." Academic Medicine 65(9): S25-S26.

Accardo, P., C. Haake and B. Whitman (1989). "The Learning-Disabled Medical Student." Developmental and Behavioral Pediatrics 10(5): 253-258.

Accardo, P., C. Haake and B. Y. Whitman (1990). "Letter to the Editor." Developmental and Behavioral Pediatrics 11(2): 103.

Ackerman-Ross, F. S. and N. Sochat (1980). "Close encounters of the medical kind: Attitudes toward male and female physicians." Social Science and Medicine 14A: 61-64.

Advisory Centre for University Education (1991). Report on the Overseas Student Questionnaire. Adelaide, South Australia, The University of Adelaide.

Aguilera, D. C. (1967). "Relationship between physical contact and verbal interaction between nurses and patients." Journal of Psychiatric Nursing 5: 2-21.

Ahmad, B., L. B. Ahmad and M. M. Al-Jouhari (1988). "Factors determining the performance of medical students of the Faculty of Medicine, University of Kuwait." Medical Education 22: 506-508.

Ahmad, W. I. U., E. E. M. Kernohan and M. R. Baker (1989). "Patients' choice of general practitioner: influence of patients' fluency in English and the ethnicity and sex of the doctor." Journal of the Royal College of General Practitioners 39: 153-155.

Ahmad, W. I. U., E. E. M. Kernohan and M. R. Baker (1991). "Patients' choice of general practitioner: importance of patients' and doctors' sex and ethnicity." British Journal of General Practice 41: 330-331.

Ainsworth, M. A., L. P. Rogers, J. F. Markus, N. K. Dorsey, T. A. Blackwell and E. R. Petrusa (1991). "Standardized patient encounters. A method for teaching and evaluation." Journal of the American Medical Association 266(10): 1390-1396.

Albanese, M., C. Prucha, J. H. Barnet and C. L. Gjerde (1997). "The Effect of Right or Left Placement of the Positive Response on Likert-type Scales Used by Medical Students for Rating Instruction." Academic Medicine 72(7): 627-630.

Aldous, C. J. H., S. R. Leeder, J. Price, A. E. Jervie Sefton and J. K. Teubner (1997). "A selection test for Australian graduate-entry medical schools." Medical Journal of Australia 166(5): 247-250.

Alfayez, S. F., D. A. Strand and J. D. Carline (1990). "Academic, social and cultural factors influencing medical school grade performance." Medical Education 24: 230-238.

Anastasi, A. (1988). Psychological Testing. New York, Macmillan Publishing Company.

Anderson, D. (1991). "Universities and the Professions." Australian Educational Researcher 18(1): 1-15.

Anderson, M. B., P. L. Stillman and Y. Wang (1994). "Growing Use of Standardized Patients in Teaching and Evaluation in Medical Education." Teaching and Learning in Medicine 6(1): 15-22.

Andrew, B. J. (1977). "The Use of Behavioral Checklists to Assess Physical Examination Skills." Journal of Medical Education 52: 589-591.

Argyle, M. (1969). Social Interaction. London, Tavistock.

Argyle, M. (1971). The Psychology of Interpersonal Behaviour. London, Cox & Wyman.

Argyle, M., A. Furnham and J. A. Graham (1981). Stressful situations. Social Situations. Cambridge, Cambridge University Press: 319-357.

Argyle, M., A. Furnham and J. A. Graham (1981). Language and speech. Social Situations. Cambridge, Cambridge University Press: 293-318.

Aubert, E. (1994). Asian-born medicine boom. Campus News: 3.

Ballard, B. and J. Clanchy (1984). Study Abroad. A Manual for Asian Students. Kuala Lumpur, Longman.

Ballard, B. (1987). "Academic Adjustment: The Other Side Of The Export Dollar." Higher Education Research and Development 6(2): 109-119.

Ballard, B. and J. Clanchy (1988). Studying in Australia. Melbourne, Longman Cheshire.

Ballard, B. and J. Clanchy (1991). Teaching Students from Overseas. Sydney, Longman Cheshire.

Ballard, B. (1992). Teaching international students: A new dimension in problem solving. Quality in International Education. An Academic Perspective., University of Southern Queensland.

Ballard, B. (1995). How Critical is Critical Thinking? In T. Crooks and G. Crewes (Eds) Language and Development: A Selection of Papers presented at the Second International Conference on Language Development. Denpasar, Bali: Indonesia Australia Language Foundation: 150-164.

Bandaranayake, R. (1978). "Can I really grade essays fairly?" Medical Journal of Australia 1(11): 595-597.

Barker, M., C. Child, C. Gallois, E. Jones and V. J. Callan (1991). "Difficulties of Overseas Students in Social and Academic Situations." Australian Journal of Psychology. 43(2): 79-84.

Barrett, R. (1993). Letter to the University of Adelaide Faculty of Medicine Curriculum Committee.

Barrett, R. (1996). The Psychiatric Team and the Social Definition of Schizophrenia. Cambridge, Cambridge University Press.

Barrows, H. S. and S. Abrahamson (1964). "The programmed patient: A technique for appraising student performance in clinical neurology." Journal of Medical Education 39: 802-805.

Barrows, H. S. (1968). "Simulated Patients in Medical Teaching." Canadian Medical Association Journal 98: 674-676.

Barrows, H. S. (1971). Simulated Patients. Springfield, Charles C. Thomas.

Barrows, H. S. (1975). "Pseudopatients." British Medical Journal 1(5958): 625-626.

Barrows, H. S., R. G. Williams and R. H. Moy (1987). "A Comprehensive Performance-Based Assessment Of Fourth-Year Students' Clinical Skills." Journal of Medical Education 62: 805-809.

Barrows, H. S. (1991). "Proper Acknowledgement." Academic Medicine 66(5): 271.

Barrows, H. S. (1993). "An Overview of the Uses of Standardized Patients for Teaching and Evaluating Clinical Skills." Academic Medicine 68(6): 443-451.

Battles, J. B., J. L. Carpenter, J. M. Wagner and D. D. McIntire (1992). Using Mean Equating to Adjust for Logistical and Chronological Variables in an OSCE Used with a Large Class of Students. Fifth Ottawa Conference: Approaches to the Assessment of Clinical Competence. Centre for Medical Education, Ninewells Hospital and Medical School, Dundee, Scotland, Page Brothers.

Baumslag, N. (1970). "Native Language Barriers: Letter to the Editor." New England Journal of Medicine 283(16): 768.

Beasley, C. (1990). "Helping E.S.L./E.F.L. Students with Language and Study Skills at Tertiary Level." HERDSA 12(1): 16-18.

Beasley, V. (1990). Can a university equity programme be made to work? Breaking the Mould: Fifteenth Annual Conference of the Higher Education Research Society of Australasia., University of Adelaide, South Australia, HERDSA.

Beckman, H. B. and R. M. Frankel (1994). "The Use of Videotape in Internal Medicine Training." Journal of General Internal Medicine 9: 517-521.

Bensing, J. (1991). "Doctor-Patient Communication and the Quality of Care." Social Science and Medicine 32(11): 1301-1310.

Bhawuk, D. P. S. (1990). Cross-cultural orientation programs. Applied Cross-Cultural Psychology. R. W. Brislin. Newbury Park, Sage: 325-346.

Biggs, J. (1996). "Enhancing teaching through constructive alignment." Higher Education 32: 247-364.

Biggs, J. (1997). Teaching International Students. Innovation in Teaching and Learning. Adelaide, South Australia, Flinders University.

Bilmes, J. and S. T. Boggs (1979). Language and Communication: The Foundations of Culture. Perspectives on Cross-Cultural Psychology. A. J. Marsella, R. G. Tharp and T. J. Ciborowski. New York, Academic Press: 47-76.

Blacket, R. B. (1990). "Foreign Medical Graduates." The Medical Journal of Australia 153(August): 125-132.

Blanck, P. D., R. Rosenthal and M. Vannicelli (1986). Talking To and About Patients: The Therapist's Tone of Voice. Nonverbal Communication in the Clinical Context. P. D. Blanck, R. Buck and R. Rosenthal. University Park, Pennsylvania State University Press: 99-143.

Block, M. R., K. F. Schaffner and J. L. Coulehan (1985). "Ethical Problems of Recording Physician-Patient Interactions in Family Practice Settings." Journal of Family Practice 21(6): 467-472.

Blondis, M. N. and B. E. Jackson (1977). Nonverbal communication with patients. New York, John Wiley and Sons.

Bloom, V. (1986). "Medical student writing." Journal of the Royal Society of Medicine 79(1): 4.

Bochner, S. and P. Wicks, Eds. (1972). Overseas Students in Australia. Sydney, New South Wales Press.

Bochner, S. (1983). Empirical studies of the social psychology of the contact situation. Cultures in Contact. Studies in Cross-Cultural Interaction. S. Bochner. Oxford, Pergamon Press: 30-31.

Bolman, W. M. (1995). "The Place of Behavioral Science in Medical Education and Practice." Academic Medicine 70(10): 873-878.

Bourhis, R. Y., S. Roth and G. MacQueen (1989). "Communication in the hospital setting: A survey of medical and everyday language use amongst patients, nurses and doctors." Social Science and Medicine 28(4): 339-346.

Bradac, J. J. and R. Wisegarver (1984). "Ascribed status, lexical diversity, and accent: Determinants of perceived status, solidarity, and control of speech style." Journal of Language and Social Psychology 3(4): 239-255.

Bradley, D. and M. Bradley (1984). Problems of Asian students in Australia. Language, culture and education. Canberra, Australian Government Publishing Service.

Brinton, D.M., M. A. Snow and M. B. Wesche (1989). Content-based Second Language Instruction. Boston, Heinle and Heinle.

Brislin, R. W. (1979). Orientation Programs for Cross-Cultural Preparation. Perspectives on Cross-Cultural Psychology. A. J. Marsella, R. G. Tharp and T. J. Ciborowski. New York, Academic Press: 287-305.

Brown, V., R. Chadwick, T. Kulski, H. Thompson, R. Palmer and K. Goldie (1997). Wicked Words and Slanguage. Perth, www.curtin.edu.au/curtin/dept/nursing/ww/wickword.html.

Buchanan, J. R. (1989). "Resources and Limitations of Medical Centers in the United States." Academic Medicine Supplement: S13-S17.

Budd, D. E. (1995). Towards literacy in Medical Science: A School-based policy for the teaching of literacy skills in context. HERDSA Conference - Higher Education: Blending tradition and technology: 183-188.

Burke, B. (1990). Meeting the reasonable expectations of overseas students. International Education Seminar on Overseas Students, Adelaide, South Australia.

Burns, R. B. (1991). "Study and Stress among First Year Overseas Students in an Australian University." Higher Education Research and Development 10(1): 61-77.

Butler, C. and S. Rollnick (1996). "Missing the meaning and provoking resistance; a case of myalgic encephalomyelitis." Family Practice 13(1): 106-109.

Butrin, J. (1992). "Clinical Diversity in the Nurse-Client Encounter." Clinical Nursing Research 1(3): 238-251.

Byrne, P. S. and C. C. Heath (1980). "Practitioners' use of non-verbal behaviour in real consultations." Journal of the Royal College of General Practitioners 30: 327-331.

Calhoun, J. G., J. O. Woolliscroft and J. D. Ten Haken (1987). "Internal Medicine House Officers' Performance As Assessed by Experts and Standardized Patients." Journal of Medical Education 62: 754-760.

Cargill, M. (1996). An Integrated Bridging Programme for International Postgraduate Students. Higher Education Research and Development 15(2): 177-189.

Cassell, E. J. and L. Skopek (1977). "Language as a Tool in Medicine: Methodology and Theoretical Framework." Journal of Medical Education 52: 197-203.

Chalmers, D. and S. Volet (1997). "Common Misconceptions about Students from South-East Asia Studying in Australia." Higher Education Research and Development 16(1): 87-98.

Chew, C. Y. C. (1972). Forward. Overseas Students in Australia. S. Bochner and P. Wicks. Sydney, New South Wales University Press: iii.

Choi, M. (1997). "Korean students in Australian universities: Intercultural Issues. Higher Education Research and Development 16(3): 263-282.

Christy, N. P. (1979). "English is our second language." New England Journal of Medicine 300(17): 979-981.

Chugh, U., E. Dillmann, S. M. Kurtz, J. Lockyer and J. Parboosingh (1993). "Multicultural issues in medical curriculum: implications for Canadian physicians." Medical Teacher 15(1): 83-91.

Chur-Hansen, A. (1985). Attitudes toward male and female doctors and nurses. Australian Association for Academic General Practice, Melbourne, Australia.

Chur-Hansen, A. and R. J. Barrett (1996). "Teaching colloquial Australian English to medical students from non-English speaking backgrounds." Medical Education 30: 412-417.

Chur-Hansen, A., J. Vernon-Roberts and S. Clark (1997). "Language background, English language proficiency and medical communication skills of medical students." Medical Education 31: 259-263.

Chur-Hansen, A. (1997). "Language background, proficiency in English, and selection for language development." Medical Education 31: 312-319.

Clark, F. (1997). Spelling Error: 2 people died. The Advertiser. Adelaide: November 29, 1-2.

Clark, J. L. D. and S. S. Swinton (1980). The Test of Spoken English as a measure of communicative ability in English-medium instructional settings. Princeton, New Jersey, Educational Testing Service.

Cleary, E. (1996). Producing and Delivering the new doctor. Adelaidean. Adelaide: 2, 6.

Cline, R. J. and C. A. Puhl (1984). "Gender, Culture, and Geography: A Comparison of Seating Arrangements in the United States and Taiwan." International Journal of Intercultural Relations 8: 199-219.

Cohen, D. S., B. Gromoff and M. H. Swartz (1992). Standardized Patients - The Advantages of the Professional Actor. Fifth Ottawa Conference: Approaches to the Assessment of Clinical Competence, Centre for Medical Education, Ninewells Hospital and Medical School, Dundee, Scotland, Page Brothers.

Cohen-Cole, S. A. and C. P. Friedman (1983). "The language problem: Integration of psychosocial variables into medical care." Psychosomatics 24(1): 52-60.

Cohen-Cole, S. A. (1991). The Medical Interview: The Three-Function Approach. St Louis, Mosby Year Book.

Cole-Kelly, K. (1994). "Cultures Engaging Cultures: International Medical Graduates Training in the United States." Family Medicine 26(10): 618-624.

Collier, J. and A. Burke (1986). "Racial and sexual discrimination in the selection of students for London medical schools." Medical Education 20: 86-90.

Collins, J. P., G. R. White and C. D. Mantell (1997). "Selection of medical students: an affirmative action programme." Medical Education 31: 77-80.

Colliver, J. A., H. S. Barrows, N. V. Vu, S. J. Verhulst, T. A. Mast and T. A. Travis (1991). "Test Security in Examinations That Use Standardized-patient Cases at One Medical School." Academic Medicine 66(5): 279-282.

Colliver, J. A., N. V. Vu, M. L. Marcy, T. A. Travis and R. S. Robbs (1993). "Effects of Examinee Gender, Standardized-patient Gender, and Their Interaction on Standardized Patients' Ratings of Examinees' Interpersonal and Communication Skills." Academic Medicine 68(2): 153-157.

Colliver, J. A. and R. G. Williams (1993). "Technical Issues: Test Application." Academic Medicine 68(6): 454-460.

Colton, F. V., L. Heun and M. Link (1994). "Helping Medical Students Speak English More Effectively." Academic Medicine 69(5): 360-361.

Commonwealth Universities Yearbook 1995-96. (1995). A Directory to the Universities of the Commonwealth and the Handbook of their Association. London, Bath Press.

Comstock, L. M., E. M. Hooper, J. M. Goodwin and J. S. Goodwin (1982). "Physician Behaviors that Correlate with Patient Satisfaction." Journal of Medical Education 57: 105-112.

Craig, P. and R. Bandaranayake (1993). "Experiences with a method for obtaining feedback on a medical curriculum undergoing change." Medical Education 27: 15-21.

Crichton, M. (1975). "Medical Obfuscation: Structure and Function." New England Journal of Medicine 293(24): 1257-1259.

Crystal, D. (1976). The diagnosis of sociolinguistic problems in doctor patient interaction. Language and Communication in General Practice. B. A. Tanner. London, Hodder and Stoughton: 40-51.

Cuff, E. C. and G. C. F. Payne (1980). Perspectives in Sociology. London, Allen & Unwin.

Davies, M. (1989). "The Way Ahead: teaching with simulated patients." Medical Teacher 11(3/4): 315-320.

Davitz, J. R. (1964). The Communication of Emotional Meaning. Westport, Connecticut, Greenwood Press.

Daw, M. and D. Penberthy (1997). Vanstone vows end to Hanson. The Advertiser. Adelaide: 7.

de Lemos, M. M. (1990). ACER Word Knowledge Test Forms E and F. Hawthorn, Australian Council for Educational Research (ACER) Ltd.

De Sweemer, C. (1972). "Language and Communication Training." Journal of Medical Education 47: 921.

Del Mar, C. B. (1994). "Communicating well in general practice." The Medical Journal of Australia 160: 367-370.

Diehl, A. K., T. J. Westwick, R. G. Badgett, N. J. Sugarek and K. H. Todd (1993). "Clinical and Sociocultural Determinants of Gallstone Treatment." American Journal of the Medical Sciences 305(6): 383-386.

Dillner, L. (1995). "Manchester tackles failure rate of Asian students." British Medical Journal 310: 209.

DiMatteo, M. R., H. S. Friedman and A. Taranta (1979). "Sensitivity to Bodily Nonverbal Communication as a Factor in Practitioner-Patient Rapport." Journal of Nonverbal Behavior 4(1): 18-26.

DiMatteo, M. R. (1979). "A social-psychological analysis of physician-patient rapport: Toward a science of the art of medicine." Journal of Social Issues 35(1): 12-33.

DiMatteo, M. R., L. M. Prince and A. Taranta (1979). "Patients' perceptions of physicians' behavior" Determinants of patient commitment to the therapeutic relationship." Journal of Community Health 4(4): 280-290.

DiMatteo, M. R. and A. Taranta (1979). "Nonverbal Communication and Physician-Patient rapport: An empirical study." Professional Psychology 10(4): 540-547.

DiMatteo, M. R., A. Taranta, H. S. Friedman and L. M. Prince (1980). "Predicting Patient Satisfaction from Physicians' Nonverbal Communication Skills." Medical Care 17(4): 376-387.

DiMatteo, M. R. and H. S. Friedman (1982). Social Psychology and Medicine. Cambridge, Massachusetts, Oelgeschlager, Gunn & Hain.

DiMatteo, M. R. and D. D. DiNicola (1982). Social Science and the Art of Medicine: From Hippocrates to Holism. Interpersonal Issues in Health Care. H. S. Friedman and M. R. DiMatteo. New York, Academic Press: 9-31.

DiMatteo, M. R., L. M. Prince and R. Hays (1986). Nonverbal Communication in the Medical Context: The Physician-Patient Relationship. Nonverbal Communication in the Clinical Context. P. D. Blanck, R. Buck and R. Rosenthal. University Park, Pennsylvania State University Press: 74-98.

Dissanayake, A. S., B. A. Ali and U. Nayar (1990). "English Proficiency in Multiple-Choice Questions." Academic Medicine 65(2): 101.

Doherty, R. L. (1988). Australian medical education and workforce into the twenty-first century. Committee of Inquiry into Medical Education and Medical Workforce. Canberra, Australian Government Publishing Service.

Drife, J. O. (1990). "Body language for beginners." British Medical Journal 300: 757.

Edge, J. (1997a). Internationalisation; 25 year vision demands shared goals and a shift in institutional culture and attitude. Adelaidean. Adelaide: 5.

Edge, J. (1997b). New arrivals welcomed. Adelaidean. Adelaide. 6: 4.

Editorial (1974). "Pseudopatients." British Medical Journal 4(5947): 729-730.

Educational Testing Service (1993). Overseas Edition for Australia, New Zealand and Papua New Guinea 1993-94 Bulletin of Information for TOEFL, TWE and TSE. Princeton, N.J., TOEFL/TSE Publications.

Edwards, H. and L. Devitt (1988). Using Simulated Patients in Clinical Education. International Health Sciences Conference, Sydney, Australia.

Egan, G. (1990). The Skilled Helper. A Systematic Approach to Effective Helping. Pacific Grove, California, Brooks/Cole.

Ekman, P. (1971). Universals and Cultural Differences in Facial Expressions of Emotion. Nebraska Symposium on Motivation. J. Cole. Nebraska, University of Nebraska Press: 207-283.

Ekman, P., W. V. Friesen and P. Ellsworth (1982). What are the similarities and differences in facial behavior across cultures? Emotion in the Human Face. P. Ekman, W. V. Friesen and P. Ellsworth. Cambridge, Cambridge University Press: 128-143.

Ekman, P., W. V. Friesen, M. O' Sullivan, A. Chan, I. Diacoyanni-Tarlatzis, K. Heider, R. Krause, W. A. LeCompte, T. Pitcairne, P. E. Ricci-Bitti, K. Scherer, M. Tomita and J. Tzavaras (1987). "Universals and Cultural Differences in the Judgements of Facial Expressions of Emotion." Journal of Personality and Social Psychology 53(4): 712-717.

El Shabrawy Ali, M. (1992). "A Study of Patient Satisfaction as an Evaluation Parameter for Utilization of Primary Health Care Services." Journal of the Royal Society of Health 112(2): 64-67.

El Shabrawy Ali, M. and M. Eisa Ali Mahmoud (1993). "A study of patient satisfaction with primary health care services in Saudi Arabia." Journal of Community Health 18(1): 49-54.

Elsass, P., H. P. Christensen, J. Falhof and A. Hvolby (1994). "Greenlanders in hospital; Lack of language understanding is not always a hindrance to intercultural communication." Artic Medical Research 53(2): 97-104.

Elwell, C. M., R. J. Brown and D. R. Rutter (1984). "Effects of Accent and Visual Information on Impression Formation." Journal of Language and Social Psychology 3(4): 297-299.

Erickson, F. and W. Rittenberg (1987). "Topic Control and Person Control: A Thorny Problem for Foreign Physicians in Interaction with American Patients." Discourse Processes 10: 401-415.

Erzinger, S. (1991). "Communication between Spanish-speaking patients and their doctors in medical encounters." Culture, Medicine and Psychiatry 15: 91-110.

Esmail, A. and S. Everington (1993). "Racial discrimination against doctors from ethnic minorities." British Medical Journal 306: 691-692.

Esmail, A., P. Nelson, D. Primarolo and T. Toma (1995). "Acceptance into medical school and racial discrimination." British Medical Journal 310: 501-502.

Evans, B. J., R. O. Stanley, G. J. Coman and G. D. Burrows (1989). "Psychological tests to measure the effects of medical education on students' interpersonal skills." Medical Education 23: 492-497.

Farmer, E. A. and D. J. Prideaux (1994). Here today gone tomorrow? Assessing the retention of patient centred behaviours in Family Medicine trainees using standardized patient encounters. Sixth Ottawa Conference in Medical Education, Toronto.

Farnill, D., S. C. Hayes and S. Barrett (1993). Origins, family background, language capacities and career motivations of students admitted to Medicine I in 1993. Sydney, New South Wales, Faculty of Medicine, University of Sydney.

Farnill, D. and S. C. Hayes (1994). Medical Students' Norms for the Screening Test of Adolescent Language., Department of Behavioural Sciences in Medicine, Faculty of Medicine, University of Sydney.

Farnill, D., S. C. Hayes and A. Chur-Hansen (1995). "Interrater reliability of the scoring of the Screening Test of Adolescent Language." Psychological Reports 76: 1027-1032.

Farnill, D. and S. C. Hayes (1995). "Retest reliability of the Woodcock Language Proficiency Battery." Perceptual and Motor Skills 81: 1147-1152.

Farnill, D. and S. C. Hayes (1996a). A User's Guide to the Australian Tertiary English Screening Test: AUSTEST. Sydney, University of Sydney Glensdale Press.

Farnill, D. and S. C. Hayes (1996b). "Screening Higher-Education Students for English Language Problems: Development of the Australian Tertiary-English Screening Test." Higher Education Research and Development 15(1): 61-71.

Farnill, D. and S. Hayes (1996c). "Do NESB University Students with Poor English Skills Make Rapid Linguistic Gains in Mainstream Studies?" Higher Education Research and Development 15(2): 261-264.

Farnill, D., J. Todisco, S. C. Hayes and D. Bartlett (1997). "Videotaped interviewing of non-English speakers: training for medical students with volunteer clients." Medical Education 31: 87-93.

Fehrsen, G. S. and R. J. Henbest (1993). "In Search of Excellence. Expanding the Patient-centred Clinical method: A Three-stage Assessment." Family Practice 10(1): 49-54.

Felix, U. (1992). Survey of Overseas Students' English Language Needs while studying at Adelaide University. Adelaide, Centre for Language Teaching and Research.

Ferguson, J. J. and J. Maclean (1988). "Introductory Course in Medical English and Primary Care for Overseas Doctors." Family Practice 5(4): 260-264.

Fessel, W. J. and E. E. Van Brunt (1972). "Assessing quality of care from the medical record." New England Journal of Medicine 286(3): 134-138.

Flaherty, J. A., A. Rezler and C. McGuire (1982). "Clinical Reading and Writing Skills of Junior Medical Students." Journal of Medical Education 57: 848-853.

Ford, J. (1977). "A Linguistic Analysis of Doctor-Patient Communication Problems." Dissertation Abstracts International 37(9(A)): 5792.

Fox, C. H. and F. Meijer (1980). "Teaching medical English to foreign-language doctors." Medical Education 14: 316-319.

Frederikson, L. G. (1995). "Exploring information-exchange in consultation: the patients' view of performance and outcomes." Patient Education and Counseling 25: 237-246.

Freemon, B., V. F. Negrete, M. Davis and B. M. Korsch (1971). "Gaps in Doctor-Patient Communication: Doctor-Patient Interaction Analysis." Pedatric Research 5: 298-311.

Frewin, D. B. (1993a). Letter - Membership of the Language Development Committee.

Frewin, D. B. (1993b). Memorandum - English Language Screening Test.

Frewin, D. (1993c). Degrees of Bachelor of Medicine, Bachelor of Surgery, Uinversity of Adelaide Faculty of Medicine.

Frewin, D. B. (1994). Heads of Pre-Clinical Departments - Language Development Program.

Friedman, H. S. (1979). "Nonverbal communication between patients and medical practitioners." Journal of Social Issues 35(1): 82-99.

Friedman, H. S., M. R. DiMatteo and A. Taranta (1980). "A Study of the Relationship between Individual Differences in Nonverbal Expressiveness and Factors of Personality and Social Interaction." Journal of Research in Personality 14(3): 351-364.

Friedman, H. S. (1982). Nonverbal Communication in Medical Interaction. Interpersonal Issues in Health Care. H. S. Friedman and M. R. DiMatteo. New York, Academic Press: 51-68.

Friedman, H. S. and M. R. DiMatteo (1982). Interpersonal Issues in Health Care: Healing as an Interpersonal Process. Interpersonal Issues in Health Care. H. S. Friedman and M. R. DiMatteo. New York, Academic Press: 3-7.

Friedman, M., A. I. Sutnick, P. L. Stillman, J. J. Norcini, S. M. Anderson, R. G. Williams, G. Henning and M. J. Reeves (1991). "Standardized Patients in Skills Assessment. The Use of Standardized Patients to Evaluate the Spoken-English Proficiency of Foreign Medical Graduates." Academic Medicine 66(9): S61-S63.

Friedman, M., A. I. Sutnick, P. L. Stillman, M. B. Regan and J. J. Norcini (1992). The Relationship Between Spoken English Proficiency of Foreign Medical Graduates and their Clinical Competence. Fifth Ottawa Conference: Approaches to the Assessment of Clinical Competence., Centre for Medical Education, Ninewells Hospital and Medical School, Dundee, Scotland, Page Brothers.

Friedman, M., A. I. Sutnick, P. L. Stillman, M. B. Regan and J. J. Norcini (1993). "Use of Standardized-Patient Examinations in Conjunction with Licensure and Certification. The Relationship of Spoken-English Proficiencies of Foreign Medical School Graduates to Their Clinical Competence." Academic Medicine 68(10): S1-S3.

Furman, G., J. A. Colliver and A. Galofre (1993). "Effects of Student Gender and Standardized-patient Gender in a Single Case Using a Male and a Female Standarized Patient." Academic Medicine 68(4): 301-303.

Furnham, A. and S. Bochner (1986). Culture Shock. Psychological reactions to unfamiliar environments. London, Methuen.

Galambos, J. and I. Simonelli (1996). Probability and its Applications. Bonferroni-type Inequalities with Applications. New York, Springer.

Gallois, C. and V. J. Callan (1981). "Personality Impressions Elicited by Accented English Speech." Journal of Cross-Cultural Psychology 12(3): 347-359.

Garfinkel, H. (1967). "Good Organizational Reasons for Bad Clinic Records." Studies in Ethnomethodology. New York, Prentice Hall: 186-207.

Gastel, B. (1995a). "Toward a Global Consensus on Quality Medical Education: Serving the Needs of Populations and Individuals: Summary of the Consultation." Academic Medicine 70(7): S3-S7.

Gastel, B. (1995b). "Toward a Global Consensus on Quality Medical Education: Serving the Needs of Populations and Individuals." Academic Medicine 70(7): S73-S75.

Gerber, P. (1991). "Overseas medical practitioners and "racial discrimination"." Medical Journal of Australia 155: 509-512.

Gerber, P. (1992). "Letter to the Editor. Overseas medical practitioners and "racial discrimination"." Medical Journal of Australia 156: 138.

Giles, H. (1977). "Social Psychology and Applied Linguistics: Towards an Integrative Approach." Review of Applied Linguistics 35: 27-42.

Giles, H., R. Y. Bourhis and D. M. Taylor (1977). Towards a Theory of Language in Ethnic Group Relations. Language, Ethnicity and Intergroup Relations. H. Giles. London, Avademic Press: 307-348.

Glendinning, E. and B. Holmstrom (1991). English in Medicine. A Course in Communication Skills. Cambridge, Cambridge University Press.

Glover, W. E. (1995). "Letter to the Editor. Communication skills of interns in New South Wales." The Medical Journal of Australia 162: 497.

Glover, J., M. Shand, C. Forster and T. Woollacott (1996). A Social Health Atlas of South Australia. Adelaide, South Australian Health Commission.

Goldbeck-Wood, S. (1996). "Choosing tomorrow's doctors." British Medical Journal 313: 313.

Golden, J. S. and G. D. Johnston (1970). "Problems of Distortion in Doctor-Patient Communications." Psychiatry in Medicine 1: 127-149.

Gollin, S. (1990). "ELSSA: Faculty based language and learning skills assistance for students of non-English speaking background." Proceedings of the HERDSA Conference, Griffiths University, Queensland, Australia, July: 183-188.

Gordon, J., R. Sanson-Fisher and N. A. Saunders (1988). "Identification of simulated patients by interns in a casualty setting." Medical Education 22: 533-538.

Gordon, J. J., N. A. Saunders and R. W. Sanson-Fisher (1989). "Evaluating interns' performance using simulated patients in a casualty department." Medical Journal of Australia 151: 18-21.

Gordon, J. J., R. W. Sanson-Fisher and N. A. Saunders (1990). Using Simulated Patients in the Evaluation of Graduates: Can They Be Detected? Innovation in Medical Education: An Evaluation of Its Present Status. Z. M. Nooman, H. G. Schmidt and E. S. Ezzat. New York, Springer: 341-347.

Gordon, J. J., N. A. Saunders, D. Hennrikus and R. W. Sanson-Fisher (1992). "Interns' Performances with Simulated Patients at the Beginning and the End of the Intern Year." Journal of General Internal Medicine 7: 57-62.

Graham, J. G. (1987). "English Language Proficiency and the Prediction of Academic Success." TESOL Quarterly 21(3): 505-521.

Grassby, A. J. (1980). "Bridging the gap." The Medical Journal of Australia 2: 701-703.

Greenberg, L. W. and L. S. Jewett (1987). "The Case Presentation: teaching medical students writing and communication skills." Medical Teacher 9(3): 281-284.

Grover, G., C. D. Berkowitz and R. J. Lewis (1994). "Parental Recall After a Visit to the Emergency Department." Clinical Pediatrics 33(4): 194-201.

Hadlow, J. and M. Pitts (1991). "The understanding of common health terms by doctors, nurses and patients." Social Science and Medicine 32(2): 193-196.

Hailstone, B. (1997). Hospital tackles langauge barrier. The Advertiser. Adelaide: 8.

Hall, J. A., D. L. Roter and C. S. Rand (1981). "Communication of Affect between Patient and Physician." Journal of Health and Social Behaviour 22: 18-30.

Hamilton, J. D. (1995). "Establishing Standards and Measurement Methods for Medical Education." Academic Medicine 70(7): S51-S56.

Hancock, L. and G. Mullins (1991). Educating Overseas Students at the University of Adelaide; How do we rate? Lumen. 20: 3-4.

Harasym, P. H. and W. Woloschuk (1992). The Halo Effect Inherent Within OSCE Checklists. Fifth Ottawa Conference: Approaches to the Assessment of Clinical Competence., Centre for Medical Education, Ninewells Hospital and Medical School, Dundee, Scotland, Page Brothers.

Harden, R. M., M. Stevenson, W. W. Downie and G. M. Wilson (1975). "Assessment of Clinical Competence using Objective Structured Examination." British Medical Journal 1: 447-451.

Harden, R. M. and F. A. Gleeson (1979). "Assessment of clinical competence using an objective structured clinical examination (OSCE)." Medical Education 13: 41-54.

Harrell, P. L., G. W. Kearl, E. L. Reed, D. G. Grigsby and T. S. Caudill (1993). "Medical Students' Confidence and the Characteristics of Their Clinical Experiences in a Primary Care Clerkship." Academic Medicine 68(7): 577-579.

Harrigan, J. A. and R. Rosenthal (1986). Nonverbal Aspects of Empathy and Rapport in Physician-Patient Interaction. Nonverbal Communication in the Clinical Context. P. D. Blanck, R. Buck and R. Rosenthal. University Park, Pennsylvania State University Press: 36-73.

Harris, G. T. and F. G. Jarrett (1990). Educating Overseas Students in Australia. Who Benefits? Sydney, Allen & Unwin.

Hart, I. R. (1986). The OSCE - Objective yes, but how useful? Newer Developments in Assessing Clinical Competence., Ottawa, Canada, Heal Publications Ltd.

Hay, R. G. (1972). "Here be Tygers . . .", The Language Barrier. Overseas Students in Australia. S. Bochner and P. Wicks. Sydney, New South Wales University Press: 129-144.

Hayes, S. C. and D. Farnill (1992). "A study of the concurrent validity of the Screening Test of Adolescent Language with recent immigrants." Psychological Reports 71: 175-178.

Hayes, S. C. and D. Farnill (1993a). "Medical training and English language proficiency." Medical Education 27: 6-14.

Hayes, S. C. and D. Farnill (1993b). English Language Proficiency Research, 1992. A two year follow-up. Sydney, New South Wales, Faculty of Medicine, University of Sydney.

Hayes, S. C. and D. Farnill (1994). English language, academic performance and medical communication 1993. A report to the Faculty of Medicine, University of Sydney., Faculty of Medicine, University of Sydney.

Hayes, S. C., D. Farnill and A. J. Sefton (1994). "Improving English Communication Skills of Preclinical Medical Students." Annals of Community-Oriented Education 7: 317-327.

Hays, R. B., B. F. Jones, P. B. Adkins and P. J. McKain (1990). "Analysis of videotaped consultations to certify competence." The Medical Journal of Australia 152: 609-611.

Helman, C. G. (1984). "The role of context in primary care." Journal of the Royal College of General Practitioners 34: 547-550.

Henbest, R. J. and M. Stewart (1990). "Patient-Centredness in the Consultation. 2: Does it Really Make a Difference?" Family Practice 7(1): 28-33.

Henbest, R. J. and G. S. Fehrsen (1992). "Patient-centredness: Is It Applicable Outside the West? Its Measurement and Effect on Outcomes." Family Practice 9(3): 311-317.

Heritage, J. (1984). Garfinkel and Ethnomethodology. Cambridge, Polity Press.

Hodges, B., J. Turbbull, R. Cohen, A. Bienenstock and G. Norman (1996). "Evaluating communication skills in the objective structured clinical examination format: reliability and generalizability." Medical Education 30: 38-43.

Holbrook, A. and S. Bourke (1989). "Assessment of the English Skills of Tertiary Students." Higher Education Research and Development 8(2): 161-179.

Holmes, J. L., A. Lin, S. Fath and L. Gray (1992). "The paperless essay: one way to teach writing and computer skills in medical school." Medical Teacher 14(1): 83-88.

Holton, R. J. and S. Salagaras (1988). The Participation of Non-English Speaking Persons in Tertiary Education. Adelaide, Flinders University of South Australia.

Hughes, B. E. (1993). Working Words. An Australian Guide to Modern English Usage. Ringwood, Victoria, Viking.

Hui, C. H. and I. W. M. Cheng (1987). "Effects of second language proficiency of speakers and listeners on person perception and behavioural intention: A study of Chinese bilinguals." International Journal of Psychology 22: 421-430.

Hui, C. H. and Y. Yam (1987). "Effects of language proficiency and physical attractiveness on person perception." British Journal of Social Psychology 26: 257-261.

Hunt, D. D., B. A. K. Khalid, S. H. Shahabudin, R. Jaafar and J. D. Carline (1995). "A Cross-Cultural Comparison of Problem Students on Clinical Rotations." Teaching and Learning in Medicine 7(1): 8-12.

Hunter, B. and M. Hayden (1990). A contextual approach to the provision of learning support services to students in the health sciences from non-English speaking backgrounds. Breaking the Mould: Fifteenth Annual Conference of the Higher Education Research Society of Australasia (HERDSA), University of Adelaide, South Australia.

Huttlinger, K., L. Krefting, D. Drevdahl, P. Tree, E. Baca and A. Benally (1992). ""Doing Battle": A Metaphorical Analysis of Diabetes Mellitus Among Navajo People." American Journal of Occupational Therapy 46(8): 706-712.

Huxham, G. J., A. Lipton and D. Hamilton (1975). "Reasons for differential performance in multiple choice and essay tests." British Journal of Medical Education 9: 264-272.

Huxham, G. J., A. Lipton and D. Hamilton (1980). "Achievement factors and personality in a cohort of medical students." Medical Education 14: 97-104.

Huxham, G. J., A. Lipton, D. Hamilton and D. Chant (1989). "What 'makes' a good doctor?" Medical Education 23: 3-13.

IELTS Handbook: International English Language Testing System (1966). University of Cambridge Local Examination Syndicate.

Ingleton, C. and B. Wake (1997). Literacy Matters. Strategies for teaching communication skills to university students. Adelaide, University of Adelaide Advisory Centre for University Education.

Ingram, D. E. (1984). Australian Second Language Proficiency ratings (ASLPR) Introduction. Canberra, Australian Government Publishing Service.

International English Language Testing System (IELTS). (1989). Canberra, International Development Program of Australian Universities and Colleges.

Inui, T. S. and W. B. Carter (1985). "Problems and Prsopects for Health Services Research on Provider-Patient Communication." Medical Care 23(5): 521-538.

Isaac, A. (1993). CSHE Research Working Papers 93.8. Equity or equality? Students with language backgrounds other than English. Melbourne, Victoria, Centre for Study of Higher Education, University of Melbourne.

Isaacs, G. (1989). "Lecture Note-taking, Learning and Recall." Medical Teacher 11(3/4): 295-301.

Jenkins, H. M. L., V. Blank, K. Miller, J. Turner and R. Stanwick (1996). "A Randomized Single-Blind Evaluation of a Discharge Teaching Book for Pediatric Patients with Burns." Journal of Burn Care and Rehabilitation 17(1): 49-61.

Jolly, B., D. Newble and T. Chinner (1993). "Learning Effect of Reusing Stations in an Objective Structured Clinical Examination." Teaching and Learning in Medicine 5(2): 66-71.

Jones, K. (1992). "Letter to the Editor. Overseas medical practitioners and "racial discrimination"." Medical Journal of Australia 156: 362.

Jordan, C. and R. G. Tharp (1979). Culture and Education. Perspectives on Cross-Cultural Psychology. A. J. Marsella, R. G. Tharp and T. J. Ciborowski. New York, Academic Press: 265-285.

Kachur, E. J., S. Green, E. L. Bruun, M. M. Philbin, J. Addis, D. K. Simonton, A. Ward, P. L. Stillman and A. I. Sutnick (1992). A Multi-Site Survey of Standardized Patients - Who are These People? Fifth Ottawa Conference: Approaches to the Assessment of Clinical Competence., Centre for Medical Education, Ninewells Hospital and Medical School, Dundee, Scotland, Page Brothers.

Katsikitis, M. (1992). The Quantification of Facial Expression Using a Mathematical Model of the Face: Validation and Extension of a Microcomputer-Based Technique. Psychiatry. Adelaide, University of Adelaide.

Kazdin, A. E. (1977). "Artifact, Bias and Complexity of Assessment: The ABCs of Reliability." Journal of Applied Behavior Analysis 10(1): 141-150.

Keats, D. (1972). New Study Patterns. Overseas Students in Australia. S. Bochner and P. Wicks. Sydney, New South Wales University Press: 105-115.

Kennedy, K. (1992). Developing a Curriculum Guarantee for International Students. Quality in international education - an academic perspective., University of Southern Queensland.

Kerr, D. (1996). "Chaos in the case notes." Journal of the Royal College of Physicians of London 30(1): 5.

Kidd, M. R. and A. Zulman (1994). "Educational support for overseas-trained doctors." Medical Journal of Australia 160: 73-75.

Klimidis, S., I. H. Minas, G. W. Stuart and C. Hayes (1997). "Cultural diversity in Australian medical education." Medical Education 31: 58-66.

Klineberg, O. and W. F. Hull (1979). At a Foreign University. An International Study of Adaptation and Coping. New York, Praeger.

Korsch, B. M., E. K. Gozzi and V. Francis (1968). "Gaps in Doctor-Patient Communication. I. Doctor-Patient Interaction and Patient Satisfaction." Pediatrics 42(5): 855-871.

Korsch, B. M. and V. F. Negrete (1972). "Doctor-Patient Communication." Scientific American 227(2): 66-74.

Korsch, B. M. (1984). "What do patients and parents want to know? What do they need to know?" Pediatrics 74((5) Supplement): 917-919.

Kraan, H., A. Crijnen, J. Zuidweg, C. van der Vleuten and T. Imbos (1990). Evaluating Undergraduate Training - A Checklist for Medical Interviewing Skills. Communicating with Medical Patients. M. Stewart and D. Roter. Newbury Park, Sage: 167-177.

Lacoste, M. (1994). Doctor-Patient Language. The Encyclopedia of Language and Linguistics. R. E. Asher and J. M. Y. Simpson. Oxford, Pergamon Press: 1055-1057.

Larsen, K. M. and C. K. Smith (1981). "Assessment of Nonverbal Communication in the Patient-Physician Interview." Journal of Family Practice 12(3): 481-488.

Leary, M. R. (1991). Introduction to Behavioural Research Methods. Belmont, Wadsworth.

Levey, G. S. (1992). "Internal Medicine and the Training of International Medical Graduates: A Time for Open Discussion and New Approaches." Annals of Internal Medicine 117(5): 403-407.

Ley, P. and M. S. Spelman (1965). "Communications in an Out-patient Setting." British Journal of Social and Clinical Psychology 4: 114-116.

Ley, P. (1979). The Psychology of Compliance. Research in Psychology and Medicine. Osborne, Gruneberg and Eiser. London, Academic Press: 187-195.

Ley, P. (1982). "Satisfaction, compliance and communication." British Journal of Clinical Psychology 21: 241-254.

Lieberman, R. J., A. M. C. Heffron, S. J. West, E. C. Hutchinson and T. W. Swem (1987). "A Comparison of Four Adolescent Language Tests." Language, Speech and Hearing in Schools 18(3): 250-265.

Light, R. L., M. Xu and J. Mossop (1987). "English Proficiency and Academic Performance of International Students." TESOL Quarterly 21(2): 251-261.

Like, R. C., P. Steiner and A. J. Rubel (1996). "Recommended Core Curriculum Guidelines on Culturally Sensitive and Competent Health Care." Family Medicine 28(4): 291-297.

Lipton, A., G. J. Huxham and D. Hamilton (1984). "Predictors of success in a cohort of medical students." Medical Education 18: 203-210.

Lipton, A., G. Huxham and D. Hamilton (1988). "School results as predictors of medical school achievement." Medical Education 22: 381-388.

Liu, S., L. W. Bassett and J. Sayre (1994). "Women's Attitudes about Receiving Mammographic Results Directly from Radiologists." Radiology 193: 783-786.

Lloyd, N. (1997). Overseas students add $90m to coffers. The Advertiser. Adelaide: 6.

Long, B. E. L. (1985). "A study of the verbal behavior of family doctors." Inetrnational Journal of the Sociology of Language 51: 5-25.

Love, R. R., P. A. Newcomb, J. H. Schiller, G. Wilding and H. L. Stone (1993). "A Comparison of Knowledge and Communication Skill Evaluations by Written Essay and Oral Examinations in Preclinical Medical Students." Journal of Cancer Education 8(2): 123-128.

Lowry, S. and G. MacPherson (1988). "A Blot on the Profession." British Medical Journal 296: 657-658.
Lucas, P., M. Lenstrup, J. Prinz, D. Williamson, H. Yip and G. Tipoe (1997). "Language as a barrier to the acquisition of anatomical knowledge." Medical Education 31: 81-86.

Luketich, G. F., J. A. Colliver and A. Galofre (1992). Two Studies of the Effects of Gender Within a Single Case Simulated by a Male and a Female Standardized Patient in a Multiple-Stations Examination. Fifth Ottawa Conference: Approaches to the Assessment of Clinical Competence, Centre for Medical Education, Ninewells Hospital and Medical School, Dundee, Scotland, Page Brothers.

Lyons, T. F. and B. C. Payne (1974). "The Relationship of Physicians' Medical Recording Performance to Their Medical Care Performance." Medical Care 12(8): 714-720.

Maclean, J. and J. C. Maher (1994). Medical Language. The Encyclopedia of Language and Linguistics. R. E. Asher and J. M. Y. Simpson. Oxford, Pergamon Press: 2431-2433.

Maher, J. (1986a). "The Development of English as an International Language of Medicine." Applied Linguistics 7(2): 206-218.

Maher, J. (1986b). "English for medical purposes. (State of the art)." Language Teaching; The International Abstracting Journal for Language Teachers and Applied Linguists. 19: 112-145.

Maher, J. C. (1987). "Editorial. English as an international language of medicine." Medical Education 21: 283-284.

Maher, J. C. (1993). "Editorial. Medical education in a multilingual and multicultural world." Medical Education 27: 3-5.

Manning, R. T. (1989). "Dication of the Discharge Resume: A Forgotten Link Between the Spoken and Written Word." Journal of General Internal Medicine 4(5): 453-456.

Marteau, T. M., C. Humphrey, G. Matoon, J. Kidd, L. M. and J. Horder (1991). "Factors influencing the communication skills of first-year clinical medical students." Medical Education 25: 127-134.

Martens, E. (1991). Strategies for Influencing Curriculum Development with an Overseas Focus. Adelaide, Flinders University.

Martin, F. J. and M. J. Bass (1989). "The Impact of Discussion of Non-Medical Problems in the Physician's Office." Family Practice 6(4): 254-258.

Martin, J. A., R. K. Reznick, A. Rothman, R. M. Tamblyn and G. Regehr (1996). "Who Should Rate Candidates in an Objective Structured Clinical Examination?" Academic Medicine 71(2): 170-175.

Matsuba, T. (1993). "A study of knowledge, attitude abd health behavior toward tuberculosis among non-immigrant Korean people in Japan." Kekkaku 68(6): 407-418.

Matsumoto, D. and M. Assar (1992). "The Effects of Language on Judgements of Universal Facial Expressions of Emotion." Journal of Nonverbal Behavior 16(2): 85-99.

McAdam, K. (1972). The Study Methods and Academic Results of Overseas Students. Overseas Students in Australia. S. Bochner and P. Wicks. Sydney, New South Wales University Press: 97-104.

McAvoy, B. R. (1988). "Teaching clinical skills to medical students: the use of simulated patients and videotaping in general practice." Medical Education 22: 193-199.

McConaghy, N. (1974). "Pseudopatients and evaluation of medical practice." Medical Journal of Australia 2(11): 383-385.

McGlinn, S. and E. W. Jackson (1989). "Predicting the Medical School Progress of Minority Students Who Participated in a Preparatory Program." Academic Medicine 64: 164-166.

McGorry, P. (1990). "Letter to the Editor. Foreign Medical Graduates." Medical Journal of Australia 153: 629-630.

McGowan, U. (1993). Metaphor in Academic Oral Discourse. International Systemic Congress, Victoria University, BC, Canada.

McGowan, U. (1994). Memorandum - Language Development Program in the Medical School. Adelaide, University of Adelaide.

McGowan, U. and M. Cargill (1994). International postgraduate students: a longitudinal study of language development and learning progress. HERDSA, Canberra, unpublished.

McGowan, U. (1995a). Modelling language for life-long learning. Adelaidean. Adelaide. 4: 2.

McGowan, U. (1995b). Memorandum - Medicine II Language Assessment Results. Adelaide, University of Adelaide.

McGowan, U. (1996a). The Language of Clinical Interviews. Oral Language Development for Students of Medicine. Adelaide, Advisory Centre for University Education Language and Learning Service.

McGowan, U. (1996b). The use of metaphor in budget reports in the media: barriers for international students of Economics and Commerce. HERDSA, Perth, unpublished.

McGowan, U. and M. Cargill (1997). "Rich tapestry" of offerings to support international student learning. Adelaidean. Adelaide: 7.

McKeigue, P. M., J. D. M. Richards and P. Richards (1990). "Effects of discrimination by sex and race on the early careers of British medical graduates during 1981-7." British Medical Journal 301: 961-964.

McKenzie, K. J. (1995). "Racial discrimination in medicine." British Medical Journal 310: 478-479.

McLaine, T. and M. Stevenson (1993). Some Implications of the Language Backgrounds of Bicultural Students for Teaching and Learning in Higher Education. HERDSA Conference, University of Wellington, New Zealand.

McManus, I. C. and P. Richards (1985). "Letter to the Editor. Admission to medical school." British Medical Journal 290: 319-320.

McManus, I. C., P. Richards and S. L. Maitlis (1989). "Prospective study of the disadvantage of people from ethnic minority groups applying to medical schools in the United Kingdom." British Medical Journal 298: 723-726.

McManus, I. C., S. L. Maitlis and P. Richards (1990). "Identifying Medical School Applicants from Ethnic Minorities." Studies in Higher Education 15(1): 57-73.

McManus, I. C., P. Richards, B. C. Winder, K. A. Sproston and V. Styles (1995). "Medical school applicants from ethnic minority groups: identifying if and when they are disadvantaged." British Medical Journal 310: 496-500.

McManus, I. C., P. Richards, B. C. Winder and K. A. Sproston (1996). "Final examination performance of medical students from ethnic minorities." Medical Education 30: 195-200.

Medical Defence Association of South Australia (no date). Problems of Medical Practice in a Multi-Cultural Society. North Adelaide, Medical Defence Association of South Australia.

Mehrabian, A. (1972). Nonverbal Communication. Chicago, Aldine-Atherton.

Menahem, S. (1987). "Teaching students of medicine to listen: the missed diagnosis from a hidden agenda." Journal of the Royal Society of Medicine 80: 343-346.

Mendez, A., J. A. Shymansky and M. Wolraich (1986). "Verbal and non-verbal behaviour of doctors while conveying distressing information." Medical Education 20: 437-443.

Miller, G. E. (1993). "Conference Summary (of the AAMC Consensus Conference on the Use of Standardized Patients in the Teaching and Evaluation of Clinical Skills)." Academic Medicine 68(6): 471-474.

Millward, C. (1970). "Familiar English for Foreign Doctors." New England Journal of Medicine 283(8): 430-431.

Milmoe, S., R. Rosenthal, H. T. Blane, M. E. Chafetz and I. Wolf (1967). "The doctor's voice: Postdictor of successful referral of alcoholic patients." Journal of Abnormal Psychology 71(1): 78-84.

Milton, G. W. (1992). "Letter to the Editor. Overseas medical practitioners and "racial discrimination"." Medical Journal of Australia 156: 138.

Minas, I. H., G. W. Stuart and S. Klimidis (1994). "Language, culture and psychiatric services: a survey of Victorian clinical staff." Australian and New Zealand Journal of Psychiatry 28(2): 250-258.

Moore, W. (1991). Introduction to Malaysia. Hong Kong, Odyssey.

Morris, D. (1982). Manwatching. Glasgow, Collins.

Morris, S. and W. Hudson (1995). "International education and innovative approaches to university teaching." Australian Universities' Review 2: 70-74.

Moss, I. (1992). "Letter to the Editor. Overseas medical practitioners and "racial discrimination"." Medical Journal of Australia 156: 138.

Mullins, G. and L. Hancock (1990). The experience of overseas students at the three South Australian universities: the University of Adelaide data. Adelaide, Advisory Centre for University Education (ACUE), University of Adelaide.

Mullins, H. and U. McGowan (1994). Assessment of Student Progress. Adelaide, University of Adelaide.

Mullins, H. (1994a). Report to the Language Development Committee on the Language Development Program: Term 1 1994. Adelaide, University of Adelaide.

Mullins, H. (1994b). Summary Report on the Language Development Program in Semester 2 1994. Adelaide, University of Adelaide.

Mullins, H. (1995a). The Introduction of a Language Development Program in the Medical School at the University of Adelaide. The Inaugural Pacific Rim First Year Experience Conference, Queensland University of Technology, Brisbane, Australia.

Mullins, H. (1995b). Memorandum from Helen Mullins. Adelaide, University of Adelaide.

Mullins, G., N. Quintrell and L. Hancock (1995). "The Experience of International and Local Students at Three Australian Universities." Higher education Research and Development 14(2): 201-231.

Munro, J. (1988). "Adapting Australian Higher Education For Overseas Students." HERDSA News 10: 5-8.

Murphy-Cullen, C. L. and L. C. Larsen (1984). "Interaction between the Socio-Demographic Variables of Physicians and their Patients: Its Impact upon Patient Satisfaction." Social Science and Medicine 19(2): 163-166.

Myerscough, P. R. (1992). Talking with Patients. A Basic Clinical Skill. Oxford, Oxford University Press.

Neame, R. L. B., D. A. Powis and T. Bristow (1992). "Should medical students be selected only from recent school-leavers who have studied science?" Medical Education 26: 433-440.

Nessa, J. and K. Malterud (1990). "Discourse Analysis in General Practice: A Sociolinguistic Approach." Family Practice 7(2): 77-83.

Nessa, J. (1995). "From a medical consultation to a written text." Scandinavian Journal of Primary Health Care 13: 89-92.

Neumann, R. (1985). "English Language Problems and University Students from a Non-English Speaking Background." Higher Education Research and Development 4(2): 193-202.

Newble, D. I., J. Hoare and R. G. Elmslie (1981). "The validity and reliability of a new examination of the clinical competence of medical students." Medical Education 15: 46-52.

Newble, D. I. and D. B. Swanson (1988). "Psychometric characteristics of the objective structured clinical examination." Medical Education 22: 325-334.

Newble, D. and G. Mullins (1990). The diagnosis and management of student learning problems. Breaking the Mould: Fifteenth Annual Conference of the Higher Education Research Society of Australasia., University of Adelaide, South Australia, HERDSA.

Newble, D. I. and E. J. Hejka (1991). "Approaches to learning of medical students and practising physicians: some empirical evidence and its implications for medical education." Educational Psychology 11(3&4): 333-342.

Newble, D. and R. Cannon (1994). A Handbook for Teachers in Universities and Colleges. A Guide to Improving Teaching Methods. London, Kogan Page.

Nieman, L. Z. and R. Thomas (1987). "Training and Evaluating Teams of Simulated Patients." Southern Medical Journal 80(9): 1164-1167.

Ninnes, P. M. (1994). Factors influencing pass rates in undergraduate courses at the University of Adelaide. Adelaide, South Australia, University of Adelaide.

Norcini, J. J., P. L. Stillman, A. I. Sutnick, M. B. Regan, H. L. Haley, R. G. Williams and M. Friedman (1993). "Scoring and Standard Setting with Standardized Patients." Evaluation and the Health Professions 16(3): 322-332.

Norman, G. R., P. L. Stillman and C. A. Woodward (1983). Simulated patients in evaluation of medical education and practice. Proceedings of the 22nd Annual Conference on Research in Medical Education.

Notzer, N., L. Weinbach and E. Lauden (1995). "Hebrew knowledge of new immigrant physicians." Harefuah 128(11): 694-696.

O'Connell, C. A. and C. Thayer-Doyle (1993). "Examining Whether Certain Myers-Briggs "Personality Preferences" Can Be Used as Criteria to Select Standardized Patients." Academic Medicine 68(12): 931.

O'Hagan, J. J., L. J. Davies and R. K. Pears (1986). "The use of simulated patients in the assessment of actual clinical performance in general practice." New Zealand Medical Journal 99: 948-951.

O'Hanlon, A., H. Winefield, E. Hejka and A. Chur-Hansen (1995). "Initial responses of first-year medical students to problem-based learning in a behavioural science course: role of language background and course content." Medical Education 29: 198-204.

Omoigui, N. A. (1991). "Clinical Skills of Foreign Medical Graduates." Annals of Internal Medicine 115(2): 158.

Ong, L. M. L., J. C. J. M. De Haes, A. M. Hoos and F. B. Lammes (1995). "Doctor-Patient Communication: A Review of the Literature." Social Science and Medicine 40(7): 903-918.

Orth, J. E., W. B. Stiles, L. Scherwitz, D. Hennrikus and C. Vallbona (1987). "Patient Exposition and Provider Explanation in Routine Interviews and Hypertensive Patients' Blood Pressure Control." Health Psychology 6(1): 29-42.

Owen, A. and R. Winkler (1974). "General practitioners and psychosocial problems. An evaluation using pseudopatients." Medical Journal of Australia 2(11): 393-398.

Parker, G. B. (1993). "On our selection: predictors of medical school success." The Medical Journal of Australia 158: 747-751.

Parsons, T. (1952). The Social System. London, Tavistock Publications Ltd.

Pattison, B. (1976). "Tests in Language Proficiency." Proceedings of the Royal Society of Medicine 69: 813.

Payne, B. (1979). "The Medical Record as a Basis for Assessing Physician Competence." Annals of Internal Medicine 91: 623-629.

Peay, M. Y. (1977). "The use of clinical observations in the teaching of behavioural science." Journal of Medical Education 52: 685-686.

Pendelton, D. A. and S. Bochner (1980). "The Communication of Medical Information in General Practice Consultations as a Function of Patients' Social Class." Social Science and Medicine 14A: 669-673.

Pendelton, D. (1983). Doctor-Patient Communication: A Review. Doctor-Patient Communication. D. Pendelton and J. Hasler. London, Academic Press: 5-53.

Petrusa, E. R., T. A. Blackwell, L. P. Rogers, C. Saydjari, S. Parcel and J. C. Guckian (1987). "An Objective Measure of Clinical Performance." American Journal of Medicine 83: 34-42.

Phillips, D. J. (1988). "Academic Adjustment to the Needs of Overseas Students." Babel: 10-17.

Phillips, D. J. (1990). Overseas students and their impact on the changing face of professional education in universities. AARE Annual Conference: The Changing Face of Professional Education., Sydney University.

Phillips, D. J. (1991). Research into the Professional Development of Tertiary Teaching for Academics: With Special Reference to Cross Cutural and Overseas Student Interaction. Belconnen, University of Canberra.

Pickering, G. (1978). "Medicine on the brink: The dilemma of a learned profession." Perspectives in Biology and Medicine(Summer): 551-560.

Pieters, H. M., F. W. W. M. Touw-Otten and R. A. De Melker (1994). "Simulated patients in assessing consultation skills of trainees in general practice vocational training: a validity study." Medical Education 28: 226-233.

Pietroni, P. (1976). Non-verbal communication in the general-practice surgery. Language and Communication in General Practice. B. A. Tanner. London, Hodder and Stoughton: 162-179.

Pilowsky, I. and H. R. Winefield (1976). "The Behavioural Science Course for Adelaide Medical Students." Medical Journal of Australia 1: 165-166.

Pilowsky, I., H. R. Winefield, M. Y. Peay, M. Augoustinos and A. Chur-Hansen (1989). Letter to the Working Party on Language and Study Skills Support. Adelaide, University of Adelaide.

Poirier, S. and D. J. Brauner (1988). "Ethics and the Daily Language of Medical Discourse." Hastings Center Report 18(4): 5-9.

Poirier, S. and D. J. Brauner (1990). "The Voices of the Medical Record." Theoretical Medicine 11(1): 29-39.

Policastro, A. M. (1990). "Letter to the Editor." Developmental and Behavioral Pediatrics 11(2): 103.

Pololi, L. H. (1995). "Standardised patients: as we evaluate, so shall we reap." The Lancet 345: 966-968.

Poole, A. D. and R. W. Sanson-Fisher (1979). "Understanding the patient: A neglected aspect of medical education." Social Science and Medicine 13A: 37-43.

Poole, A. D. and R. W. Sanson-Fisher (1980). "Long-term effects of empathy training on the interview skills of medical students." Patient Counselling and Health Education: 125-127.

Posen, S. (1968). "The Asian Undergraduate in Australia." The Medical Journal of Australia 2(12): 483-489.

Pousada, L. (1995). "Hispanic-American Elders. Implications for Health-Care Providers." Ethnogeriatrics 11(1): 39-52.

Powers, D. E. and C. W. Stansfield (1983). The Test of Spoken English as a measure of communicative ability in the health professions: Validation and standard setting. Princeton, New Jersey, Educational Testing Service.

Powis, D. A., R. L. B. Neame, T. Bristow and L. B. Murphy (1988). "The objective structured interview for medical student selection." British Medical Journal 296: 765-768.

Powis, D. (1996). Oral Assessor for Medical Student Selection Training Session. Presented at the Faculty of Medicine Division of Health Sciences, University of Adelaide.

Poyatos, F. (1984). Linguistic Fluency and Verbal-Nonverbal Cultural Fluency. Nonverbal Behavior. Perspectives, Applications, Intercultural Insights. A. Wolfgang. Lewiston, New York, C.J. Hogrefe: 431-459.

Prather, E. M., S. V. A. Breecher, M. L. Stafford and E. M. Wallace (1981). Screening Test of Adolescent Language (STAL). Seattle, University of Washington Press.

Prather, E. M., A. C. Brenner and K. S. Hughes (1981). "A Mini-Screening Language Test for Adolescents." Language, Speech and Hearing Services in Schools 12(2): 67-73.

Presswell, N. and J. Stanton (1992). "Does the doctor listen?" Medical Journal of Australia 156: 189-191.

Prideaux, D. (1992). 'We Are Only Interested in the Exam'. Assessment and Self Directed Learning in Medical Education. Fifth Ottawa Conference: Approaches to the Assessment of Clinical Competence., Centre for Medical Education, Ninewells Hospital and Medical School, Dundee, Scotland, Page Brothers.

Quintrell, N. (1991). Comments on surveys of overseas students carried out by Flinders University, the University of Adelaide and the University of South Australia, 1990-1991. Adelaide, Flinders University.

Quintrell, N. and M. Westwood (1994). "The Influence of a Peer-Pairing Program on International Students' First Year Experience and Use of Student Services." Higher Education Research and Development 13(1): 49-57.

Radford, M. H. B. and L. K. Wong (1986). Medical Care and Counselling Services and the Malaysian Overseas Student. Adelaide, Flinders University of South Australia.

Razali, S. M. (1996). "Medical school entrance and career plans of Malaysian medical students." Medical Education 30: 418-423.

Reiderer, R. E. (1970). "Native Language Barriers: Letter to the Editor." New England Journal of Medicine 283(16): 768.

Remland, M. S., T. S. Jones and H. Brinkman (1991). "Proxemic and Haptic behavior in three European countries." Journal of Nonverbal Behavior 15(4): 215-232.

Remland, M. S., T. S. Jones and H. Brinkman (1995). "Interpersonal Distance, Body Orientation and Touch: Effects of Culture, Gender and Age." Journal of Social Psychology 135(3): 281-297.

Rethans, J. J. E. and C. P. A. van Boven (1987). "Simulated patients in general practice: a different look at the consultation." British Medical Journal 294: 809-812.

Rethans, J., F. Sturmans, R. Drop and C. Van der Vleuten (1991). "Assessment of the performance of general practitioners by the use of standardized (simulated) patients." British Journal of General Practice 41: 97-99.

Rethans, J., R. Drop, F. Sturmans and C. Van der Vleuten (1991). "A method for introducing standardized (simulated) patients into general practice consultations." British Journal of General Practice 41: 94-96.

Reynolds, J. F., D. C. Mair and P. C. Fischer (1992). Writing and Reading Mental Health Records. Issues and Analysis. California, Sage Publications.

Rhodes, P. (1983). "Some hurdles for the overseas doctor." British Medical Journal 286: 1136-1137.

Richards, J., J. Platt and H. Weber, Eds. (1992). Longman Dictionary of Applied Linguistics. England, Longman.

Richards, T. (1994). "The Overseas Doctors Training Scheme: failing expectations." British Medical Journal 308: 1627-1631.

Robins, L. S. and F. M. Wolf (1988). "Confrontation and politeness strategies in physician-patient interactions." Social Science and Medicine 27(3): 217-221.

Roessler, R., J. W. Lester, W. T. Butler, B. Rankin and F. Collins (1978). "Cognitive and Noncognitive Variables in the Prediction of Preclinical Performance." Journal of Medical Education 53: 678-680.

Roland, C. G. (1970). "Clarifying the Literature: Letter to the Editor." New England Journal of Medicine 283(17): 933.

Rolfe, I. E. and R. Pearson (1994). "Communication skills of interns in New South Wales." Medical Journal of Australia 161: 667-670.

Rolfe, I. and J. McPherson (1995). "Formative Assessment: how am I doing?" Lancet 345: 837-839.

Rolfe, I. E., S. Pearson, D. A. Powis and A. J. Smith (1995). "Time for a review of admission to medical school?" Lancet 346: 1329-1333.

Rolfe, I. E., J. M. Andren, S. Pearson, M. J. Hensley and J. J. Gordon (1995). "Clinical competence of interns." Medical Education 29: 225-230.

Rolfe, I. E. and D. A. Powis (1997). "Selecting Australian doctors of the future." Medical Journal of Australia 166(5): 229-230.

Rosenhan, D. L. (1973). "On being sane in insane places." Science 179: 250-258.

Roter, D. L. and J. A. Hall (1993). Doctors Talking with Patients/Patients Talking with Doctors. Westport, Auburn House.

Rowland-Morin, P. A. and J. G. Carroll (1990). "Verbal Communication Skills and Patient Satisfaction. A Study of Doctor-Patient Interviews." Evaluation and the Health Professions 13(2): 165-185.

Rusek, W. (1992). IELTS: Does it Predict Success at University? ELICOS Association 5th Annual Education Conference, Adelaide, South Australia.

Rush, W. E. (1972). "Medical Orientation U.S.A. Program: A preliminary report." Journal of Medical Education 47(6): 440-442.

Rustemli, A. (1986). "Male and female personal space needs and escape reactions under intrusion: A Turkish sample." International Journal of Psychology 21: 503-511.

Rutala, P. J., D. B. Witzke, E. O. Leko, J. V. Fulginiti and P. J. Taylor (1991). "Sharing of Information by Students in an Objective Structured Clinical Examination." Archives of Internal Medicine 151: 541-544.

Rutala, P. J., D. B. Witzke, E. O. Leko and J. V. Fulginiti (1992). Gender Effect at High and Low Cut-off Points in an Objective Structured Clinical Examination. Fifth Ottawa Conference: Approaches to the Assessment of Clinical Competence., Centre for Medical Education, Ninewells Hospital and Medical School, Dundee, Scotland, Page Brothers.

Samuelowicz, K. (1987). "Learning Problems of Overseas Students: Two Sides of a Story." Higher Education Research and Development 6(2): 121-133.

Sanders, D. (1993). "Talking about sex." The Practitioner 237: 674-679.

Sanson, A., M. Augoustinos, H. Gridley, M. Kyrios, J. Reser and C. Turner (1997). "Racism and Prejudice: Psychological Perspectives." Position Paper, Australian Psychological Society.

Sanson-Fisher, R. W. and A. D. Poole (1980). "Simulated patients and the assessment of medical students' interpersonal skills." Medical Education 14: 249-253.

Sanson-Fisher, R. and J. Cockburn (1997). "Effective teaching of communication skills for medical practice: selecting an apropriate clinical context." Medical Education 31: 52-57.

Scarpaci, J. L. (1988). "Help-Seeking Behavior, Use, and Satisfaction Among Frequent Primary Care Users in santiago de Chile." Journal of Health and Social Behavior 29: 199-213.

Schoonover, S. C., E. L. Bassuk, R. Smith and D. Gaskill (1983). "The Use of Videotape Programs to Teach Interpersonal Skills." Journal of Medical Education 58: 804-810.

Schutz, A. (1976). Collected papers II: Studies in Social Theory. The Hague, Martinus Nojhoff.

Schwenk, T. L., J. T. Marquez, D. Lefever and M. Cohen (1989). "Physician and Patient Determinants of Difficult Physician-Patient Relationships." Journal of Family Practice 28(1): 59-63.

Scott, N. and M. F. Weiner (1984). "Patientspeak: An exercise in communication." Journal of Medical Education 59: 890-893.

Seedhom, B. B., J. E. Smeathers and D. T. Thompson (1985). "Devising a course for overseas vistors who don't speak English well." British Medical Journal 291: 802-804.

Sefton, A. J. (1990). "Selection of Medical Students." The Medical Journal of Australia 153(October): 440-441.

Segaran, J. (1993). The Malaysian Australian Tertiary Education Scheme. Adelaide, University of Adelaide.

Shapiro, M. (1980). Doctor Talk. The Jargon of Medicine. Getting Doctored. Ontario, Between the Lines.

Shapiro, J. and P. Lenahan (1996). "Family Medicine in a Culturally Diverse World: A Solution-oriented Approach to Common Cross-cultural Problems in Medical Encounters." Family Medicine 28(4): 249-255.

Siegal, S. (1956). Nonparametric statistics for the behavioral sciences. New York, McGraw-Hill.

Singh, S. (1994). "Cultural adjustment and the overseas trainee." British Medical Journal 308: 1169.

Skakun, E. N. (1986). Construct validation: process and methodology. Newer Developments in Assessing Clinical Competence., Ottawa, Canada, Heal Publications Ltd.

Smith, K. F. and J. R. Baldauf (1982). "The concurrent validity of self-rating with interviewer rating on the Australian Second Language Proficiency Ratings Scale." Educational and Psychological Measurement 42: 1117-1124.

Smith, R. (1987). "Prejudice against doctors and students from ethnic minorities." British Medical Journal 294: 328-329.

Smith, R. C. and R. Hoppe (1991). "The Patient's Story: Integrating the Patient- and Physician-centered Approaches to Interviewing." Annals of Internal Medicine 115(6): 470-477.

Smith, A. J., J. D. Hamilton, I. E. Rolfe and S. Pearson (1995). "Letter to the Editor. Communication skills of interns in New South Wales." The Medical Journal of Australia 162: 498.

Snowden, C. (1995). Successful Medical School Program speaks for itself. Adelaidean. 4: 7.

Solomon, D. J., A. J. Speer, M. A. Ainsworth and D. J. DiPette (1993). "Investigating Gender Bias in Preceptors' Ratings of Medical Students." Academic Medicine 68(9): 703.

Sommers, R. K. (1985). Review of Screening Test of Adolescent Language. The Ninth Mental Measurements Yearbook. J. V. Mitchell. Lincoln, Nebraska, The Buros Institute of Mental Measurements.: 1332.

Spack, R. (1988). "Initiating ESL Students Into the Academic Discourse Community: How Far Should We Go?" TESOL Quarterly 22(1): 29-51.

Stephens, M. I. and A. A. Montgomery (1985). "A critique of recent relevant standarised tests." Topics in Language Disorders 5(3): 21-45.

Stevens, R. A. (1995). "International Medical Education and the Concept of Quality: Historical Reflections." Academic Medicine 70(7): S11-S18.

Stewart, M. A. (1984). "What is a successful doctor-patient interview? A study of intercations and outcomes." Social Science and Medicine 19(2): 167-175.

Stiles, W. B. (1978). "Verbal Response Modes and Dimensions of Interpersonal Roles: A Method of Discourse Analysis." Journal of Personality and Social Psychology 36(7): 693-703.

Stiles, W. B. and S. M. Putnam (1990). Analysis of Verbal and Nonverbal Behavior in Doctor-Patient Encounters. Communicating with Medical Patients. M. Stewart and D. Roter. Newbury Park, Sage: 211-222.

Stillman, P. L., J. S. Ruggill and D. Sabers (1978). "The use of practical instructors to evaluate a complete physical examination." Evaluation and the health professions 1(1): 49-54.

Stillman, P. L., J. S. Ruggill, P. J. Rutala and D. L. Sabers (1980). "Patient Instructors as Teachers and Evaluators." Journal of Medical Education 55(March): 186-193.

Stillman, P. L., P. J. Rutala, G. I. Nicholson, D. L. Sabers and A. E. Stillman (1982). Measurement of Clinical Competence of Residents Using Patient Instructors. Twenty-first Annual Conference on Research in Medical Education.

Stillman, P. L., M. Y. Burpeau-Di Gregorio, G. I. Nicholson, D. Sabers and A. E. Stillman (1983). "Six Years of Experience Using Patient Instructors to Teach Interviewing Skills." Journal of Medical Education 58(December): 941-946.

Stillman, P. L. (1986). Evolution of the standardized patient program. Newer Developments in Assessing Clinical Competence, Ottawa, Canada, Heal Publications Ltd.

Stillman, P. L., D. B. Swanson, S. Smee, A. E. Stillman, T. H. Ebert, V. S. Emmel, J. Caslowitz, H. L. Greene, M. Hamolsky, C. Hatem, D. J. Levenson, T. Levin, G. Levinson, B. Ley, G. J. Morgan, T. Parrino, S. Robinson and J. Willms (1986). "Assessing Clinical Skills of Residents with Standardized Patients." Annals of Internal Medicine 105: 762-771.

Stillman, P. L. and D. B. Swanson (1987). "Ensuring the Clinical Competence of Medical School Graduates Through Standardized Patients." Archives of Internal Medicine 147: 1049-1052.

Stillman, P. L., M. B. Regan, D. B. Swanson and the Fourth-year Performance Assessment Task Force Group (1987). "A Diagnostic Fourth-Year Performance Assessment." Archives of Internal Medicine 147: 1981-1985.

Stillman, P. L., M. B. Regan, M. Philbin and M. Haley (1990). "Results of a Survey on the Use of Standardized patients to teach and Evaluate Clinical Skills." Academic Medicine 65(5): 288-292.

Stillman, P. L., M. B. Regan, D. B. Swanson, S. Case, J. McCahan, J. Feinblatt, S. R. Smith, J. Willms and D. V. Nelson (1990). "An Assessment of the Clinical Skills of Fourth-Year Students at Four New England Medical Schools." Academic Medicine 65(5): 320-326.

Stillman, P. L., H. A. Hayley, A. I. Sutnick, M. M. Philbin, S. R. Smith, J. O'Donnell and H. Pohl (1991). "Is Test Security an Issue in a Multistation Clinical Assessment?-A Preliminary Study." Academic Medicine 66(9 (Supplement)): S25-S27.

Stillman, P. L., D. B. Swanson and M. B. Regan (1991). "In Response: Letter from N.A. Omoigui "Clinical Skills of Foreign Medical Students"." Annals of Internal Medicine 115(2): 158-159.

Stillman, P. L., D. Swanson, M. B. Regan, M. M. Philbin, V. Nelson, T. Ebert, B. Ley, T. Parrino, J. Shorey, A. Stillman, E. Alpert, J. Caslowitz, D. Clive, J. Florek, M. Hamolsky, C. Hatem, J. Kizirian, R. Kopelman, D. Levenson, G. Levinson, J. McCue, H. Pohl, F. Schiffman, J. Schwartz, M. Thane and M. Wolf (1991). "Assessment of Clinical Skills of Residents Utilizing Standardized Patients. A Follow-up study and Recommendations for Application." Annals of Internal Medicine 114(5): 393-401.

Stillman, P. L. and W. D. Sawyer (1992). "A New Program to Enhance the Teaching and Assessment of Clinical Skills in the People's Republic of China." Academic Medicine 67(8): 495-499.

Stillman, P. L., M. B. Regan, H. A. Haley, J. J. Norcini, M. Friedman and A. I. Sutnick (1992). "The Use of a Patient Note to Evaluate Clinical Skills of First-year Residents Who Are Graduates of Foreign Medical Schools." Academic Medicine 67(10): S57-S59.

Stillman, P. L. (1993). "Technical Issues: Logistics." Academic Medicine 68(6): 464-470.

Stone, G. C. (1979). "Patient Compliance and the Role of the Expert." Journal of Social Issues 35(1): 34-59.

Street, R. L., R. M. Brady and W. B. Putman (1983). "The Influence of Speech Rate Stereotypes and Rate Similarity or Listeners' Evaluations of Speakers." Journal of Language and Social Psychology 2(1): 37-56.

Sussman, N. M. and H. M. Rosenfeld (1982). "Influence of Culture, Language, and Sex on Conversational Distance." Journal of Personality and Social Psychology 42(1): 66-74.

Sutnick, A. I. (1970a). "First Symposium on the Problems of Foreign Medical Graduates." JAMA 213(13): 2241-2246.

Sutnick, A. I. (1970b). "Clarifying the Literature: Letter to the Editor." New England Journal of Medicine 283(17): 933-934.

Sutnick, A. I., J. F. Reichard and A. P. Angelides (1971). "Orientation of Foreign Medical Graduates." International Education and Cultural Exchange 6: 91-99.

Sutnick, A. I., P. R. Kelley and D. Knapp (1972). "The English Language and the FMG." Journal of Medical Education. 47: 434-439.

Sutnick, A. I., P. L. Stillman, J. J. Norcini, M. Friedman, M. B. Regan, R. G. Williams, E. K. Kachur, M. A. Haggerty and M. P. Wilson (1992). ECFMG Large Scale Clinical Competence Assessment Study. Fifth Ottawa Conference: Approaches to the Assessment of Clinical Competence., Centre for Medical Education, Ninewells Hospital and Medical School, Dundee, Scotland, Page Brothers.

Sutnick, A. I., M. L. Shafron and M. P. Wilson (1992). "Impact of the New United States Medical Licensing Examination on the Certification Process of the Educational Commission for Foreign Medical Graduates." International Journal of Dermatology 31: 798-799.

Sutnick, A. I., P. L. Stillman, J. J. Norcini, M. Friedman, M. B. Regan, R. G. Williams, E. K. Kachur, M. A. Haggerty and M. P. Wilson (1993). "ECFMG Assessment of Clinical Competence of Graduates of Foreign Medical Schools." Journal of the American Medical Association 270(9): 1041-1045.

Sutnick, A. I., P. L. Stillman, J. J. Norcini, M. Friedman, R. G. Williams, D. A. Trace, M. A. Schwartz, Y. Wang and M. P. Wilson (1994). "Pilot Study of the Use of the ECFMG Clinical Competence Assessment to Provide Profiles of Clinical Competencies of Graduates of Foreign Medical Schools for Residency Directors." Academic Medicine 69(1): 65-67.

Swanson, D. B. and J. J. Norcini (1989). "Factors Influencing Reproducibility of Tests Using Standardized Patients." Teaching and Learning in Medicine 1(3): 158-166.

Swanson, D. B. and P. L. Stillman (1990). "Use of standardised patients for teaching and assessing clinical skills." Evaluation and the health professions. 13(1): 79-103.

Swartz, M. H. (1992). ASPIRE: Advancement of Standardized Patients In Research and Education. Fifth Ottawa Conference: Approaches to the Assessment of Clinical Competence., Centre for Medical Education, Ninewells Hospital and Medical School, Dundee, Scotland, Page Brothers.

Sweetman, K. (1996). Asian student backlash to Hanson. The Advertiser. Adelaide: 4.

Tamblyn, R., G. Schnabl, D. Klass, M. Kopelow and M. Marcy (1988). How Standardized are Standardized Patients? Research in Medical Education.

Tamblyn, R. M., D. J. Klass, G. K. Schnabl and M. L. Kopelow (1991). "The accuracy of standardized patient presentation." Medical Education 25: 100-109.

Tan, K. M. (1977). "Foreign Medical Graduate Performance-a Review." Medical Care 15(10): 822-829.

Tertiary Multicultural Education Committee (1988). Multiculturalism in Tertiary Education, Office of Tertiary Education.

Thomas, J. B. and B. N. Steele (1966). "A Linguistic Approach to Medical Education." Journal of Medical Education 41: 697-700.

Thomas, D. R. (1974). "Social Distance in Fiji." Journal of Social Psychology 93: 181-185.

Thompson, R. M. (1990). "Writing-Proficiency Tests and Remediation: Some Cultural Differences." TESOL Quarterly 24: 99-102.

Thompson, A. N. (1990). "Case specificity of performance with simulated patients." New Zealand Medical Journal 103: 372-374.

Thompson, A. N. (1994). "Reliability of consumer assessment of communication skills in a postgraduate family practice examination." Medical Education 28: 146-150.

Thong, Y. H. (1990). "Medical immigration and the medical workforce." Medical Journal of Australia 153(August 6): 124-125.

Tiller, J. W. G. and K. V. Jones (1984). "Countries of origin and languages of Australian medical students." Medical Education 18: 435-438.

Tiller, J. W. G. and K. V. Jones (1985). "Languages spoken by medical students." The Medical Journal of Australia. 142: 601-602.

Torda, T. A. (1995). "Letter to the Editor. Communication skills of interns in New South Wales." The Medical Journal of Australia 162: 497-498.

Tran-Dinh, H. D. (1990). "Letter to the Editor. Foreign medical graduates." Medical Journal of Australia 153: 569-570.

Truax, C. B. and R. R. Carkhuff (1967). Toward Effective Counseling and Psychotherapy: Training and Practice. Chicago, Aldine.

Tse, C. (1990). Teaching students from non English speaking backgrounds. The lost NESBS! Breaking the Mould: Fifteenth Annual Conference of the Higher Education Research Society of Australasia., University of Adelaide, South Australia, HERDSA.

Tutton, P. J. M. and S. J. Wigg (1990). "The influence of last secondary school attended, subjects taken in last year of secondary education and gender on preclinical performance of medical students." Australian Journal of Education 34(2): 168-173.

Uba, L. (1992). "Cultural Barriers to Health Care for Southeast Asian Refugees." Public Health Reports 107(5): 544-548.

University of Adelaide (1996). Annual Report 1995. Adelaide, University of Adelaide.

University of Adelaide (1996). Calendar: Handbook of Courses. Adelaide, University of Adelaide.

University of Adelaide (1997). Undergraduate Courses for International Students 1998. Adelaide, University of Adelaide.

University of Adelaide Department of Psychiatry (1987). Memorandum: Behavioural Science course for first year Medical and Dental students. Students with difficulties with English. Adelaide, University of Adelaide.

University of Adelaide Faculty of Medicine Curriculum Committee (1993). Report of Meeting of 13th December, 1993. Adelaide, University of Adelaide.

University of Adelaide Faculty of Medicine Curriculum Committee (1993). Report of Meeting of 5th November, 1993. Adelaide, University of Adelaide.

University of Adelaide Faculty of Medicine Curriculum Committee (1993). Report of Meeting of 21st June, 1993. Adelaide, University of Adelaide.

Van der Vleuten, C. P. M. and S. J. van Luyk (1986). A validity study of a test for clinical and technical medical skills. Newer Developments in Assessing Clinical Competence, Ottawa, Canada, Heal Publications Ltd.

Van der Vleuten, C. P. M. and D. B. Swanson (1990). "Assessment of clinical skills with standardized patients: State of the art." Teaching and Learning in Medicine 2(2): 58-76.

Van Naerssen, M.M. (1985). "Medical records: one variation of physicians' language." International Journal of Society and Language 51:43-73.

Varki, A. (1992). "Of Pride, Prejudice, and Discrimination." Annals of Internal Medicine 116(9): 762-764.

Vernon-Roberts, J. (1991). Application for a Clinical Skills Assistant. Internal document, University of Adelaide Faculty of Medicine.

Vernon-Roberts, J. and D. Yeatman (1994). "Bridging the Gap." Australian Nursing Journal 1(8): 30-31.

Vernon-Roberts, J. M. and A. Chur-Hansen (1995). "Letter to the Editor. Communication skills of interns in New South Wales." The Medical Journal of Australia 163: 112.

Vore, A. L. (1991). "Enhancing verbal communication skills and promoting effective socialization of newly hired Spanish-speaking registered nurses." Journal of Nursing Staff Development 7(6): 286-289.

Vroeijenstijn, A. I. (1995). "Quality Assurance in Medical Education." Academic Medicine 70(7): S59-S67.

Vu, N. V., M. L. Marcy, S. J. Verhulst and H. S. Barrows (1990). "Generalizability of Standardized Patients' Satisfaction Ratings of their Clinical Encounter with Fourth-year Medical Students." Academic Medicine 65(9): S29-S30.

Vu, N. V., M. M. Marcy, J. A. Colliver, S. J. Verhulst, T. A. Travis and H. S. Barrows (1992). "Standardized (simulated) patients' accuracy in recoding clinical performance check-list items." Medical Education 26: 99-104.

Vu, N. V., H. S. Barrows, M. L. Marcy, S. J. Verhulst, J. A. Colliver and T. Travis (1992). "Six Years of Comprehensive, Clinical, Performance-based Assessment Using Standardized Patients at the Southern Illinois University School of Medicine." Academic Medicine 67(1): 42-50.

Waitzkin, H. (1985). "Information giving in medical care." Journal of Health and Social Behavior 26: 81-101.

Wakeford, R., A. Farooqi, A. Rashid and L. Southgate (1992). "Does the MRCGP examination discriminate against Asian doctors?" British Medical Journal 305: 92-94.

Webb, J. (1993). CSHE Research Working Papers 93.7. Academic support for non-traditional students. Melbourne, Victoria, Centre for Study of Higher Education, University of Melbourne.

Weinberger, M., J. Y. Greene and J. J. Mamlin (1981). "The Impact of Clinical EncounterEvents on Patient and Physician Satisfaction." Social Science and Medicine 15E: 239-244.

Wible, D. S. and C. H. Hui (1985). "Perceived Language Proficiency and Person Perception." Journal of cross-Cultural Psychology 16(2): 206-222.

Wilkes, G. A. (1993). Exploring Australian English. Sydney, ABC.

Williams, B. and G. Wilkinson (1995). "Patient Satisfaction in Mental Health Care." British Journal of Psychiatry 166: 559-562.

Winefield, H. R. (1988). Language capabilities in incoming medical and dental students. University of Adelaide 1988. Adelaide, University of Adelaide.

Winefield, H. R. and T. G. C. Murrell (1991). "Speech patterns and satisfaction in diagostic and prescriptive stages of general practice consultations." British Journal of Medical Psychology 64: 103-115.

Winefield, H. R. and M. Y. Peay (1991). Behavioural Science in Medicine. Adelaide, The University of Adelaide.

Winefield, H. R. and T. G. C. Murrell (1992). "Verbal interactions in general practice. Information, support and doctor satisfaction." The Medical Journal of Australia 157: 677-682.

Winefield, H. R. (1993). Language skills and progress in Medicine at the University of Adelaide. Vocabulary and Behavioural Science, 1992. Adelaide, South Australia, University of Adelaide.

Winefield, H. R., T. G. Murrell and J. Clifford (1995). "Process and outcomes in General Practice consultations: Problems in defining high quality care." Social Science and Medicine 41: 969-975.

Winefield, H. R., T. G. C. Murrell, J. V. Clifford and E. A. Farmer (1995). "The usefulness of distinguishing different types of general practice consultation, or are needed skills always the same?" Family Practice 12(4): 402-407.

Winefield, H. R. (1996). "Counselling Work in Normal General Practice: An Analysis of Australian Consultations." Journal of Health Psychology 1(2): 223-234.

Winefield, H. R. and J. A. Weinman (1996). "Communication in Health Care: Process and Outcomes." Psychology and Health 11: 769-771.

Winefield, H. R., T. Murrell, J. Clifford and E. Farmer (1996). "The search for reliable and valid measures of patient-centredness." Psychology and Health 11: 811-824.

Winkler, R. C. (1974). "Research into Mental Health Practice Using Pseudopatients." Medical Journal of Australia 2: 399-403.

Woods, D. (1981). "Good English is good medicine." Canadian Medical Association Journal 125: 624-629.

Yanoff, K. L. and F. D. Burg (1988). "Types of Medical Writing and Teaching of Writing in U.S. Medical Schools." Journal of Medical Education 63: 30-37.

Young, J. W. (1979). "Symptom disclosure to male and female physicians: Effects of sex, physical attractiveness and symptom type." Journal of Behavioural Medicine 2: 159-169.

Young, J. A. (1995). "Letter to the Editor. Communication skills of interns in New South Wales." The Medical Journal of Australia 162: 497.

Zinn, W. (1993). "The Empathic Physician." Archives of Internal Medicine 153: 306-312.

Zwingmann, C. A. A. and A. D. G. Gunn (1983). Uprooting and Health. Psycho-social problems of students from abroad. Geneva, World Health Organisation.

09 PH
C 5595
c.2

UNIVERSITY OF ADELAIDE LIBRARY
13.11.98

Revisions to thesis based on examiners' comments:
"An Investigation of the English Language Proficiency and Academic and Clinical Performance of University of Adelaide Medical School Undergraduates".

Anna Chur-Hansen.

I would like to thank the examiners for their helpful comments and suggestions. They have clearly spent considerable time and effort in reading and assessing my work.

Largely, the first nine chapters of this thesis have remained unaltered. However, after consultation with my supervisors, I have amended the text of this thesis as described below. Note that where page numbers are indicated, they refer to the original version of the thesis.

The definition of the terms "clinical performance" and "academic performance" have been clearly stated, both in the Abstract and again in the General Discussion and Conclusions (Chapter X). This should clarify ambiguity about whether the studies conducted addressed the stated aims of the research.

The first section of Chapter X, which summarised the conclusions of each of the seven studies presented in the thesis, has been completely rewritten. Specifically, I have demonstrated how each study addressed the two primary aims of the thesis, why the study was necessary and how the study has made a contribution to knowledge. I have also indicated a study's methodological strengths and/or weaknesses. However, the reader should note that a more comprehensive critique of each study can be found in the Discussion section that follows each study.

Within this section, I have clearly stated the strengths and weaknesses of the LRS. A definitive statement about the utility of the LRS cannot be made until further research is conducted, and this is made clear both in the body of the thesis and again in Chapter X. Many of the "masses of statistical analyses" in the Chapters reporting Studies V and VI were conducted in order to assess the reliability and validity of the LRS and the WLRS, in an attempt to make clear and objective statements about the utility of these instruments for measuring spoken and written English language proficiency.

The final section of Chapter X, subheaded "Conclusion", has been deleted.

I feel that Chapter X, along with the Discussion sections of each study, clearly demonstrate my ability for critical thought.

One examiner indicated that "p293 - Rater bias last sentence - the argument given per se does not exclude "this as a possible explanation for the results." This criticism is accepted, and the sentence has been changed.

I have now indicated that in establishing alternate-forms reliability (p 107) the two tests must be administered in quick succession.

All grammatical suggestions have been considered and the appropriate alterations made. Similarly, there were several inaccuracies identified by the examiners, which have been altered (Table 4 to Table IV, eight studies not seven, Study VII not VI, Farnill and Hayes in the next chapter, not this chapter, "most useful" to "useful", "spoken" to read "written").

Some of the examiners' comments have been carefully considered but in consultation with my supervisors, alterations to the original thesis have not been made. For instance;

"p 128-138 - the multitude of tables makes it hard to assess differences in cohorts studied. It would be more useful to compare cohorts regarding various criteria on the same tables." I have considered this as an option, but on compiling tables found them to be rather "busy". Thus, I have left the tables as originally presented.

"Table XXXVI needs further clarification as it seems beyond the realms of coincidence that the number of comments, relating to the number of students by a number of tutors should be absolutely identical for the 1992 III year students when they progress to the 1993 IV year status." In fact, this finding was coincidental. The comments pertained to some of the same students over the two years, but not all of the same students, and some of the same clinicians made comments in both years, but again, not all of the same clinicians did so. This is explained in the text (p 190 to p 191).

"Chapter IV uses the author as rater for STAL but Chapter III used rater 1. Consistency would be better unless the variation can be justified". Rater 1 refers to the collective raters who were Faculty staff members (see footnote 14). Rater 2 refers to myself. In Chapter III Rater 1 scores were used because these were the scores with which decisions about compulsory participation in the Language Development Programme (LDP) were based. In the following study, the third year cohort of students who were tested on the STAL were not eligible to participate in the LDP. Thus, there were no Rater 1 scores for them, only Rater 2 scores. For consistency within the Study, I chose to compare my scores for the two groups of students.

"Scores for the WKT were normally distributed while scores for STAL were skewed - this could suggest that the two tests are measuring different things. This possibility is not discussed." In fact, whether the two tests are tapping into the same skills is discussed on p 174 to p 175, where the correlation coefficients are presented for the STAL and WKT, and a consideration of what these correlations might indicate is made.

"p171-175 - females did better than males, therefore what was the gender breakdown of MATES and non-MATES?" I have not gone further into this, as on p 171 I have stated that method of entry into the course and gender were not related.

"Study V - it would be interesting to compare the ratings for each individual standardised patient with the other raters who examined the identical group of students." Whilst I concur that it would be interesting, I do not believe that this information would be helpful in further informing us about the reliability or validity of the LRS, given the other, extensive analyses conducted.

"p248 - it seems quite appropriate for a medical student to use a medical register including terms such as "bowel movement", "exacerbate", "localised pain" when addressing a patient. These words are in general usage. Whilst this is a matter of opinion, obviously, I do not agree that students or practitioners should use such language until they have established that the patient does understand these words. Research shows that terms that might be considered common are not necessarily understood by patients (or indeed by medical and nursing staff) (Hadlow, J. and Pitts, M. (1991) "The understanding of common health terms by doctors, nurses and patients", *Social Science and Medicine*, 32 (2), 193-196).

"p259 - instructions to students were to report in writing "so that another health care professional would understand". The later instruction, to "avoid use of specialised medical terminology" is therefore somewhat contradictory." I refer the examiner to the point made above, in answer to this. Furthermore, I would point out that this instruction to students was written by the Language Development Committee, of which I was one member. Thus, I did not have complete control over this aspect of the research.

"p306 - paragraph 1 - the newspaper article (Appendix XXXVI) does not really "highlight the importance of adequate writing skills." This is a debatable point, I admit, but I do think that a simple spelling mistake being implicated in the death of two people would indicate that accurate written language is important for the effective delivery of health care.

Anna Chur-Hansen

July 2, 1998.