Auditory measures of intelligence and intelligence-related functions:

Where have we come from and where to from here?

Ian Taylor Zajac

BHlthSci (Hons) (Psych), BSocSci

A Thesis Submitted in Fulfilment of the Requirements for the Degree of

Doctor of Philosophy

School of Psychology, The University of Adelaide

September, 2011

DECLARATION

I, Ian Zajac certify that this work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institute of higher learning and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in text.

I give consent to this copy of my thesis when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

I acknowledge that copyright of published work contained within this thesis (as listed below*) resides with the copyright holder(s) of those works. I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library catalogue and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

* Zajac, I. T., & Burns, N. R. (2011). Relationships between three auditory inspection time tasks and processing speed. *Australian Journal of Psychology*, 63, 163-172. doi: 10.1111/j.1742-9536.2011.00020.x

* Zajac, I. T., & Burns, N. R. (*in press*). Do auditory temporal discrimination tasks measure temporal resolution of the CNS? *Psychology*

Ian T. Zajac

Signed……………………………………Date……………………

ACKNOWLEDGEMENTS

Many people have encouraged me throughout this adventure. From kind words through to stern words; from well thought out advice through to generous editorial revisions: without all of these people, this dissertation wouldn't have been achieved. So, I thank all my friends in their various shapes and sizes. Thanks must go specifically to Daniel King and Anthony Venning. They opened their office space and accommodated my personality. More importantly, we shared our lives. Thank you both for your continued friendship.

My supervisor, Nicholas Burns, always encouraged my journey into academia and supported my sometimes abstract ideas. He accommodated my needs and provided expert advice on the studies presented herein. Thank you, Nick, for the laughs but more importantly, your understanding of my personal circumstances. I hope our relationship transitions from supervisor - student to life-long friends. Gratitude must also go to my co-supervisors Vanessa Danthiir and Ted Nettelbeck who have both inspired me and supported my endeavours when requested. I hope to collaborate with all of you as my peers in the future.

Three other extremely important people have been instrumental in my success. My boys, Taylor and Zachary, I hope I have been an example for you. Never settle for second-best. Always strive to achieve the things you desire but most importantly, do not waste your lives pursuing things that will not make you happy. You are the only people in control of your futures. My better-half Jodie: Thank you for always supporting me in the past, but also into the future. You have never questioned the logic behind my undertaking these arduous tasks. You make me proud to be a father and husband. You and the boys are the centre of my universe. I love you all.

Finally, I dedicate this dissertation to my late Dziadek. You were a very smart man, born in the wrong era filled with war and hate. You never achieved your full potential but you understood the importance of higher education. You always expressed your desires for my future during the countless hours we spent discussing this world. And, to my mother and father, you were the first to heed Dziadek's words and break-the-mould. You have always taught me to be strong willed and independent but most importantly, you have taught me that the only thing standing between myself and a successful future is a little bit of effort.

I sincerely hope I have made each and every one of you proud. Much love.

SUMMARY

Much of what is known of the nature of human intelligence derives from research concerned primarily with visually presented tests. This thesis, on the other hand, sought to extend this knowledge by exploring intelligence as expressed in the auditory modality. More specifically, the dissertation reports on a series of studies designed to broadly examine the extent to which existing broad ability constructs defined in modern intelligence taxonomies are expressed in performance on auditory tasks. The reason for doing so reflects the largely untested assumption that constructs like fluid and crystallised intelligence, for example, are not modality specific but cognitively general.

Study 1 aimed to purposely design auditory tests to measure the broad construct general speed of processing (Gs). N=96 university undergraduates completed these new auditory tasks together with a selection of existing putative auditory Gs measures, and also a selection of established visual Gs marker tests. The new auditory tasks were found to display good reliability and, together with the visual tests, they defined moderately correlated broad Gs and RT speed factors.

Study 2 extended the findings of Study 1. In particular, in light of the under-representation of visual tasks in the previous study, it sought to increase the number of visual Gs tests to investigate the presence of modality specific speed factors. N=80 university undergraduates completed the test battery. Several structural models were tested in which modality specific speed factors were specified; however, these models were not supported. Instead, results supported those found in Study 1: auditory and visual tests combine to define broad Gs and RT speed factors.

Study 3 examined existing measures of temporal discrimination. Although it has recently been suggested that these tasks provide a direct measure of neural efficiency they appear in fact to be cognitively complex, possibly relying on memory functions. Therefore, N=66 university undergraduates completed a battery of tests measuring temporal discrimination, memory (Gm) and speed (Gs). Results showed that temporal tasks related more strongly and consistently with Gm than Gs. Further re-analysis of previously published data supported these findings with Gm functions mediating the relationship between temporal tasks and general intelligence.

Study 4 explored Auditory Inspection Time (AIT) tasks and their relationship to measures of Gs. N=96 university undergraduates completed these tasks and measures of Gs. Of the three variations of AIT, only the spatial version related well to Gs. A distinct and independent AIT factor was also identified which supports previous suggestions that these tasks largely reflect auditory perceptual processes.

Based on the results of all studies it is concluded that broad constructs defined in intelligence theories can be indexed auditorily provided that sufficient effort is devoted to ensuring the auditory tasks emphasise the cognitive processes underpinning the constructs of interest. Further, it is recommended that future studies of the auditory modality consider their tasks in relation to existing broad constructs and that auditory tasks be incorporated into intelligence testing.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

AC	Audio Code
ACoR	Auditory Cognition of Relations
AFF	Auditory Flutter Fusion
AIT	Auditory Inspection Time
AIT-L	Auditory Inspection Time – Loudness
AIT-P	Auditory Inspection Time – Pitch
AIT-S	Auditory Inspection Time – Spatial Localisation
AP_d	Auditory Pitch Discrimination
APM	Raven's Advanced Progressive Matrices
ART	Auditory Reaction Time
AT_d	Auditory Temporal Discrimination
CCFT	Cattell Culture Fair Test
CD_A	Chasing Digits Auditory
CD_V	Chasing Digits Visual
CFA	Confirmatory Factor Analysis
DaSP	Discrimination among Sound Patterns
DD	Duration Discrimination
DM	Dot Matrix
ECT	Elementary Cognitive Task
EFA	Exploratory Factor Analysis
FAs	Finding (Letter) As
FSIQ	Full Scale IQ (Intelligence Quotient)
g	General Intelligence
Ga	Broad Auditory Reception
Gc	Crystallised Intelligence

Gf	Fluid Intelligence
Glr	Learning and Retrieval; Long-term Memory
GRT	General Reaction Time
Gs	Speed of Processing
Gsm	Short-term Memory
Gt	General Timing Ability/Factor
Gv	Visualisation
HAs	Hears (Letter) As
IT	Inspection Time
NC	Number Comparison
RP	Rhythm Perception
RPM	Raven's Progressive Matrices
RT	Reaction Time
SD	Symbol Digit
SEM	Structural Equation Modeling
SPuD	Speech Perception under Distraction
TC	Tone Comparisons
TG	Temporal Generalisation
TOJ	Tonal Order Judgement
VIT	Visual Inspection Time
VL_d	Visual Line-Length Discrimination
VRT	Visual Reaction Time
VT_d	Visual Temporal Discrimination
WM	Working Memory

Chapter 1: Introduction and Literature Review

Preamble

Much of what is known regarding the nature of human intelligence results from research concerned primarily with visually presented cognitive tests. This thesis, on the other hand, concerns itself with advancing this knowledge through exploring the measurement of intelligence via modalities other than visual. More specifically, it reports on a series of studies utilising both existing as well as purpose-designed tests engaging the auditory modality. Broadly speaking, the studies aimed to explore whether tasks can be purpose-developed to measure the broad ability General Speed of Processing (Gs), and also to examine existing auditory tasks and their relationships to Gs and general memory processes (Gm).

The research program was incremental in nature, leading to a series of four separate but interrelated studies. This dissertation centres on these studies which are reported across four manuscripts. These manuscripts are encompassed either side by chapters providing broader context and discussion relevant to the research program as a whole. The manuscripts (one published, one in- press, and two currently submitted for publication) are presented in typeset herein in order to aid ease of reading. Tables and figures have been labelled chronologically across this dissertation rather than within the manuscript chapters individually for the same reason.

The present chapter (Chapter 1) provides a broad overview of research relevant to the measurement of intelligence via the auditory modality in order to establish why this research programme and future studies with a similar aim are important. It includes a review of several key areas of auditory research which have generally not been concerned with the measurement of intelligence per se. It also examines previous literature in order to review findings regarding the extent to which auditory abilities relate to other established intellectual constructs.

Chapter 2 provides an exegesis for the four studies. It highlights the reasoning behind each in the context of the broader research program and provides information that is largely considered out-of-scope of the submitted/published versions of these manuscripts. This information includes theoretical and methodological reasoning relating to the development of new auditory tasks.

The middle chapters (3, 4, 5 and 6) present the four manuscripts and contain statements describing contributions by me and my co-authors. These manuscripts outline the rationale more specifically for each of the studies and they each contain a discussion of that study's findings. Chapter 7, the final chapter, is reserved for a summary of all results and provides a concluding discussion concerning the research program as a whole and outlines future directions for auditory-focussed human intelligence research.

***

The beginnings of intelligence research

In order to understand the current state of intelligence research and the general, although not necessarily intentional neglect of alternate modalities to vision, it is wise to briefly consider its rather short history. The reason for this is that two fields of differential psychology emerged in the late 19th and early 20th centuries and the influence of these fields is quite different as they have taken alternate approaches to the investigation of intellectual abilities.

Historically, there has long been an interest in human intellectual functioning, predating the modern tests with which it is measured. Such interest can be traced back to ancient philosophers including Plato and Aristotle who were amongst the earliest ponderers of the nature of human intelligence. Aristotle was amongst the first to initiate discussions concerning the heritability of intelligence (Sternberg, 2000) and in the late 19th century Sir Francis Galton expanded on this idea (Galton, 1865, 1892). In order to link biology and cognition he hypothesised that elementary sensory processes like discrimination were related to intellectual functioning. He reasoned that brighter people were able to make better discriminations in relation to sounds, colour, and weight than those of duller intelligence (Deary, 2000b). Rather unfortunately, Galton's evidence was inconclusive but Charles Spearman further developed these ideas empirically (Spearman, 1904). Spearman essentially equated higher-order intelligence with sensory discrimination, and considered discrimination in visual and auditory modalities to be the simplest mental operation which was clearly intellective in nature (Acton & Schroeder, 2001). James McKeen Cattell was also influenced by Galton's work yet he considered that the elementary basis of intelligence was not sensory discrimination but information processing speed (Acton & Schroeder, 2001). Cattell's efforts focussed on the time taken to react to visual and auditory stimuli (now referred to as Reaction Time) and the time taken to inspect and judge the colour of a visual stimulus; the latter is considered to be one of the earliest measures of what is now termed Inspection Time (Deary, 2000b). Cattell's ideas were not supported by his own data and it was again Spearman who published results in support of both theories (Brody, 2000).

Spearman reported positive correlations between different sensory measures, reaction time measures, different estimates of intelligence (such as school grades), as well as between all three of these (Brody, 2000). The correlations formed what was termed a positive manifold (positive correlations between all tests). This finding was replicated several times and led to the well-known theory of General Intelligence (Spearman, 1904, 1923). According to Spearman, there exists a “something” termed “General Sensory Discrimination” and similarly a “General Intelligence” for which “the functional correspondence between the two was not appreciably less than absolute” (Spearman, 1904 p.272). Spearman believed the basis of intelligence was the ‘general’ ability to make discriminations, whether they be at the sensory level, or a higher cognitive level (Deary, 2000b).

The work of Galton, Cattell and Spearman marks some of the earliest explorations into different sensory functions and their relationships to higher-order cognition. Both auditory and visual tasks were included in this early research because both modalities were acknowledged as being highly relevant to cognitive functioning (Deary, 2000b). On the contrary, modalities like the tactile were considered to be largely unrelated to intelligence estimates (Burt, 1909/1910). According to Jensen, this early work marks the birth of what has come to be known as the reductionist approach to understanding intelligence (Jensen, 1998). This approach has aimed to reduce complex cognitive processes manifest in tasks requiring reasoning and arithmetic, for example, into more simple, elementary processes. Although reductionism died-off in the early 20th Century it has experienced a re-birth in the last few decades. Moreover, it has provided some of the most complete research programmes that have explored the relation of auditory processes to higher order intellectual functioning. Examples of this research include the study of Auditory Inspection Time and Auditory Temporal Discrimination. This research is discussed in more detail later in the present chapter.

One factor or more? The birth of psychometrics

Around the time that reductionism and Spearman's *g* factor theory emerged, other psychologists including Binet (Binet & Henri, 1896) and Ebbinghaus (1897) voiced opposition to the idea of reducing complex cognition to sensory and other elementary processes. Binet believed the focus should be on higher-order mental operations because individual differences seemed to be more varied in the complex rather than elementary mental processes (Binet & Henri, 1896). Ebbinghaus (1897) noted that tests of simple visual and auditory sensory processes did not exhibit clear relationships to intelligence expressed in academic performance.

Thurstone developed a method of 'multiple factor analysis' (Thurstone, 1931) and applied it to his data to empirically test Spearman's ideas. He concluded that there were in fact around seven different ability factors (see Thurstone, 1936; Thurstone, 1938). These factors include number facility, spatial reasoning, memory, deductive and inductive abilities, and verbal comprehension (Brody, 2000). The identification of these other abilities challenged Spearman's claims that intelligence and elementary discrimination processes were one-and-the-same, and was one of the primary reasons for the demise of the reductionist agenda. Another reason for the reductionist demise was the growing interest in intelligence testing. Binet and others believed that intelligence testing was a means of moving away from laboratory based research to an applied setting (Brody, 2000). The first intelligence test was designed to understand and address the educational needs of children with an intellectual disability as opposed to normal children (Binet & Simon, 1905) and due to its success, a short time later ability testing was introduced in the military setting (Yerkes, 1921) where it remains today (e.g., see ASVAB, 2000).

The identification of abilities other than *g* (Spearman's term for general ability) and the introduction of intelligence testing marked the birth of psychometrics which has certainly dominated differential psychology for most of the 20th century. This research agenda was primarily concerned with reliability theory, factor analysis, intelligence theories and intelligence test construction and validation (Brody, 2000). Unfortunately, contrary to reductionism, the field of psychometrics saw an almost complete departure from studying abilities (sensory or otherwise) in alternate modalities such as audition and how these abilities relate to the different constructs measured by intelligence tests. Because of this, auditory abilities – and indeed abilities in modalities other than vision – have almost never been integrated into intelligence tests or theories arising from this period of psychometrics, and this does not do justice to the significance these abilities have (Seidel, 2008, p. 82). Instead, research on the auditory modality diverged to a focus on developing musical ability measures for use in selecting musicians. Although this research was theoretically significant in its own right, such measures have seldom been related to different intelligence constructs. The contribution of auditory focussed psychometric research is reviewed more thoroughly later in this chapter.

Modern taxonomies of human intelligence

Fast-forward 100 odd years, and debate still exists in terms of the nature and structure of human intelligence. However, one theory in-particular is regarded at present as being the most comprehensive and empirically valid. The Cattell-Horn-Carroll (CHC) model of intelligence is an amalgamation of two prominent theories concerning the structure of intelligence. The first of these is Gf-Gc theory, proposed by Cattell (1941, 1943, 1963) and Horn (Horn, 1968, 1988; Horn & Cattell, 1966).

This theory argues that a single *g* factor is not responsible for all intellectual behaviour and that intelligence is actually a myriad of perhaps 100-odd abilities which interact in different ways in different people. Gf-Gc theory separates these abilities into two broad but distinct sets of abilities having different trajectories over the human lifespan. The ability to think quickly, encode short-term memories and solve novel problems is considered to reflect fluid intelligence (Gf). Gf is relatively independent of education and acculturation and is the driving force behind acquiring new information. On the other hand, crystallized intelligence (Gc) is a result of learning and acculturation. It is manifest in tests of general knowledge, vocabulary, and other 'acquired' skills.

The second theory forming the CHC model is Carroll's Three Stratum theory (Carroll, 1993). This model was based on the re-analysis of some 400 data sets of cognitive ability test scores spanning most of the 20th Century. The theory proposes that abilities relate to each other at one of three levels, or strata. The lowest level encompasses group or test-specific factors that subsume highly specific abilities. These are much alike the 100-odd primary abilities proposed in Gf-Gc theory. At the second level are broad cognitive ability factors of which there are eight in total: these are fluid intelligence, crystallized intelligence, general memory and learning, broad visual perception, broad retrieval ability, broad cognitive speediness, processing speed, and broad auditory reception. At the third level, there exists a sole factor which was tentatively termed G (possibly akin to Spearman's *g*). Carroll suggested that this factor represents an ability, or abilities that are so general that they are manifest over the total domain of cognitive abilities (see Carroll, 1993, p. 591).

According to McGrew the CHC taxonomy has provided scholars and practitioners "the first empirically-based consensus Rosetta stone from which to organize research and practice" (2005, p. 171). The theory recognises that human intelligence is not unitary but clearly multidimensional, and it has become the definitive psychometric theory upon which standardised intelligence tests can be developed (McGrew, 2005). Throughout this dissertation continual reference will be made to broad cognitive abilities, or distinct intellectual constructs, and unless otherwise acknowledged they are envisaged to represent those defined in the CHC taxonomy (such as Gf and Gc, etc).

The problem with taxonomies and intelligence tests

As briefly highlighted, the early reductionist agenda provided very limited data suggesting a relationship between auditory processes and intelligence measures but the more influential field of psychometrics did not. It is the case, however, that psychometrics is primarily responsible for informing modern intelligence theories including the CHC model introduced above (Brody, 2000). Consequently, because of the neglect of alternate modalities by psychometricians, these taxonomies have been constructed almost entirely on the basis of performance on intelligence tests presented via the visual modality. This raises several theoretical questions including most importantly whether or not these taxonomies are structurally invariant across alternate modalities. In other words, if, hypothetically speaking, a battery of auditory tasks was able to be developed so as to be analogous to existing visual tasks, would the same broad abilities represented in the CHC taxonomy subsume them?

Importantly, the CHC framework is an open ended empirical theory that allows for future tests to be developed that may or may not identify new factors and abilities; either broad or narrow (McGrew, 2005). Indeed, there is emerging factorial evidence that broad perceptual abilities encompassing olfaction (Go) and tactile/kinaesthesia (Gh, Gk) exist, much like the broad factor general auditory perception (Ga; McGrew & Evans, 2004). However, although distinct perceptual factors have been acknowledged there exists an absence of systematic investigations of whether *other* broad abilities including Gf, Gc, and the like, are clearly manifest in tasks designed for modalities other than visual. At present there is scattered but limited evidence that this might be the case. What might be particularly problematic for intelligence theories is if this hypothesis was to be proved incorrect. It could be the case that these broad abilities might not be successfully indexed via alternate modalities. If this were so then, given that tests in other modalities generally define broad *perceptual* factors, it would be entirely plausible that the current taxonomy and abilities within it might also be at least to some degree visual-specific.

There are several imperative reasons for exploring whether the broad abilities at the second stratum of CHC theory are manifest in different modalities. Firstly, for taxonomic models of human cognitive abilities to be considered complete, all sensory modalities and the abilities manifest in them must be encompassed within its framework (Danthiir, Roberts, Pallier, & Stankov, 2001). The use of intelligence tests should also be informed by knowledge of what kinds of abilities they measure and what kinds of tasks – inclusive of all modalities – these abilities pertain to (Carroll, 1992). Unfortunately, it is an implicit assumption of modern intelligence theories that no knowledge of importance can be gained through employing tests of complex abilities that utilise alternate modalities (Roberts, Stankov, Pallier, & Bradley, 1997). This assumption must, however, be empirically tested during the course of arriving at a full and complete taxonomy. According to Carroll, the lack of knowledge about the variety and structure of cognitive abilities is a hindrance to the development of an adequate theory of cognitive abilities and to the construction of satisfactory and *comprehensive* measures of these abilities (Carroll, 1992).

A second reason for exploring this hypothesis was pointed out by Carroll (1992) when he stated: "...the technology of intelligence testing offers little in the way of novelty over procedures developed in the early years of the 20th century" (p. 267). In other words, save for a few revisions and occasional re-standardising, current intelligence test batteries resemble those used over 100 years ago by Binet, for example (Binet & Simon, 1905). Whilst the absence of adequate technology has historically hindered research on cognitive testing in alternate modalities, advances in this area have now certainly made it feasible (Stankov, 1994b). Furthermore, if additional modalities were to be incorporated into the measurement of cognitive abilities – whether perceptual-specific or otherwise – then the more ecologically valid these intelligence batteries would likely be. After all, humans do not only '*see*' their environment; they touch, smell, listen-to *and* visualise it each day of their lives.

The nature of broad cognitive abilities

In order to question whether broad abilities are manifest in alternate modalities, it is necessary to briefly outline what exactly is meant by this nomenclature. It is matter-of-fact that humans possess an infinite number of intellectual abilities typically referred to with terms like reasoning, abstracting, problem solving, memory, and concept attainment (Horn & Noll, 1994). It is also the case that some abilities tend to rise and fall together. For example, an individual's performance on a spelling test relative to others is likely to resemble their performance on a test of comprehension, or vocabulary. A similar pattern would emerge for tests of different mathematical skills, like multiplication and division.

These categories of tests have something in common and there appears some degree of overlap in what they measure (Cooper, 1999).

Factor analysis is a statistical procedure providing a method via which researchers have been able to explore the extent to which these tests group together according to the degree to which they measure a common factor or ability. Depending upon study design, factors emerge at different strata as apparent in the CHC theory (see Carroll, 1993). Primary factors represent specific abilities, such as multiplication, for example, and if appropriately measured they emerge at the first stratum. Broader, more general factors emerge at the second stratum, subsuming groups of primaries. Consider a study measuring various primary reasoning factors such as induction and quantitative reasoning. Each primary factor indexes processes not common to the others, but they nevertheless all rely on '*the ability to reason'* in a more general sense. Consequently, these primary, lower-order abilities are subsumed by a broader factor reflecting the ability to reason generally. It is these latter, more general factors which have come to be known as broad cognitive abilities and these account for a substantial amount of covariation in performance across infinitely large batteries of tests (Carroll, 1993). Each of these broad abilities, or constructs, shares a different relationship with external measures (such as age), and each also relies on different workings of different cognitive and neuropsychological functions (Roberts & Stankov, 1999).

The term broad cognitive ability thus recognises groups of abilities that can seemingly be drawn on by any number of tests provided they are demanding of similar mental operations (Cooper, 1999). There appears, therefore, no logical reason to assume that tests presented via the auditory modality cannot index these broad abilities. According to Carroll (1993) tasks should be designed by considering exactly what aspects of cognitive performance are tapped by them. This in turn guides hypotheses regarding how the test will relate to the broad abilities defined in the CHC model. According to this framework then, a test needs only to depend on the abilities subsumed by a given broad factor in order to relate to it. An auditory comprehension task, for example, would be expected to load on Gc given its dependence on acquired vocabulary knowledge, which also underpins performance on visual measures of this construct.

The current status of auditory abilities

Before exploring whether broad cognitive abilities are manifest in performance on auditory tasks it is necessary to review the areas of auditory research that have provided the most information regarding performance in this modality. Although I have suggested that auditory focussed research is lacking, this modality is actually the second-most researched in the study of human intelligence (Stankov, Seizova-Cajic, & Roberts, 2001). One reason for this increased interest is that the prime channel for communicating and acquiring information during childhood is audition. This reliance shifts to the visual system only after about five years of age (Sloutsky & Napolitano, 2003). The auditory domain is therefore recognised as being vital for the normal development of intelligence. Children with hearing impairments and those with total hearing loss are disadvantaged because of the lack of exposure to abstract concepts and as a consequence the global IQ of deaf people is typically lower on average than for blind people (Stankov, 1994a). The auditory modality thus plays an integral role in human intellectual functioning and likely does so not only during the developmental years but throughout the lifespan.

The fact that auditory research is the second-most common after visual by no means implies that such research is either common or comprehensive. As briefly noted in the preceding sections, broad auditory reception, or Ga, is a recognised cognitive ability in CHC theory. Its presence in this model though does not establish that auditory tasks defining Ga relate directly to other broad abilities such as Gf and Gc, which is the crux of this thesis. The factorial analyses of Carroll (1993) and others (e.g., Stankov, 1971) quite clearly demonstrated that Ga is linearly-independent of these other abilities and this is the reason for its emergence as a distinct broad factor at the second strata. I will now introduce the larger auditory-focussed research programmes relevant to this thesis prior to reviewing whether auditory tasks relate to visual measures of other broad abilities.

Auditory-focussed 'psychometric' research

In keeping with the earlier distinction between psychometrics and reductionism, I will split my reviews of auditory research according to these headings. Dedicating a section to psychometric auditory-focussed research seems at odds with my earlier statement that this field led to an almost complete departure from interest in this modality. However, psychometrics is a broad discipline concerned primarily with test construction and, consequently, theory development. Although it has its origins in intelligence research, it is not exclusive to this field and permeates throughout other research areas including studies of personality and musical ability (see e.g., Rust & Golombok, 2009 for a comprehensive history of psychometrics). Thus, there are auditory research agenda that are broadly psychometric in nature, but which have seldom considered their tasks in relation to other broad intellectual abilities (as opposed to general estimates of intellectual ability like Grade Point Average).

Tests of Musical Ability

When psychometrics began to dominate intelligence-research and concentrate purely on intelligence testing it spawned interest in the development of auditory test batteries for the purpose of selecting musicians and military personnel such as sonarmen and radio operators (Stankov, 1994a). Thus, it was motivated by practical needs and consequently there exists an impressive body of literature concerned with musical ability testing dating back to the start of the 20th century. The earliest standardised version of a musical screening test was the Seashore Measures of Musical Talent (Seashore, 1919). All but one of the subtests – tonal memory – required little more than the comparison of simple tones (tonal memory is the only subtest which does not). This led to the conclusion that the test battery was not complex enough to be classified as a measure of musicality given it essentially measured discrimination processes, whereas good musicianship seemed to reflect complex abilities like musical creativity (see Seidel, 2008, pp. 92-95).

Other subsequent and notable tests of musical intelligence include the Wing Musical Aptitude Test (Wing, 1948, 1962), the Bentley test (Bentley, 1966), the Drake Musical Aptitude Tests (Drake, 1933a, 1933b), the Gordon Tests (Gordon, 1965, 1989) and those of Buttsworth, Fogarty and Rorke (Buttsworth, Fogarty, & Rorke, 1993). These test batteries vary in scope from assessing simple discrimination, as was the case for the Seashore measures (Seashore, 1919), through to more complex abilities like intonation; the ability to play in-tune (Buttsworth, et al., 1993). Theories regarding the structure of musical abilities also vary between these researchers. For example, although Drake (1933a, 1933b) and Wing (1948) proposed the presence of a general musical ability factor, others ascribed to the idea that musical abilities depend on basic capacities of time, intensity/volume discrimination, and pitch discrimination ability (Bentley, 1966; Seashore, Lewis, & Saetveit, 1960) . Holmstrom (1969) proposed that three factors underpinned performance on musical ability tests reflecting 1) primary perceptual abilities (termed Alpha); 2) tonal memory (Beta); and 3) a broad musical factor loaded on by tasks measuring rhythm, pitch and memory (Gamma). Franklin (1956, as cited in Seidel, 2008) argued elsewhere that there are only two parts to musical ability termed 'mechanical-acoustic' talent, which reflects basic discrimination abilities regarding pitch and time, and 'judicious-musical' talent which reflects the ability to work with complex musical sequences and musical context.

Factor analytic studies of musical abilities tests support each of the aforementioned theories to a different degree. For example, Franklin's analyses (1956, as cited in Seidel, 2008) clearly showed that tests of mechanical-acoustic and judicious-musical abilities separated across factors. Further analysis of the Seashore and Wing tests showed the emergence of factors denoting tonal memory, pitch discrimination, melody and harmony, and rhythm. It has been suggested that tasks group according to whether they reflect basic auditory capabilities or more advanced musical abilities, and that the former more basic abilities are a prerequisite for good performance on more complex musical tasks (Seidel, 2008).

Broad Auditory Reception (Ga)

The pièce de résistance of psychometric auditory research which had a clear 'intelligence' focus is the PhD dissertation of Lazar Stankov (Stankov, 1971). Without this work, the Ga factor might never have been convincingly incorporated into CHC theory. When Stankov commenced his research he noted that there was nothing in the field of differential psychology even remotely concerned with auditory intelligence. Twenty-two years later when Carroll (1993) undertook his re-analysis he echoed this sentiment, noting that “evidence on the [factor] structure of auditory abilities is very meagre because of the little attention that has been given to this domain” (p.609).

Stankov’s research was driven by the intention to identify a broad auditory ability in the taxonomy of intelligence. It was borne out of the ideas of his supervisor John Horn (Horn, 1968) who proposed that auditory abilities would define a distinct broad auditory perceptual factor (Ga). Horn suggested that if tests were constructed that were otherwise like existing cognitive abilities tests, although involved processing of auditory material, they would be shown to share variance with existing intellectual tasks. Confirmation of this hypothesis would imply that measures of intelligence, as they currently stood, were biased, and that truly balanced measures of intelligence should contain items emphasising the use of audition as well as visualisation; tactile intelligence was also discussed (p.253).

Stankov (1971) assembled a large battery of auditory tasks (reported in Horn & Stankov, 1982; Stankov & Horn, 1980) that were classifiable into three categories: First, some tests were the same as existing visual tests but were delivered auditorily. For instance, in a disarranged sentences test, words are spoken (rather than written) in haphazard order and the participant is to arrange them correctly. Second, some tests were devised to closely resemble existing visual tests. In an auditory test of inductive reasoning, for example, three notes are played followed by three answer choice notes. The participant indicates which note, out of the answer choices, is the same tonal interval away from the third note as the second was from the first note. The third category consists of the existing auditory measures mentioned previously; that is, musical tests and acuity tests.

Stankov and Horn (1980) reported acceptable reliability for the auditory measures (split-half reliability M_r=.75; *SD*=.11) and it appeared that the constructs indexed were comparable in stability to those indexed by visual tasks. An analysis of the primary factor structure of the auditory tasks uncovered seven factors which accounted for approximately 50% of the variability in task performance. The seven primaries were 1) tonal memory; 2) speech perception under distraction; 3) verbal comprehension; 4) immediate memory; 5) cognition of relationships; 6) discrimination among sound patterns; and 7) maintaining and judging rhythm. Two other primaries were also hypothesised to exist and included loudness discrimination and temporal tracking. They did not emerge as distinct abilities due to problems with the data and tests.

Correlations between the auditory primaries were generally positive, leading to the emergence at the second order of a general auditory factor (Ga), which would be located at the same strata as the broad visual ability factor, Gv. Interestingly, the primary factors comprising Ga are distinguishable from the abilities defining auditory acuity. Ga measures require holistic comprehension of sound and patterns, for example, whereas acuity tests relying on mutilated and incomplete sounds rely more on elementary sensory processes representing organisation among sensory detector functions of hearing (Seidel, 2008). Therefore, Ga is considered to represent more than just a sensory factor and instead reflects higher-order perceptual processes (Stankov, 1994a).

Test of Basic Auditory Capabilities

Around the time that Stankov's research on Ga was being published, Watson, Johnson, Lehman, Kelly and Jensen (1982) had also recognised the importance of the auditory modality. They developed the Test of Basic Auditory Capabilities (TBAC; Suprenant & Watson, 2001; Watson, et al., 1982) in an attempt to document the abilities underpinning this modality. Thus, their motivation differed from studies concerned with musical abilities testing, and also from that of Stankov (Horn & Stankov, 1982; Stankov, 1971; Stankov & Horn, 1980) who was additionally concerned with locating these abilities in the taxonomy of human intelligence.

The first TBAC battery (Watson, et al., 1982) comprised six subtests measuring abilities such as pitch, intensity, and temporal discrimination, as well as a seventh subtest measuring nonsense syllable identification. A later revision expanded only on the number of speech subtests incorporated (Seidel, 2008). The auditory ability factors that have been identified based on analysis of the revised edition (Suprenant & Watson, 2001) include temporal order discrimination, which resembles Stankov and Horn's (1980) temporal tracking factor, as well as a non-speech ability factor subsuming pitch and intensity discrimination; factors also previously identified by Stankov (1971). In addition to these, a speech-identification factor has also been established, which is again remarkably similar to the Speech Perception Under Distraction factor identified by Stankov (1971). Interestingly, speech factors have been shown to be largely independent of other auditory abilities. This suggests either that a speech-specific processing mechanism exists (Kidd, Watson, & Gygi, 2007), or that speech processing occurs at a higher cognitive level than the processing of basic auditory stimuli such as pitch and intensity, which is largely perceptually based (see e.g., Stankov, 1994a). Temporal discrimination does appear to relate reliably and moderately to speech processing (see Kidd, et al., 2007). However, its status as an auditory-specific ability is questionable given that it has been shown to load strongly on factors dominated by academic measures including Grade Point Average (GPA) and Scholastic Aptitude Test scores (SAT; Suprenant & Watson, 2001).

General Conclusions

Research on musical ability testing, auditory abilities, and the TBAC are usually reviewed in distinction from one another, as I have again done herein (see e.g., Seidel, 2008; Stankov, 1971). However, subtests comprising the measures used in each of these fields of research are markedly similar. The Seashore (Seashore, 1919) measures are almost indistinguishable from those used by Stankov (1971) because many of Stankov's tasks were adapted from existing musical ability batteries. Similarly, discrimination tasks in the TBAC bear strong resemblance to those used in musical abilities research. Although on first impressions the TBAC, Musical Abilities, and Auditory Abilities research programmes appear to be quite distinct from each other, the difference is really only due to the focus of these individual agendas.

The results from all of these studies in terms of the structure of auditory tasks are surprisingly similar. More specifically, Holmstrom's (1969) and Franklin's analyses (1956, as cited in Seidel, 2008) of musical abilities tests demonstrate the existence of primary factors that subsume tasks involving pitch discrimination, tonal memory, and the analysis of rhythm, respectively. These factors are almost certainly analogous to the temporal tracking/rhythm, temporal discrimination, pitch discrimination and intensity factors identified later in auditory abilities studies by Stankov (1971) and by Watson and colleagues during TBAC research (Kidd, et al., 2007; Suprenant & Watson, 2001; Watson, et al., 1982; Watson & Miller, 1993). Thus, each of these works is supportive of a series of primary auditory-specific factors that are pivotal to performance in this modality.

Whilst studies concerning the TBAC and musical abilities are seldom concerned with how auditory abilities locate themselves within intelligence taxonomies, they again all appear to provide general evidence in support of the broad auditory factor (Ga) identified by Stankov (1971) and later confirmed by Carrol (1993). For example, a factor analysis of the non-speech subtests of the TBAC performed by Watson and Miller (1993) uncovered a single factor accounting for over 57% of the variance in test performance. Similarly, Kidd et al. (2007) employed an extended TBAC battery and found that the four first order factors of loudness duration, amplitude modulation, familiar sounds and pitch and time, were subsumed by a single broad higher-order general auditory processing factor. Thus, it can be considered relatively well established that there does exist a broad auditory perceptual factor, which subsumes primary abilities such as pitch discrimination and which is likely located on the same stratum of CHC theory as factors including Gf and Gc. The extent to which primary factors like the processing of speech can be considered entirely auditory-specific is yet to be established, as is the extent to which other broad cognitive abilities are manifest in auditory task performance.

Auditory-focussed 'reductionist' research

As pointed out previously, reductionists have also contributed to knowledge of the auditory modality but have done so from a completely different perspective than that of psychometricians. Contrary to the research discussed in the preceding section, which has taken a very broad approach to auditory testing, this discipline has concerned itself with specific classes of tasks or paradigms, often utilising a single task across numerous studies in an attempt to elucidate the relationship between higher-order cognition and elementary processes. As such, it has seldom considered its auditory measures in light of existing knowledge regarding the structure of auditory abilities. Where it has contributed most is to the understanding of how auditory processes relate to other intelligence measures and this is discussed in more detail in a later section. However, I will first introduce some of the notable reductionist research programmes relevant to this thesis and briefly review how their findings relate to the structure of auditory abilities outlined above.

Auditory Inspection Time

A particularly common and well-known example of reductionist auditory focussed research concerns the measurement of Auditory Inspection Time (AIT). Several variations of AIT have been developed (see Zajac & Burns, 2007 for a discussion of these) yet they were all borne out of a desire to measure an Inspection Time (IT) analogous to that measured in the visual modality. Vickers was the first to propose the theory behind visual IT (VIT) and its measurement (Vickers, Nettelbeck, & Willson, 1972), commonly instantiated as a critical stimulus onset asynchrony on a visual pattern backward masking task (Zajac & Burns, 2007). In its most common form, the task requires a discrimination regarding which of two high-contrast lines presented vertically, side-by-side, is the shorter (or longest; i.e., left or right). The discrimination is easy given unrestricted viewing time but the use of pattern backward masking and reducing critical stimulus onset asynchrony for correct responses allows demonstration of individual differences which correlate reliably and moderately with measures of intelligence (see e.g. Nettelbeck, 2001, 2003).

The discovery of a reliable and moderate relationship between complex cognition manifest in intelligence measures and the supposedly elementary processes thought to underpin VIT prompted attempts to develop an analogous auditory measure of IT. The first AIT task was devised by Brand and Deary (1982) and required participants to judge the temporal order of two tones differing only in pitch (hereafter referred to as AIT-P). The suggestion that performance on AIT-P and VIT tasks reflects a common information processing mechanism has been challenged because AIT-P relies significantly on pitch discrimination ability which, as highlighted previously, is inherently auditory by nature (Irwin, 1984). Moreover, the absence of a reliable and significant correlation between AIT-P and VIT is further evidence that the tasks measure very different processes (Brand & Deary, 1982, r=0.05; Deary, Caryl, Egan, & Wight, 1989, r=0.24; Irwin, 1984, r=0.17; Nettelbeck, Edwards, & Vreugdenhil, 1986, r=0.39). Consequently, Olsson, Bjorkman, Haag and Juslin (1998) developed a revised measure, which required a discrimination concerning loudness, rather than pitch (hereafter referred to as AIT-L). Again, performance on the task was found to relate significantly to pitch discrimination as well as to AIT-P (Olsson, et al., 1998) and, although never directly examined, it is likely that AIT-L would rely significantly on loudness discrimination ability as identified in studies of auditory abilities. Thus, both AIT-L and AIT-P would appear to rely heavily on relatively elementary auditory-specific processes not relevant to performance on IT in the visual modality.

The most recent instantiation of AIT (Parker, Crawford, & Stephen, 1999; hereafter referred to as AIT-S) bears no formal similarity to its predecessors because the discrimination involved concerns spatial localisation, which is also the case for VIT. The results of several studies suggest that AIT-S is better than AIT-P and AIT-L at measuring processing mechanisms analogous to those underpinning performance on VIT given more consistent relationships with this measure (Campbell, 1995; Parker, et al., 1999; Zajac & Burns, 2007). Moreover, unlike for the other AIT tasks, performance on AIT-S does not rely on basic auditory abilities such as pitch discrimination (Bates, 2005). Therefore, spatial discrimination may not be an auditory specific ability.

Auditory Temporal Discrimination

Another example of reductionist research concerns auditory temporal discrimination. As already noted, temporal tracking and discrimination has previously been identified as a primary ability subsumed by Ga (see e.g., Stankov & Horn, 1980; Suprenant & Watson, 2001). However, in recent times it has been revisited from a completely different perspective to the Ga focussed research already reviewed. More specifically, given the association found between higher-order cognition and low-level processes exhibited in tasks such as IT, researchers have needed to formulate theories explaining this observation. Several theories revolve around the concept of efficiency of the central nervous system (CNS) as the determinant of intelligence (see e.g., A. E. Hendrickson, 1982; D. E. Hendrickson, 1982; Vernon, 1993). For example, Surwillo (1968) appealed to a hypothetical oscillatory, or ‘clock’, mechanism in the CNS to explain individual differences in speed of information processing and intelligence. It was proposed that a slower clock rate resulted in requiring more time to complete mental operations and because of this, the probability of interfering events would increase (Rammsayer & Brandler, 2007).

Rammsayer and colleagues have sought to demonstrate that presumed measures of clock rate differ between individuals of low and high intelligence (Helmbold, Troche, & Rammsayer, 2006, 2007; Rammsayer & Brandler, 2002, 2004, 2007). They proposed that temporal resolution as measured in a temporal discrimination task was a sufficient measure of CNS resolution, because faster neural oscillations would result in finer temporal resolution. Audition was the modality of choice in this instance because it is purported to have finer temporal resolution than the visual system (Rammsayer & Brandler, 2002). The research programme of Rammsayer and his colleagues is impressive, with generally well designed studies. Unfortunately, much like research on AIT, the focus has been on relating these tasks to typical measures of intelligence and only two studies have considered the relationship between temporal discrimination and more typical sensory discrimination measures involving pitch and intensity (Helmbold & Rammsayer, 2006; Troche & Rammsayer, 2009b). In one of these studies (Helmbold, et al., 2006) temporal and pitch discrimination tasks shared close to 20% of their variance (r=-.41). In the other, a temporal discrimination and a sensory discrimination factor – defined by un-speeded pitch and intensity tasks – correlated almost perfectly (r=.94). Although these studies glean information on how auditory abilities might relate to other broad abilities, by aligning temporal discrimination with CNS resolution, they have completely disregarded findings that suggest temporal discrimination is actually a relatively complex auditory ability.

Auditory Reaction Time

A third example of reductionist research relates to measures of reaction time. The requirement of RT tasks is that participants respond to stimuli as quickly as possible, generally by pressing a response key. The majority of studies concerned with RT and its relationship to intelligence have used visual stimuli (VRT). Many alternate versions have been devised, progressing from very simple in task requirements to more complex (such as in the Hick paradigm; Hick, 1952), yet RT indices tend to consistently share a statistically significant amount of variance with measures of intelligence (Deary, 2003). Shorter reaction times (i.e., faster information processing) coincide with higher scores on intelligence tests.

Reaction time tasks are very simple and, unsurprisingly, have been translated into auditory format (ART). Unlike the temporal discrimination and AIT research noted above, however, ART has rarely been related to intelligence measures, with researchers showing a preference for its visual counterpart, possibly because of the assumption that these tasks measure the same thing. Similarly to AIT and temporal discrimination though, it has also not been examined in terms of the structure of auditory abilities. Instead, the most extensive use of ART has occurred outside of intelligence research in areas like neuropsychological assessment following traumatic brain injury (see e.g., Duncan, Kosmidis, & Mirsky, 2005; Stålnacke, Elgh, & Sojka, 2007; Whyte, Polansky, Fleming, Coslett, & Cavallucci, 1995). It is also used to assess factors such as ageing effects on attention (Sparrow, Bradshaw, Lamoureux, & Tirosh, 2002), to explore the neurophysiological effects of Schizophrenia (Pfefferbaum, Ford, White, & Roth, 1989), and to evoke event-related potentials in order to examine brain function (Jausovec & Jausovec, 2000). In all of these alternate fields, however, theory regarding what decision processes RT tasks measure certainly derives from its history in mental chronometry – the assessment of speed of cognitive processing – which has at its heart interest in human intellectual functioning.

General Conclusions

Findings in each of these areas highlight several issues relevant to auditory research in general and the development of auditory tests. As already alluded to, reductionist studies often show disregard for findings elsewhere in the literature, specifically those concerning Ga abilities. This is the case for AIT research, where despite the seeming reliance of AIT-P and AIT-L on well replicated primary auditory abilities like pitch and loudness discrimination, researchers have regarded IT as being a direct measure of *g* (Brand & Deary, 1982). It is also the case for the work of Rammsayer and colleagues (Helmbold & Rammsayer, 2006; Helmbold, et al., 2007; Rammsayer & Brandler, 2002, 2004, 2007) because there currently exists evidence to suggest that temporal discrimination is at least partially auditory-specific (Stankov, 1983). Thus, it is overly ambitious to consider such tasks as direct measures of temporal resolution of the CNS. These studies – particularly AIT – also highlight the difficulty of developing auditory tasks that neatly and reliably measure constructs that are not auditory specific, such as that manifest in VIT performance. It seems that the more these tasks rely on innate basic auditory capabilities including pitch and loudness discrimination, the more confounded by Ga 'type' variance the measures become.

Relationships between auditory tasks and 'other' broad abilities

It should now be apparent to the reader that although there is a significant amount of literature concerned with auditory testing, much of the research has been conducted independently with little regard given to findings from other auditory studies. Evidence of this was apparent in a conversation between myself and Lazar Stankov when he highlighted that he was somewhat perplexed by the general disregard of his auditory abilities studies by AIT researchers (personal communication, November 25, 2009). For the most part, despite evidence of a clear dependence on auditory abilities like pitch discrimination, that agenda held onto the idea that AIT tasks measure something 'special', something largely independent of the auditory modality.

As indicated earlier, an important question for intelligence theorists is whether broad abilities like Gf are indexed via numerous modalities. In accepting the belief that broad abilities have the potential to be engaged by all cognitive tasks, it certainly appears 'likely' that this would be the case. Unfortunately, this specific question has been given very little attention. Therefore, one is left to scour existing literature for sporadic evidence of such relationships. The purpose of the present section is to review evidence of links with broad cognitive abilities, and the psychometric and reductionist studies are reviewed independently only for ease of reading.

Auditory abilities research

There is little difference between many of the auditory tasks used across the musical abilities, auditory abilities and TBAC research programmes. Consequently, the abilities identified by these different studies can be regarded as being largely one-and-the-same. For example, Suprenant & Watson's (Suprenant & Watson, 2001) temporal order discrimination factor very closely resembles Stankov and Horn's (1980) temporal tracking factor. Thus, it is unnecessary to intentionally split a review of relationships between auditory measures and other abilities according to these areas of research, and I instead review them as a collective.

Not surprisingly, there has long been an interest in whether musical ability is related to general intellectual ability (Shuter-Dyson, 1968). For the most part, the two constructs do appear correlated but the estimate of shared variance between them has been placed at a weak 9%; a correlation of approximately r=0.30 (Wing, 1948). Indeed, those lower on IQ scores tend to perform poorly on tests of musical ability, but there is disagreement where high IQ individuals also score poorly. Thus, it has been suggested that the relationship is linear only up to a score of about 90 IQ points for secondary students, after which intelligence no longer relies significantly on underlying musical abilities (Shuter-Dyson, 1968; Stankov, 1971). Other notable exceptions are the reported cases of idiot-savants who often exhibit an outstanding memory for tunes as well as superior recall powers which enable fluent replication of complex musical pieces (Shuter-Dyson, 1968).

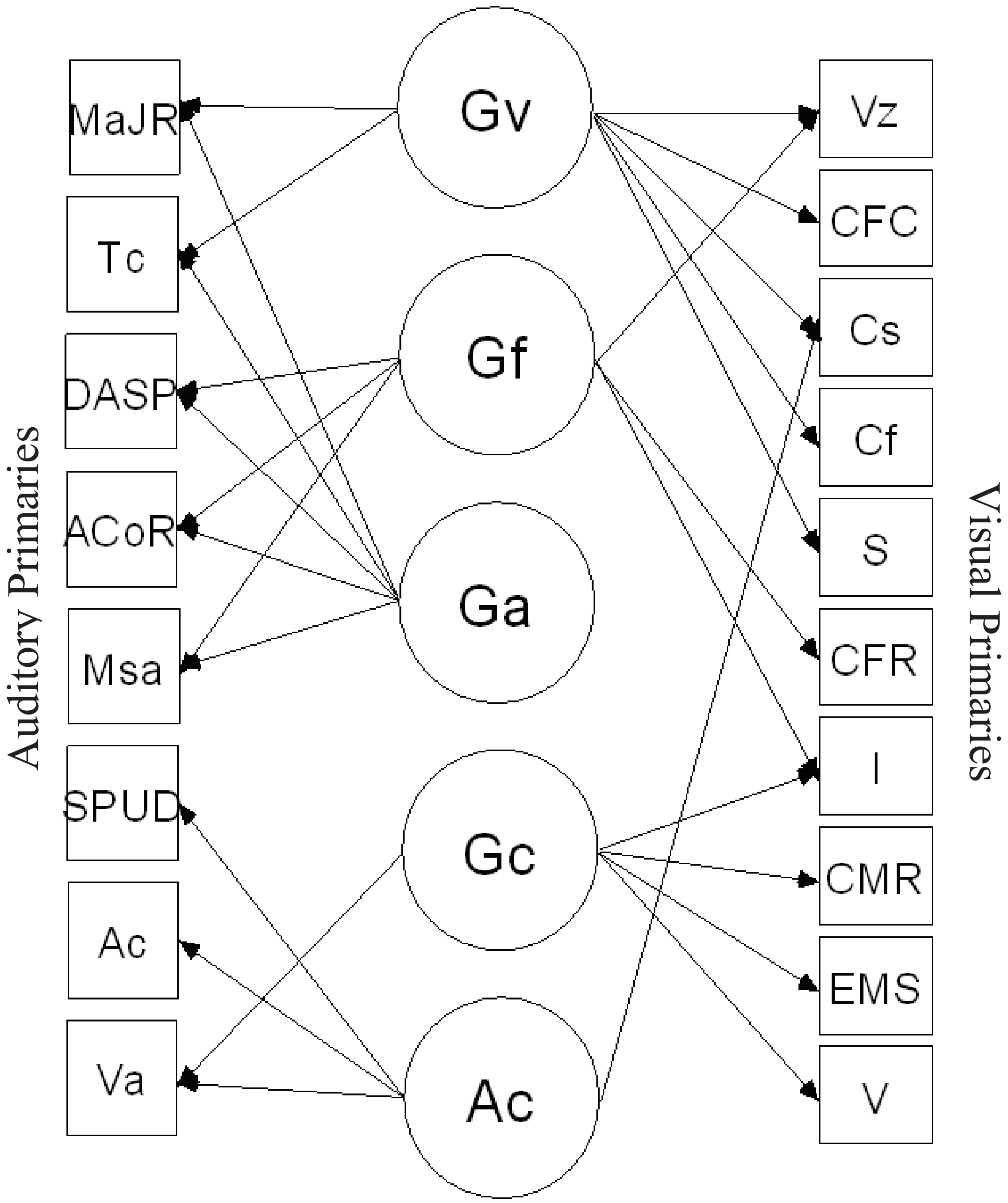
In the present thesis we are interested in whether musical abilities relate to other ‘specific’ intellectual constructs like Gf and Gc rather than simply full scale IQ estimates, or measures of academic performance. McLeish (1950) found evidence of a relationship between Seashore’s pitch and memory measures and Cattell’s timed intelligence test (otherwise referred to as Cattell’s culture fair test). This led to the suggestion that cognitive speed – otherwise referred to as General Speediness (Gs) – might be more crucial to musical ability than other abilities like arithmetic, given the need to maintain tempo in fluent musical performance. Woodrow (1939, cited in Stankov, 1971) also found links between musical performance and speed. More specifically, of the 10 orthogonal factors extracted by Woodrow, one of them was loaded on by all of the Seashore measures as well as visual tests of copying, reaction time, and speed. The work of Crawford and Stankov (1983) also supports a relationship between audition and speed. They ascertained measures of speed only for a selection of the visual tests used in their battery, yet three auditory tasks loaded on a second order Gs factor together with these. The association was suggested to reflect that the auditory tasks required rapid registration of the auditory stimuli. Roberts and Stankov (1999) found evidence for an auditory-Gs link in the time taken to complete auditory and visual tasks, which itself defined a first order factor that was subsumed by a general timed performance factor (Gt), thought to broadly reflect cognitive speed.

Elsewhere in the literature, tentative links between auditory measures and other broad constructs are also apparent. For instance, Franklin (1956, cited in Stankov, 1971) reported that Seashore's memory test, but not the pitch test, correlated with a vocabulary measure at r=.28. Vocabulary, broadly speaking, would be considered in CHC theory a marker of crystallised intelligence (or Gc) rather than Gs. Interestingly, a possible auditory x Gc link is apparent in other studies also. Holmes (1954) showed that variation in spelling test scores was associated with performance on musical tests. Pelletier (1963) tested this association in an experimental framework by exploring whether instrumental instruction would improve vocabulary in third grade students. After six months of instrument instruction reading comprehension, but not reading vocabulary nor spelling, had improved significantly over a control group.

Studies of the TBAC have also revealed relationships between various auditory abilities and distinct non-auditory constructs. Deary, Bell, Bell, Campbell and Fazal, (2004) conducted two independent studies using TBAC measures. In the first, they found that loudness discrimination correlated weakly-to-moderately with a measure of Gc (Mill-Hill Vocabulary, r=.28), a measure of Gf (Cattell culture fair, r=.44) and a measure of Gs (Digit Symbol, r=.40). In the second study, the TBAC pitch discrimination task was used. Although the first order correlations were not reported, pitch discrimination related to a general intelligence factor via a general sensory discrimination factor, resulting in a relationship with *g* of around .25 (.31 x .68). Interestingly, of the 13 cognitive tests defining *g* in this study only one of them was a marker of verbal ability. Thus it is possible that this intelligence factor might have been better considered as a reasoning, or Gf type factor. Acton and Schroeder (2001) also utilised the TBAC pitch measure and reported a weak but significant relationship with a factor termed structural visualisation, presumably reflecting the broad ability General Visualisation (Gv). It should be noted that it is often difficult to distinguish Gv from Gf (see Lohman, 2000) and it is entirely plausible that this visualisation factor might also better resemble reasoning, or Gf.

Suprenant and Watson (2001) studied the full TBAC battery and its relationship to Scholastic Aptitude Test scores. The SAT verbal score (SAT-V) is a measure of Gc and was reported to relate strongest with the auditory temporal order tones and syllables tasks (r=.43 and r=.65, respectively). SAT math (SAT-M) – a broad measure of Gf type abilities – related similarly well to the temporal order tones and syllables tasks (r=.42 and r=.45, respectively) but also to single-tone duration discrimination and embedded test-tone loudness (r=.36 and r=.40, respectively). Kidd et al. (2007) also studied the TBAC along with SAT scores and reported that SAT-V related weakly to a TBAC speech factor defined by auditory words and sentences measures (r=.18), as well as a pitch and time/temporal processing factor (r=.14). SAT-M, on the other hand, related only to the pitch and time/temporal processing factor (r=.15).

Stankov's research on auditory abilities was also designed to glean information concerning the relation of auditory abilities to other constructs. More specifically, Horn and Stankov (1982) extended their earlier analysis of auditory tasks (Stankov & Horn, 1980) by incorporating 15 visual tasks chosen to define specific primary visual factors. The results of this subsequent analysis are presented graphically as Figure 1. In support of relationships already noted, the factors of Gf and Gc are both loaded by visual and auditory factors. This supports the idea that Gf and Gc reflect cognitively central processes rather than processes associated with a particular modality (i.e., visual).



Second order factors are: general visualisation ability (Gv); fluid intelligence (Gf); general auditory ability (Ga); crystallised ability (Gc); auditory acuity (Ac). Primary auditory factors are: maintaining and judging rhythm (MaJR); temporal tracking of sounds (Tc); discrimination among sound patterns (DASP); auditory cognition of relations (ACoR); auditory immediate memory (Msa); speech perception under distraction (SPUD); auditory acuity (Ac); auditory verbal comprehension (Va). Primary visual factors are: visualisation (Vz); figural classes (CFC); closure speed (Cs); closure flexibility (Cf); spatial orientation (S); figural relations (CFR); induction (I); semantic relations (CMR); semantic systems (EMS); verbal comprehension (V).

Figure 1. Horn and Stankov's (1982) factorial model of auditory and visual tests.

Interestingly, similar to Acton and Schroeder's (2001) findings, Horn and Stankov's Gv also appears related to auditory tasks (two in this instance). The fact that some of the auditory primary factors appear related to Gv supports the belief that broad perceptual factors are not necessarily modality specific as does the separation of Ga from auditory acuity in Figure 1. Gv and Ga are possibly more representative of broad perceptual processes whilst acuity factors (both visual and auditory) reflect strictly sensory processes (Horn & Stankov, 1982).

Stankov (1978) reported evidence of similar links in a sample of primary school children. The number of auditory tests administered was reduced in comparison to the battery used in his PhD studies. They were specifically chosen to measure the seven auditory primaries previously identified (Stankov & Horn, 1980). Visual markers of Gf, Gc and Gv were also completed. The findings were generally congruent with the study discussed above. Specifically, four factors emerged at the second-order of analyses; Gf, Gc, Gv and Ga (auditory acuity was not measured). The Gf and Gc factors were defined by both visual and auditory primaries, but no auditory primaries loaded on Gv. Another study by Stankov, Horn and Roy (1980) again demonstrated similar relationships, with the higher-order factors of Gf and Gc being defined by both auditory and visual tests. A second-order memory factor termed short-term acquisition and retrieval (SAR) was also loaded on by auditory tasks. The findings of Fogarty and Stankov (1988) and Crawford and Stankov (1983) are similarly congruent with these; auditory tasks appear to relate to some degree with Gf and Gc factors.

Reductionist Research

As noted, reductionists have always been concerned with relating their tasks to measures of intelligence and thus there are numerous reports of correlations between presumed low-level tasks like AIT and intelligence constructs. Indeed, it was the relationship between VIT and markers of general intelligence that sparked the first studies of AIT and, subsequently, each of the three AIT tasks described above has been shown to correlate with performance on various cognitive tests. Initially, the hypothesis regarding the VIT x intelligence relationship supposed that IT captured something of fluid intelligence (Gf), requiring the ability to reason both abstractly and with speed (Kranzler & Jensen, 1989). Consequently, the first publications concerning AIT-P reported on correlations between it and both Gf and Gc constructs. Brand and Deary (1982) reported that, in their study, AIT-P correlated strongly with the Gf marker test (Raven's Progressive Matrices, RPM; r=-.70), and nearly as strongly with the Gc measure, Mill-Hill vocabulary (r=-.66). Irwin (1984) conducted a similar study and reported significant but markedly weaker correlations between AIT-P and these same marker tests of Gf and Gc (r=-.24 and r=-.42 respectively), as did Nettelbeck et al. (1986), who reported an AIT-P x RAPM correlation of r=-.38, and a correlation of r=-.33 between AIT-P and an alternate verbal ability (Gc) measure. Interestingly, Nettelbeck et al. also reported that AIT-P related significantly (r=-.36) to a measure of Associative Memory.

The two alternate versions AIT-L and AIT-S have similarly been related to these constructs. Olsson et al. (1998) reported that AIT-L related to RAPM significantly (r=-.36) but that AIT-P did not (r=-.21). McCrory and Cooper (2005) reported that both these versions of AIT related to RPM performance (r=-.42 and *r*=-.24, respectively), but that only AIT-P related to verbal ability (*r*=-.40). The newer version, AIT-S, has in all reported studies related significantly to RPM: Parker et al. (*r*=-.37 and *r*=-.57, 1999); Campbell (*r*=-.71, 1995); Bates (*r*=-.52, 2005); and Zajac and Burns (*r*=-.34, Modified Progressive Matrices, 2007). However, its correlation with verbal ability (or Gc), has been less reliable. More specifically, Parker et al. (1999) reported a statistically non-significant AIT-S x Gc (WAIS-R) correlation (*r*=-.16) as did Zajac and Burns (*r*=-.07, Information Scales). McCrory and Cooper (2005) on the other hand, reported a significant relationship between AIT-S and WAIS Verbal IQ (-.26).

Over approximately the last decade, the hypothesis regarding which of the intelligence constructs Visual IT measures has changed. Burns, Nettelbeck and colleagues (Burns & Nettelbeck, 2003, 2005, 2008; Burns, Nettelbeck, & Cooper; O'Connor & Burns, 2003) have provided increasing evidence that IT is a marker of Gs. Interestingly, AIT-P and AIT-L tasks do not exhibit this pattern. For instance, McCrory (2005) and Cooper noted near-zero relationships between a marker of Gs and both AIT-P (*r*=-.08) and AIT-L (*r*=-.08). Conversely, AIT-S has displayed a consistent relationship with Gs measures in studies that have used them (*r*=-.25, McCrory & Cooper, 2005; *r*=-.36, Zajac & Burns, 2007), as well as with VIT (see Zajac & Burns, 2007 for a discussion of this).

Findings relating auditory RT measures to intelligence constructs are particularly rare, despite the prevalence of RT studies in the general literature. Much data exist for its visual counterpart, and like for VIT, it has been shown that VRT loads consistently on Gs. Danthiir (2005) found that VRT tasks as well as traditional psychometric measures were subsumed by a common speed factor and similar findings were reported by Roberts and Stankov (1999), and O'Connor and Burns

(2003). Sen, Jensen, Sen and Arora (1983) published one of the few studies to have considered auditory RT in an intelligence framework. They reported a very strong correlation between simple VRT and ART tasks (*r*=.86), suggesting the two tasks tap an almost identical construct. This supports a review by Brebner and Welford (1980) who concluded that variations between RT in alternate modalities are due primarily to peripheral processes and not central cognitive processes. These differences reflect variables like afferent conduction velocity. In Sen et al. (1983) ART also displayed significant and moderately strong relationships with various other visual choice reaction time measures, ranging in strength from *r*=.43 to *r*=.67. The intelligence constructs in the study included a marker of Gf (RPM) and two of memory function (forward and backward digit span). ART correlated significantly with RPM (-.36) but weaker with the memory measures (*r*=.17 and *r*=.18, respectively). Poon, Yu and Chan (1986) also reported significant relationships between various auditory RT parameters and RPM scores and also noted the general disregard of alternate modalities in previous studies of RT and intelligence. Contrary to these findings, however, Swani (1995) reported no relationship between Gf or Gc marker tests and visual and auditory measures of simple RT.

Temporal processing, as researched in the field of auditory abilities studies is expected to reflect concepts like working memory (Seidel, 2008). As noted, however, the reductionist approach has postulated that such tasks (or at least particular versions of them) measure temporal resolution of the CNS. Under the latter framework, Rammsayer and Brandler (2002) have reported that auditory Duration Discrimination (DD) was significantly better for higher-IQ groups than for low-IQ groups, and that it explained around 20% of the total variance in a single measure of Gf (Cattell's culture fair test). Other temporal tasks including Temporal Order Judgement (TOJ) and

Auditory Flutter Fusion (AFF) displayed statistically non-significant correlations with the Gf task (-.19 and -.04 respectively). Following this work, Helmbold and Rammsayer (2006) measured duration discrimination, temporal generalisation, temporal order judgement, and rhythm perception. They also measured performance on Gf and Gs marker tests. All of the temporal measures shared variance with the Gs tasks and Gf task, in the vicinity r=.15 to r=.47.

Given the limited psychometric intelligence tests used in these publications, later studies expanded to include full intelligence batteries measuring several distinct, reliable abilities. Helmbold et al. (2006) measured speed, reasoning and memory factors. Simple correlations were not reported in the publication of this study. However, according to the structural models reported, the measure of temporal discrimination (Temporal Generalisation, or TG) related to a reasoning factor (Gf) at approximately .33 (.38 x .86), a memory factor at approximately .26 (.38 x .68), and to a Gs factor at .31 (.38 x .81). Independent of these effects was a direct relationship with pitch discrimination performance of r=.41. Helmbold et al. (2007) later used a larger battery of temporal discrimination tasks which defined a general timing factor, termed temporal *g*. Correlations were again absent from the publication, but in the full SEM model reported, temporal *g* related to speed at .50 (.59 x .85), memory at .31 (.59 x .52) and with a capacity/Gf factor at .53 (.59 x .90). Temporal *g* also related strongly to a general RT factor (r=-.65).

Two other studies concerning temporal discrimination have reported the full correlation matrix for all measures used. In Rammsayer and Brandler (2007) all measures of temporal discrimination – Duration Discrimination, Temporal Generalisation, Temporal Order Judgement, Rhythm Perception, and Auditory Flutter Fusion – correlated weakly but significantly with almost all of the 15 intelligence measures used to tap constructs including Gf, Gc, Gs, Gv and Gm (memory). They also correlated weakly with various Hick RT measures. This pattern was again evident in Troche and Rammsayer (2009b) where all temporal measures correlated weakly but significantly with various capacity/reasoning and speed measures. Interestingly, in their structural models, the temporal g factor related primarily to speed rather than reasoning (r=.40 versus r=.22), whilst a non-temporal sensory discrimination factor related to capacity/reasoning and not speed (r=.43 versus r=-.02). The relation between temporal g and general sensory discrimination factors was r=.94 suggesting that whilst primarily sensory processes are involved in both tasks, they are certainly distinguishable in terms of the extent to which speed and/or cognitive efficiency are important.

General conclusions and position of this thesis

The foregoing review of auditory research highlights that, although there does exist a reasonable amount of literature concerned with auditory testing, it is most definitely fractured. Only the studies of Stankov and Horn were suitably positioned to properly examine auditory abilities in an intelligence framework (Horn & Stankov, 1982; Stankov & Horn, 1980). Others considered audition only because of a desire to select suitable candidates for musical courses (Seashore, 1919; Seashore, et al., 1960), or simply to 'map' the auditory domain independently of studies of other intellectual abilities (Suprenant & Watson, 2001; Watson, et al., 1982). It is apparent that most investigators concerned with auditory abilities, and even reductionists concerned with specific classes of tasks, have generally not committed themselves to careful selection of intelligence tests and the variations in correlations between the auditory tests and intelligence estimates are partly explained by differences in the criteria (Stankov, 1971).

The foregoing section regarding specific relationships between auditory abilities and other broad CHC constructs clearly demonstrates the confusion and apparent contradictory findings in this area. For example, during the course of auditory abilities studies, pitch discrimination was shown to relate to Gs (McLeish, 1950), to Gf (Acton & Schroeder, 2001), and to Gc (Holmes, 1954; Franklin, 1956, cited in Stankov, 1971). Similarly, in the reductionist studies, temporal discrimination has been shown to relate to estimates of *g* (Rammsayer & Brandler, 2007), Gf (Rammsayer & Brandler, 2002), Memory (Helmbold, et al., 2007), and Gs (Troche & Rammsayer, 2009b). Auditory RT has correlated with Gf and memory tasks (Sen, et al.) and AIT tasks have related to Gf (Brand & Deary, 1982), Gc (Irwin, 1984), Memory (Nettelbeck, et al., 1986), and Gs (McCrory & Cooper, 2005; Zajac & Burns, 2007).

There are several explanations for the varied relationships between specific auditory tasks and various intelligence constructs. One argument might be that the auditory task is measuring *g* and in the absence of a well defined Ga factor, it relates to the intellectual tasks because they are also confounded with some form of *g* variance. A specific example of this explanation could be the study of Nettelbeck et al. (1986) in which AIT-P related equally well to each of the Gf, Gc and associative memory tasks. Another explanation for the contradictory relationships is that inadequate attention is often given to the selection of intelligence tests. More specifically, Cattell's culture fair test (CCFT), for example, is often used as a measure of Gf but it has been shown to correlate as highly as r=.79 with full scale Verbal IQ, as measured by the WISC (Downing, 1965). Thus, the relationship between the auditory task and the CCFT might not reflect Gf or reasoning at all, but simply *g*. Another explanation is that most auditory tasks are relatively low-level in terms of complexity and the elementary processes required for performance permeate through all intellectual activities including performance on intelligence tests (Seidel, 2008).

There are only two conclusions that can be drawn from discussion thus far. First, it can be considered well established that there exists a series of distinct and reliable auditory abilities reflecting pitch discrimination, loudness discrimination and speech perception, for example, and that these converge under a broad auditory perceptual factor within the taxonomy of intelligence tests. Second, it can be concluded that auditory tasks do correlate with broad intelligence estimates, but probably only weakly on average at around r=.30 (Stankov, 1971). In regard to the hypothesis concerning whether other distinct broad ability constructs are indexed auditorily, no conclusions can be drawn at this time. There is insufficient evidence to support that auditory tasks can *consistently* and *reliably* relate to other specific broad CHC abilities, independently of Ga, because of the varied relationships between similar auditory tasks and numerous abilities reported across a myriad of independent publications. This hypothesis, however, is the main concern of this thesis for reasons that were outlined earlier in this chapter. Whilst some evidence does exist that auditory tasks can relate to other constructs, a common characteristic is the absence of systematic efforts to establish a detailed program of study concerning such relationships (Stankov, 1971); in other words, the correlations have been investigated ad-hoc.

The purpose of the present thesis is to redress this problem. Through a series of four separate but interrelated studies using auditory tasks, the thesis explores the relationship of auditory tasks to distinct broad abilities defined in CHC theory. The following exegesis outlines in more detail the reasoning and aims motivating each of the studies and their relevance to this hypothesis.

Chapter 2: Exegesis

The broad rationale for this thesis originated from my general interest in both intelligence testing and the auditory modality. I had previously formed an interest in Auditory Inspection Time which was consequently the topic of my Psychology Honours dissertation. During that research programme I realised the general neglect of auditory testing in relation to the structure of intelligence; that is, studies were concerned with whether auditory tasks related to broad intelligence estimates as opposed to the extent to which – if at all – auditory tasks could move 'out' of the auditory domain per se, and load on other abilities instead. Much research existed on AIT, but the story concerning the processes these tasks measure was, and still is, largely incomplete.

After concluding my honours studies, I began to read Carroll's (1993) book on the taxonomy of intelligence. During this time I noticed a pivotal problem with the data sets that formed the basis of his impressive reanalysis. More specifically, the tasks employed by nearly all of these studies utilised the visual modality only. Not surprisingly, the constructs of Gf, Gc and the like emerged as broad factors. However, despite the presence of only visual tests in these analyses, quite general remarks were made regarding the nature of these broad constructs. For example, in regard to his 2F (or Gf) factor, Carroll (1993) stated that "the types of variables that had high salient loadings on ... factors identified as Gf ... confirm the characterization of this factor as one involving difficult tasks of induction, reasoning [and] problem solving" (p.599). Similarly, for Gc, he noted that the primary factors subsumed by it involved language or verbal content, and that Gc reflected learning, experience and acculturation. Interestingly, however, Carroll also suggested that many of the auditory primary factors of Stankov and Horn (Horn & Stankov, 1982; Stankov & Horn, 1980) that defined Gf and Gc in their studies had probably loaded incorrectly and would be better regarded as measures of Ga instead.

This last point presents something of a contradiction. In the first instance, concluding remarks regarding broad constructs are so general as to imply that all tasks requiring reasoning and induction, for example, should load on Gf. However, the auditory tasks that do appear to require such processes are subsequently all but discounted as markers of this very ability. This is not a criticism of Carroll's (1993) work and, indeed, he notes the shortcomings of intelligence research in terms of how intellectual abilities are expressed in all of our modalities. Nevertheless, this raises the question of when exactly will an auditory task be regarded as a measure of Gf as opposed to Ga? Similarly, at what point will we accept that an auditory task is measuring vocabulary, or Gc – or any of the other constructs – and not Ga? As previously noted by Roberts et al. (1997) it is largely an assumption that existing broad constructs like Gf and Gc can be measured via all modalities. Historically, many assumptions have been proven wrong. One of the largest errors in history was the belief that the earth was flat but Greek astronomy and ocean explorations later proved this theory incorrect. Simply put, assumptions are not facts and differential psychology does itself no favours by neglecting to test such beliefs systematically and empirically.

In my previous investigations of AIT, I asked a similar question to that posed above. More specifically, I considered how exactly it might be established and subsequently concluded that AIT tasks measured the same processes as VIT. Without being able to identify exactly what processes underpin VIT and subsequently manipulate them, the only real method to explore this is by way of correlational studies. Correlations can be examined in terms of the extent to which AIT and VIT task performance are related, and factorial studies can be considered in light of whether such tasks load on the same broad factor(s) identified. Accordingly, as the tasks rely more on the same or similar processes, the higher the correlations and factor loadings will become. Another method used in conjunction with this is to deconstruct and consider what processes task performance actually relies upon. If there is no congruence between tasks in different modalities in terms of required processes, then it would seem unlikely that such tasks would index the same construct.

According to Carroll (1993), knowledge of broad abilities and the processes they require should guide the development of tests and hypotheses regarding how those tests will relate to existing constructs. Inadequate attention to task requirements has been a very real shortcoming of auditory research and this is evident in AIT studies as well as those concerned with temporal discrimination (this will be expanded upon in the following sections regarding Studies 3 and 4, respectively). One exception to this criticism is the work of Stankov and Horn (1982; 1980). In addition to exploring the structure of auditory abilities, Stankov considered the hypothetical relationships each auditory task would have with existing ability constructs based on task requirements. In some instances, these hypotheses were correct, with a primary auditory verbal comprehension factor loading well on Gc together with verbal comprehension as measured by visual tests (primary factor loadings of r=.43 and r=.50 respectively). The only other moderate loading of an auditory primary ability was for auditory immediate memory with Gf (r=.55). Both the memory tasks and verbal tasks did not primarily make use of processes that would be considered specifically auditory (like pitch discrimination). On the contrary, tasks that appeared to rely heavily on auditory specific processes did not relate well to other broad factors. An example of this is the *auditory cognition of relations* (ACoR) primary factor, which was measured by tasks including chord series, tonal analogies, chord decomposition and pitch changes in chords. Each of these tasks certainly depended on reasoning ability, yet ACoR loaded only weakly on Gf (r=.24), suggesting the tasks are saturated primarily with Ga variance.

My observations regarding the assumptions that apparently underpin intelligence taxonomies and the advice of Carroll (1993) in terms of test development, prompted me to explore these issues as part of my PhD studies. Specifically, I aimed to examine whether auditory tasks can be developed from the outset with the intention of measuring a specific construct other than Ga. To achieve this, careful consideration must be given to the processes that underpin the construct of interest and the tasks subsequently designed to emphasise those processes. If the auditory tasks are successful, then one can expect moderate to strong correlations between the measures in alternate modalities (see studies 1 and 2). Secondly to this, I wanted to extend this same approach and examine the processes required by existing auditory tasks in order to develop a-priori hypotheses regarding exactly how the tasks would relate to intelligence constructs defined by visual tasks based on the extent to which they indexed the same processes. Thus, I examined existing measures of AIT and temporal discrimination using this approach (studies 3 and 4). This is a far more structured and guided approach than simply 'mining' for relationships between these tasks and intelligence tests, which is typical of previous studies that have used these tasks, as highlighted in the introduction to this thesis.

One of the broad constructs relevant to all of the studies in this dissertation is that of general speediness, or Gs. It is outside the scope of this exegesis to thoroughly review the history of this construct because doing so would constitute a thesis in itself and excellent reviews of Gs have already been provided elsewhere (see e.g., Danthiir, Roberts, Schulze, & Wilhelm, 2005). However, to review briefly, the term Gs loosely defines a group of abilities considered to reflect speeded performance. Gs tasks are inherently simple to complete and this is what distinguishes speed factors from level (traditional accuracy based) factors; sometimes referred to as power tests. Performance on the latter is a function of difficulty; that is, items progress from simple to very difficult and regardless of time limits some people will be unable to solve certain items. In Gs tasks, on the other hand, item difficulty is controlled and the outcome is usually the number of items completed in a specified time. Pure tests of speed ensure that performance reflects the rate at which examinees complete the items (see Carroll, 1993, pp.458-461 for a discussion of this), not the difficulty of the items themselves.

Another specific construct of interest in the present thesis (Study 3, Chapter 5) reflects memory processes, or Gm. Again, it is beyond the scope of this exegesis to provide a complete review of this construct. However, according to Carroll (1993), Gm is reflected in all tasks that involve the learning, memorising, and recollection of content and there is evidence of rather distinct memory factors reflecting short-term memory, long-term memory and memory span, for example. I will now provide a refined overview of the rationale for each of the four studies comprising the dissertation because the background for each study is discussed in more detail in the relevant chapter.

Study 1

The purpose of this study – presented as Chapter 3 – was to begin to fill the void in research concerning how auditory abilities relate to other broad CHC factors.

Typically, auditory task batteries have been administered alongside limited batteries of visual intelligence measures and correlations are observed and interpreted post-hoc. Thus, these correlations have received little attention because the studies have been primarily concerned with auditory abilities themselves, or whether these auditory tasks relate to intelligence in a very general sense (often using notions of Spearman's *g* factor). As highlighted already, the relation between auditory abilities and other constructs is definitely not clear.

Rather than administer a large battery of auditory tasks together with a handful of intelligence measures and then look at how these tasks relate post-hoc, I decided to take a different approach. This study aimed to purpose-design auditory tasks to measure a single, specific broad factor. This method is advantageous because it provides an empirical test of the assumption that broad constructs are manifest in all modalities. It also provides a test of Carroll's (1993) ideas regarding designing tests based on the construct to be measured, rather than develop a task and subsequently explore *what* it measures. If auditory tasks are developed to maintain the cognitive requirements of existing visual tasks but no relations are found between them, then it becomes necessary to question the extent to which the processes that underpin the visual tasks are cognitively general as opposed to specific to the visual modality.

As already highlighted, I focussed on the broad ability Gs for a number of reasons. First, some existing tasks such as ART have been suggested as measuring Gs but have never been systematically investigated in this regard, and this study would provide the framework to do so. Second, traditional visual Gs tests are quick and easy tasks and many lend themselves readily to adaptation to the auditory modality. When developing the new tasks used in this study the nature of Gs was carefully considered and preserved in the auditory tasks, as recommended by Carroll (1993). More specifically, Gs subsumes intellectually easy tasks that require rapid scanning and responding (Horn & Noll, 1994). Preserving the elementary nature of Gs tasks is paramount because this is essentially what distinguishes them from higher level (accuracy-based) factors like Gf; where performance is a function of difficulty. Thus it would prove necessary to control item difficulty in the auditory speed tasks to ensure that performance only reflects the rate at which participants complete the items in the test (Carroll, 1993).

Some of the tests used in Study 1 were adaptations and extensions of those used in previous studies – such as ART – but particular interest was given to developing new ones. Regarding the new measures, I selected visual tasks that I thought could be translated into an auditory format, thereby ensuring consistency in the cognitive requirements across modalities. Examples of these visual Gs tasks that I adapted include Symbol Digit (McPherson & Burns, 2005), Number Comparisons (Ekstrom, French, Harman, & Derman, 1976), and Finding As (Danthiir, Wilhelm, Schulze, & Roberts, 2005).

Two auditory tasks in-particular required very careful consideration during their development. The first of these was based on the Symbol Digit task in which participants are presented with a code table pairing nine symbols to nine digits, and this table guides the response required by the participant on a given trial. For example, if a '+' sign is presented during a trial then the participant selects the number '2' in a numerical response pad because this symbol is paired to this number in the code table. I chose to adapt this task by presenting an auditory stimulus as opposed to a visual stimulus, and it was aptly named Audio Code. Pictures of musical instruments were paired to digits in a code table shown on the computer screen, in a similar fashion to Symbol Digit. A computer screen-shot of how this task looked is shown in Figure 2.

Two factors were of particular importance when choosing the instruments and their sounds. First, the instrument itself had to be easily identifiable. For this reason I chose typically familiar instruments like a drum, a guitar and a piano, with which the majority of people would be likely familiar. Second, the sounds presented for the instruments needed to be suitably distinct from each other so as not to invoke complex decision making processes. For example, discriminating the sound of a flute from the sound of a recorder is arguably more difficult than deciding between the sound of a drum and a piano, respectively. On any given trial, the sound of one of the instruments was presented auditorily via headphones and the participant clicked the digit corresponding to the instrument heard in the numerical response grid using the computer mouse. Like symbol digit, the outcome measure was the number of items completed correctly in 2 minutes. It is acknowledged that this test also requires a visual matching component and the implications of this are considered in more detail in Chapter 7.

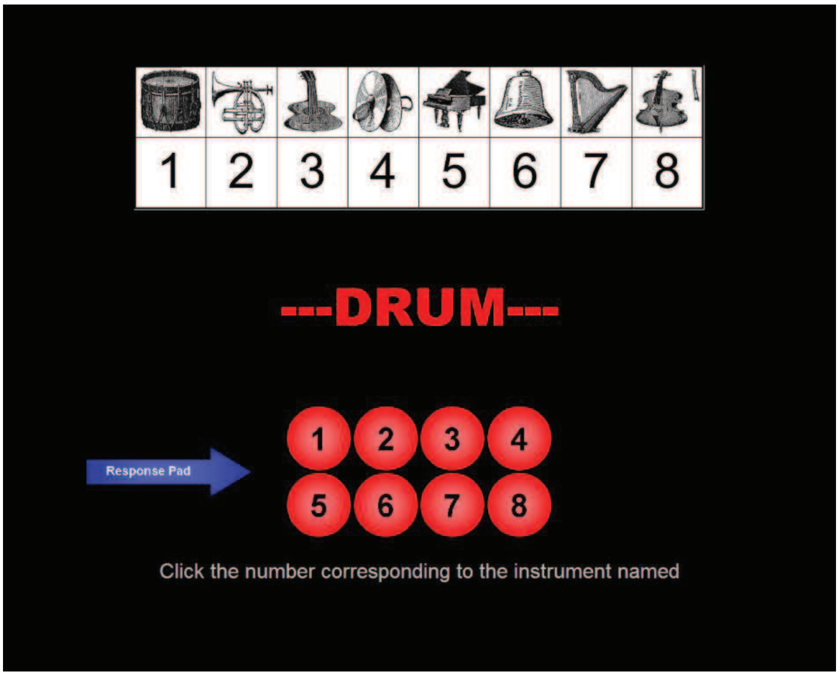


Figure 2. Computer screen-shot of the Audio-Code task

The Finding As task also presented some challenges during its adaptation to the auditory modality. Essentially, it was adapted by presenting words auditorily rather than visually. Participants had to indicate whether the spoken word contained the letter 'A', or not, and the outcome was the number of items completed correctly in 90 seconds. However, several important factors were considered during its development. For example, I identified a potential spelling confound which might occur if words contained the short vowel pronunciation of the letter A, such as in the word 'dial', or silent A in the case of words like 'head'. To alleviate this I selected words which only contained the long vowel sound such as 'grace', or 'chase'. Word length also posed as a possible confounding factor, as did word familiarity; less familiar words might require more complex consideration than more familiar words. In an attempt to combat these, nouns of five-to-eight letters in length were chosen and they all had a concreteness-of-imagery value of 600 or over on a scale ranging from 100 to 700: lower values indicate maximum abstractness and higher, maximum concreteness.

The purpose of Study 1, then, was to pilot the new auditory tasks (see Chapter 3 for details of the additional tone comparisons and chasing digits tasks) as well as some existing auditory measures including several variations of ART and AIT. Primary motivations were to examine the test reliability of the new auditory tasks, and to examine in a preliminary way the structure of these auditory tasks in relation to existing visual measures of Gs.

Study 2

Study 2 – presented as Chapter 4 – was a definite extension of the first study. The results of Study 1 are discussed in detail in Chapter 3. However, it is necessary to note them briefly here, to elucidate the reasoning for Study 2. As noted, the auditory tasks had been chosen and designed in light of the processes which underpin the broad Gs construct and, as hypothesised, the correlations were generally moderate where expected and statistically significant. Furthermore, factor analysis of the data uncovered two broad factors. One was tentatively termed broad RT, because it was loaded primarily by the auditory and visual RT tasks. The second factor was interpreted as a broad Gs factor and it was loaded on by the remaining tasks. It also appeared as though the factors were not modality specific given the mix of both visual and auditory tasks loading on each of them in a comparable fashion (loadings were generally moderate). Had the auditory tasks been saturated with auditory-specific variance, I would have expected them to define a distinct Ga-type speed factor, with only low loadings from the visual tasks.

The results of the study were thus promising and suggested that it might be possible to index Gs auditorily if adequate attention is given to test design. However, I regarded the findings of that study as tentative only. The reason for this was that an insufficient number of visual tasks had been employed and this was due primarily to time constraints. More specifically, the test battery already took at least 1.5 hours for participants to complete. Participation was largely voluntary and extending the test battery would have detracted from my ability to recruit a sufficient number of people. Therefore I concentrated primarily on the auditory tasks at the expense of visual ones.

Elsewhere, it has been suggested that at least three variables are required to permit the emergence of any one factor (Carroll, 1993). Thus, I reasoned that if the number of visual tasks were increased so that there was a ratio of 1:1 for auditory and visual tasks, then the idea that modality specific variance might underpin these tasks could be better explored. To accommodate the extra visual tasks, I excluded the AIT tasks from the test battery. I also reduced the number of trials in the ART tasks from 60 down to 30, given the tasks had displayed good reliability in Study 1. The final test battery included a visual speed task matched to every auditory speed task to ensure that these modalities were appropriately balanced.

One additional aim of Study 2 was to examine the nature of the Gs construct defined by the tasks. More specifically, it was of interest to see whether latent Gs related differently to more complex ability measures depending upon whether it was defined solely by visual tasks as opposed to a mixture of visual and auditory tasks. To achieve this, participants also completed Raven's Advanced Progressive Matrices (Raven, Raven, & Court, 1998) which is considered a complex measure of fluid reasoning ability (Gf). The intention was to regress Gf scores onto the visual speed and visual+auditory speed factors independently, and compare differences in the associations between them.

Study 3

For Study 3 aims departed somewhat from the preceding two studies. It was not concerned with developing new measures, as were studies 1 and 2, but was instead concerned with examining existing tasks in light of the processes required by them. According to Carroll (1993), this in turn would guide hypotheses regarding how these tasks should relate to broad intelligence measures and elucidate the nature of the relationship evident between these tasks and estimates of general intelligence. Study 3 was specifically concerned with the temporal discrimination tasks discussed in the introductory chapter.

While perusing the literature on auditory tasks during my first year of PhD candidature I stumbled across a paper concerned with temporal discrimination

(Helmbold, et al., 2006). I was familiar with the notion of timing discrimination being related to factors like working memory and speech abilities via the auditory and musical abilities studies reviewed earlier. Thus, I was somewhat perplexed by the conclusions being based on the data reported by Helmbold et al. I therefore tracked down the remaining papers in this series, reviewed them thoroughly and, being familiar with the pitfalls of working with auditory tasks, was not overly surprised at what I considered to be rather ambitious conclusions in terms of what the temporal discrimination tasks measure.

Rammsayer and colleagues (Gibbons, Brandler, & Rammsayer, 2002; Helmbold & Rammsayer, 2006; Helmbold, et al., 2006, 2007; Rammsayer & Brandler, 2002, 2004, 2007; Troche & Rammsayer, 2009a, 2009b) have referenced the mental speed approach to human intelligence as a key motivation for their studies of temporal discrimination. This approach has provided converging evidence for a linear relationship between efficiency and speed of information processing as measured by elementary cognitive tasks (ECTs) and psychometric measures of intelligence (see Deary, 2000a). Explanations of this association have focussed on the concept of neural efficiency and any number of related variables including reliability of neuronal transmission, cortical activation, neural pruning and myelination of neurons (Helmbold, et al., 2006). An additional theory (see Surwillo, 1968), termed the master clock theory, concerns a hypothetical neural oscillation process. According to this, temporal resolution power of the CNS is related to psychometric intelligence because of faster speed of information processing.

The studies of temporal discrimination referred to have essentially attempted to consolidate this theory by way of auditory temporal discrimination tasks. The authors have argued that such tasks provide an estimate of CNS resolution and that the correlations between these tasks or timing ‘factors’ and psychometric intelligence support the importance of speed of neural oscillations to intellectual functioning. My earlier work on AIT (Zajac & Burns, 2007) had highlighted that researchers did not adequately consider what processes underpinned performance on AIT and I was beginning to see a similar trend emerging for research on temporal discrimination. More specifically, many of these temporal tasks are purportedly ‘elementary’, but a close inspection of task requirements shows that they have the potential to actually be rather difficult to complete. This is discussed in more detail in Chapter 3, but an example of such tasks is that of Temporal Generalisation (TG). Essentially, this task requires a participant to learn a standard stimulus and compare it to stimuli presented throughout a lengthy auditory test. Even if the task does index temporal resolution, performance appears to be most certainly confounded by higher order cognitive operations, including memory processes. The authors of these studies never openly acknowledge this as a possible explanation for the relationship between these tasks and intelligence. If it cannot be confirmed that these tasks *only* measure CNS resolution, then conclusions that purport temporal discrimination as reflecting neuronal efficiency, and that this itself relates to intelligence, are speculative.

The two independent studies reported in Chapter 3 of this thesis sought to address these issues. The first study was concerned primarily with the TG task and the extent to which it relates to existing ECT/speed tasks such as RT and IT as opposed to a more complex working memory task. Analysis of the processes required in completing the task suggests that TG should load more strongly with memory measures, and that this might explain the relationship between this task and estimates of *g*. In addition to this, the TG task was adapted to the visual modality. The purpose of this was to assess the extent to which timing discrimination reflects general, as opposed to modality specific, processes. Results of this study established a strong relationship between temporal discrimination performance in both visual and auditory modalities. Moreover, regression analyses implied that temporal discrimination was better related to the memory measure than the speed tasks, supporting the hypothesis that it measures memory processes.

Study 1 used a relatively small sample and only one measure of memory. To address this short coming and further explore the relationship between temporal tasks and memory, I undertook a reanalysis of Rammsayer and Brandler's (2007) data. Specifically, I tested the extent to which memory, as measured by intelligence tests, mediated the relationship between temporal tasks and *g* using structural equation modeling. The analyses showed that around 34% of the relationship between temporal discrimination and *g* was attributable to shared processes with the memory tasks.

Study 4

The final study in this thesis capitalised on the data collected during Study 1. Similar to Study 3, it was not concerned with the new auditory tasks that I had developed but instead aimed to examine existing measures of AIT in more detail because this was outside the scope of Study 1. Study 4 followed somewhat from my honours research regarding these tasks. In light of existing literature reporting on correlations between these tasks, VIT, and intelligence constructs, I have concluded previously (see Zajac & Burns, 2007) that AIT measures have largely failed because during their development insufficient attention has been paid to the discriminations required by, and therefore the cognitive processes which underpin, VIT performance. These conclusions accord with Carroll's (1993) view that tasks should be designed based on the cognitive processes required. Unfortunately, AIT-P and AIT-L are fundamentally different to VIT in terms of the discriminations required by them. In addition to this, there is sufficient evidence that pitch and loudness discrimination are auditory-specific, and even if these AIT tasks measured a similar construct to VIT, the variance in them would be confounded by auditory processes. This would make it difficult to disentangle how much of the relationships between these measures and intelligence tests is due to common IT processes as opposed to Ga type processes.

AIT-S on the other hand is remarkably similar to VIT in terms of cognitive operations given that both require spatial discriminations concerning the target stimulus. Spatial information has been shown to be shared between these modalities (Teng & Whitney, 2008) and it is therefore unlikely that spatial localisation constitutes a purely auditory ability. In each published study reporting on AIT-S and VIT, these measures have correlated significantly and both also correlate significantly with measures of Gs, supporting the notion that they are at least partly measuring a similar construct. Furthermore, whilst neither AIT-L nor AIT-P appear on face-value to be inherently difficult, data concerning completion rates show that many participants struggle with these tasks (see Zajac & Burns, 2007 for a review). This might reflect the typically large variability in auditory abilities in samples classified as being of ‘normal hearing’ (Suprenant & Watson, 2001) and highlights that – in light of this variability – the parameters applied to auditory tests adapted from the visual modality might be restrictive. For example, the shortest stimulus duration at which most participants can resolve an AIT task is significantly longer than average VIT estimates and some participants cannot complete AIT at the longest stimulus duration permitted in the traditional VIT task. The fact that these tasks are inordinately difficult essentially violates the requirements of ECTs and speeded tasks in general. Ultimately then, there are numerous factors that can prove problematic when developing auditory tasks to measure specific constructs like AIT.

Study 4 thus considered AIT tasks in more detail. More specifically, although both AIT-P and AIT-L share variance with pitch discrimination, this does not fully account for the correlation between these tasks and measures of intelligence. Thus, this shared variance might reflect temporal limitations in processing, or Gs. To date there has been only one study (McCrory & Cooper, 2005) that has reported on all three AIT tasks and measures of speed and therefore the data collected during the course of Study 1 represented a opportunity to replicate this work. Moreover, the study reported on modifications to the psychophysical procedure used in AIT tasks which were designed to overcome some of the issues inherent in earlier investigations of AIT. More detailed information regarding these modifications is provided in Chapter 6.

Chapter 3: Study 1 – Auditory Measures of Gs

DEVELOPING AUDITORY MEASURES OF GENERAL SPEEDINESS

Ian T. Zajac[1], Nicholas R. Burns[1], Vanessa Danthiir[2] & Ted Nettelbeck[1]

[1]School of Psychology, University of Adelaide, South Australia

[2]CSIRO, Adelaide, South Australia

Manuscript submitted for publication

SAGE Open

Statement of Contributions

Ian Zajac (Candidate)

I was responsible for the conception and primary authorship of the paper. I was responsible for the development and programming of software based assessment tools and collection of all data. I conducted the statistical analyses independently with advice from the co-authors. I am corresponding author and primarily responsible for responses to reviewers and revisions to the paper.

Nicholas R. Burns (Co-author)

I was the supervisor (advisor) for the research programme that led to this publication. In terms of conceptualisation of the programme, there was extensive and

ongoing collaboration between Mr. Zajac and me in developing the direction of the research. The realisation of the study and the programming of assessment tools used, were the work of Mr. Zajac. I had an advisory role with respect to selection of the test battery used and on the direction and specifics of the data analyses.

Mr. Zajac was responsible for writing this paper; my role was to comment on drafts, make suggestions on the presentation of material in the paper, and to provide editorial input. I also provided advice on responding to comments by the journal reviewers and editor.

I hereby give my permission for this paper to be incorporated in Mr. Zajac's submission for the degree of PhD in the University of Adelaide.

Vanessa Danthiir (Co-author)

I was the co-supervisor for the research programme that led to this publication. The realisation of the study and the programming of assessment tools used, were the work of Mr. Zajac. I had an advisory role with respect to selection of the test battery used.

Mr. Zajac was responsible for writing this paper; my role was to comment on drafts and to provide editorial input.

I hereby give my permission for this paper to be incorporated in Mr. Zajac's submission for the degree of PhD in the University of Adelaide.

Ted Nettelbeck (Co-author)

I was the co-supervisor for the research programme that led to this publication. The realisation of the study and the programming of assessment tools used, were the work of Mr. Zajac. I had an advisory role with respect to selection of the test battery used.

Mr. Zajac was responsible for writing this paper; my role was to comment on drafts and to provide editorial input.

I hereby give my permission for this paper to be incorporated in Mr. Zajac's submission for the degree of PhD in the University of Adelaide.

Abstract

This study examined whether the broad ability General Speediness (Gs) could be measured via the auditory modality. Existing and purpose-developed auditory tasks that maintained the cognitive requirements of established visually presented Gs markers were completed by *N*=96 university undergraduates. Analyses showed that the auditory tasks combined with established visual Gs measures to define latent Gs and Reaction Time (RT) factors. Moreover, variance in the Gs and RT factors was not related to peripheral auditory processes as measured by auditory inspection time tasks. This study provides preliminary evidence which suggests that if auditory tasks are developed that maintain the same cognitive requirements as existing visual measures, then they will index those cognitive processes.

Introduction

Little research has explored whether known broad cognitive abilities can be identified when tests are presented via modalities other than visual. The abilities referred to are those identified in current taxonomies of human intelligence – including Gf-Gc theory (see e.g., Horn & Noll, 1994), and Carroll's (1993) three-stratum theory – and the issue is of interest because the extent to which these abilities may reflect modality specific versus general cognitive processes is not entirely clear. Moreover, it was previously noted that an adequate description of intelligence will reflect knowledge of the basic forms in which it is manifest within each of our sensory-perceptual modalities (Horn & Stankov, 1982). This knowledge is still lacking.

There is reason to think that intelligence is revealed other than just via the visual system. Consider the auditory modality: This is the prime channel for communicating and acquiring information during infancy and this reliance only shifts to the visual system at about five years of age (Sloutsky & Napolitano, 2003). Thus, auditory abilities are crucial for the normal development of intelligence and children with hearing impairments, or total hearing loss, are disadvantaged because of lack of exposure to abstract concepts (Stankov, 1994a). Research which shows that the average IQ of deaf people is lower than that for people with normal hearing (Vernon, 2005) and for the blind (see Stankov, 1994a) supports the integral role audition plays in human intellectual functioning.

Despite this, audition has been referred to as the neglected domain in intelligence models and tests (Seidel, 2005). More specifically, Stankov (1971) long ago noted that prior to his PhD research there was nothing in the field of differential psychology concerned with general auditory (Ga) intelligence. More than two decades later Carroll (1993) echoed this sentiment, commenting that evidence on the structure of auditory abilities was scant due to the little attention given to that domain. These concerns have gone unheeded, however, and Seidel (2008) recently reiterated the ongoing neglect of putative auditory cognitive abilities.

It is true that our understanding of general auditory abilities (Ga) has progressed, and a hierarchy that distinguishes between sensory, perceptual and thinking abilities has been proposed (Stankov, 1994a). However, research on auditory abilities as they relate to traditional measures of intelligence remains fractured; as is research concerning how they relate to established broad intellectual factors (e.g., fluid intelligence; Gf). We next provide a brief overview of research on auditory cognitive abilities, before outlining the aims of the present study.

Horn (1968) was the first to propose that if tests were constructed that were like existing cognitive abilities tests yet involved auditory processes, then they would define a broad Ga factor in the taxonomy of human intelligence. Confirmation of this hypothesis would imply that intelligence test batteries – as they currently stood – were biased, and that truly balanced measures should contain items emphasising the use of audition as well as vision. Stankov's (1971) doctoral research tested Horn's (1968) ideas. Performance was measured on a large battery of existing and purpose-developed auditory tests (reported in Horn & Stankov, 1982; Stankov & Horn, 1980). Exploratory factor analysis identified seven primary auditory factors accounting for approximately 50% of the variability in the tasks. As proposed by Horn (1968), these primaries were subsumed in hierarchical analyses by a single broad Ga factor (Stankov & Horn, 1980).

Prior to Stankov and Horn's work, 'practical need' motivated research on auditory abilities. The aim was to identify those abilities that might prove useful in the selection of musicians and of military personnel, including radio and sonar operators (Stankov, 1994a). This early research can be broadly classified as either musical-or speech-focused. Musical tests measure the ability to discriminate and make judgments with respect to all attributes of musical sounds (Carroll, 1993), whilst speech tests – such as comprehension – depend on knowledge of language (Seidel, 2008). Studies have shown that auditory speech abilities are relatively independent of non-speech (or musical) abilities (Seidel, 2008; Suprenant & Watson, 2001). Moreover, speech abilities appear to relate more strongly to traditional intelligence measures and particularly to the verbal tests subsumed by general crystallized intelligence (Gc; Seidel, 2008).

This finding of distinct speech and non-speech factors accords with the view that auditory abilities form three hierarchically organised layers comprising sensory, perceptual and thinking processes (Stankov, 1994a). Abilities at the sensory level relate weakly with each other as well as to the higher-order processes (Seidel, 2008); and they relate weakly to abilities in other sensory modalities. They are best measured using acuity tasks that determine thresholds for frequency discrimination, loudness discrimination, spatial localization, and the like. At the highest level, however, the abilities are intellective in nature and the reliance on audition is incidental. Stankov (1994a) has proposed that such higher-order abilities can be measured in all modalities.

Carroll (1993) suggested that auditory speech abilities extend beyond sensory level factors and it is likely that they can be located at the thinking level because they rely not on the sensory aspects of the tasks but rather on the ability to grasp the complex and demanding relations among elements of the problem (Stankov, 1994a). In this sense, they are similar to traditional Gc marker tests where performance does not rely so much on the perception of two visually presented words, for example, but rather on the cognitive processing required to decide whether they are synonyms or antonyms. Thus, it can be hypothesised that any broad cognitive ability could be measured via the auditory modality provided that the tasks developed to do so maintain the same cognitive requirements as corresponding visual tasks, but use auditory material.

Some of Stankov and Horn's studies (Horn & Stankov, 1982; Stankov, 1978; Stankov & Horn, 1980) support this hypothesis. Correlations between certain auditory tasks and broad factors defined by traditional visual measures suggest that fluid intelligence (Gf) and crystallised intelligence (Gc) can be measured auditorily. Listening verbal comprehension, for example, loads moderately on Gc (Horn & Stankov, 1982); whilst tonal series and chord series load strongly on Gf (Stankov, et al., 1980). There is also evidence that auditory tasks measure memory functions (Gibbons, et al., 2002; Stankov, et al., 1980; Troche & Rammsayer, 2009a; Zajac & Burns, 2009) and these findings have been integrated into the Woodcock-Johnson III Tests of Cognitive Abilities (Woodcock, McGrew, & Mather, 2001), which includes auditory measures of short-and long-term memory, which are considered independent from the Ga tasks in that battery.

Research on so-called elementary cognitive tasks (ECTs) suggests there might also be a link between auditory ECTs and general speed of processing (Gs). Specifically, auditory reaction times (ART) have been shown to share variance with cognitive abilities measures (Poon, et al., 1986). Whilst their relationships to Gs appear not to have been addressed, they relate strongly to visual RT tasks (VRT; Sen, et al., 1983), which do relate to Gs (Danthiir, 2005; Danthiir, Wilhelm, et al., 2005). Thus, it is probable that ART will also relate to broad Gs.

Measures of Auditory Inspection Time (AIT) were proposed as indexing information processing speed and there now exists several alternate versions of AIT (Bates, 2005; Deary, 1992, 1994; Deary, Caryl, et al., 1989; Deary, Head, & Egan, 1989; Irwin, 1984; McCrory & Cooper, 2005; Parker, et al., 1999; Zajac & Burns, 2007). However, these AIT tasks relate differentially to visual Inspection Time (VIT) and inconsistently with Gs (see Zajac & Burns, 2007, 2011). Loudness and pitch versions of AIT share considerable variance with un-speeded sensory tasks (Irwin, 1984; Olsson, et al., 1998; Zajac & Burns, 2011) and frequency discrimination ability can account for the observed relationship between intelligence measures and the pitch task (Olsson, et al., 1998). On the other hand, spatial AIT relates more consistently to VIT and speed tasks, but variation in performance is still largely accounted for by a general AIT factor (Zajac & Burns, 2011). Thus, Zajac and Burns (2011) have argued that AIT tasks are not measures of Gs. Instead, they suggest that the processes that underpin AIT performance are largely perceptual and auditory specific. In terms of the hierarchy of auditory functions already discussed, these abilities are characteristic of the middle level; above sensory processes but below thinking processes.

An important issue concerning these putative relationships between auditory tasks and broad abilities other than Ga is that there have been few such reports, and they have received little attention because the studies have been primarily interested in auditory abilities themselves. Moreover, single marker tests have often been used and the broad abilities to which the auditory tasks relate have generally been poorly defined, leaving the associations open to different interpretations. Thus, putative relationships are only speculative and yet to be systematically explored.

The purpose of the current study was to begin to fill the void in research concerning the relation of auditory abilities to other broad factors. Rather than take the approach of administering a battery of auditory tasks and then assessing their correlations with traditional ability measures, this study aimed to explore whether auditory tasks can be purpose-designed to measure a single, specific broad factor. We felt this method was advantageous because it potentially permits an understanding of the relation of auditory abilities to a specific broad ability, as well as providing new understanding of the broad ability itself. More specifically, if auditory tasks are developed that maintain the same cognitive requirements as visual tasks, then the visual and auditory tasks should correlate well. On the other hand, if no relationships are found then it would be necessary to question the generality of the relevant broad ability and why it was not possible to measure it auditorily; for instance, should this broad ability be regarded as vision-specific given that only visual tests could define it?

The present study focused on the broad ability Gs for two reasons: First, tasks such as ART have been suggested as measuring Gs but have never been systematically investigated; second, traditional Gs tests are quick and easy tasks and many lend themselves readily to adaptation to presentation in the auditory modality. When developing new auditory tasks the nature of Gs must be carefully considered and preserved if the tasks are to prove successful. In short, Gs subsumes intellectually easy tasks that require rapid scanning and responding (Horn & Noll, 1994). The elementary nature of these tasks is paramount because this is what distinguishes Gs from higher level (accuracy-based) factors, where performance is a function of difficulty. Controlling item difficulty in pure tests of speed ensures that performance only reflects the rate at which participants complete the items in the test (Carroll, 1993).

The rationale for this research dictated that auditory tests be obtained or designed to represent processes analogous to existing visual Gs marker tests. Therefore, some tasks have been adapted from previous work – such as ART – whilst others were newly developed. Regarding the latter, we selected visual tasks that we thought could be translated into an auditory format, thereby ensuring consistency in the cognitive requirements across modalities. Examples of these visual Gs tasks include Symbol Digit (McPherson & Burns, 2005), Number Comparisons (Ekstrom, et al., 1976), and Finding As (Danthiir, Wilhelm, et al., 2005).

The purpose of the study reported herein was to pilot the new auditory tasks. We sought to examine their test re-test reliability and their factor structure in relation to existing visually presented Gs measures. At the core of this exploration is the intention to index the highest level of the auditory hierarchy already noted; the thinking level. Therefore it was considered necessary to examine whether latent factors defined by the new auditory tasks reflect auditory processes representative of the lower peripheral and sensory levels of the hierarchy. To accomplish this, AIT tasks were used because as already outlined a growing body of research suggests these tasks measure auditory specific functions such as frequency discrimination and loudness discrimination.

Method

Participants

Participants were N=96 undergraduate psychology students of the University of Adelaide. There were 69 females and 27 males, and the mean age of the sample was 20.0 years (*SD*=4.0 years). All participated as part of their first year psychology course requirements.

Apparatus

The presentation of all tasks and recording of responses were controlled by one of three identical Pentium 4 class computers. Visual stimuli were presented on 17 inch LCDs. Auditory stimuli were presented via Sony MDR-XD100 stereo headphones. All auditory tones were calibrated prior to the study using a Radio Shack 33-4050 Sound Level Meter.

Auditory Processing Measures

Auditory Inspection Time-Pitch (AIT-P). AIT-P required participants to make a judgement regarding the temporal order of two tones of markedly different pitch. Stimuli were generated in real time and trials consisted of a cue-tone (832 Hz) lasting 500 ms; 1000 ms of silence; a pair of target tones of different frequencies (880 Hz/784 Hz or 784 Hz/880 Hz), with zero inter-stimulus interval; and a backward mask consisting of alternating 10 ms bursts of both target tones. Target tone pairs were presented equiprobably and the intensity of all tones was held at 69 dB. Participants indicated whether the high-pitched (880 Hz) tone or the lower-pitched (784 Hz) tone was presented first by clicking either of two on-screen buttons marked ‘high first’ and ‘low first’, respectively.

Auditory Inspection Time-Loudness (AIT-L). AIT-L was identical to AIT-P except that target tone pairs were of a constant frequency (832 Hz) but differed in loudness (69 versus 64 dB). Participants indicated the presentation order of target tone pairs by clicking ‘faint first’ or ‘loud first’, respectively.

Auditory Inspection Time-Spatial (AIT-S). AIT-S required participants to make a spatial judgement concerning the apparent location of a target tone that appeared to originate from the left or right hand side of the saggital mid-line axis of the head. Stimuli were generated in real time and each trial consisted of: a cue-tone of 500 ms, followed by 1000 ms silence; a target tone which was a stereo tone with one channel phase-shifted by 40º; followed by a square wave tone delivered to both ears for 500 ms and which acted as a backward mask. The frequency of all tones was held constant at 450 Hz and intensity levels were 68 dB for the target and cue, and 77 dB for the backward mask. To offset artifacts due to phase shifting, the target tone was ramped from 0 dB through 68 dB over the first 1/3rd of the SOA. Participants indicated on which side the target originated by clicking either of two on-screen buttons marked 'left' and 'right', respectively.

Reaction Time Measures

Simple Auditory Reaction Time (ART). To begin each trial, the participant pressed the number '5' key in the numeric keypad of the keyboard. This key acted as a non-functional home button only and ensured that responses originated from the same position on each trial. After 300 ms a cue-tone (100 ms at 880 Hz) was presented followed, after a silent interval of variable duration (1300 ms, 1700 ms, 2100 ms or 2500 ms), by the target tone; a 500 ms 'bell' sound centered on a frequency of 800 Hz. Participants lifted their finger off the number '5' key and pressed the number '8' key as quickly and as accurately as possible. Participants were required to complete 10 correct trials out of 10 before they proceeded to the test. The outcome measure was mean RT – time between onset of target and pressing of response key – calculated after the removal of outliers (±3 SD) and errors.

Two-Choice Auditory Reaction Time (ART_2). The target tone in this task was presented to the left or right ear only. Participants responded by pressing the number

‘4’ key if the target tone was played to the left ear, or number ‘6’ if it was played to the right ear. All other aspects of the task were identical to ART.

Three-Choice Auditory Reaction Time (ART$_3$). During this task the target tone was presented to the left ear only, both ears, or right ear only. Participants responded by pressing the number ‘4’ key in the numeric keypad for the left ear, number ‘8’ if it was played to both ears, or number ‘6’ for the right ear. All other aspects of the task were identical to ART.

Visual Reaction Time (VRT). This task was functionally equivalent to ART but required participants to respond upon the illumination of an empty circle, 4 cm in diameter, presented against the black background of the computer screen. The white outline of the circle was presented at the onset of each trial and acted as a cue, and the circle illuminated red after a variable duration of 1300 ms, 1700 ms, 2100 ms, or 2500 ms.

General Speed of Processing Measures

Symbol Digit (SD). A computerised coding task was employed as a measure of Gs (see McPherson & Burns, 2005, for a detailed description of this task). A code table was presented at the top of the computer screen throughout the task. This comprised nine symbols arranged horizontally, to which nine digits, presented directly beneath them, were paired. For each item, one symbol was presented in the centre of the computer screen and participants responded by left clicking the mouse on its corresponding digit in a 3 x 3 numerical grid positioned at the bottom of the screen. Subsequent items did not commence until a correct response was registered. Participants were required to complete two practice trials correctly before they proceeded to the test. The outcome measure was the number of items correctly completed in 2 minutes.

Audio Code (AC). This task was developed to be an auditory analogue of the symbol digit task described above. A code table is displayed at the top of the computer screen for the duration of the task, comprising of pictures of eight musical instruments arranged horizontally, to which one of the numbers one through eight was paired. The instruments include a snare drum, trumpet, guitar, cymbals, piano, bell, harp and violin. For each item, the sound of one of the instruments was presented via headphones at an intensity of 65 db. Participants responded by left clicking the mouse on its corresponding digit in a 2 x 4 numerical response grid positioned at the bottom of the screen. Subsequent items commenced after a response was registered. Participants completed two familiarization phases: in the first, instrument names were presented and participants clicked on the corresponding instrument (2 trials each); in the second, instrument sounds were presented instead of text (2 trials each). Following this, participants were required to complete four test trials for each instrument correctly before they could proceed to the test phase. The outcome measure was the number of items correctly completed in 2 minutes.

Chasing-Digits Auditory (CD_A). This was designed to be an auditory analogue of the Digit-Digit task used by McPherson and Burns (2005), which was found to share substantial variance with Gs marker tests. It incorporated a 3 x 3 numerical response grid positioned in the centre of the computer screen against a black background. For each item, one of the digits 1-through-9 was presented auditorily to participants via headphones at an intensity of 65 dB; trial order was pseudo-randomised with the restriction that no digit could be presented on successive trials. The participant responded as quickly and accurately as possible by left clicking the mouse on the corresponding number in the response grid. Subsequent items commenced 200 ms following the response. Participants were required to complete 10 correct trials out of 10 before they proceeded to the test. The outcome measure was the number of items correctly completed in 60 seconds.

Tone Comparisons (TC). This task was developed to be an auditory analogue of Number Comparisons from the ETS Factor Reference Kit (Ekstrom, et al., 1976). Trials consisted of two sequentially presented tones which were identical, or differed by either a semi-tone or tone. Tones were presented at intensity of 65 dB and participants clicked the on-screen response button ‘Yes’, if they thought the tones were identical, or ‘No’ if they were different. Participants were required to complete five correct trials out of five before they proceeded to the test. The outcome measure was the number of items correctly completed in 90 seconds.

Hearing As (HA_S). This task was an auditory version of the search task, Finding As (see e.g., Danthiir, Wilhelm, et al., 2005; Roberts & Stankov, 1999). In this version, stimuli were nouns, five to eight letters in length. All words had a concreteness-of-imagery value of 600 or over, on a scale ranging from 100 to 700: the lowest value indicated maximum abstractness and the highest maximum concreteness. Fifteen of the 60 words contained the letter ‘A’ and all of these used the long vowel pronunciation. The purpose of this was to reduce any potential spelling confound that might occur for the short vowel sound – where the pronunciation of the ‘A’ is not as distinct – or in the case of silent ‘A’. Participants were informed of this restriction.

Words were presented pseudorandomly – one at a time – at an intensity of 65 dB. To respond, participants pressed the onscreen button “yes”, if the word contained an “A”, or “no” if it did not. Participants were required to complete five correct trials out of five before they proceeded to the test, and practice trials used different stimuli to the test phase. The outcome measure was the number of items correctly completed in 90 seconds.

Visual Inspection Time (VIT).

The vertical lines inspection time task was used to estimate VIT. Stimuli were presented on a video monitor at a viewing distance of approximately 60 cm. Preceding the target figure was a warning cue of approximately 520 ms; the cue was a small white plus (+) sign measuring 6 x 6 mm, presented in the centre of the computer screen. The target figure consisted of two vertical lines; one measured 15 mm and the other 30 mm. These were joined at the top by a horizontal line of approximately 18 mm. The shorter line appeared on either side of the target figure equiprobably. A 'flash mask' (see Evans & Nettelbeck, 1993) of 375 ms immediately replaced the target figure and consisted of two vertical lines 35 mm in length, shaped as lightning bolts. Participants indicated on which side the short line appeared by clicking either the left or right mouse button, respectively.

Procedure

Upon arriving at the testing session participants were seated in a cubicle in our laboratory and they were guided through the test battery automatically by the computer. Detailed instructions and practice phases were presented prior to the onset of each task, and the first author was present to answer any questions. Participants completed the tasks in the following order: AIT-S; ART; CD; HA_S; AIT-P; VIT; AC; ART_2; AIT-L; TC; ART_3; SD; and VRT. Each RT task consisted of 60 trials and the total testing session lasted 1hr and 45 minutes.

For the AIT tasks, instructions emphasised accuracy rather than speed of responding. Three practice phases preceded the test: the first required 10 correct trials out of 10 with SOA of 800 ms; the second required 10 correct trials out of 10 with SOA of 600 ms; and the third presented 12 trials with SOA of 300 ms, of which 10 were required to be scored correctly. If a participant did not meet this criterion in the third phase the task was aborted and the computer moved on to the next test. The estimation process began with SOA of 300 ms which was reduced by 10 ms following each successive correct response. After the first error the task began to follow an adaptive staircase algorithm (Wetherill & Levitt, 1965) and SOA increased by 10 ms following each error, and decreased by 10 ms following three successive correct responses. The average SOA was calculated over eight reversals of direction on the staircase, giving an estimate of the SOA with an associated probability of 79% of making a correct response.

A computerised tutorial was also completed for the VIT task and the instructions again emphasised accuracy rather than speed of responding. Practice trials required 10 correct trials out of 10 with SOA of approximately 835 ms; 10 correct trials out of 10 with SOA approximately 420 ms; and nine correct trials out of 10 with SOA approximately 250 ms. The estimation process began with SOA approximately 250 ms and followed the same adaptive staircase algorithm as the AIT tasks (Wetherill & Levitt, 1965), although SOA step size was approximately 17 ms. Average SOA was calculated as per AIT tasks.

To obtain test-retest reliability estimates for the new auditory tasks, N=17 participants returned six weeks after the initial testing session. Because of limited time, participants completed ART_2 only, along with the new tests Audio Code (AC), Chasing Digits (CD), Hearing As (HA_S), and Tone Comparisons (TC). Test-retest reliability estimates were not obtained for the IT tasks because of time restrictions and the fact that they are typically good (see Kranzler & Jensen, 1989 for VIT; Olsson, et al., 1998; Parker, et al., 1999 for AIT).

Results

Data preparation

Outliers (±3 SD) and errors were removed from individual RT data files and the average number of trials used to calculate RT scores was: ART (*M*=58.49, *SD*=1.20); ART_2 (*M*=58.33, *SD*=1.98); ART_3 (*M*=57.35, *SD*=2.04); and VRT (*M*=58.58, *SD*=.81). Following this the data for all tasks were collated and assessed for outliers and missing values. Immediately, we noted that the completion of AIT tasks was problematic and that IT thresholds were not available for *N*=18 AIT-L, *N*=7 AIT-P, and *N*=13 AIT-S participants. Given the proportion of missing values for these tasks, we did not impute values for these missing data. Across the remaining 15 variables there were three missing values and eight outliers which were subsequently removed (i.e., ±3 SD). These values were imputed using the Expectation Maximization (EM) algorithm in SPSS v.15.

Reliability estimates

Table 1 presents reliability estimates for the tasks. Split-half reliability is presented where possible and test-retest estimates are provided for the auditory measures. As can be seen, reliability was generally good, with split-half and test-retest estimates being acceptable for the new auditory measures.

Table 1: Split-half and test-retest reliability estimates

	Split-half	Test-Retest*
VRT	.90	-
ART	.72	-
ART_2	.91	.63
ART_3	.87	-
CD_A	.52	.69
SD	.84	-
AC	.75	.89
HA_S	.84	.80
TC	.79	.76

* (*N*=17)

CD_A, SD, AC, TC, HA_S = items. All else = msec.

VRT, Visual Reaction Time; ART, Simple Auditory Reaction Time; ART_2, Two-Choice Auditory Reaction Time; ART_3, Three-Choice Auditory Reaction Time. CD_A, Chasing Digits; AC, Audio Code; HA_s, Hearing As; TC, Tone Comparisons;

Descriptive statistics and correlations

Table 2 presents the descriptive statistics for the test battery. Of interest is that, although AC maintained very similar cognitive requirements to SD, the number of items completed was markedly fewer [$t(95)=18.34$, $p<.001$, $d=1.80$]. The simple RT tasks showed the opposite effect, however, with ART performance being significantly faster than for VRT [$t(95)=7.62$, $p<.001$, $d=0.66$].

Table 3 shows the full correlation matrix for all measures. To summarise: AC and SD correlated moderately, suggesting that they measure a similar construct. The simple RT tasks – ART and VRT – also correlated strongly. HA_S, CD, and TC were developed based on existing Gs marker tests and they related weakly to AC and SD, and shared variance with the RT measures also. The RT tasks generally correlated moderate-to-strongly with each other.

Table 2: Descriptive statistics for all measures

		M	SD	*Min*	*Max*
Items Correct	SD	91.6	16.0	59	132
	AC	67.5	10.2	46	100
	CD	64.0	4.9	49	78
	HA_S	43.4	5.8	28	56
	TC	49.8	6.0	34	61
Milliseconds	VRT	349.5	56.2	242	526
	ART	309.8	64.7	153	508
	ART_2	462.6	96.5	306	752
	ART_3	579.9	105.6	379	859
	VIT	45.6	13.2	14	93
	AIT-L[a]	155.7	64.6	56	295
	AIT-S[b]	87.0	72.9	6	293
	AIT-P[c]	109.7	64.4	25	291

[a] *N*=78; [b] *N*=82; [c] *N*=89

SD, Symbol Digit; AC, Audio Code; CD, Chasing Digits; HA_s, Hearing As; TC, Tone Comparisons; VRT, Visual Reaction Time; ART, Simple Auditory Reaction Time; ART_2, Two-Choice Auditory Reaction Time; ART_3, Three-Choice Auditory Reaction Time; VIT, Visual Inspection Time; AIT-L, Auditory Inspection Time Loudness; AIT-S, Auditory Inspection Time Spatial; AIT-P, Auditory Inspection Time Pitch.

Table 3: Full correlation matrix

	SD	AC	CD	HA_S	TC	VRT	ART	ART_2	ART_3	VIT	AIT-L	AIT-S
AC	0.59											
CD	0.23	0.25										
HA_S	0.26	0.24	0.27									
TC	0.24	0.44	0.16	0.24								
VRT	-0.14	-0.08	-0.24	-0.04	-0.17							
ART	-0.09	-0.01	-0.23	-0.21	-0.08	0.65						
ART_2	-0.28	-0.18	-0.44	-0.21	-0.3	0.48	0.54					
ART_3	-0.35	-0.28	-0.41	-0.25	-0.39	0.46	0.46	0.78				
VIT	-0.45	-0.36	-0.13	-0.28	-0.29	0.3	0.3	0.23	0.34			
AIT-L	0.00	-0.02	0.00	-0.1	-0.25	0.15	0.19	0.01	0.19	0.02		
AIT-S	-0.34	-0.18	-0.33	-0.16	-0.26	0.06	0.03	0.09	0.32	0.12	0.49	
AIT-P	-0.11	-0.29	-0.05	-0.23	-0.45	0.24	0.21	0.09	0.23	0.16	0.49	0.23

SD, Symbol Digit; AC, Audio Code; CD, Chasing Digits; HA_s, Hearing As; TC, Tone Comparisons; VRT, Visual Reaction Time; ART, Simple Auditory Reaction Time; ART_2, Two-Choice Auditory Reaction Time; ART_3, Three-Choice Auditory Reaction Time; VIT, Visual Inspection Time; AIT-L, Auditory Inspection Time Loudness; AIT-S, Auditory Inspection Time Spatial; AIT-P, Auditory Inspection Time Pitch.

Factor Analysis

Because we had developed the auditory measures to be analogous to existing visual speed tasks we had explicit expectations about the factorial structure underpinning the test battery. Specifically, we expected that RT tasks would define a latent RT factor, and that AC, SD, HA_S, TC, CD and VIT would define a latent Gs factor.

Rather than test these assumptions using confirmatory factor analysis (CFA) we instead employed exploratory factor analysis (EFA) to assess whether the expected structure would emerge naturally. We explored the suitability of the data for EFA using Bartlett's Test of Sphericity [$x^2(340) = 45.00$, $p<.001$] and the Measure of Sampling Adequacy (.75), and both measures showed the data were acceptable for further analysis. Next, the visual and auditory speed tasks were entered into a principal components analysis. There were two components with eigenvalues greater than 1 (3.75 and 1.69) accounting for 37.5% and 16.9% of the variance respectively. An inspection of the scree plot confirmed two dominant factors. Using the eigenvalue and scree criteria as a guide, we extracted two factors using maximum likelihood estimation with promax rotation. Table 4 presents the results of these analyses.
As can be seen, the expected structure emerged well. There is a clear reaction time factor (RT) defined by both the visual and auditory RT measures. There is also a clear speediness (Gs) factor defined by the remaining visual and auditory speed tasks. Of interest, although it was assumed that CD would share more in common with the Gs-type measures it appears to load more strongly on the G_{RT} factor. Given that this task requires participants to react quickly when one of the digits one-through-nine is presented auditorily, this is acceptable and suggests the task may be a measure of complex RT.

Table 4: Item loadings on the first un-rotated principal component, the two promax rotated factors, and the correlation between rotated factors.

	PCA[a]	**ML**[b]	
	F1	*RT*	*G*s
AC	-.44	.20	**.90**
CD	-.49	**-.37**	.20
HA_s	-.33	-.12	**.31**
TC	-.44	-.13	**.46**
SD	-.48	.02	**.72**
VIT	.44	.13	**-.45**
VRT	.55	**.63**	.06
ART	.56	**.70**	.14
ART_2	.85	**.87**	-.04
ART_3	.86	**.76**	-.20
G_{RT}			*-.41*

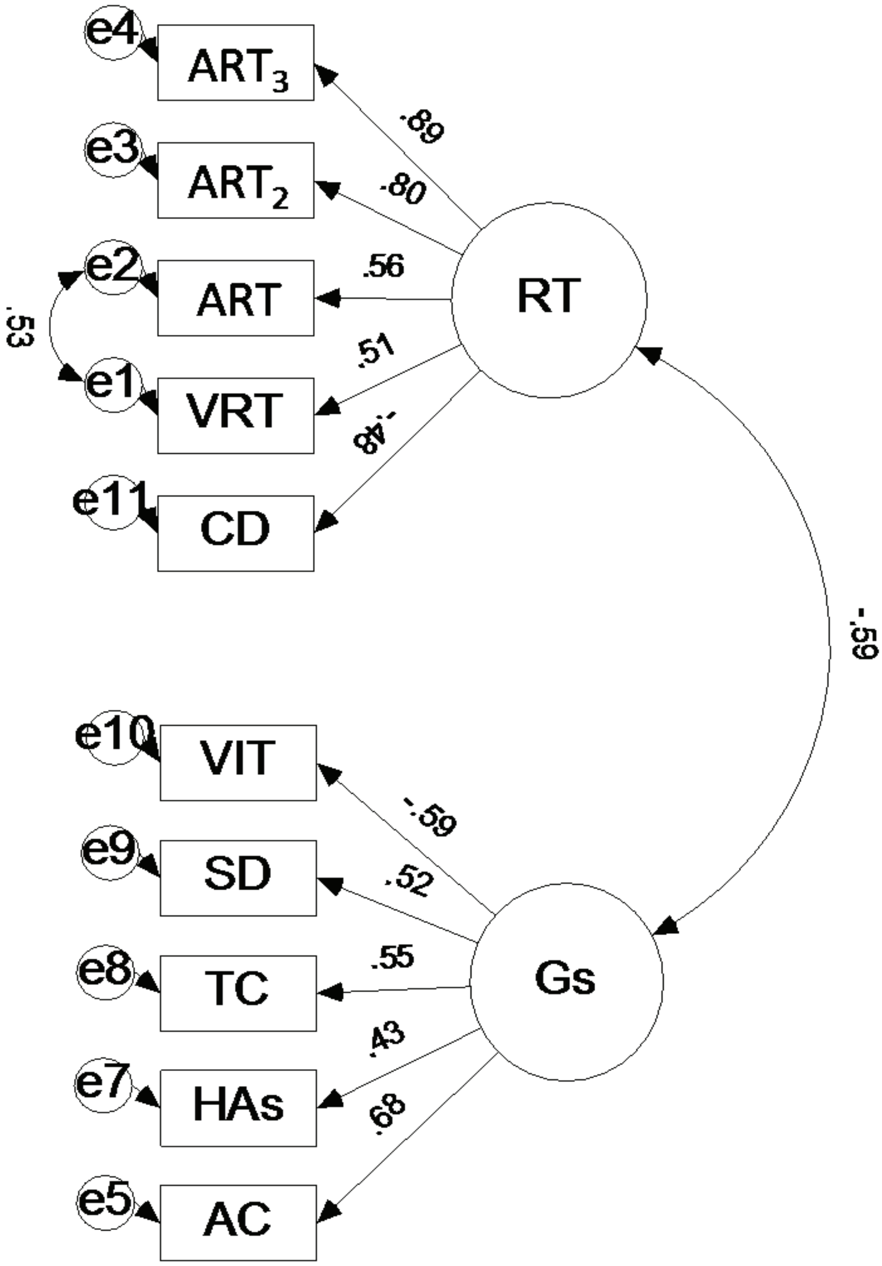
[a] PCA = Principal Componenets Analysis
[b] ML = Maximum Likelihood Solution
AC, Audio Code; CD, Chasing Digits; HA_s, Hearing As; TC, Tone Comparisons; SD, Symbol Digit; VIT, Visual Inspection Time; VRT, Visual Reaction Time; ART, Simple Auditory Reaction Time; ART_2, Two-Choice Auditory Reaction Time; ART_3, Three-Choice Auditory Reaction Time

Following EFA we tested the fit of this solution in MPlus (Muthen & Muthen, 1998) using CFA with maximum likelihood estimation. In our first model we tried to confine CD to the Gs factor as per our initial expectations but despite all tasks having moderate loadings, results showed the fit of the model was not adequate

[$\chi^2(34)=88.49$, $p<.001$, CFI=.80, RMSEA=.13, SRMR=.09]. Therefore, we followed the EFA solution with regards to CD as well as the theoretically sensible suggestion in the modification indices (MIs) to permit the residuals of ART and VRT to covary. Permitting these changes in a subsequent model resulted in adequate fit [$\chi^2(33)=53.13$, $p=.01$, CFI=.93, RMSEA=.08, SRMR=.06]. This model is shown in Figure 3.

As can be seen, the correlation between the RT and Gs factors was moderately strong and in the expected direction. Importantly, loadings of the visual tasks – SD, VIT and VRT – were as good as those for the auditory tasks suggesting that the latent factors may not be reflecting modality specific processes. Despite the problem of missing values noted earlier, we incorporated the AIT tasks to explore this further.

As noted, AIT tasks have been found to relate inconsistently and weakly to speed tasks with most studies suggesting that at least AIT-P and AIT-L variance can be attributed to auditory perceptual processes captured in the second level of the auditory hierarchy (see Zajac & Burns, 2011). Given that our intention was to develop auditory speed tasks that index cognitive processes (i.e., the thinking level of the auditory hierarchy) we examined the extent to which the perceptual processes underpinning AIT performance were responsible for variance in the latent speed factors. To do this, we incorporated AIT tasks into the model shown in Figure 3. *N* for this model decreased because of the missing values noted earlier.



$\chi^2(33)=53.13$, $p=.01$, CFI=.93, RMSEA=.08, SRMR=.06

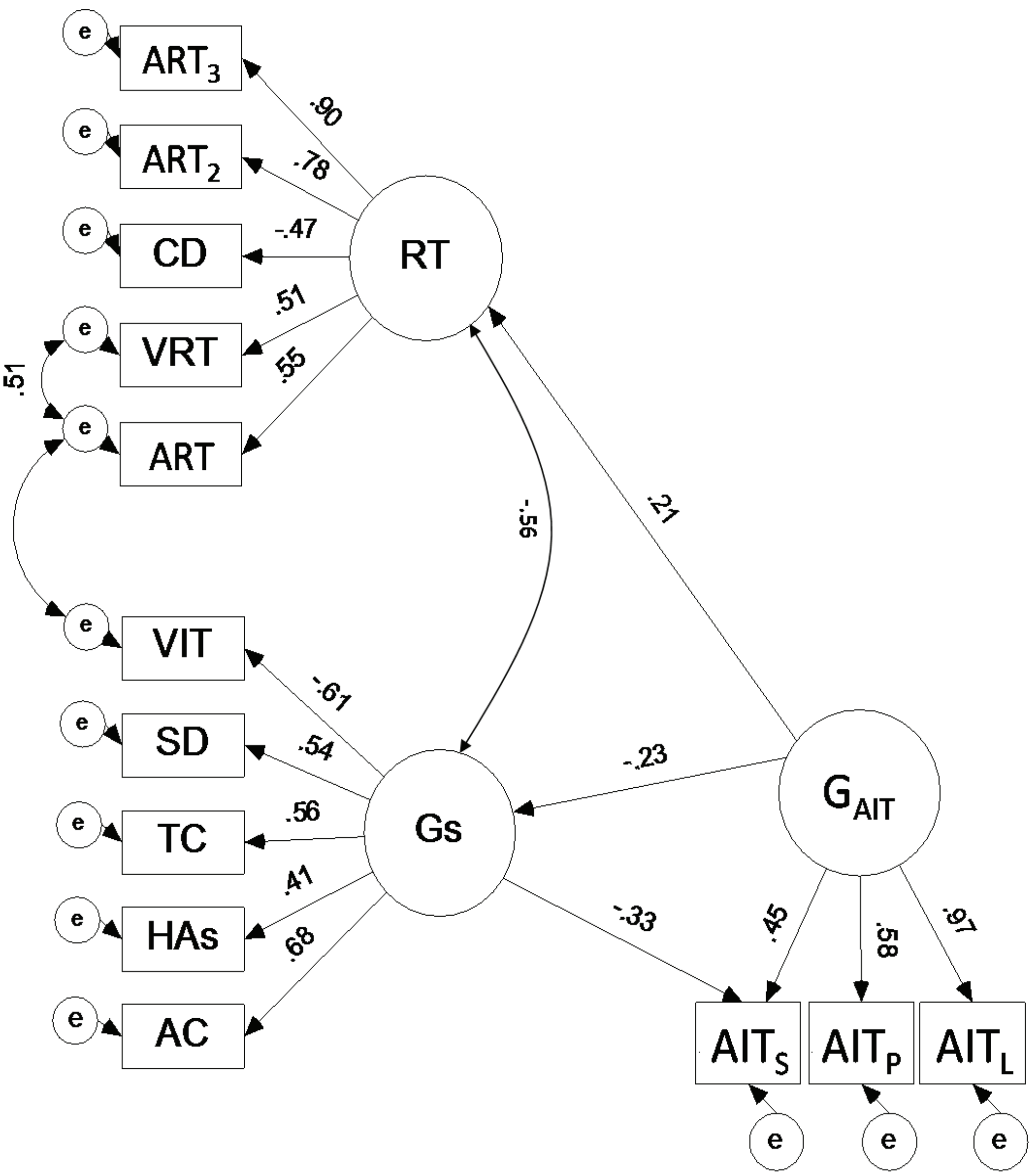
RT, Reaction Time; Gs, General Speediness; AC, Audio Code; CD, Chasing Digits; HA_s, Hearing As; TC, Tone Comparisons; VRT, Visual Reaction Time; ART, Simple Auditory Reaction Time; ART_2, Two-Choice Auditory Reaction Time; ART_3, Three-Choice Auditory Reaction Time; VIT, Visual Inspection Time

Figure 3. Latent structure of auditory and visual speed tasks.

First, RT and Gs were regressed onto a latent AIT factor. In this initial model AIT tasks loaded moderately on a general AIT factor (G_{AIT}) which itself, related weakly but significantly to RT (r=.30, p=.02), and moderately to Gs (r=-.42, p=.005). However, the fit of the model was not considered adequate [$\chi^2(61)$=107.67, p<.001, CFI=.86, RMSEA=.09, SRMR=.08]. Modification indices suggested that AIT-S should be allowed to load Gs. This covariance is consistent with our previous investigation which shows that AIT-S shares variance with Gs that is independent of its relationship to other AIT tasks (Zajac & Burns, 2011). Therefore, we permitted AIT-S to load both G_{AIT} and Gs and this resulted in a significant improvement in fit [$\Delta\chi^2(1)$=3.85, p=05] which, although better, was still not adequate [$\chi^2(60)$=103.83, p<.001, CFI=.88, RMSEA=.08, SRMR=.08]. A final modification based on MIs involved permitting the residuals of VIT and ART to covary, and this achieved good fit ([$\Delta\chi^2(1)$=14.63, p<.001]; Model Fit [$\chi^2(59)$=89.20, p=.007, CFI=.92, RMSEA=.07, SRMR=.07]). The need to correlate these residuals likely reflects the simplicity of these tasks relative to the others: together with VRT they are arguably the easiest tasks in the test battery. This model is shown in Figure 4 and, as can be seen, G_{AIT} did not account significantly for variance in either RT or Gs (r=.21, p=.12; and r=-.23, p=.20, respectively).

Discussion

The purpose of the present study was to examine whether the broad ability Gs could be measured via the auditory modality. Tasks were derived from existing paradigms or were purpose-developed to maintain the cognitive requirements of established visually presented Gs marker tests. To the extent that the auditory tasks



χ²(59)=89.20, *p*=.007, CFI=.92, RMSEA=.07, SRMR=.07

RT, Reaction Time; Gs, General Speediness; G_{AIT}, General Auditory Inspection Time; AC, Audio Code; CD, Chasing Digits; HA_s, Hearing As; TC, Tone Comparisons; VRT, Visual Reaction Time; ART, Simple Auditory Reaction Time; ART_2, Two-Choice Auditory Reaction Time; ART_3, Three-Choice Auditory Reaction Time; VIT, Visual Inspection Time; AIT-L, Auditory Inspection Time Loudness; AIT-S, Auditory Inspection Time Spatial; AIT-P, Auditory Inspection Time Pitch.

Figure 4. The relationship of General Auditory Inspection Time (G_{AIT}) abilities to General Speediness (Gs) and Reaction Time (RT).

were successful measures of Gs, we expected to find correlations between them and the visual Gs tasks that defined common speed factors.

First, we assessed the reliabilities of the auditory measures and found these to be acceptable. Of particular interest were the new auditory measures, with the strong split-half reliability estimates indicating that the constructs measured by these tasks did not change over the course of the individual tests. This provides some evidence that item difficulty was consistent within the individual measures, a characteristic of utmost importance because it ensures that performance reflects the rate at which examinees complete the items in the test, rather than item difficulty (Carroll, 1993). Test-retest reliability estimates were collected approximately six weeks after initial testing and they were also acceptable, showing that the measurement constructs did not alter significantly over this period.

Second, exploratory factor analysis suggested that two latent factors underpinned performance on the auditory and visual speed tasks. Based on factor loadings these were identified as latent RT, and Gs. Subsequent confirmatory analysis supported this solution and latent RT and Gs factors shared 35% of their variance. Importantly, the two visual tests loading on Gs (VIT, SD) had moderate loadings of a similar magnitude to the auditory measures, suggesting that the latent factor is not likely to be modality specific. Similarly, VRT loaded on RT moderately and comparably with the three ART tasks, suggesting that latent RT is not likely to be modality specific either. The hypothesis that these factors reflect speed rather than auditory abilities is supported by the finding of weak and statistically non-significant relationships between the speed factors and peripheral auditory processes captured by G_{AIT}.

Regarding G_{AIT}, it was found to be significantly related to Gs and RT in the first model tested, sharing around 17% of its variance with Gs. However, model fit was poor and modification indices showed AIT-S should be permitted to define both Gs and G_{AIT}. This modification is consistent with our research which shows AIT-S reflects both speed and AIT specific processes, whilst the other AIT tasks reflect only the latter (Zajac & Burns, 2011). Subsequent inclusion of this AIT-S x Gs path reduced the G_{AIT} x Gs relationship by 70% (from R^2=.17 to R^2=.05), and the relation of G_{AIT} to both speed factors became weak and not statistically significant. The absence of a significant effect in this latter model reflects a lack of statistical power. However, if an adequately powered sample was used and the strength of the relationships remained consistent, the effects are only small (*r*=.20). It can be concluded therefore that the auditory processes underpinning AIT performance are largely independent of the processes responsible for performance on the auditory Gs and RT tasks.

The fact that we used AIT tasks to measure auditory perceptual processes in this study might be considered a serious limitation. However, we have argued (Zajac & Burns, 2007) that AIT is a poor and unreliable measure of Gs abilities (see e.g., Bates, 2005; Deary, 2000a, 2000b; Irwin, 1984; Olsson, et al., 1998; Parker, et al., 1999; Raz, Willerman, & Yama, 1987; Zajac & Burns, 2007, 2011). Current results provide further support for this hypothesis because it was not necessary to include direct paths from either of AIT-P or AIT-L to Gs or RT. If anything, the weak relation of perceptual processes – as measured by AIT – to Gs and RT reported here is likely inflated because of the speeded aspects of AIT tasks. It is therefore necessary to resolve this issue in future research by measuring perceptual processes (i.e., pitch-discrimination and loudness discrimination) using unspeeded tasks.

As noted, auditory and visual tasks combined to define latent factors and the loadings of visual tasks relative to auditory tasks were comparable. On the basis of this we have proposed that the speed factors may not be modality specific but rather reflect cognitive processes which underpin performance on all tasks. It is the case, however, that visual tasks are under-represented in this study and the present authors acknowledge that this is a serious limitation. Given that the identification of latent factors generally requires three or more measures (Carroll, 1993, p. 52), it is entirely plausible that distinct visual speed factors did not emerge because of the relatively few visual measures used. It is essential that this issue be addressed in future research by ensuring that a sufficient number of tasks from each modality are employed to permit the identification or otherwise of modality specific speed factors. Furthermore, although the relation of G_{AIT} to speed factors was assessed in this study, auditory sensitivity was not and future research should explore this relationship. Another limitation in the present study is the use of undergraduate university students as well as the relatively modest sample size. Future research should employ a more representative sample and increase *N* to improve statistical power.

The present study represents a pilot study of new, relatively novel auditory tasks and the findings provide preliminary evidence that broad cognitive abilities such as General Speediness (Gs) may be measured using auditory tasks. Results suggest that the auditory tasks combine with existing visual measures to define Gs and RT factors and that these might not reflect auditory specific processes. In accord with Stankov's (1994a) hypothesis it appears that if auditory tasks are designed to engage the same cognitive processes as visual tasks, then they will index the same broad cognitive ability. These findings should encourage research into other broad abilities and the extent to which they might be measured auditorily also. Research such as this might ultimately result in the development of more 'truly balanced' measures of intelligence (Horn, 1968).

Chapter 4: Study 2 – Auditory Measures of Gs, Part 2

DO PURPOSE-DESIGNED AUDITORY TASKS MEASURE GENERAL SPEEDINESS?

Ian T. Zajac and Nicholas R. Burns

School of Psychology, University of Adelaide, Australia

Manuscript submitted for publication

Journal of Individual Differences

Statement of Contributions

<u>Ian Zajac (Candidate)</u>

I was responsible for the conception and primary authorship of the paper. I was responsible for the development and programming of software based assessment tools and collection of all data. I conducted the statistical analyses independently with advice from the co-author. I am corresponding author and primarily responsible for responses to reviewers and revisions to the paper.

<u>Nicholas R. Burns (Co-author)</u>

I was the supervisor (advisor) for the research programme that led to this publication. In terms of conceptualisation of the programme, there was extensive and

ongoing collaboration between Mr. Zajac and me in developing the direction of the research. The realisation of the study and the programming of assessment tools used, were the work of Mr. Zajac. I had an advisory role with respect to selection of the test battery used and on the direction and specifics of the data analyses.

Mr. Zajac was responsible for writing this paper; my role was to comment on drafts, make suggestions on the presentation of material in the paper, and to provide editorial input. I also provided advice on responding to comments by the journal reviewers and editor.

I hereby give my permission for this paper to be incorporated in Mr. Zajac's submission for the degree of PhD in the University of Adelaide.

Abstract

This study was concerned with the measurement of General Speediness using the auditory modality. Existing and purpose-developed auditory tasks that maintained the cognitive requirements of established visually presented Gs marker tests were completed by N=80 university undergraduates. Analyses confirmed the results of a previous unpublished study of ours (Zajac, Burns, Danthiir, & Nettelbeck, 2009) and auditory and visual tasks combined to define latent RT and Gs factors. Moreover, the analysis found no evidence of modality specific speed factors. Overall, the study provides evidence suggesting that auditory tasks might be able to successfully measure existing broad abilities defined in intelligence theories (i.e., Gf, Gc etc) provided they maintain the same cognitive requirements as existing visual measures of these constructs.

Introduction

The auditory perceptual domain is the second-most researched modality in the study of human intelligence (Stankov, et al., 2001). Research has shown that auditory tasks combine to define a broad auditory perceptual factor (Ga) thought to exist at the second stratum of intelligence hierarchies – i.e., Gf Gc theory (Horn & Noll, 1994) – alongside other broad constructs including fluid reasoning (Gf) and crystallized ability (Gc). Although the existence of this broad perceptual factor is relatively well replicated, our understanding of auditory abilities and how they relate to other broad constructs remains poor. According to Roberts, Pallier and Goff (1999), conclusions regarding the auditory modality have been drawn on the basis of only a handful of data sets, none of which offers a satisfactory account of this modality.

Stankov (1971) was the first to test the hypothesis that auditory tasks would define a broad Ga factor and empirically validated its presence and importance in intelligence theories. Carroll (Carroll, 1993) noted however, that although the existence of Ga was undeniable, defining the domain of auditory abilities was difficult because there had been "no trustworthy or extensive factor-analytic studies of musical talent"(p. 364) and the majority of factor analytic studies of intelligence had "totally neglected the domain of auditory abilities"(p. 365). The need to distinguish between abilities that are and are not strictly auditory abilities was also stressed. In order to be considered an auditory ability, Carroll (1993) requires there to be a reliance on the characteristics of the auditory stimulus and the ability to recognize and discriminate those characteristics. On the other hand, abilities such as speech comprehension rely on the knowledge of language structure and the reliance on audition is incidental. Speech comprehension might only be considered an auditory ability when the speech is distorted or manipulated, and in which case increased attention is devoted to the processing and discrimination of the incoming auditory stimuli.

These guidelines adopted by Carroll (Carroll, 1993) for the classification of auditory abilities accord with the hierarchy of auditory processes proposed by Stankov (1994a). This hierarchy incorporates three layers comprising sensory, perceptual and thinking processes, respectively. Abilities at the lowest, sensory level relate weakly with each other as well as to higher-order processes (Seidel, 2008), and they relate weakly to abilities in other sensory modalities. Auditory hearing threshold is characteristic of this sensory level. The second level encompasses abilities involved in making fine frequency and tonal discriminations. The highest level of the hierarchy encompasses abilities that are intellective in nature and not reliant on audition. Auditory abilities as outlined by Carroll (1993) are most characteristic of the perceptual level of the hierarchy whereas speech comprehension ability reflects the highest, thinking level.

Stankov (1994a) suggests that mode of stimulus presentation at the highest level of the hierarchy is incidental and has proposed that such thinking abilities should be measurable via all modalities. These higher-order abilities are thought to rely on cognitively complex mechanisms and unfortunately, according to Roberts, Stankov, Pallier and Bradley (1997), an implicit assumption of many extant theories of intelligence is that no knowledge of importance can be gained by employing tests of these complex abilities which utilize alternate modalities. Possibly as a result of this mindset, there appears to have been relatively little research that has explored whether it is, in fact, possible to measure complex cognitive abilities via alternate modalities such as audition, olfaction or tactile-kinesthesia.

One study of olfactory abilities (Danthiir, et al., 2001) and two of tactile kinesthetic abilities (Roberts, et al., 1997; Stankov, et al., 2001) appear to have addressed this issue but have provided somewhat mixed results. More specifically, Danthiir et al. (2001) developed a single olfactory task for each of Gf, Gc, Short-term memory (SAR) and long-term memory (TSR) constructs. For the most part, these broad factors emerged during factor analysis of the test battery inclusive of validated visual marker tests but not all of the olfactory measures loaded as predicted. Only Olfactory Swaps and Multiple Choice Smell Identification loaded as hypothesised on Gf and Gc, respectively. Open Ended Smell Identification and Odor Memory did not load on Tertiary Storage and Retrieval (TSR) and Short Term Memory as expected (TSR did not emerge during factor analysis at all). On the other hand, Roberts et al.'s (1997) and Stankov et al.'s (2001) studies of tactile-kinesthetic abilities do suggest that tasks presented in these modalities can index constructs including Gf and broad visualization (Gv). Roberts et al. (1997) proposed that the tactile-kinesthetic tasks in their study were cognitively complex and therefore they would relate to Gf and Gv more so than to Gc. Factor analysis of their data confirmed this: tactile-kinesthetic measures loaded on Gv rather than a modality specific factor. Stankov et al. (2001) replicated this finding and reported that complex tactile-kinesthetic tasks were again difficult to differentiate from Gv. Interestingly, Stankov et al.'s study also included "cognitively simpler"(p.25) tactile-kinesthetic tasks and these were found to define modality specific factors analogous to broad Ga.

When considering the auditory modality, there does not appear to have been any studies specifically concerned with the extent to which auditory tasks might index established constructs like Gf and Gc. Some of Stankov and Horn's studies (Horn & Stankov, 1982; Stankov, 1978; Stankov & Horn, 1980) have found relationships between certain auditory tasks and broad second stratum constructs. For instance, listening verbal comprehension does appear to correlate moderately with Gc (Horn & Stankov, 1982) and tonal series and chord series have been found to relate well to Gf (Stankov, et al., 1980). In the case of tonal series and chord series, however, these tasks usually share more variance with broad Ga than Gf when sufficient auditory measures are present for Ga to be defined. This is not surprising given these latter tasks depend primarily on the ability to make tonal comparisons whilst listening comprehension relies on prior knowledge of language. Findings such as these appear to support Carroll's (1993) classification of auditory abilities noted earlier.

Other studies broadly concerned with auditory tasks provide limited data concerning the extent to which such tasks index predefined, existing constructs. For example, attempts have been made to measure an auditory inspection time (AIT) analogous to that measured by the vertical lines visual IT task (VIT; see Burns & Nettelbeck, 2003). A series of auditory Inspection Time (AIT) measures have been developed, based on pitch discrimination (Brand & Deary, 1982), loudness discrimination (Olsson, et al., 1998) and spatial localisation (Parker, et al., 1999), respectively. As is the case with VIT, each of these measures has been shown to relate to performance on intelligence tests (McCrory & Cooper, 2005). However, the pitch and loudness discrimination tasks do not appear to rely on the constructs underpinning VIT. Instead, they have been shown to relate to measures of fine perceptual resolution including pitch discrimination ability (see Zajac & Burns, 2007, for a discussion of these findings). The spatial task relates more consistently with the intended VIT construct and other associated abilities (Parker, et al., 1999; Zajac & Burns, 2007), but it still shares considerable variance with the other AIT measures that is independent of its relationship to VIT (Zajac & Burns, 2011). Thus, despite the intention of measuring the IT indexed by VIT, the auditory tasks have generally not succeeded.

Overall, findings from tactile, olfactory and auditory research appear in some instances to challenge the notion that existing constructs can be measured via all modalities. The crux of Stankov's (1994) hypothesis, however, is that alternate tasks must maintain the cognitive requirements that underpin the ability/construct in question. Thus, although findings regarding AIT and olfactory measures do not entirely support this theory, there is a distinct possibility that this reflects inattention to this necessary detail. We have argued elsewhere (Zajac & Burns, 2007) that the reason why pitch and loudness AIT tasks are unsuccessful measures of IT might be due to a disregard of the cognitive demands imposed by the VIT task. For example, pitch and loudness tasks are based on temporal discriminations whereas VIT and the putatively more successful spatial AIT task require spatial discriminations. On the other hand, complex tactile tasks which appear to rely on visualisation display consistent relationships with visual tasks that rely on this same ability. Similarly, listening verbal comprehension and speech comprehension, as already outlined, rely on prior knowledge of language and this underpins performance on visual measures of these abilities, hence their strong relationships.

Recently, Zajac et al. (2009) considered the notion that thinking abilities could be measured via different modalities more directly by examining whether it was possible to purpose-develop auditory tasks to measure the specific, established broad cognitive ability, General Speediness (Gs). To achieve this, they compiled a battery of auditory tasks, some of which were developed so as to be analogous to existing visual Gs marker tests including Digit Symbol (McPherson & Burns, 2005), Number Comparisons (Ekstrom, et al., 1976) and Findings As (Danthiir, Wilhelm, et al., 2005). During task development the important characteristics of speed tasks were recognised and maintained. For example, it is particularly important that items be relatively easy to complete and that item difficulty be maintained within each task to ensure that performance reflects only the speed with which participants complete the items (see Carroll, 1993).

Zajac et al. (2009) reported that the auditory and visual tasks in their test battery combined to define moderately related factors termed Reaction Time (RT) and Gs, respectively. RT was loaded on by the visual and auditory RT tasks whilst Gs subsumed the new auditory speed tasks. Zajac et al. suggested that the latent factors might not reflect modality specific functions because visual speed tasks loaded comparably to the auditory tasks. This conclusion was endorsed by the finding that low-level perceptual processes captured in AIT performance were not significantly related to either of the speed factors. Zajac et al. suggested that it is possible for auditory tasks to measures Gs provided the auditory measures utilise the cognitive processes underpinning this construct and minimise reliance on auditory specific processes.

Despite Zajac et al.'s (2009) study providing promising results a number of limitations render them tentative only. In particular, although the auditory and visual tasks combined to define two related speed factors only three visual tasks had been employed. Therefore, the visual modality was under-represented and this may have precluded the extraction of modality specific speed factors. The purpose of the present study then was to address this limitation. In particular, it seeks to increase the number of visual marker tests used. The purpose of this is to examine the possibility that modality specific speed factors may exist. Additionally, this study will examine whether the latent Gs construct changes as a result of the inclusion of auditory tasks by regressing a measure of Gf onto both a visual-specific factor and a visual/auditory speed factor, respectively.

Method

Participants

Participants were *N*=80 undergraduate psychology students of the University of Adelaide. There were *N*=53 females and *N*=27 males with a mean age of 21.0 years (*SD*=3.6 years). All participated as part of their first year psychology course requirements.

Apparatus

The presentation of all tasks and recording of responses was controlled by one of three identical Pentium 4 class computers. Visual stimuli were presented on 17 inch LCDs. Auditory stimuli were presented via Sony MDR-XD100 stereo headphones. All auditory tones were calibrated prior to the study using a Radio Shack 33-4050 Sound Level Meter.

Raven's Advanced Progressive Matrices- Short Form (APM).

This was a computerised 12 item version of the original test (Raven, et al., 1998). The 12 items used were the same subset validated by Bors and Stokes (1998) for use as a brief form. Items were presented by displaying the original stimuli on a computer screen with the response options numbered from 1 to 8. Responses were made by using the computer mouse to select the number corresponding to the option they thought was correct and then clicking an 'Enter' button to enter their response. The 'Enter' button was included to allow participants to review their answer first. This was included as a measure of fluid reasoning ability (Gf). The outcome was the number of items completed correctly.

Auditory Reaction Time Measures

Simple Auditory Reaction Time (ART). To begin each trial, the participant pressed the number '5' key in the numeric keypad of the keyboard. After 300 ms a cue-tone (100 ms at 880 Hz) was presented followed, after a silent interval of variable duration (1300 ms, 1700 ms, 2100 ms or 2500 ms), by the target tone; a 500 ms 'bell' sound centered on a frequency of 800 Hz. Participants lifted their finger off the number '5' key and pressed the number '8' key as quickly and as accurately as possible. Participants were required to complete 10 correct trials out of 10 before they proceeded to the test. The outcome measure was mean RT – time between onset of target and pressing of response key – calculated after the removal of outliers (±3 SD) and errors.

Two-Choice Auditory Reaction Time (ART_2). The target tone in this task was presented to the left or right ear only. Participants responded by pressing the number '4' key if the target tone was played to the left ear, or number '6' if it was played to the right ear. All other aspects of the task were identical to ART.

Three-Choice Auditory Reaction Time (ART_3). During this task the target tone was presented to the left ear only, both ears, or right ear only. Participants responded by pressing the number '4' key in the numeric keypad for the left ear, number '8' if it was played to both ears, or number '6' for the right ear. All other aspects of the task were identical to ART.

Visual Reaction Time Measures

Visual Reaction Time (VRT). This task was functionally equivalent to ART but required participants to respond upon the illumination of an empty circle, 4 cm in diameter, presented against the black background of the computer screen. The white outline of the circle was presented at the onset of each trial and acted as a cue, and the circle illuminated red after a variable duration of 1300 ms, 1700 ms, 2100 ms, or 2500 ms.

Two-Choice Visual Reaction Time (VRT2). This task was the same as VRT but used two circles presented side by side, and with a space of 12 mm between them. Participants responded by pressing the number '4' key if the left circle illuminated, or the number '6' key if it was the right circle.

Three-Choice Visual Reaction Time (VRT3). This task utilized three circles presented side by side, and with a space of 12 mm between them. Participants responded by pressing the number '4' key if the left circle illuminated red, the number '8' key if the centre circle illuminated, or the number '6' key if it was the right circle.

Speed of Processing Measures

Symbol Digit (SD). A computerised coding task was employed as a measure of Gs (see McPherson & Burns, 2005, for a detailed description of this task). A code table was presented at the top of the computer screen throughout the task. This comprised nine symbols arranged horizontally, to which nine digits, presented directly beneath them, were paired. For each item, one symbol was presented in the centre of the computer screen and participants responded by left clicking the mouse on its corresponding digit in a 3 x 3 numerical grid positioned at the bottom of the screen. Subsequent items did not commence until a correct response was registered. Participants were required to complete two practice trials correctly before they proceeded to the test. The outcome measure was the number of items correctly completed in 2 minutes.

Audio Code (AC). This task was developed to be an auditory analogue of the symbol digit task described above. A code table is displayed at the top of the computer screen for the duration of the task, comprising of pictures of eight musical instruments arranged horizontally, to which one of the numbers one through eight was paired. The instruments include a snare drum, trumpet, guitar, cymbals, piano, bell, harp and violin. For each item, the sound of one of the instruments was presented via headphones at an intensity of 65 db. Participants responded by left clicking the mouse on its corresponding digit in a 2 x 4 numerical response grid positioned at the bottom of the screen. Subsequent items commenced after a response was registered. Participants completed two familiarization phases: in the first, instrument names were presented and participants clicked on the corresponding instrument (2 trials each); in the second, instrument sounds were presented instead of text (2 trials each). Following this, participants were required to complete four test trials for each instrument correctly before they could proceed to the test phase. The outcome measure was the number of items correctly completed in 2 minutes.

Chasing Digits Visual (CD_v). This was designed to be similar to the Digit-Digit task used by McPherson and Burns (2005), which was found to share substantial variance with Gs marker tests. It incorporated a 3 x 3 numerical response grid positioned in the centre of the computer screen against a black background. For each item, one of the digits one through nine illuminated green; trial order was pseudo-randomised with the restriction that no digit could be presented successively. The participants responded as quickly and accurately as possible by left clicking the mouse on the corresponding number in the response grid. Subsequent items commenced 200 ms following the response. Participants were required to complete ten correct trials out of ten before they proceeded to the test. The outcome measure was the number of items correctly completed in 1 minute.

Chasing Digits Auditory (CD_a). This task was developed to be an auditory analogue of, and was functionally equivalent to, CD_v. However, for each item one of the digits one through nine was presented auditorily to participants via headphones at an intensity of 65 dB. To respond, participants clicked the corresponding number in the response grid.

Number Comparisons (NC). This task was based on the test with the same name, from the ETS Factor Reference Kit (Ekstrom, et al., 1976). Trials consisted of two digit-strings of equal length (3 to 12 digits long) presented side by side and with a 6 cm space between them. Participants clicked the on-screen response button 'Yes', if they thought the strings were identical, or 'No' if they were different. Participants were required to complete five correct trials out of five before they proceeded to the test. The outcome measure was the number of items correctly completed in 90 seconds.

Tone Comparisons (TC). This task was developed to be an auditory analogue of NC and was functionally equivalent to it. Trials consisted of two sequentially presented tones which were identical, or differed by either a semi-tone or tone. Tones were presented at intensity of 65 dB and participants clicked the on-screen response button 'Yes', if they thought the tones were identical, or 'No' if they were different.

Finding As (FA_s). There exist numerous variations of this search task (see e.g., Danthiir, Wilhelm, et al., 2005; Roberts & Stankov, 1999). In this version, stimuli were two nouns, five to eight letters in length. All words had a concreteness-of-imagery value of 600 or over, on a scale ranging from 100 to 700: the lowest value indicated maximum abstractness and the highest maximum concreteness. Fifteen of the 60 words contained the letter A. Words were presented pseudorandomly – one at a time – and participants were to press the onscreen button 'yes', if the word contained an "A" and the 'no' key if it did not. Participants were required to complete five correct trials out of five before they proceeded to the test, and practice trials used different stimuli to the test phase. The outcome measure was the number of items correctly completed in 90 seconds.

Hearing As (HA_s). This task was an auditory version of Finding As and was functionally equivalent. In this version, stimuli were nouns, five to eight letters in length. All words again had concreteness-of-imagery values of 600 or over. Fifteen of the 60 words contained the letter 'A' and all of these used the long vowel pronunciation. The purpose of this was to reduce any potential spelling confound that might occur for the short vowel sound – where the pronunciation of the 'A' is not as distinct – or in the case of silent 'A'. Participants were informed of this restriction.

Visual Inspection Time (VIT). The vertical lines inspection time task was used to estimate VIT. Stimuli were presented on a video monitor at a viewing distance of approximately 60 cm. Preceding the target figure was a warning cue of approximately 520 ms; the cue was a small white plus (+) sign measuring 6 x 6 mm, presented in the centre of the computer screen. The target figure consisted of two vertical lines; one measured 15 mm and the other 30 mm. These were joined at the top by a horizontal line of approximately 18 mm. A 'flash mask' (see Evans & Nettelbeck, 1993) of 375 ms immediately replaced the target figure and consisted of two vertical lines 35 mm in length, shaped as lightning bolts. The shorter line appeared on either side of the target figure equiprobably. Participants indicated on which side the short line appeared by clicking either the left or right mouse button, respectively.

Procedure

Upon arriving at the testing session participants were seated in a cubicle in a quiet laboratory and they were guided through the test battery automatically by the computer. Detailed instructions and practice phases were presented prior to the onset of each task, and the first author was present to answer any questions. Participants completed the tasks in the following order: APM; VRT; VRT_2; VRT_3; ART; ART_2; ART_3; CD_A; $CD_{V;}$ SD; AC; HA_S; FA_S; TC; SC; and VIT. Simple and two-choice RT tasks consisted of 32 trials whilst the three-choice tasks consisted of 36 trials.

For the VIT task, the instructions emphasised accuracy rather than speed of responding. Practice trials required 10 correct trials out of 10 with SOA of approximately 835 ms; 10 correct trials out of 10 with SOA approximately 420 ms; and nine correct trials out of 10 with SOA approximately 250 ms. The estimation process began with SOA approximately 250 ms and followed an adaptive staircase algorithm (Wetherill & Levitt, 1965). The average SOA was calculated over eight reversals of direction on the staircase, giving an estimate of the SOA with an associated probability of 79% of making a correct response

Results

Data preparation

Outliers (±3 SD) and errors were removed from individual RT data files and the average number of trials used to calculate RT scores was: ART (M=31.2, SD=1.0); ART_2 (M=31.2, SD=0.9); ART_3 (M=34.5, SD=1.3); VRT (M=31.2, SD=0.6); VRT_2 (M=31.4, SD=0.6); and VRT_3 (M=35.0, SD=0.9). Following this the data for all tasks were collated and assessed for outliers (i.e., ±3 SD) and missing values. No missing values were present, but 7 outliers were found randomly distributed across six variables. These values were replaced using the Expectation Maximization (EM) imputation procedure in SPSS v.15.

Table 5: Descriptive statistics and split-half reliability estimates.

	M	*SD*	Split-Half
ART	274.0	62.2	.88
VRT	304.3	46.1	.89
ART_2	368.0	73.3	.90
VRT_2	360.6	45.0	.90
ART_3	505.0	91.1	.90
VRT_3	424.5	52.5	.81
VIT	46.9	11.2	-
APM	6.0	2.5	.42
CD_A	65.1	4.4	.71
CD_V	89.0	6.6	.92
FA_S	55.7	5.9	.73
HA_S	43.5	6.7	.84
NC	24.7	4.2	.67
TC	53.1	4.7	.81
AC	68.7	10.4	.79
SD	92.8	15.5	.80

RT & IT = Msec; All else = N correct

Descriptive statistics and correlations

Table 5 presents descriptive statistics for all measures and split half reliability estimates. As can be seen, reliabilities were generally acceptable except for APM. Regarding test performance, it is interesting to note that simple ARTs are faster than for VRT [$t(79)=5.99$, $p<.001$] yet this pattern alters as the number of alternatives increases, with VRT_3 performance becoming better than for ART_3 [$t(79)=9.99$, $p<.001$]. Furthermore, for the Gs tasks, the number of completed items is generally higher for the visually presented tasks except for NC, where the average is lower than for TC.

Presented in Table 6 are the correlations between the tasks. As expected, the RT tasks relate moderate to strong with each other, suggesting they all tap a similar construct. The Gs measures correlate well with one another also and the auditory measures correlate well with their visual analogues except for NC and TC, which do not correlate significantly. This may partly reflect differences in the complexities of these tasks, which would also explain the marked difference between the average numbers of items completed in each, as already noted.

Exploratory Structural Equation Modeling

In order to explore the latent factors underpinning performance across the speed tasks, we generated eigenvalues and a scree plot using principal components analysis. There were four components with eigenvalues greater than 1, accounting for 33.3%, 18.5%, 8.4% and 7.1% of the variance, respectively. Inspection of the scree plot suggested, however, that there were two dominant factors and possibly a third, with the scree commencing after this. Therefore, we performed exploratory structural equation modeling using maximum likelihood (ML) estimation in MPlus version 5.0 (Muthen & Muthen, 2007). We modeled three related factors and allowed the tasks to

Table 6: Correlations between cognitive tests

	ART	VRT	ART2	VRT2	ART3	VRT3	VIT	APM	CDA	CDV	FAS	HAS	NC	TC	AC
VRT	.69**														
ART_2	.52**	.32**													
VRT_2	.43**	.65**	.53**												
ART_3	.46**	.28*	.77**	.38**											
VRT_3	.25*	.37**	.49**	.61**	.61**										
VIT	-.19	.01	-.03	.10	.01	.26*									
APM	.13	-.07	-.05	-.13	-0.20	-.22*	-.20								
CD_A	-.06	-.21	-.27*	-.26*	-.34**	-.49**	-.21	.22							
CD_V	-.09	-.15	-.37**	-.34**	-.43**	-.59**	-.30**	.26*	.68**						
FA_S	.09	-.10	-.12	-.07	-.22	-.30**	-.23*	.25*	.52**	.49**					
HA_S	-.03	-.05	-.16	-.08	-.32**	-.30**	-.25*	.06	.20	.16	.43**				
NC	.18	-.06	.07	-.08	-.05	-.20	-.36**	.45**	.40**	.32**	.53**	.10			
TC	-.02	-.21	-.28*	-.41**	-.37**	-.54**	-.27*	.20	.33**	.35**	.15	.38**	.18		
AC	.11	.04	-.11	-.14	-.14	-.40**	-.16	.16	.35**	.32**	.33**	.24*	.34**	.40**	
SD	.10	-.11	-.05	-.17	-.06	-.33**	-.20	.41**	.41**	.45**	.40**	.06	.50**	.21	.44**

**. Correlation is significant at the 0.01 level (2-tailed). *. Correlation is significant at the 0.05 level (2-tailed).

load freely across them. The model terminated but the residual covariance matrix was not positive definite due to VRT having a loading of r=1.30 on a factor that was otherwise unrelated to the remaining visual and auditory tasks (average loading of r=.07). Therefore, we decided to remove VRT from subsequent models as well as ART, given it was its auditory analogue. We also reduced the number of latent factors to two given the third was clearly defined by VRT only.

Following these modifications the model was re-estimated and it terminated successfully. The fit of the model was considered quite acceptable given that no paths were constrained [$\chi^2(53)=115.66$, $p<.001$; CFI=.84; RMSEA=.12; SRMR=.06], resulting in a large number of near zero factor loadings. The first factor was identified as an RT factor having strong loadings from both visual and auditory RT tasks, and weak but significant loadings from CD_A, CD_V and TC. The second factor was interpreted as a Gs factor, defined by generally moderate loadings from the other tasks, and with a weak loading from VRT_3. The correlation between the latent factors was weak but significant (r=-.31, p=.01).

We modified this two factor model by constraining the non-significant paths to zero. Then, based on statistical significance of loadings and modification indices the model went through several permutations to achieve adequate fit. The final model is presented as Figure 5 and as can be seen the two factors of RT and Gs remain distinct and the fit of the model is adequate [$\chi^2(60)=93.44$, p=.003; CFI=.92; RMSEA=.08; SRMR=.08]. It was possible to achieve better fit through correlating the residuals of RT measures and other auditory / visual pairs (e.g., SD and AC) but in the interest of parsimony we chose this more restrictive model. The one discrepancy noted is that although TC was envisaged as a marker of Gs, it was better placed as an indicator of the RT factor as it shared more variance with this latent construct.

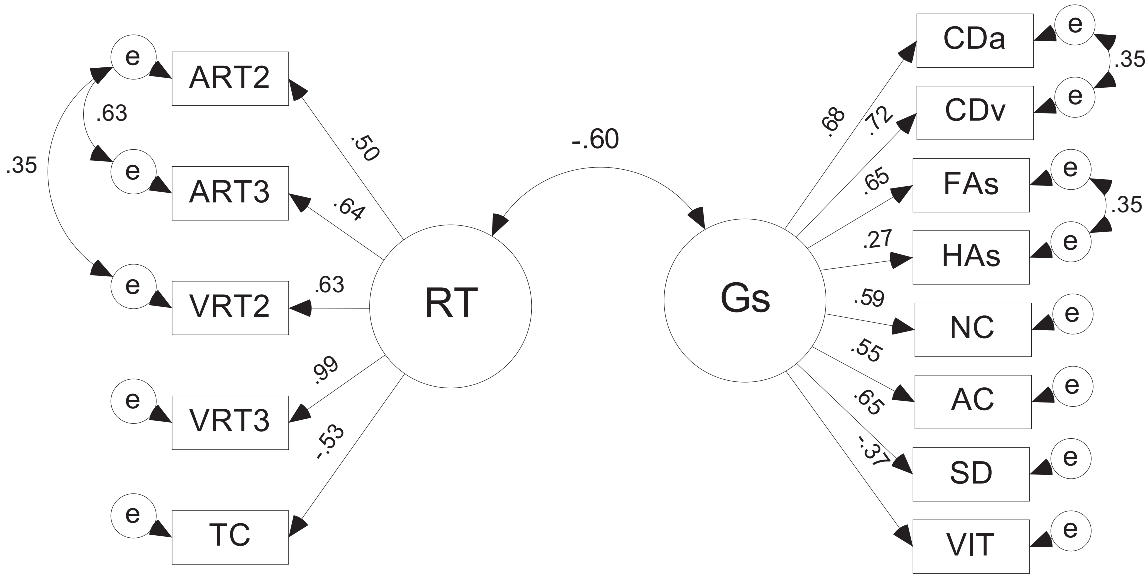


Figure 5. Latent structure of auditory and visual speed tasks.

Modality Specific Speed Factors

Although the factors resulting from ESEM herein conform well to those found in our earlier study (Zajac, et al., 2009), we chose to specifically test the possibility of modality specific speed factors using confirmatory modeling. We modified the model shown in Figure 5 to include four latent variables: auditory RT, visual RT, auditory Gs and visual Gs. To ensure sufficient marker tests were available to define the latent RT constructs, simple ART and VRT were reintroduced in this model. We also allowed TC to define auditory Gs rather than auditory RT given it was developed to be an auditory analogue of NC.

This model terminated normally [$\chi^2(84)=238.14$, $p<.001$; CFI=.69; RMSEA=.15; SRMR=.11]. However, fit was poor due to the latent variable residual covariance matrix not being positive definite. This issue was traced to a linear dependency between visual Gs and auditory Gs ($r=1.09$). Therefore, we modified the model to include only modality specific RT factors, and a single latent Gs defined by tasks of both modalities. This subsequent model terminated successfully and all tasks loaded moderately and significantly on their respective factors. Moreover, the relationship between latent factors RT_A and RT_V was significant and strong ($r=.74$), and the relationship of Gs to each of these factors was moderate and significant also ($r=.42$ and $r=.67$, respectively). The fit of this model, however, was again not adequate [$\chi^2(87)=242.99$, $p<.001$; CFI=.69; RMSEA=.15; SRMR=.11]. Short of correlating the residuals of nearly all indicator variables, we could not get this model to fit adequately. Thus, these analyses support the model which resulted from exploratory analyses shown in Figure 5 because adequate fit is only obtained when latent auditory and visual RT factors are merged.

Construct Validity of Latent Factors

Our final analysis involved examining whether the addition of auditory tasks to the measurement of Gs altered the extent to which Gs predicted a criterion measure. To accomplish this, we used exploratory factor analysis and generated: 1) visual Gs scores (Gs_V), which only included the visual Gs tasks; 2) visual/auditory Gs scores (Gs_{VA}) by adding auditory Gs measures (CD_A, HA_S, TC, AC); 3) visual RT scores (RT_V); and 4) visual/auditory RT scores (RT_{VA}). In each instance we employed ML estimation and extracted a single general factor.

APM was regressed onto each of these factor scores independently. RT_V did not explain any APM variance [$F(1,78)=1.25$, $p=.266$; $R=.13$; $R^2=.02$] and neither did

RT_{VA} [$F(1,79)=1.25$, $p=.267$; $R=.12$; $R^2=.02$], with the models being essentially identical. On the other hand, Gs_V predicted 20% of APM variance [$F(1,78)=19.66$, $p<.001$; $R=.45$] and Gs_{VA} explained 15% of APM variance [$F(1,78)=13.38$, $p<.001$; $R=.38$]. The difference between the two regression coefficients was not significant ($z=.44$, $p=.65$). To assess whether the same variance was being accounted for, we performed stepwise regression. Gs_V was entered on the first step, after which the model terminated. Results were as per the Gs_V only model above. The model did not include Gs_{VA} as it could not explain any additional variance. It appears then, that the addition of auditory tasks to the measurement of latent RT and Gs does not alter these constructs in any way. Both factors appear to be adequately measured with the use of visual tasks only.

Discussion

The purpose of the present study was to replicate and extend the findings of our earlier research (Zajac, et al., 2009) concerning whether broad Gs could be measured auditorily. The purpose-developed auditory tasks were employed again and in line with our previous findings we expected to find relationships between the visual and auditory Gs tasks. Furthermore, by including more visually presented tasks in this study we sought to establish whether their absence in Zajac et al. (2009) precluded the identification of modality specific speed factors.

Overall, findings from the present study support those of our earlier investigation. The auditory tasks generally correlated well with their visual analogues except for Tone Comparisons (TC) and Number Comparisons (NC), which were not significantly correlated. It is probable that this finding reflects the marked differences in the difficulty of each of these tasks. Whilst TC involves a simple comparison of two successive stimuli (single tones of different frequency) the stimuli in NC are complex, involving up to 12 digits in each number string. Thus, comparing the two digit strings on any given trial is arguably more complex than comparing two simple tones. It would be interesting to correlate performance on TC and NC inclusive only of the simpler three-digit strings. Unfortunately, given the format of our test our data do not allow this.

The inclusion of more visual marker tests in this study has answered an important question. Specifically, we were concerned that the dominance of auditory tasks in the previous study (Zajac, et al., 2009) may have prevented the appearance of distinct auditory and visual speed factors. In the present study we undertook exploratory structural equation modeling in order to allow the latent factors to be naturally defined. Despite the differences in analytical techniques between our previous study and the present one, the models generated are congruent. Distinct Gs and RT factors again emerged and they were moderately related, sharing about 36% of their variance.

The inclusion of more visual marker tests allowed us to explore whether specific modality speed factors could be extracted from the data. In the case of Gs, there appears rather convincing evidence that it does not distinguish between the auditory and visual modalities. If this were the case, then modality factors should have emerged naturally in the exploratory analysis used. Forcing the tasks to define modality specific Gs factors using confirmatory modeling provides further support that they are not representative of the underlying ability. The correlation between latent visual and auditory Gs was r=1.09 demonstrating a perfect linear dependence. Evidence for a single latent RT factor is less convincing because it was possible to define distinct auditory and visual RT factors. However, it was not possible to achieve adequate fit statistics for this model without correlating most residuals – with many resulting correlations not being theoretically sensible – and the resulting model would certainly not have been parsimonious. Moreover, the correlation between the modality specific factors was high (r=.74) and they shared approximately 55% of their variance. It can be concluded at this time that it is possible that there exists distinct auditory and visual RT factors but that our two datasets are not supportive of this hypothesis.

Both the study reported herein and our previous (Zajac, et al., 2009) investigation fill an important gap in intelligence research. We have explicitly sought to test Stankov's (1994a) hypothesis that higher-order thinking abilities, otherwise referred to as broad cognitive abilities, can be indexed regardless of stimulus modality. Until now, this hypothesis had not been deliberately tested from an auditory perspective. However, previous work on tactile-kinesthesia (Roberts, et al., 1997; Stankov, et al., 2001) has supported this hypothesis with some tactile tasks consistently relating to performance on similar visually presented tasks. Elsewhere auditory research has identified consistent relationships between auditory speech and verbal comprehension measures and visual tasks that also rely on knowledge of language and its structure (see e.g., Horn & Stankov, 1982). Our research further supports this idea. Specifically, broad Gs can be measured via both visual and auditory pathways and this has been replicated twice-over. Additionally, despite allowing for modality specific speed factors in our analysis and study design, we do not find this to be the case.

Research on Olfactory abilities and on Auditory Inspection Time has not seen the same success. More specifically, two tests used by Danthiir et al. (2001) were hypothesized to load on short-term memory (SAR) and long-term memory (TSR) but they did not exhibit these relationships. Similarly, of three AIT tasks developed, none display a sufficiently stable or strong enough relationship to warrant the conclusion that they measure the same as does VIT (Zajac & Burns, 2007, 2011). A possible reason why the olfactory tasks used by Danthiir et al. (2001) did not index the intended constructs reflects that only one task was devised for each of the noted constructs. Furthermore, it is plausible that these olfactory tasks placed increased emphasis on perceptual information and, in line with Carroll's (1993) conception of auditory abilities encompassing tasks requiring careful discrimination of auditory content, it is not surprising that they cluster together to define a latent olfactory-specific factor. In analogy to the auditory hierarchy noted earlier, these olfactory tasks are likely to be confounded by processes at the sensory and perceptual levels.

In the case of AIT and quite possibly the olfactory measure also, it might be hypothesised that the tasks have not adequately maintained the cognitive requirements of the constructs they seek to measure. This is a most important facet of the idea that general cognitive abilities are not modality specific. Obviously, in order to measure a specific ability it must be engaged in task performance. Weak and/or non-existent correlations are the first indication that such errors have occurred. Our studies depart from the former in this regard as the auditory tasks were carefully and purposely designed to maintain the cognitive requirements of existing visual Gs tasks and have shown moderate relationships with latent speed factors on two occasions, both in exploratory and confirmatory frameworks.

As Roberts et al. (1997) have noted, an implicit assumption of modern intelligence theories is that no knowledge of importance can be gained through employing tests of complex abilities which utilise alternate modalities. Indeed this might be the case, and we have found no evidence that latent Gs as measured by visual and auditory tasks respectively, differs. However, regardless of whether more information is gained by broadening intelligence models and subsequently intelligence tests to encompass measurement of complex abilities via alternate modalities, in doing so the measurement itself becomes more ecologically valid. People touch, smell, listen *and* visualise their environment. The persistent neglect of alternate modalities needs to be overcome and intelligence measures should be broadened to include multiple modalities. Only when this happens can we claim to have achieved 'truly balanced' measures of intelligence (Horn, 1968).

Chapter 5: Study 3 – Auditory Temporal Discrimination Tasks

DO AUDITORY TEMPORAL DISCRIMINATION TASKS MEASURE

TEMPORAL RESOLUTION OF THE CNS?

Ian T. Zajac and Nicholas R. Burns

School of Psychology, University of Adelaide, Australia

Article in Press

Psychology

Statement of Contributions

Ian Zajac (Candidate)

I was responsible for the conception and primary authorship of the paper. I was responsible for the development and programming of software based assessment tools and collection of all data. I conducted the statistical analyses independently with advice from the co-author. I am corresponding author and primarily responsible for responses to reviewers and revisions to the paper.

Nicholas R. Burns (Co-author)

I was the supervisor (advisor) for the research programme that led to this publication. In terms of conceptualisation of the programme, there was extensive and

ongoing collaboration between Mr. Zajac and me in developing the direction of the research. The realisation of the study and the programming of assessment tools used, were the work of Mr. Zajac. I had an advisory role with respect to selection of the test battery used and on the direction and specifics of the data analyses.

Mr. Zajac was responsible for writing this paper; my role was to comment on drafts, make suggestions on the presentation of material in the paper, and to provide editorial input. I also provided advice on responding to comments by the journal reviewers and editor.

I hereby give my permission for this paper to be incorporated in Mr. Zajac's submission for the degree of PhD in the University of Adelaide.

Abstract

Rammsayer & Brandler (2002) have proposed that auditory temporal discrimination tasks provide a measure of temporal resolution of the CNS which is argued to be partly responsible for higher order cognitive functioning. We report on two studies designed to elicit the nature of the functions underpinning these auditory tasks. Study 1 assessed whether temporal generalisation (TG) might be better considered as a measure of working memory rather than of temporal resolution of the CNS. In N=66 undergraduates TG did not predict speed of processing tasks; however, there was evidence of a relationship between TG and working memory. Study 2 reanalysed previously published data on temporal discrimination tasks and showed that the relationship between auditory temporal tasks and intelligence reflects memory functions and processing speed. Auditory temporal discrimination tasks are confounded by speed and memory and caution should be taken when interpreting them as measures of temporal resolution of the CNS.

Introduction

The last few decades have seen a shift in focus from the taxonomic study of cognitive abilities to the identification of lower-order cognitive and physiological correlates of human intelligence (Neubauer & Fink, 2005). This shift has been driven by a desire to identify the biological roots of higher order cognition (Stankov, 2005). The exploration of biological correlates of intelligence has been aided by advances in the measurement of brain activity. Studies employing electroencephalograms, for instance, have reported that peripheral nerve conduction velocity and event related potentials share variance with cognitive ability measures (Burns, Nettelbeck, & Cooper, 2000; Reed & Jensen, 1993). On the other hand, in order to measure lower-order cognitive processes, researchers have turned to a class of tasks termed Elementary Cognitive Tasks (ECTs). The impetus for this is that ECTs are characteristically easy tasks which putatively rely on a limited number of mental processes or operations (Carroll, 1993). Thus, they supposedly provide cleaner measures of biological processes than traditional, more complex tests (Stankov, 2005).

The two most commonly researched ECTs are reaction time (RT) and inspection time (IT): RT tasks measure the speed with which an individual is able to respond to a particular reaction stimulus; and IT tasks measure the minimum exposure duration required to accurately discriminate stimuli that differ on some dimension. Both classes of tasks are held to reflect information processing speed (Jensen, 2005). However, it has been found that performance in these tasks is relatively independent. The correlation between these ECTs is seldom more than r=.30, with the strength of the correlation appearing to increase as complexity of the RT task increases (Burns &

Nettelbeck, 2003; O'Connor & Burns, 2003; Petrill, Luo, Thompson, & Detterman, 2001).

Despite their relative independence, IT and RT tasks have been found to share a statistically significant amount of variance with measures of psychometric intelligence. People with higher speed of information processing – faster average RTs and shorter ITs – perform better on tests of cognitive ability than those who are slower. It has been proposed that RT and IT could account for as much as 25% of the variance in intelligence test performance (Grudnik & Kranzler, 2001; Jensen, 1982, 2005, 2006; Nettelbeck, 1987, 2001, 2003). However, a more recent meta analysis which based its conclusions on 1146 correlations between speed of processing measures and intelligence measures proposes a much smaller effect: around 10% shared variance for RT and intelligence and about 8% between IT and intelligence (Sheppard & Vernon, 2008).

Regardless of the size of these effects, and returning to the idea of identifying the biological basis of intellectual functioning, it is necessary to explain the observed relationship between intelligence and performance in ECTs. Many explanatory models appeal to the concept of "neural efficiency" as the determinant of both information processing speed and intelligence (see e.g., A. E. Hendrickson, 1982; D. E. Hendrickson, 1982; Vernon, 1993). Jensen's (1982) model of neural oscillations, for example, proceeds from the assumption that RT provides an index of the efficiency of the central nervous system (CNS). Individual differences in both processing speed – as measured by ECTs – and intellectual functioning are attributed to differences in the rate of oscillation between refractory and excitatory states of neurons. The transmission of neurally encoded information is assumed to be more efficient as well as faster at a higher rate of neural oscillations. This is because it takes less time for a neuron to re-enter its excitatory phase when processing information than when oscillations are slow.

An alternative theory linking higher-order cognitive processes to elementary functions has recently been revisited by Rammsayer and others (Helmbold & Rammsayer, 2006; Helmbold, et al., 2006, 2007; Rammsayer & Brandler, 2002, 2004, 2007). Originally proposed by Surwillo (1968), this theory also appeals to a hypothetical oscillatory, or 'clock', mechanism in the CNS to explain individual differences in speed of information processing and intelligence. Thus, "if the hypothesised internal master clock of individual A works at half the clock rate as the one of individual B, then A does not only need twice as long as B to perform a specific sequence of mental operations, but also the occurrence probability of interfering incidents will be increased" (Rammsayer & Brandler, 2007, p.124). According to the theory this results in both slower performance on speed of processing tasks and lower intelligence. The central features of the internal clock mechanism are a pacemaker and an accumulator (Rammsayer & Brandler, 2002).

Jensen (2006) cites findings from ECT research to support his neural oscillation model. In order to obtain empirical support for the master clock theory, Rammsayer and colleagues have sought to demonstrate that presumed measures of clock rate differ between individuals of low and high intelligence (Helmbold & Rammsayer, 2006; Helmbold, et al., 2006, 2007; Rammsayer & Brandler, 2002, 2004, 2007). They have argued that accuracy on psychophysical timing tasks – by analogy with performances on ECTs –reflects basic processes related to neural efficiency (Helmbold, et al., 2007). According to this theory the number of neural oscillations generated by the pacemaker during a timed interval is recorded by the accumulator and becomes the internal representation of that interval. Thus, the higher the frequency of oscillations, the finer the temporal resolution.

Because audition has finer temporal resolution than vision (Rammsayer & Brandler, 2002), attempts to measure temporal resolution of the CNS have focussed on auditory tasks. Rammsayer and Brandler (2002) found that auditory duration discrimination was significantly better for a high-IQ group than for a low-IQ group and that it explained around 20% of the total variance of a single fluid intelligence (Gf) measure. A later study concluded that a general pacemaker based interval timing mechanism is involved in auditory temporal order judgement, duration discrimination, and temporal generalisation and that performances on these tasks is independent of general auditory discrimination ability (Rammsayer & Brandler, 2004). Factor scores on this general timing (Gt) mechanism have subsequently been shown to share substantial variance (about 25%) with psychometric '*g'* (Rammsayer & Brandler, 2007).

Whether Gt solely reflects temporal resolution is arguable and Helmbold et al. (2006) have explored whether sensory discrimination abilities rather than temporal resolution of the CNS account for the relationship between Gt and intelligence. Temporal generalisation and pitch discrimination performance was measured and regression analyses showed these tasks combined to predict 25% of the variance in *g* factor scores. The unique contributions of temporal and pitch tasks were 9% and 6%, respectively. The shared and presumably sensory processes accounted for the remaining 10% of predicted *g*. Helmbold et al. (2006) concluded that the unique contribution of temporal discrimination to the prediction of *g* supports the notion that it measures specific aspects of neuronal information processing related to intellectual capacity but independent of non-temporal aspects of sensory discrimination.

As noted, it has been proposed that auditory psychophysical timing tasks are analogous to existing ECTs in terms of measuring basic processes related to neural efficiency (Helmbold, et al., 2007). Therefore, one should expect these tasks to correlate at least moderately with existing ECTs, including RT and IT, but evidence regarding this hypothesis is equivocal (Helmbold, et al., 2007; Rammsayer & Brandler, 2007). Although the correlation between latent RT and latent temporal discrimination factors appears moderately strong (r=.65: Helmbold, et al., 2007), the correlation between individual temporal tasks and RT parameters is markedly weaker and in many cases not statistically significant. The average correlation between eight temporal tasks and different RT parameters in Helmbold and Rammsayer (2006) was only r = -.19 (SD_r=.07). This absence of notable and statistically significant correlations between temporal tasks and RT suggests that temporal discrimination tasks may not be measuring the elementary speed processes reflected in RT tasks.

Nonetheless, these findings have been interpreted as providing evidence that auditory temporal discrimination tasks index neural efficiency. There are, however, several issues with these tasks which question whether the observed correlation between temporal performance and intelligence is a result of neural efficiency and, by extension, temporal resolution of the CNS, as opposed to other more complex cognitive processes.

First, although it has been supposed that these tasks are elementary (Rammsayer & Brandler, 2004) some do not appear to be. The Temporal Generalisation (TG) task appears rather more complex than archetypal ECTs. TG requires participants to judge whether a test stimulus is the same as a standard stimulus learnt in a pre-exposure phase. Thus, the task requires: 1) accurate learning of the standard stimulus; 2) accurate registration of the test stimulus; 3) accurate retrieval of the learnt standard; and 4) a successful comparison of the test and learnt standard, in order to complete each test item. It appears that the cognitive operations required in this test are complex, and even if temporal resolution of the brain itself is independent of higher order cognitive operations (Rammsayer & Brandler, 2002, p509), performance on this task is not likely to be. It is plausible that the observed relationship between TG and intelligence reflects the shared cognitive operations common to TG and general intelligence tests, rather than temporal resolution of the brain.

This hypothesis applies to other discrimination tasks used in these studies. Duration Discrimination (DD), for example, requires participants to compare two successively presented time intervals to decide which was longer. Thus, an internal representation of each interval must be formed and, given the length of the intervals – 1 sec or longer in one condition – and the ISI (900 ms), these representations need to be accessible for up to three seconds after presentation; at least for the first presented interval. Unless the accumulator in the master clock theory incorporates an information storage component, performance on this task is also likely to rely on more complex cognitive functions. Considering the requirements of both TG and DD, it appears that the cognitive operations involved may reflect memory functions.

A second issue regarding these findings is that most have focused on factor scores. As such, little information is gleaned in terms of the relationship between specific temporal discrimination tests and intelligence measures. Moreover, the nature of the latent construct defined by the temporal tasks is merely surmised based on theory of what the tests have in common. In order to accurately assess whether internal clock rate – or temporal resolution of the CNS – is related to intellectual capacity, it must first be established that each of the individual tests used provides some measure of this. Only then can a latent variable defined by these tests be taken to represent the internal clock rate.

This paper presents the findings of two studies on temporal discrimination tasks. The impetus for these studies was to investigate whether temporal discrimination tasks provide a measure of elementary functions such as temporal resolution of the CNS, or whether they might better be conceptualized as measures of more complex cognitive operations like memory functions.

Study 1

The Temporal Generalisation (TG) task has been shown to relate to *g* but its relationship with specific cognitive abilities and ECTs has not been considered. Thus, the relationship may reflect executive cognitive functions utilised in task performance and not neural efficiency, as proposed. The purpose of this study was to provide a test of this hypothesis by exploring the relationship between TG and measures of processing speed (Gs) and working memory (WM). Importantly, speed of processing was measured by traditional speed tasks and ECTs, including RT and IT, because of the considerable evidence that these ECTs are reliable measures of elementary functions (Jensen, 2006; Nettelbeck, 2001). If TG measures elementary functions as opposed to executive cognitive functions, then the relationship between TG and Gs will be stronger than that between TG and WM.

We used the same dissociation paradigm as Helmbold et al. (2006). The purpose was to assess the direct relationship of TG to Gs after partialling out variance due to general sensory discrimination processes; reflected in the pitch discrimination task (AP_d). If TG measures elementary processes related to intelligence and which are independent of general sensory discrimination, then TG should make a direct contribution to the prediction of intelligence test performance.

The superior temporal discrimination in audition has been the motivation for the use of auditory tasks. However, if the master clock which determines performance on these tasks is a general feature of the neural system, it should also be responsible for temporal discrimination in other modalities. Therefore, the current study sought also to measure temporal resolution in the visual modality. This was achieved through adapting the dissociation paradigm for the visual system to include a temporal and a line-length discrimination task. Line-length discrimination ability would be measured to assess variance in visual temporal performance reflecting general sensory processing. The correlation between the visual and auditory temporal discrimination tasks should be at least moderately strong if they reflect the same elementary timing processes.

Rammsayer and Brandler (2002) have reported that temporal resolution of the CNS is independent of cognitive operations. We tested this assumption by introducing a backward masking condition for the discrimination tasks. Masking has previously been used in visual and auditory modalities to investigate temporal processes underpinning perception, and which operate at a precognitive level (Breitmeyer, 2007). If temporal tasks measure temporal resolution of the CNS then their relationship with cognitive ability measures should not be negatively affected by the introduction of a masking stimulus; because it emphasises pre-cognitive functions. In fact, the strength of the relationships might be expected to increase.

Methods

Participants

Participants were *N*=66 undergraduate students of the University of Adelaide, South Australia. There were 7 males and 26 females in each of the masked and unmasked conditions. All participated as part of their Level I Psychology course requirements.

Apparatus

The presentation of all tasks and recording of responses was controlled by one of two identical computers. Visual stimuli were presented on 17 inch LCDs. Auditory stimuli were presented via Sony MDR-XD100 stereo headphones. Auditory tones were calibrated prior to the study using a Radio Shack 33-4050 Sound Level Meter.

Discrimination Tasks

Auditory and visual discrimination abilities were assessed using the experimental dissociation paradigm developed by Gibbons, Brandler, & Rammsayer (2002); stimuli varied on two dimensions simultaneously. The first dimension was temporal: there were seven levels of stimulus duration. The second dimension for the auditory modality was pitch and for the visual modality was line-length; there were seven levels of each (see Appendix 1, p.219). Line length dimensions were piloted on a small number of colleagues to be at a comparable level of difficulty to the duration levels.

The design of the set of stimuli for the dissociation paradigm is based on the requirement that: 1) for duration, as well as pitch/line length, there should be a probability for the standard stimulus of .33 in the total number of trials; 2) within each level of one stimulus dimension, each level of the other dimension should be represented; and 3) for each of the seven levels of one stimulus dimension, there should be a probability of .33 for the occurrence of the standard of the other stimulus dimension. Simultaneous variation on two dimensions according to these requirements results in a set of 81 stimuli for each of the visual and auditory tasks, resulting in the frequency distribution presented in Appendix A. The test phase for each of the discrimination tasks comprised 81 trials, including 27 presentations of the standard and nine presentations of each nonstandard stimulus. Presentation order within each task was pseudo-randomised, with the restriction that there were no more than two successive presentations of the standard. The outcome measure for each of the discrimination tasks was percentage of standard stimuli correctly identified.

Auditory Temporal and Pitch Discrimination Tasks. In each task, participants were required to identify the standard tone among the set of nonstandard tones. Participants were instructed to attend solely to tone duration in the temporal task, and to tone frequency in the pitch task. All tones were presented at an intensity of 67 db. Each task was preceded by a learning phase in which participants were asked to learn the standard tone. For the temporal task, a standard tone duration (i.e., 200 ms) with a pitch (900 Hz) not administered during the test phase was presented five times. For the pitch task, the learning phase consisted of five presentations of the standard tone (i.e., 1000 Hz) for 260 ms, a duration which was not included in the test period. The testing phase immediately followed and the onset of each trial was marked by the presentation of a visual fixation point (small white cross) in the centre of the computer screen. After a foreperiod of 1000 ms the trial stimulus was presented and the cross remained on the screen. In the masking condition, a burst of white noise immediately followed the trial stimulus for 500 ms, otherwise the trial terminated. Following each trial the participant mouse-clicked one of the onscreen buttons ("standard" or "nonstandard") to indicate whether they thought the trial stimulus matched the frequency or duration of the standard tone, depending upon which task was being completed. Feedback was given for each trial in the form of a “correct” or “incorrect” on-screen message which was displayed for 500 ms. Subsequent trials commenced immediately after the feedback.

Visual Temporal and Line-Length Discrimination Tasks. The requirements of these tasks were similar to the auditory tasks. White horizontal lines presented against a black computer screen were used analogously to tones in the auditory tasks. For the temporal task, participants were asked to attend solely to stimulus duration whilst in the line-length task they were asked to attend solely to the length of the line. The learning phase for the temporal task consisted of five presentations of the standard duration (i.e., 200 ms) with a line length (6 cm) not administered during the test period and for the line-length task, consisted of five presentations of a standard 10 cm line for a duration (260 ms) not included in the test. The testing phase immediately followed and each trial was marked by the onset of a visual fixation point (small white cross) in the centre of the computer screen. After a foreperiod of 1000 ms the visual fixation point was replaced by the trial stimulus. In the masking condition, a 4 x 8 grid of 16 cm wide by 6 cm high lines immediately followed the trial stimulus for 500 ms (see Figure 6) otherwise the trial terminated. The response format was the same as for the auditory tasks with participants indicating whether they thought the

Standard stimulus Masking stimulus

Figure 6. Target and masking stimuli used in the visual discrimination tasks.

test stimulus matched the duration or line-length of the standard, depending upon the task. Feedback was given after each trial.

Working Memory

Dot Matrix Test (DM). A computer-administered version of the Dot Matrix Task (Law et al., 1995) was used as a measure of working memory (WM). Participants verified a series of simple matrix equations whilst simultaneously remembering the locations of dots on a 5 X 5 grid. Matrix equations were either addition or subtraction equations presented as lines drawn on 3 X 3 dot matrices. Participants verified each equation by mouse-clicking either the “True” or “False” buttons displayed on the screen within 10 seconds, otherwise they received a prompt (“response required”). Following an incorrect response the message “No, look again closely” was displayed, and the equation remained until a correct was response was given.

Following correct responses a 5 X 5 grid was displayed for 1500 ms with a dot presented in one of the squares. There were four levels during the test (2, 3, 4, and 5 equation-grid pairs each with 4 items) and this equation-grid sequence was repeated according to the level. At the end of each equation-grid sequence, a blank 5 X 5 grid was displayed on the screen. Participants were required to mouse-click the spaces on the blank grid which had contained the dots during the trial sequence. Participants could not select more grid spaces than there were equation-grid pairs but they could select fewer grid spaces (e.g., 3 of 5 dot locations). An ‘enter’ button was clicked after locations were selected. Three practice questions consisting of two equation-grid pairs preceded the test. The measure for the task was the number of dot positions, out of a total of 56, correctly recalled.

Speed of Processing.

Symbol Digit (SD). A computerised coding task was employed as a measure of Gs (see McPherson & Burns, 2005, for a detailed description of this task). A code table was presented at the top of the computer screen throughout the task. This comprised of nine symbols arranged horizontally, to which nine digits were paired; digits were presented directly beneath the symbols so that they were aligned. For each item, one symbol was presented in the centre of the computer screen and participants responded by left clicking the mouse on its corresponding digit in a 3 x 3 numerical grid positioned at the bottom of the screen. Subsequent items did not commence until a correct response was registered. Participants were required to complete two practice trials correctly before they proceeded to the test. The outcome measure was the number of items correctly completed in 120 seconds.

Audio Code (AC). This task was developed in our laboratory to be an auditory analogue of the symbol digit task described above. It has good reliability (r=.89) and correlates well with other speed measures (Zajac, et al., 2009). In this task, a code table is displayed at the top of the computer screen for the duration of the task. This comprised of pictures of eight musical instruments arranged horizontally, to which one of the numbers one through eight was paired; the instruments were a snare, trumpet, guitar, cymbals, piano, bell, harp and violin. For each item, the sound of one of the instruments was presented via headphones at an intensity of 65 db. Participants responded by left clicking the mouse on its corresponding digit in a 2 x 4 numerical response grid positioned at the bottom of the screen. Subsequent items commenced after a response was registered. Participants were required to complete four practice trials correctly before they proceeded to the test. The outcome measure was the number of items correctly completed in 120 seconds.

Visual Inspection Time (VIT). Stimuli were presented on a video monitor at a viewing distance of approximately 60 cm. Preceding the target figure was a warning cue of approximately 520 ms; the cue was a small white plus (+) sign measuring 6 x 6 mm, presented in the centre of the computer screen. The target figure consisted of two vertical lines; one measured 15 mm and the other 30 mm. These were joined at the top by a horizontal line of approximately 18 mm. A ‘flash mask’ (see Evans & Nettelbeck, 1993) of 375 ms immediately replaced the target figure and consisted of two vertical lines 35 mm in length, shaped as lightning bolts. The shorter line appeared on either side of the target figure equiprobably.

A computerised tutorial preceded the test phase and the instructions emphasised accuracy rather than speed of responding. What was required was explained using diagrams, along with unmasked target stimuli. Practice trials required 10 correct trials out of 10 with a stimulus onset asynchrony (SOA) of approximately 835 ms; 10 correct trials out of 10 with SOA approximately 420 ms; and nine correct trials out of 10 with SOA approximately 250 ms. The estimation process began with SOA approximately 250 ms and followed an adaptive staircase algorithm (Wetherill & Levitt, 1965). The algorithm required three correct responses at any SOA before reducing the SOA by approximately 17 ms. The average SOA was calculated over eight reversals of direction on the staircase, giving an estimate of the SOA with an associated probability of 79% of making a correct response. Participants indicated on which side the short line appeared by clicking either the left or right mouse button, respectively.

Auditory Reaction Time (ART). This task required participants to respond as quickly as possible to an auditory target stimulus. To start each trial, the participant pressed the number ‘5’ key in the numeric keypad on the computer keyboard. After

300 ms a short beep (100 ms at 880 Hz) was presented to confirm the trial had started. The target stimulus was then presented after a silent interval of variable duration (1300 ms, 1700 ms, 2100 ms, 2500 ms), and it was a 500 ms 'bell' sound centered on a frequency of 800 Hz. Participants were instructed to lift their finger off the number '5' key when they heard the target sound and press the number '8' key in the numeric keypad as fast and as accurately as possible. The test phase consisted of 32 trials before which participants had to complete five practice trials correctly. Mean RT was calculated after removing errors and outliers (±3 SD). The average number of trials remaining after these removals – and from which Mean RT was derived – was M=31.30 (SD=.63, Min=30, Max=32).

Procedure

Upon attending the testing session participants were assigned to either the masked or unmasked condition depending on whether they were an odd or even numbered participant. They were seated in a quiet room in the laboratory and were guided through the tasks by the computer. The four discrimination tasks (see below) were interspersed with cognitive ability measures, which were ordered as they are set out below. The discrimination tasks were ordered so as to switch between modalities (auditory pitch / visual length / auditory temporal / visual temporal), and the discrimination tasks were counterbalanced within conditions to reduce fatigue effects (visual length / auditory pitch / visual temporal / auditory temporal). The ordering of cognitive ability measures remained constant. The testing session took 60 minutes to complete.

Results

After collating the data it was apparent that two participants did not complete Dot Matrix (DM), one participant failed to complete Auditory Temporal discrimination (AT_d) and another participant failed to complete Symbol Digit (SD). These missing data were replaced using the Expectation Maximization (EM) method in Missing Values Analysis in SPSS v.15. Following this an outlier analysis was performed by standardizing scores on each variable. The only identified outlier was for Audio Code (AC; z = 3.14), which was deleted and subsequently replaced using EM.

Table 7: Descriptive statistics for discrimination tasks, cognitive measures, VIT and RT

		M[a]	SD	Min	Max	d[b]
VT_d	*Unmasked*	.58	.18	.07	.89	.34
	Masked	.52	.19	.15	.89	
AT_d	*Unmasked*	.68	.16	.26	.93	.53
	Masked	.58	.19	.11	.89	
VL_d	*Unmasked*	.81	.13	.52	1	.33
	Masked	.77	.11	.56	.96	
AP_d	*Unmasked*	.64	.15	.26	.89	.71
	Masked	.53	.16	.19	.81	
SD		90.8	16.1	64	133	
AC		63.8	7.9	49	82	
DM		38	6.3	19	51	
VIT (ms)		45.3	11.6	19.5	76.4	
ART (ms)		502.6	115.4	312.4	768.1	

VT_d = Visual Temporal Discrimination; VL_d = Visual Length Discrimination; AP_d = Auditory Pitch Discrimination; AT_d = Auditory Temporal Discrimination; SD = Symbol Digit; AC = Audio Code; DM = Dot matrix; VIT = Visual Inspection Time; ART = Auditory Reaction Time.
[a] Temporal tasks = percent correct; SD, AC & DM = *N* correct; VIT and RT = msec
[b] Cohen's *d*

Descriptive statistics for the cognitive measures and discrimination tasks are presented in Table 7. As can be seen, performance in the masked condition was poorer for all of the discrimination tasks, with small to large effects. The difference was only statistically significant for the auditory temporal discrimination task (AT_d [t(64) = 2.11, p = .038]) and auditory pitch discrimination task (AP_d [t(64) = 2.77, p=.007]).

Table 8 presents the correlations between the cognitive tests for the total sample and Table 9 displays the correlations between the discrimination tasks for the masked and un-masked conditions. As can be seen the correlations between the cognitive tests are small-to-moderate and the correlations between the discrimination tasks are moderate-to-strong. Of particular note is the correlation between AT_d and Visual Temporal (VT_d) discrimination. As hypothesized, the correlation between them is notably strong indicating that to a large degree these tasks index the same construct.

Table 8: Correlations between cognitive measures

	SD	AC	DM	VIT
AC	.47**			
DM	.25*	.34**		
VIT	-.17	-.15	-.24*	
ART	-.39**	-.21*	-.20	.00

SD = Symbol Digit; AC= Audio Code; DM = Dot Matrix; VIT = Visual Inspection Time; ART= Auditory Reaction Time
* *p*<.05 (1-tailed) ** *p*<.01 (1-tailed)

In order to assess the extent to which the temporal tasks predict performance in the speed tasks and working memory task (DM), linear regression was used. Rather than regress each speed task onto the discrimination tasks, a composite speed measure was calculated by averaging standardised scores on the speed variables (SD, AC, VIT, ART). A series of models were subsequently run in which either speed or DM was the dependent variable. The visual discrimination or auditory discrimination tasks were used as predictor variables, respectively.

Table 9: Correlations between discrimination tasks for masked (above diagonal) and unmasked conditions (below diagonal)

	VT_d	VL_d	AT_d	AP_d
VT_d	-	.36*	.65**	.27
VL_d	.34*	-	.35*	.17
AT_d	.64**	.20	-	.50**
AP_d	.59**	.30*	.66**	-

VT_d = Visual Temporal Discrimination; VL_d = Visual Length Discrimination; AT_d = Auditory Temporal Discrimination; AP_d = Auditory Pitch Discrimination
* *p*<.05 (1-tailed) ** *p*<.01 (1-tailed)

The results of these analyses are presented in Table 10. As can be seen, none of the models was statistically significant. The association between discrimination tasks and DM does however appear to be stronger than for the composite speed measure as well as more consistent. It is of a comparable magnitude in each of the modalities and in the different masking conditions.

Because the discrimination tasks are effectively identical in both conditions – they differed only in terms of the addition of a backward-masking stimulus – the regressions with DM as the dependent variable were repeated using the total sample (i.e., participants in both conditions were combined). According to these models, visual discrimination tasks and auditory discrimination tasks predicted a statistically significant amount of variance in DM (visual model [R^2=.09, F(2,63) = 3.26, p=.045] and auditory model [R^2=.11, F(2,63) = 3.74, p=.029]), and the sizes of the effects remained consistent with those in Table 4. The standardised coefficients for the auditory temporal and pitch tasks were β=.30 and β=.04, and β=.22 and β=.15 for the visual temporal and line length tasks. Thus, in both modalities – and particularly the auditory model – the temporal task is the stronger predictor of DM as opposed to the sensory measures (pitch and line length tasks).

Table 10: Regression models for masked and unmasked conditions

Dependent Variable	Predictor Variables	Condition	R^2	F	df	p
Composite Speed	VT_d & VL_d	*Masked*	.02	.22	*2, 30*	0.80
		Unmaksed	.02	.23	*2, 30*	0.80
	AT_d & AP_d	*Masked*	.02	.33	*2, 30*	0.72
		Unmasked	.09	1.51	*2, 30*	0.24
Dot Matrix	VT_d & VL_d	*Masked*	.12	1.96	*2, 30*	0.16
		Unmaksed	.08	1.33	*2, 30*	0.28
	AT_d & AP_d	*Masked*	.09	1.45	*2, 30*	0.25
		Unmasked	.10	1.73	*2, 30*	0.20

VT_d = Visual Temporal Discrimination; VL_d = Visual Length Discrimination; AP_d = Auditory Pitch Discrimination; AT_d = Auditory Temporal Discrimination

Discussion – Study 1

The relationship between Temporal Generalisation (TG) and markers of specific cognitive abilities was explored. The analyses suggest TG relates more strongly to the marker of Working Memory (WM) than to the composite speed measure. This result provides only limited support for the hypothesis that TG measures executive cognitive functions and not temporal resolution of the CNS because of the lack of statistical power and the limited number of marker tests.

Study 2

In light of the limited evidence provided in Study 1, the purpose of the second study was to explore further whether temporal tasks rely on memory functions by reanalysing previously published data. Rammsayer and Brandler (2007) reported on five temporal discrimination tasks; the Hick RT task (Hick, 1952); and a well defined battery of cognitive ability tasks measuring different aspects of intelligence corresponding to Thurstone's (1938) primary mental abilities. These tasks were completed by a large sample (N=100). The temporal tasks included: 1) Duration Discrimination (DD), requiring a decision concerning which of two successively presented timed intervals was longer; 2) Rhythm Perception (RP), requiring a decision concerning which of five beat-to-beat silent intervals – marked by 3 ms clicks – deviated from the constant 150 ms duration; 3) Temporal-order Judgment (TOJ), in which participants decide whether the onset of a Visual LED preceded that of an auditory stimulus, or vice versa; 4) Auditory Flutter Fusion (AFF), which derives an estimates of the ISI at which two successively presented auditory noise bursts appear fused; and 5) Temporal Generalisation (TG). These tasks comprise the battery used in previous investigations of temporal discrimination (see Helmbold & Rammsayer, 2006; Helmbold, et al., 2007; Rammsayer & Brandler, 2002, 2004, 2007).

Rammsayer and Brandler (2007) reported that a temporal *g* (Gt) factor defined by the discrimination tasks predicted 31% of variance in psychometric *g*, as defined by the cognitive ability measures. Combining Gt and a Hick *g* factor increased the proportion of explained psychometric *g* by only 2%. The unique contribution of temporal *g* was 20.5%, the shared contribution of temporal and Hick *g* was 10.5%, and the unique contribution of Hick *g* was only 1.5%. The authors concluded that temporal discrimination reflects an aspect of brain functioning that is stronger and more comprehensively related to *g* than parameters derived from the Hick RT task.

As already noted, temporal discrimination tasks may invoke demands on executive cognitive functions. To the extent that this is so, one would expect a Gt factor to relate strongly with *g* – and to a greater degree than RT tasks – because it would be saturated with variance reflecting cognitive functions underpinning both Gt and *g*. We have argued that the processes underpinning performance on temporal discrimination tasks might best align with memory functions, and previous research has established a strong and consistent relationship between WM and reasoning ability; as measured by intelligence tests (e.g., Burns, Nettelbeck, & McPherson, 2009; Kyllonen & Christal, 1990). Thus, tasks relying on memory functions should relate strongly to measures of intelligence, and temporal discrimination tasks may be an example of such tasks. Put more concisely, memory functions rather than temporal resolution of the CNS may be responsible for the relationship between temporal discrimination and intelligence. We present a reanalysis of Rammsayer and Brandler's (2007) data with the aim being to test whether memory mediates the relationship between temporal discrimination and intelligence.

Methods

Listed in Table 11 are the cognitive ability measures and temporal discrimination tasks used by Rammsayer and Brandler (2007) which are relevant to our aims. Participants in their study were 40 male and 60 female volunteers ranging in age from 18 to 45 years *(M* and *SD* of age: 26.0±6.8 years). The cognitive measures are composed of subtests of the Leistungsprüfsystem (Horn, 1983), Berliner Intelligenztruktur-Test (Jäger, Süβ, & Beauducel, 1997), and the German adaptation of Cattell's Culture Free Test Scale 3 (CCFT; Cattell, 1961; Weiss, 1971). Three of these subtests measure memory functions (Verbal, Numerical and Spatial Memory). The temporal tasks and their requirements are described briefly above and more detailed explanations can be found in the original publication.

A data file containing the correlations, means and standard deviations reported in Rammsayer and Brandler (2007) was created for analysis using MPlus 5.2 (Muthen & Muthen, 1998). Confirmatory Factor Analysis (CFA) was then performed on the covariance matrix using Maximum Likelihood estimation. Using this approach, different models were able to be tested which either included or omitted a relationship between temporal discrimination and memory functions, and these were compared using the model chi-square difference test. The fit of CFA models was assessed using the chi-squared test of model fit (χ^2), the comparative fit index (CFI), the root mean squared error of approximation (RMSEA), and the standardised root mean squared residual (SRMR).

Table 11: Intelligence scales and discrimination tasks used in Rammsayer and Brandler (2007) and the broad ability constructs measured

Intelligence Scale	Broad Ability	Temporal Discrimination Tasks	Broad Ability
Verbal Comprehension (VC)	Gc	Duration Discrimination 1 (DD1)	Gt
Word Fluency (WF)	Gc	Duration Discrimination 2 (DD2)	Gt
Perceptual Speed (PS)	Gs	Duration Discrimination 3 (DD3)	Gt
Number 1 (N1)	Gs	Temporal Generalisation 1 (TG1)	Gt
Number 2 (N2)	Gs	Temporal Generalisation 2 (TG2)	Gt
Space 1 (SP1)	Gf	Rhythm Perception (RP)	Gt
Space 2 (SP2)	Gf	Tonal-order Judgment (TOJ)	Gt
Flexibility Of Closure (CLO)	Gf	Auditory Flutter Fusion (AFF) [a]	Gt
Series (SE)	Gf		
Classifications (CL)	Gf		
Matrices (MA)	Gf		
Topologies (TO)	Gf		
Verbal Memory (VM)	Gm		
Numerical Memory (NM)	Gm		
Spatial Memory (SM)	Gm		

Gc, Crystallised Intelligence; Gs, General Speed of Processing; Gf, Fluid Intelligence; Gm, General Memory; Gt, General Temporal Discrimination

[a] AFF excluded from Study 2 analyses

Results

We attempted to confirm the presence of the general timing (Gt) factor reported in Rammsayer and Brandler (2007). However, we excluded the Auditory Flutter Fusion (AFF) task from our analysis because it has typically loaded poorly on Gt and might better be considered a sensory rather than temporal measure. CFA results indicate that the temporal tasks defined a Gt factor adequately [$\chi^2(14) = 20.26$, p=.122; CFI=.965; RMSEA=.067; SRMR =.047]. Modification indices suggested that the residuals of TG1 and TG2 should be allowed to co-vary. Therefore, in an additional model we added this path and it resulted in a significant improvement in fit [$\Delta\chi^2(1) = 9.33$, p=.002]. Rhythm Perception (RP) had a weak but significant loading (r=.37, p<.001) whilst the remaining tasks loaded strongly with an average of r=.64 (*Min*=.50, *Max*=.75, SD_r=.09).

Next, we confirmed the presence of a memory factor by specifying the three memory tasks to define a single latent Gm factor. Fit statistics are not available for this model because degrees of freedom are equal to zero. However, all three tasks loaded moderately supporting the presence of latent Gm: VM (r=.56), NM (r=.48), and SM (r=.45).

Rammsayer and Brandler extracted a single psychometric *g* (G) factor from the cognitive measures in their study. Instead, we used a hierarchical model in which specific broad cognitive factors were defined (Gc, Gs, and Gf; see Table 11) as well as *g*. First we attempted to define the lower order factors but statistics showed the model's fit was not adequate [$\chi^2(51) = 107.83$, p<.001, CFI=.885; RMSEA=.106; SRMR =.072]. Therefore, in consultation with modification indices, we correlated the residuals of the Series and Matrices tests. This resulted in a significant improvement in fit [$\Delta\chi^2(1) = 32.53$, p<.001], which was now considered adequate [$\chi^2(50) = 75.29$, p=.01, CFI=.949; RMSEA=.071; SRMR =.059]. The average loading of the tasks across all factors in this improved solution was r=.69 (*min*=.51, *max*=.86, SD_r=.11) and the correlations between the three first order factors were strong. In a subsequent model the first order factors were used to defined a *g* factor and the fit of this hierarchical model was also adequate [$\chi^2(50) = 75.29$, p=.01, CFI=.949;

RMSEA=.071; SRMR =.059]. The loading of each factor on *g* was strong: Gf (*r*=.84), Gc (*r*=.78), Gs (*r*=.94).

Having confirmed the presence of temporal, memory and psychometric factors, we were able to address the extent to which Gt and Gm are related and predict variance in psychometric *g*. To accomplish this, we first ran an *unrelated* predictor model in which *g* was regressed onto the independent factors Gt and Gm. In this model, both Gt (*r*=.62) and Gm (*r*=.55) predicted a significant but comparable amount of variance in *g*. Model statistics showed that the fit was not quite adequate [$\chi^2(202) = 262.99$, $p<.001$, CFI=.921, RMSEA=.055, SRMR=.086]. Therefore, we tested a *related* predictor model in which Gm was regressed onto Gt; whilst still maintaining regression paths from each of these to *g*. This related predictor model resulted in a significant improvement in fit [$\Delta\chi^2(1) = 7.24$, p=.01; $\chi^2(201) = 255.74$, p=.005, CFI=.930, RMSEA=.052, SRMR=.069].

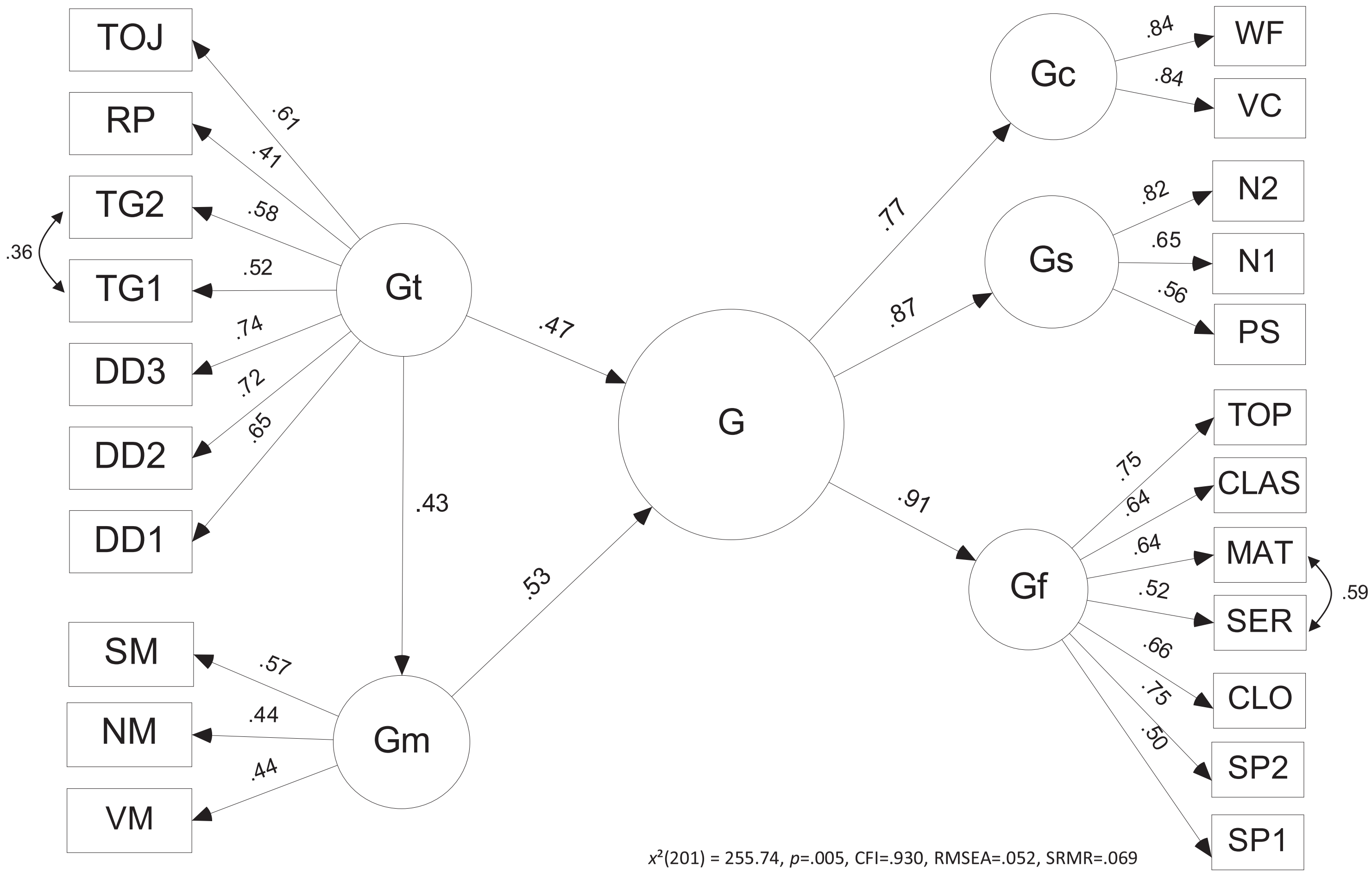


Figure 7. Hierarchical g model with related Gt and Gm predictors and standardised parameter estimates.

This hierarchical g with related predictors model is presented as Figure 7. The relationship between Gt and Gm is moderately strong, with the latent variables sharing approximately 20% of their variance. This path was necessary for satisfactory fit and its addition resulted in a marked decrease in the size of the coefficient between Gt and *g* (.47 compared to .62), but not between Gm and g (.53 versus .55). The standardised direct effect of Gt on *g* is .47 and the indirect effect is .23 (.425*.531). Thus, 34% of the total effect of Gt on G appears to reflect memory abilities captured in latent Gm.

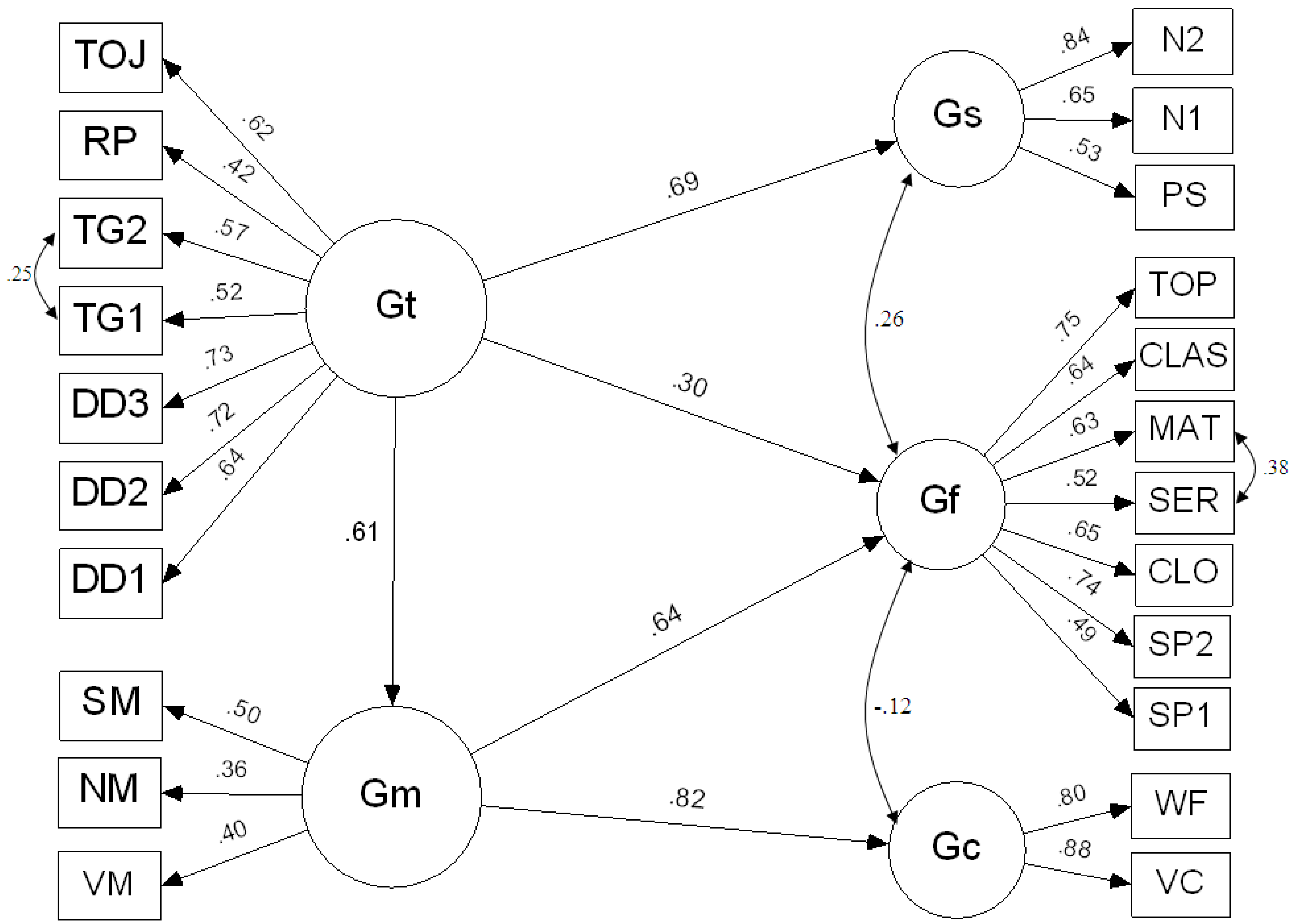
In light of the smaller yet significant path between Gt and *g* in this related predictor model, we defined a model which excluded *g* and instead regressed each of the lower order factors onto related Gt and Gm factors. The purpose of this was to better understand the moderate relationship between Gt and *g* after accounting for Gm functions. The fit of this model was good [$\chi^2(197) = 244.91$, p=.01, CFI=.938, RMSEA=.049, SRMR=.067]. The path from Gm to Gc was significant (r=.57, p<.001) but not from Gt to Gc (r=.22, p=.11). The path from Gm to Gs was not significant (r=.19, p=.24) but it was from Gt to Gs (r=.55, p<.001). Gf was predicted significantly by Gm (r=.56, p<.001) and to a weaker degree by Gt (r=.39, p=.003). In a subsequent model we dropped these non-significant paths, as well as the covariance between Gs and Gc because of their relative independence. The fit of this model decreased significantly [$\Delta\chi^2(3)$=22.34, p<.001]. However, overall model fit remained statistically adequate [$\chi^2(200) = 267.25$, p<.001, CFI=.913, RMSEA=.058, SRMR=.078] and in the interest of parsimony, this more restrictive model – shown in Figure 8 – should be favoured over the former. As can be seen, Gt relates strongly to Gm (R^2=.37) and Gs (R^2=.48). The relationship between Gt and Gf is markedly weaker and these constructs share only 9% of variance. Gm on the other hand, relates strongly to both Gc (R^2=.67) and Gf (R^2=.40).

Discussion – Study 2

The reanalysis in Study 2 has provided a more rigorous assessment of the hypothesis that temporal discrimination tasks reflect memory functions than Study 1 because of a larger test battery of cognitive measures. The CFA models show that the relation of latent Gt to Gm must be incorporated into these structural models to achieve adequate fit. Moreover, it appears that around 35% of the relationship between Gt and general intelligence estimates (*g*) can be explained by memory functions shared with Gm. Of the three broad cognitive factors extracted in the second model (Figure 3), Gt appeared to relate more strongly to speed of processing (Gs) and Gm than either Gf, or Gc.

General Conclusions

Recent research has proposed that temporal resolution of the CNS is partly responsible for intelligent functioning and that auditory temporal discrimination tasks provide a valid measure of this resolution (Helmbold & Rammsayer, 2006; Helmbold, et al., 2006, 2007; Rammsayer & Brandler, 2002, 2004, 2007). This paper has questioned this notion and has presented the results of two studies designed to elicit the nature of the functions underpinning performance on temporal discrimination tasks.



$X^2(200) = 267.25$, $p<.001$, CFI=.913, RMSEA=.058, SRMR=.078

Figure 8. Broad ability factors model with related Gt and Gm predictors and standardised parameter estimates

Study 1 showed that the construct measured by the auditory TG task is not modality specific. The correlation between visual and auditory TG was strong; the tasks shared 42% of variance. Strong relationships have generally not been evident when adapting ECTs across modalities. For example, the relationship between visual IT and auditory IT seldom exceeds r=.30 (see e.g., Deary, 2000) and the variance in these tasks has been largely attributed to peripheral sensory type processes (Burns, Nettelbeck, McPherson, & Stankov, 2007; Burns, Nettelbeck, & White, 1998; White, 1996; Zajac & Burns, 2007). Contrary to this, the strong correlation between visual and auditory TG suggests the processing required by these tasks might not be sensory but rather cognitively based. This would explain the relative independence of sensory and temporal discrimination factors reported previously (Rammsayer & Brandler, 2004).

Study 1 measured the distinct constructs, Gs and WM, to better understand the observed correlation between TG and intelligence. Speed of processing was measured using traditional speed tasks as well as widely researched ECTs (RT and IT). The impetus for including RT and IT was the proposition that auditory temporal tasks might be analogous to ECTs in terms of providing an estimate of neural efficiency (Helmbold, et al., 2007b). The present study does not support this hypothesis. Neither the visual nor auditory TG tasks predicted a statistically significant amount of variance in the composite speed measure with the shared variance near zero.

The regressions of DM – a measure of working memory – onto visual and auditory TG tasks were not statistically significant. However, given the near equivalence of the TG tasks across experimental conditions and the consistency of the effect size, the samples were combined across conditions and visual and auditory TG did predict a significant amount of variance in DM; the size of the effect again remained consistent (about 10% shared variance). This shows that the absence of a significant effect within experimental conditions reflects a lack of statistical power.

The reanalysis of Rammsayer and Brandler's (2007) data in Study 2 provided a more rigorous test than Study 1 of the hypothesis that temporal discrimination tasks reflect memory functions because of a more comprehensive test battery and larger sample size. The results provide further evidence that temporal discrimination tasks rely, at least to some extent, on memory functions. In the hierarchical *g* model, the path between latent Gt and Gm factors was both necessary and significant, with the latent factors sharing 20% of their variance. Furthermore, it was found that around 35% of Gt's relationship to *g* could be attributed to memory functions represented by latent Gm.

In the second model (Figure 8), the regression of Gm on Gt was stronger, with the constructs sharing around 37% of their variance. Of further interest was that of the three broad cognitive ability factors, Gt only predicted a substantial amount of variance in Gs (47%); Gt predicted only 9% of Gf. This finding is somewhat consistent with earlier studies in which Gt has been found to share variance with RT factors. The analysis in earlier studies, however, has been framed to explore which of Gt and RT explains more *g* variance. Not surprisingly, Gt emerges as the stronger predictor and it almost wholly accounts for the relationship between RT and *g* (Helmbold & Rammsayer, 2006; Helmbold, et al., 2007; Rammsayer & Brandler, 2007).

In the hierarchical model in Study 2 it appeared that Gt measured functions over and above memory which predicted *g* variance. The non-hierarchical model shows, however, that this significant Gt x *g* path essentially reflects Gt's relationship to Gs, and it is because of this relationship that Gt can account for the correlation between RT and *g*. Such a finding does not, as proposed by Rammsayer and Brandler (2007), imply that Gt measures anything more fundamental to intelligence than do RT tasks; it suggests that they measure the same functions.

The relationship reported herein between Gt and memory measures are consistent with the requirements of the temporal tasks. Duration Discrimination (DD) requires internal representations of timed intervals to remain accessible for several seconds following their presentation; TG requires accurate learning – and thus memorizing – of a standard stimulus, as well as accurate retrieval of the learnt standard and comparisons with trial stimuli. Interestingly, the pitch and line-length discrimination tasks in Study 1 had the same requirements as TG but did not contribute substantially to the prediction of the working memory task. One explanation for this finding is that functions involved in TG are more complex than for pitch and line-length tasks. Auditory sensory memory, for instance, can retain information concerning dimensions like intensity and frequency for four-to-ten seconds (Jaaskelainen, Hautamake, Naatanen, & Ilmoniemi, 1999). In the pitch task, then, it is plausible that the comparisons of stimuli rely heavily on these sensory memory traces. Conversely, time is not a perceptual dimension but a cognitively derived entity (Michon, 1990) and therefore the comparisons of stimuli in auditory and visual TG tasks rely on cognitive representations of the durations which appear to be memorised and rehearsed.

Previous research supports this hypothesis. It has been found in a number of studies that temporal processing of durations involves the prefrontal cortex (Elbert, Ulrich, Rockstroth, & Lutzenberger, 1991; Harrington, Haaland, & Knight, 1998) which is the brain region thought to play a critical role in the distributed neural systems which achieve working memory (Engle, Kane, & Tuholski, 1999; Gibbons, et al., 2002). Furthermore, an event-related potential (ERP) study which compared temporal and pitch discrimination tasks showed enhanced prefrontal activation in the temporal task (Gibbons, et al., 2002). This finding was interpreted as indicating a much stronger contribution of executive memory functions to temporal as opposed to pitch discrimination and it was concluded that "to perceive time and to evaluate temporal properties of a given stimulus, formation of cognitive temporal representations is required – a process primarily based on executive working memory functions" (Gibbons, et al., 2002, p.963).

In summary, the findings herein and those of previous studies raise questions regarding the extent to which auditory temporal discrimination tasks should be considered measures of neural efficiency and by extension, temporal resolution of the CNS. It appears that the observed relationship between auditory temporal discrimination tasks and measures of *g* may be explained almost entirely in terms of memory functions and speed of processing. More specifically, temporal discrimination performance is confounded by *both* memory and speed functions and its relationship to the latter does not automatically imply that temporal resolution of the CNS is involved. Even if temporal resolution of the CNS is independent of higher order cognitive operations (Rammsayer & Brandler, 2002, p509) temporal discrimination tasks are not. Therefore, attempts to gauge the strength of the relationship between CNS resolution and intelligence – if indeed there is one – using such tasks will not be unequivocal.

Chapter 6: Study 4 – Auditory Inspection Time Tasks

RELATIONSHIPS BETWEEN THREE AUDITORY INSPECTION TIME TASKS AND PROCESSING SPEED

Ian T. Zajac and Nicholas R. Burns

School of Psychology, University of Adelaide, Australia

Published Manuscript

Australian Journal of Psychology

Statement of Contributions

Ian Zajac (Candidate)

I was responsible for the conception and primary authorship of the paper. I was responsible for the development and programming of software based assessment tools and collection of all data. I conducted the statistical analyses independently with advice from the co-author. I was corresponding author and primarily responsible for responses to reviewers and revisions to the paper.

Nicholas R. Burns (Co-author)

I was the supervisor (advisor) for the research programme that led to this publication. In terms of conceptualisation of the programme, there was extensive and

ongoing collaboration between Mr. Zajac and me in developing the direction of the research. The realisation of the study and the programming of assessment tools used, were the work of Mr. Zajac. I had an advisory role with respect to selection of the test battery used and on the direction and specifics of the data analyses.

Mr. Zajac was responsible for writing this paper; my role was to comment on drafts, make suggestions on the presentation of material in the paper, and to provide editorial input. I also provided advice on responding to comments by the journal reviewers and editor.

I hereby give my permission for this paper to be incorporated in Mr. Zajac's submission for the degree of PhD in the University of Adelaide.

Zajac, I.T. & Nicholas R. Burns, N.R, (2011) Relationships between three auditory inspection time tasks and processing speed.
Australian Journal of Psychology, v. 63 (3), pp. 163-172

NOTE:
This publication is included on pages 154-175 in the print copy of the thesis held in the University of Adelaide Library.

It is also available online to authorised users at:

http://dx.doi.org/10.1111/j.1742-9536.2011.00020.x

Chapter 7: Discussion

The broad rationale for this thesis was to examine auditory cognitive tests and their relationships to other broad intelligence constructs typically measured by visual tests. Two studies aimed to examine whether auditory tasks could be purpose-developed to index the broad ability: general speed of processing (Gs). A further two studies sought to re-examine existing auditory tasks and the abilities/processes engaged by them in order to clarify their relationships to broad constructs including Gs and memory (Gm). The results of these studies have gleaned substantive information regarding cognitive performance in the auditory modality and its relationship to more traditional visual intelligence measures. The studies are now briefly revisited prior to considering the combined findings in a broader context and outlining viable directions for future research.

Auditory measures of Gs: Studies 1 and 2

The purpose of Studies 1 and 2 – reported across Chapters 3 and 4 – was to test whether auditory tasks could be purpose-designed to measure the broad construct, Gs. The motivation for undertaking these studies derived from the largely untested assumption that broad abilities in the taxonomy of intelligence, like Gf, Gc and Gs, are not modality-specific (see e.g., Roberts, et al., 1997), but also from the recommendations of Carroll (1993) regarding designing tests specifically for particular constructs. In other words, tests should be carefully thought-out and developed to engage particular cognitive processes and these processes, theoretically, will govern the relationships these tests display with existing intellectual abilities. In order to explore these issues, tasks were borrowed from existing literature where it seemed sensible that they might tap Gs constructs. More importantly, others were purpose-developed to maintain the same cognitive requirements as existing visual Gs measures. It was expected that the new and existing tasks would correlate well with their visual analogues and, consequently, define broad Gs-type constructs.

The data collected during Study 1 provided preliminary evidence suggesting that auditory tasks can measure broad Gs. Given the links reported elsewhere in the literature between reaction time tasks and speed constructs (Danthiir, 2005), we expected auditory RT tasks (ART) to relate well to visual RT (VRT) and other speed measures. This hypothesis was supported because simple VRT and ART tasks correlated strongly and shared 42% of their variance. The slightly more complex ART tasks (two and three choice versions) also correlated moderately, at around r=.50, with VRT. Furthermore, these measures defined a broad RT factor that emerged clearly during exploratory factor analysis of the data. Similarly, the newly developed auditory tasks – chasing digits, audio code, hearing As, and tone comparisons – were expected to share variance with speed measures, given they had been designed to engage similar mental operations as those required by existing visual measures. The correlations between these and other tests in the battery were generally significant where expected and the tasks combined to define a broad latent Gs type factor. The Gs and RT factors in this study shared an impressive 35% of their variance (see Figure 3, p.82). Reliability measures for the new and existing auditory tasks – split-half and test re-test – were also acceptable, implying that the construct measured by the tasks remained stable over the duration of the individual tests but also over different testing occasions.

Study 1 used a limited number of visual tests because the primary purpose was to pilot the new auditory tasks, but this also reflected time constraints arising from recruiting participants on a voluntary basis. It was speculated though, that the speed factors recovered from the data were not likely to be modality specific given the visual tasks loaded moderately on them together with the auditory measures. Nevertheless, this issue was specifically addressed in Study 2 because the test battery was enlarged to include a visual marker test corresponding to every auditory measure used.

The results of the second study supported and extended the earlier results concerning the extent to which purpose-designed auditory tasks can index Gs. Overall, the auditory tasks correlated well with their visual analogues except for Tone Comparisons (TC) and Number Comparisons (NC), which correlated weakly. There are several plausible explanations for these poorer than expected correlations, which are discussed in more detail in the coming section regarding design issues concerning auditory tests. Despite this, the tasks still defined strong latent Gs and RT factors, which correlated in a comparable fashion to Study 1.

The inclusion of equal numbers of visual and auditory marker tests in Study 2 enabled us to explore the pertinent question of whether the tasks, and thus factors, might be saturated with modality specific variance. During exploratory structural equation modeling modality specific speed factors did not clearly emerge. In order to explore the issue further, we adopted a confirmatory framework and attempted to specifically define these modality factors. The model would not converge when distinct auditory and visual Gs factors were specified due to perfect linear dependency between them. The model did terminate when a broad Gs factor and distinct RT_A and RT_V factors were specified, but the statistical fit was not adequate and the RT factors correlated strongly at r=.74 suggesting they essentially reflected a similar construct.

The final examination of latent Gs and RT factors involved regressing a criterion measure onto visual and visual+auditory speed factors, respectively. It was reasoned that the relationship of these factors to the criterion measure would change significantly if the latent construct changed after incorporating auditory tasks into its measurement. Neither the visual nor visual+auditory RT factors predicted variance in the criterion (Advanced Progressive Matrices). On the other hand, both visual and visual+auditory Gs factors did predict variance but the difference between the regression coefficients was not statistically significant (R=.45 and R=.38, respectively). Essentially then, the notion of modality specific speed factors is not supported by these data.

Auditory Temporal Discrimination and Inspection Time

Studies 1 and 2 were primarily concerned with developing new auditory tasks to measure the broad construct, Gs. A somewhat different but related aim of Studies 3 and 4 was to reconsider existing auditory tasks in terms of their cognitive requirements and assess whether these resulted in observed correlations between these tasks and existing constructs representing similar processes.

Study 3 concerned itself with temporal discrimination tasks. As already highlighted, the work of Rammsayer and colleagues has appealed to these tasks as simple measures of elementary processes reflecting neural efficiency and, by extension, temporal resolution of the CNS (see Helmbold & Rammsayer, 2006; Helmbold, et al., 2006, 2007; Rammsayer & Brandler, 2002, 2004, 2007). Contrary to these works, elsewhere in the literature temporal tasks have been regarded as being cognitively complex (Stankov, 1983). According to Michon (1990) time is not a perceptual dimension but a cognitive entity and performance on temporal discrimination tasks is associated with enhanced activity in the prefrontal cortex

(Gibbons, et al., 2002) which encompasses the complex distributed neural systems which underpin working memory (Engle, et al., 1999).

Careful decomposition of the tasks used in the temporal discrimination studies revealed that many of the tasks have an apparent memory load. For instance, duration discrimination (DD) requires internal representations of timed intervals to remain accessible for several seconds following their presentation. Temporal generalisation requires accurate learning – and thus memorising – of a standard stimulus, as well as accurate retrieval of the learnt standard and comparisons with trial stimuli. Thus, in line with Carroll's (1993) recommendations in terms of understanding the processes that underpin task performance, it was proposed that these temporal tasks would relate primarily to measures of working memory, and that this might explain the reported correlation between these tasks and intelligence estimates (see e.g., Helmbold & Rammsayer, 2006; Rammsayer & Brandler, 2002, 2007).

The first of the two studies presented in Chapter 5 explored the temporal generalisation (TG) task and also adapted it to the visual modality. Results showed that the visual and auditory TG tasks shared around 42% of their variance, implying that they largely index the same construct. Further regression analyses showed that the tasks related to a measure of working memory as opposed to speed of processing, or Gs. Importantly, in both the visual and auditory modality, the temporal tasks – as opposed to pitch and line-length sensory discrimination measures – were the strongest predictor compared to the more sensory tasks, pitch discrimination and line-length discrimination. The second of the two studies undertook a reanalysis of a larger, previously published data set to explore these tasks further. Structural equation modeling showed that a general temporal discrimination factor (Gt) related to a broad memory construct (Gm). Moreover, Gm accounted for around 34% of the relationship observed between Gt and *g*. Further analysis revealed that of the factors Gf, Gc, Gs and Gm – defined by typical intelligence measures – Gt related notably to Gm and Gs only (r=.61 and r=.69, respectively).

Study 4 (Chapter 6) concerned itself with measures of auditory inspection time (AIT). During the course of AIT research the tasks have been refined and reinvented. Unfortunately, little thought has been given to the processes which underpin performance on the various measures and it has been argued elsewhere (Zajac & Burns, 2007) that this explains the poor relationships pitch and loudness versions of AIT (AIT-P an AIT-L) have displayed with VIT and Gs tasks; the latter are closely related to VIT performance (see e.g., Burns & Nettelbeck, 2003, 2008; Burns, et al., 1999; Burns, et al., 2007; O'Connor & Burns, 2003). On the other hand, the most recent iteration relying on spatial discrimination (AIT-S) appears cognitively similar to VIT and moreover previous studies suggest it is more consistent in its relationships with both VIT and measures of Gs.

Findings presented in Study 4 glean further information concerning these AIT tasks. More specifically, of the three versions of AIT used in the study only AIT-S appeared to relate to a broad Gs construct defined by tasks including VIT, VRT and ART. This is consistent with previous research (see e.g., McCrory & Cooper, 2005; Zajac & Burns, 2007). The other AIT measures defined a general AIT factor, which was also loaded on by AIT-S, but which was essentially unrelated to the broad Gs construct (Gs x G_{AIT} r=-.15). Thus, it was concluded that the AIT tasks by-and-large measure modality specific processes unrelated to Gs, and that out of all versions of the task, AIT-S behaves ‘most’ alike its visual analogue, VIT.

Summary and overall significance of main findings

When considered together, the findings of the studies presented across this dissertation indicate that auditory tasks might be able to measure existing broad constructs defined in the CHC taxonomy of human intelligence. A key construct of interest was that of general speed of processing or Gs but the relation of auditory tasks to latent memory functions, or Gm was also considered. In all instances, where it was sensible to assume relationships based on an analysis of auditory task requirements, those relationships were indeed evident. More specifically, the newly developed auditory tasks and purposely selected existing auditory measures related to Gs and there was no evidence of modality specific factors underpinning these tasks. Similarly, the temporal tasks that appear to engage memory processes relate moderate-to-strongly with latent Gm which partly mediates their relationship with estimates of *g*. Although not specifically hypothesised in Study 3 (Chapter 5) temporal tasks also related to Gs. This is not surprising given that most tasks require the auditory stimuli to be encoded quickly and correctly for successful resolution. Further examination of AIT established that AIT-S is a more similar measure to VIT in terms of a more consistent relationship with VIT and consequently, Gs type tasks, but also that all of the auditory IT tasks rely significantly on similar, presumably modality-specific type processes.

Although broad constructs other than Gs and Gm were not specifically considered, there are indications in the literature that similar results could be obtained for constructs such as fluid intelligence (Gf) and crystallised intelligence (Gc), for example. As reviewed in the Introduction (Chapter 1), these relationships have previously but sporadically presented themselves in the literature. Aside from Stankov’s (1971) work and the present studies, however, there have been no other systematic attempts to document a detailed programme of research concerning the relation of auditory tasks to visually defined constructs. Indeed, Stankov considered Gf, Gc and Gv relationships during his work on auditory tasks but there are several limitations to the approach that was taken in that study which will be further elaborated in the coming sections. Regardless, the present studies, when taken together particularly with those of Stankov (1971), provide increasing evidence that the constructs defined in CHC theory, for example, are likely cognitively general and manifest in at least the visual and auditory modalities on the condition that tasks are suitably designed.

The studies presented as part of this dissertation mark a step forward in the area of auditory focused intelligence research and establish a framework for future studies. More specifically, rather than administer and subsequently interpret auditory tests post-hoc on the basis of observed correlations between them and other intelligence estimates – typical of research in this area – the present research programme shows that more accurate information can be obtained by carefully considering auditory task requirements a-priori. Subsequently, based on this consideration, criterion measures of intellectual constructs can be appropriately referenced during auditory test development and measured during the course of interpreting performance on said tasks. This approach is superior to most previous studies and will help to improve our understanding of the relationships between these measures and other largely non-specific constructs like *g*, for example. Attempting to understand tasks based on their relationships with poorly-understood or non-specific factors, like *g*, leads to guess-work and erroneous hypotheses regarding what processes the auditory tasks reflect. Future research on AIT and temporal discrimination should consider these auditory tasks according to this framework to further elucidate the expression of intellectual constructs in alternate modalities. Future attempts to develop auditory tasks to measure distinct constructs other than Gs can benefit from previous attempts to do so (i.e., the present thesis) that have been generally successful when the intended construct was appropriately considered.

Interestingly, a reviewer of the manuscript presented as Chapter 1 (Developing Auditory Measures of Gs) questioned the overall significance and relevance of the findings to an understanding of human intelligence. In particular, the reviewer noted that the absence of modality specific factors and consequent zero-difference between the Gs constructs defined by visual as opposed to visual and auditory tasks, respectively, resulted overall in no improvement in measurement of the construct itself. However, the purpose of Study 1 and this research programme as a whole was not to test whether measurement of the construct would improve. Rather, in light of the assumption that broad CHC constructs are cognitively general, but that this idea is largely untested in the literature, the purpose was to assess whether measurement of Gs, and by extension other constructs, via alternate modalities was at all possible. In a sense then, these are somewhat a series of validation studies designed to test just how ‘general’ these constructs are. It might be that case that measurement of distinct broad cognitive abilities like Gs will not improve through use of additional auditory tasks – as suggested by the findings of Study 2 – and this idea is captured in the ‘indifference of the indicator’ hypothesis (Spearman, 1927). However, this fact alone does not discount the utility of auditory tests of intellectual ability in both the applied and research settings.

General Limitations

The studies presented within this thesis are not without methodological limitations. One of the most obvious empirical problems, pertinent in much psychological research in general, was the sole use of relatively homogeneous university samples (first and second-year psychology undergraduates). This was largely unavoidable due to the nature of institutional research participation programmes, whereby students receive a proportion of credit toward their course grade in exchange for participation in internal research studies. One effect of using such samples is attenuated correlations due to restricted variability in cognitive performance. Thus, particularly in regards to the new auditory tasks developed for the present thesis, future studies should recruit samples ranging widely in both age and cognitive ability to ensure the strength of the associations are generalisable.

An additional problem that arises from using a voluntary participation pool is borne out in time limitations. It is difficult to recruit participants into studies that require in-person attendance – as opposed to internet based studies – and not many will voluntarily sit through hours of cognitive testing. In Studies 1 and 2 (Chapters 3 and 4), the testing session was around 1.5 to 2 hours in duration. The desire to include many auditory tests in the first study resulted in visual tasks being under-represented. This issue was resolved in the second study by pairing visual and auditory tasks, but to ensure the study session remained within a suitable time-length some auditory tasks were shortened (i.e., ART trials were cut by 50% from the first study). This might have impacted the reliability of the auditory tasks, but no conclusions can be made regarding this as test reliability estimates were not collected as part of Study 2. Future studies should address the issue of reducing trials in RT tasks on test reliability.

Time constraints also resulted in the omission of marker tests for other broad constructs. This is a very notable limitation of the studies reporting on auditory measures of Gs and AIT. It is entirely possible that the latent constructs recovered from the data that were tentatively named Gs and RT do not entirely reflect Gs-type processes. Given the absence of alternate constructs it is possible that the tests were constrained to load together. Certainly, constructs other than Gs would be expected to relate to some of these measures to varying degrees. However, several points of note suggest these factors and task loadings would probably emerge even in the presence of other ability constructs.

First, the models generated were theoretically sensible and the solutions were supported by acceptable model-fit statistics. If the auditory tests were measuring a markedly different construct to the visual tests then fit statistics would not achieve acceptable levels because the model would not account for sufficient covariance in test scores. Second, when an attempt was made to purposely define modality specific latent Gs and RT constructs, the model was insufficient because the factors were essentially perfectly related. If the auditory tasks were measuring an alternate construct much weaker associations between these modality specific factors would have been expected. Moreover, when the AIT tasks were introduced to the model, latent AIT did not relate well at all to the latent Gs and RT type factors. Thus, it is probable that the auditory speed tasks within the battery are not overly reliant on peripheral or sensory type processes because the model would not have achieved adequate fit. Regardless, it is entirely necessary that future studies attempt to incorporate sufficient marker tests to allow several alternate broad cognitive constructs to emerge from the data in order to explore the relationships between these auditory tasks and other cognitive abilities.

In a similar vein, future studies should better define a Ga factor in order to further explore the processes underpinning the auditory speed tasks. In the first study AIT-L and AIT-P were used as a proxy for peripheral auditory processes. This was due primarily to the time constraints already discussed, but also the mounting evidence in the literature that these tasks reflect peripheral processes reflecting pitch discrimination, and probably intensity-discrimination abilities. To a large degree, this latter aspect was supported by our data because the latent AIT factor was essentially unrelated to the speed measures; only AIT-S appears dependent upon speeded processes. However, more complex auditory abilities should be measured in order to assess the extent to which they underpin performance on these tasks also.

Finally, the sample sizes in each of the studies were considered sufficient, but small. This is a particular problem when employing structural modeling with some authors suggesting sample sizes of at least N=200 be used (Barrett, 2007). Obviously, as the number of measured constructs increases, and consequently the number of tests utilised, the number of participants will also need to be increased. This will result in better statistical power and a more robust test of model fit statistics. Consequently, the relationships observed in future studies will also be more robust in terms of estimates of shared variance.

Design issues facing auditory tests

Results of the present studies as well as findings reported in past literature highlight two pertinent issues relevant to the construction of auditory tests. The first of these issues concerns task difficulty. It is not the case that tasks translated from the visual modality to the auditory maintain comparable levels of difficulty. Auditory IT tasks are a prime example of this difference. More specifically, although completion of the VIT task is trivially easy for anyone (Deary, 2000a), difficulties inherent in performing pitch discriminations have proved problematic for AIT-P. Completion of AIT-L and AIT-S has been typically better than for AIT-P, but it still suffers from being too difficult for some participants; only 81% (18/96), 86% (13/96) and 92% (7/96), respectively, of participants in Study 4 completed these tasks. Another example of this problem is borne out in Study 2 (Chapter 4) in the case of the auditory speed tasks, tone comparisons (TC) and hearing As (HAs). Although the auditory tests were modeled on visual number comparisons (NC) and finding As (FAs), they did not relate well to these measures and the number of items completed differed markedly between the modalities.

It was noted in the discussion of Study 2 that the absence of notable correlations between these tasks and differences in item completion might reflect differences in difficulty levels. For example, whilst TC involves a relatively simple comparison of two successively presented tones, the comparison in NC is made on the basis of two number strings presented simultaneously, involving up to 12 digits in each string. Thus, comparing the digit strings in the visual task is more complex, requiring searching and scanning type processes. Finding As might similarly be more complex than its auditory analogue because it requires participants to visually check for the presence of the letter A in a presented word, whereas the pronunciation of the letter A in the auditory tasks was always clearly audible.

When developing auditory tasks to index specific constructs attention should be paid to task difficulty because alterations in this across modalities might affect what processes the tasks ultimately measure. For example, if the auditory version of the NC task presented number strings instead of single tones, participants would likely need to encode both number strings into short-term memory in order to be able to compare them. This is because, unlike for many tests in the visual modality, the stimuli are not readily available for a second or third examination and thus accurate representations of them must be registered when they are heard. Obviously, the longer the auditorily presented number strings are, the more complex the memory processes will become. In effect then, the task might change from measuring the intended construct Gs, to Gm, for example. It might be possible to present shorter digit strings auditorily – consisting of say four digits – as opposed to longer ones, and this may rely less on alternate cognitive processes that confound task performance. Regardless, the difficulty of auditory tasks needs to be carefully considered and evaluated in order to understand the processes involved in completing them. Attempting to maintain a comparable level of difficulty across tasks will help to ensure that latent constructs derived from factor analysis, for example, reflect common cognitive processes rather than difficulty factors.

A second issue related to developing auditory tasks concerns moving away from relying on discriminations that are inherently auditory. More specifically, the idea of an auditory hierarchy as envisaged by Stankov (1994a) was introduced in Chapter 4 and is displayed graphically as

Figure 9. The highest level of the hierarchy denotes thinking abilities that are intellectual in nature. It is at this level that the broad constructs in CHC theory would reside, and although they are likely manifest in alternate modalities – like the auditory modality – it can become difficult to distinguish these processes from perceptual or even sensory ones. Carroll (1993) regarded auditory abilities as being reliant on the first two levels of this hierarchy. Thinking abilities, on the other hand, reflect other expertise – like knowledge of language structure – and reliance on auditory processes would be incidental. Thinking abilities are different to those involved in a tonal

analogies task, for instance, in which perceptual abilities are magnified because of the comparison of complex tone sequences.

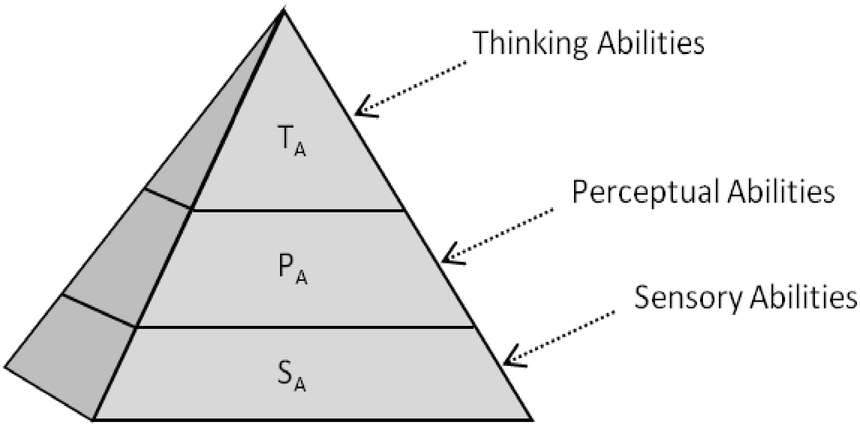


Figure 9. Hierarchy of auditory abilities

Unfortunately, in the case of auditory research this distinction has not always been satisfied. I noted previously that Stankov's (1971) approach to exploring auditory performance was not without shortcomings. More specifically, although one intention in those studies was to explore the relation of auditory tasks to existing visual constructs, the primary aim was to identify auditory abilities. Thus, it was not necessary to refrain from amplifying the involvement of sensory and perceptual processes in task performance. Indeed, some of the auditory tasks used did relate to constructs like Gf, and Gc. However, the most notable loadings were from auditory primaries defined by tasks that relied only minimally on perceptual processes. Auditory verbal comprehension tasks, for example, relied on knowledge of language and required participants to identify spoken words. The primary factor defined by these loaded moderately (*r*=.43) on Gc together with verbal comprehension as measured by visual tests (*r*=.50). On the other hand, discrimination among sound patterns (DaSP) – another auditory primary ability – most certainly involved reasoning abilities similar to those subsumed by Gf, but it loaded only weakly (*r*=.21) on the broad Gf factor marked by visual tasks. The loadings of DASP and auditory verbal comprehension on Ga were *r*=.50 and *r*=.11 respectively. Thus, auditory tasks seem to be cleaner measures of broad constructs other than Ga if they refrain from emphasising perceptual processes.

The present studies were not concerned with auditory abilities per se and made no attempt to specifically engage auditory processes like, for example, intensity discrimination or rhythm perception. Therefore, it is not surprising that modality factors did not emerge because reliance on the auditory modality was incidental. Many of Stankov's (1971) tasks, on the other hand, tended to overlap modality specific and other broad thinking constructs because of the emphasis on auditory processing. Indeed, comparison of tones was employed in the speeded tone comparisons task used in Studies 1 and 2 herein. However, the smallest absolute difference between tones presented in that task was 15 Hz and a difference of as much as 10 Hz has been cited as a threshold at which tone discriminations are trivially easy for most people (Roederer, 1973). Thus, because of the absence of difficult tone discriminations even this task loaded well with visual measures of Gs. If we had reduced the difference in frequency between tones, results would likely have been quite different for this task.

Practical implications of design issues

The design issues facing auditory tests may prove problematic in a practical setting if they are not given due consideration. For example, as has been the case for

AIT research, task difficulty can be a significant problem. In the research setting, poor completion rates prove problematic for power calculations and empirical results are severely compromised by the fact that they are based on sub-samples of samples. Similarly, in an applied setting there would appear to be little application for an auditory task which is too difficult for people to complete, unless this aspect of performance alone provides the information sought.

Another design issue having practical implications relates to the constructs measured by auditory tasks. As already highlighted, there is a real potential for auditory specific processes to confound performance on auditory tasks. Consider as an example a research project concerned with the broad construct, Gc. To measure this construct auditorily, the researcher uses the Speech Perception under Distraction (SPuD) task because it relies somewhat on knowledge of language structure and, therefore, verbal ability. The results of this hypothetical study suggest that the auditory task does not measure Gc as the measures do not correlate well. Unfortunately, the fact that SPuD relies heavily on the ability to separate spoken language from background noise is not acknowledged but it is because of this reliance on auditory processing that the SPuD task and Gc do not relate. Essentially then, when auditory tasks overlap several constructs it becomes difficult to disentangle which of the observed effects reflect auditory specific variance as opposed to the actual construct of interest.

If auditory tasks are eventually extended to an applied setting in the form of sub-tests in intelligence test batteries, for example, then this issue would be even more pertinent. Consider the Woodcock-Johnson III Tests of Cognitive Abilities (WJ-III; Woodcock, et al., 2001). Three potentially useful scores can be derived from this battery including: 1) Full Scale IQ – FSIQ –scores; 2) Cluster Scores; and 3) Broad factor scores. Although the fact that an auditory task might overlap several constructs is not overly problematic for FSIQ scores, it is when interpreting cluster and broad factor scores. More specifically, the WJR-III measures three clusters reflecting Verbal Ability, Thinking Ability and Cognitive Efficiency. If the SPuD task noted above was employed in this battery and was considered a measure of cognitive efficiency, scores on this cluster would be confounded unnecessarily by verbal ability, given that SPuD also relies at least to some extent on this. In both the practical and research settings then, when distinct constructs are the topic of interest careful selection of auditory tests used to measure them is paramount.

Future Directions

There are many directions that auditory research can and should take during the exploration of cognitive abilities. In a similar vein to the studies reported herein, research should explore other broad constructs defined in CHC theory. For some constructs, there has been very little systematic effort devoted to measuring them across alternate modalities. Before outlining the reasons for undertaking this work in the future, I will consider each of the main constructs in CHC theory – not including Ga – and speculate on the success of such attempts to index them auditorily.

Crystallised Intelligence, or Gc, should not prove difficult to index auditorily. Horn and Stankov (Horn & Stankov, 1982) have shown that auditory verbal ability measured by tasks like cloze and rapid spelling relates moderately to Gc as measured by visual tests. Importantly, these tasks do not emphasise auditory processing as is the case with SPuD, for instance. This is reflected in the absence of a significant loading of auditory verbal ability on Ga. Instead, these tasks rely on other cognitive areas of expertise like spelling and deconstructing disarranged sentences. Attempts to measure Gc with little confound from auditory or other processes should prove successful.

Glr, or long-term retrieval, is a memory type factor reflecting the ability to store information and accurately retrieve it after long durations. The WJ-III (Woodcock, et al., 2001) currently measures this construct using a visual-auditory learning task which requires participants to learn to associate rebus symbols with words, and later read aloud a sentence comprised only of these symbols. It therefore measures a complex interplay of auditory and visual processing and memory. It should, however, be possible to index Glr using a purely auditory task. For example, prior to beginning a large test battery participants could be presented with a series of auditory sounds including that of a car, a cat meowing, or a dog barking. The encoding of the stimuli into memory depends on the auditory modality and at different time intervals during the course of a longer test battery, the participant could be asked to recall the sounds heard earlier. The number of stimuli recalled correctly, even days later, would be an index of this construct. One problem with a test of this type would be familiarity of the sounds presented. Some sounds are more familiar than others and will be recalled more easily. Therefore, there should be control for this across the stimuli used. Attempts to measure Glr with little confound from auditory or other processes should prove successful.

Short-term memory, or Gsm, is related to working memory and reflects the ability to hold information in immediate awareness to be used within a short time period. Auditory Gsm is currently measured in the WJ-III battery by having participants repeat in reverse order, a series of auditorily dictated number sequences of increasing length. The task relies almost entirely on memory processes as opposed to auditory perceptual processes. Tests used previously by Horn and Stankov (Horn &

Stankov, 1982; Stankov & Horn, 1980) might also prove useful in the measurement of Gsm and include tonal figures and memory for emphasis. In the tonal figures task, participants select which of four choices of tone sequences matches one heard previously. As the tones used differ markedly in pitch, frequency discrimination does not appear to be a confounding factor. In the memory for emphasis task participants identify which of the words in a spoken discourse were markedly emphasised. Again, as long as the intensity of the emphasised words is marked, intensity discrimination should not confound performance. Attempts to measure Gsm with little confound from auditory or other processes should prove successful.

As explored in the present thesis Gs, or speed of processing, appears able to be measured via auditory pathways. Audio Code appears particularly promising as a measure of Gs, although it does involve a visual coding element reflecting that the reference code is presented visually on a computer screen. As is the case with Symbol Digit (McPherson & Burns, 2005), however, the code is quickly learned and the response becomes automatic. Thus, the visual coding aspect is reduced. Other potentially useful auditory Gs tasks that do not include a visual component include hearing letter As and tone comparisons. If the tone comparison task changed to frequency comparisons and the difference in frequencies throughout the test trials were reduced, the task would likely tend to relate better to Ga than to Gs. Thus, as already recommended, it is important to refrain from engaging complex perceptual discrimination processes. As evidenced herein, attempts to measure Gs with little confound from auditory or other processes should prove successful.

Two CHC constructs remain to be discussed and I regard them as being probably the most difficult to measure auditorily. It is likely that when designing tasks to index these constructs other factors will inadvertently be involved in performance on them. First, the construct broad visualisation, or Gv, is a complex ability thought to reflect visual perceptual abilities. Thus, intuitively speaking, it should not be possible to measure this ability via auditory pathways just like visual tasks should not measure Ga. However, there is evidence in the literature that Gv might not be modality specific. For example, Horn and Stankov (1982) have reported weak but significant loadings of auditory primary factors reflecting maintaining and judging rhythm and temporal tracking on broad Gv. Similarly, Roberts et al (1997) and Stankov et al (2001) have reported significant associations between tactile-kinesthesia abilities and Gv, and Danthiir et al (2001) have published evidence of a relationship between performance on olfactory tasks and established Gv marker tests. A plausible explanation of these relationships is that in many of these tasks a visual representation of presented stimuli is conjured by participants to aid task performance. For example, in a tactile object identification task, participants might visualise the object as they are attempting to identify, by touch only, what it is. Similarly, for rhythm and temporal tasks, participants may form visual representations of durations or rhythmic patterns to aid in solving them. Regardless, the loadings of tasks in alternate modalities on Gv are often weak and the tasks display stronger relationships to other broad constructs. One example is the tonal memory task which relates more strongly to memory and Ga constructs although it has previously related to Gv, presumably because participants visualise the tonal series as a sequence of steps (see e.g., Danthiir, et al., 2001). Thus, whilst it might be possible to find weak but consistent relationships between auditory tasks and Gv, attempts to solely measure Gv without other constructs confounding performance are probably not likely to be successful.

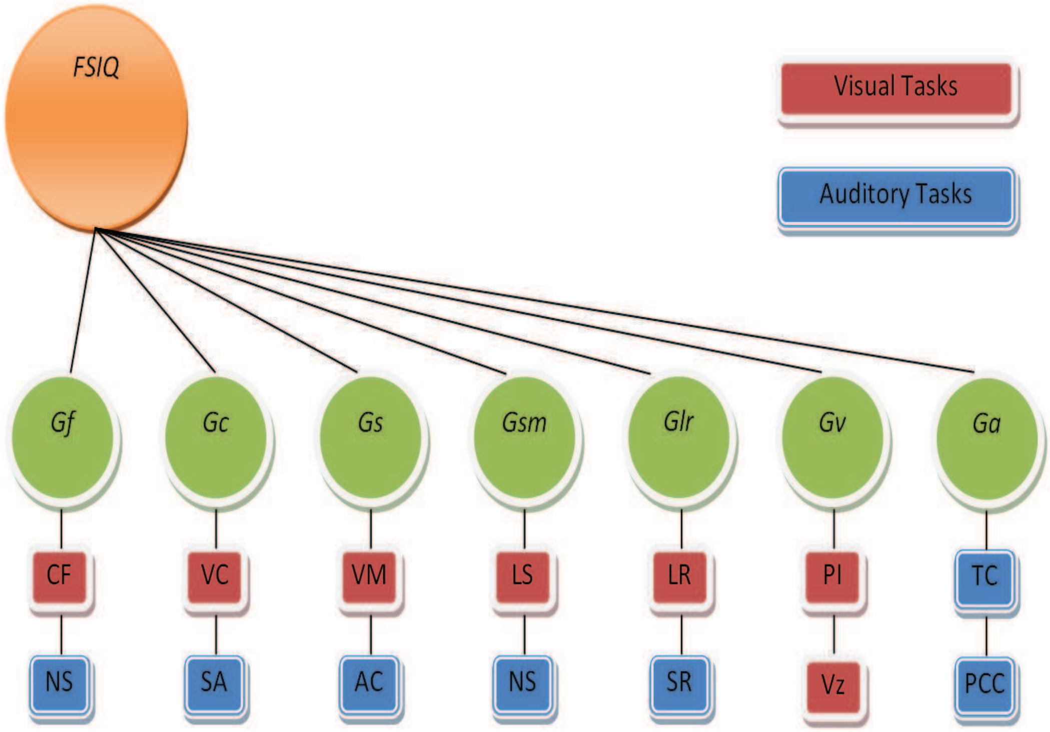
Fluid intelligence, or Gf, might also prove problematic to measure accurately via auditory pathways. Because this construct reflects complex reasoning abilities it is generally necessary to present rather complex stimuli and in the case of auditory tasks this can exacerbate the influence of Ga processes on task performance. For example, Stankov and Horn (1980) attempted to measure inductive reasoning and auditory cognition of relations with tasks such as chord matching, tonal classifications, tonal series and loudness series. These tasks required relatively demanding discriminations of the auditory specific stimuli presented and although they did relate weakly to Gf, they loaded primarily on Ga. These difficulties might be overcome by developing auditory tasks that require reasoning but do not rely on auditory abilities like intensity or frequency discrimination. It might be possible to adapt the visual number series task to the auditory modality simply by dictating sequences of numbers to participants and asking them to complete the series auditorily. For instance, on a given trial a participant could hear the number "30, 33.5, 37, 40.5 ..." and be required to identify what number comes next (i.e., 44). A potential confound for these tasks are memory processes that might involve themselves when the number sequence, for example, needs to be memorised in order to assist in reasoning and arriving at a correct solution. Ways of avoiding this confound might include allowing the participant to hear the sequence as many times as necessary rather than restrict them to hearing only one presentation of the sequence. Despite the difficulties inherent in measuring Gf auditorily, future attempts to measure this construct with little confound from auditory or other processes could prove successful if the nature of the auditory tasks is carefully considered and the tasks are appropriately designed.

As already noted, future studies focussing on auditory tasks and their measurement of the broad constructs just discussed are important to the study of intelligence for several reasons. One area that these studies could contribute is distinguishing Gf and Gv type abilities in order to understand them better. A common problem in factorial studies is that Gf and Gv tasks often combine to define a single latent reasoning, or Gf type factor. According to Lohman (2000) one explanation for this phenomenon reflects that most spatial tasks are solvable at least in part through the application of reasoning abilities. Similarly, many Gf tasks like RPM (Raven, et al., 1998), for example, also have a spatial visualisation component (Guttman, 1974). Thus, these overlapping requirements make it difficult to empirically separate these factors. If auditory tasks can be successfully designed to primarily index Gf then including such tasks in factorial studies may result in Gf and Gv factors becoming more distinguishable. In a similar vein to reducing the influence of auditory processes in auditory task performance, it would also be beneficial to reduce the dependence on spatial reasoning in visual Gf tasks, as this would seem more related to Gv. Alternative visual measures of Gf not requiring complex visualisation processes include tasks involving series completion, analogies, seriation and classifications (Lohman, 2000).

Further research on the auditory modality and its relation to existing broad constructs is also potentially useful in the research areas of cognitive ageing and neurodegenerative disease. For example, there exists the potential to establish cognitive marker tests to identify individuals at risk for less favourable future cognitive ageing (see e.g., Gregory, et al., 2011; Gregory, Nettelbeck, & Wilson, 2009). Auditory measures might prove useful in this regard as it has been proposed that both auditory and cognitive processes are affected by widespread neural degeneration, and that cognitive performance is compromised when the auditory system begins to fail (Gallacher, 2004). The auditory modality might be more sensitive to neural degeneration and therefore provide earlier indications of cognitive decline than tests in the visual modality and this possibility should be explored.

Similarly, auditory tasks might prove useful in exploring the cognitive affects of neurodegenerative disease. Neuropsychological assessment currently incorporates auditory tasks like the oddball task (see e.g., Walhovd & Fjell, 2001) but a key problem is the lack of adequate theory underpinning test development as well as interpretation of the processes they measure. Thus, developing auditory tasks on the basis of well established intelligence taxonomies like the CHC model will offer theoretically driven alternatives in these areas.

A final area which holds much promise for the development of auditory tests is the assessment of intelligence in the applied setting. As Carroll (1993) has noted, intelligence testing is not novel but cumbersome and based on the same techniques used over a century ago. Although the WJR-III incorporates a limited number of auditory tests, there is no reason why auditory tasks cannot be developed and used in addition to visual tasks to measure each of the broad constructs defined in CHC theory. Figure 10 provides a graphical illustration of what future test batteries could look like if researchers devote more attention to cognitive testing in the auditory modality. The only constructs not measured by a combination of tasks from both modalities are Gv and Ga because these factors should be considered to reflect largely perceptual processes. There are several advantages to developing a test battery similar to that illustrated. Some advantages include: 1) generating a better understanding of the expression of broad abilities in alternate modalities across the human lifespan; 2) possibly having greater potential to identify learning or other difficulties in educational settings if multiple modalities are incorporated into testing; 3) provide a more novel testing session; and 4) achieve arguably more ecologically valid intelligence estimates given the inclusion of multiple modalities.



FSIQ = Full Scale IQ; Gf = Fluid Ability; Gc = Crystallised Ability; Gs = Speed of Processing; Gsm = Short Term Memory; Glr = Long-term Storage and Retrieval; Gv = Visualisation Ability; Ga = Broad Auditory Reception; CF = Concept Formation; NS = Number Series; VC = Verbal Comprehension; SA = Synonyms Antonyms; VM = Visual Matching; AC = Audio Code; LS = Letter Span; NS = Number Span; LR = List Recall; SR = Sound Recognition; PI = Picture Identification; Vz = Visualisation; TC = Tonal Classifications; PC = Pitch Change in Chords.

Figure 10. Graphical representation of a possible future intelligence test battery

Concluding statement

Much of what is known of the nature of human intelligence reflects research concerned primarily with visually presented tests. This thesis, on the other hand, sought to explore the expression of intellectual abilities in an alternate modality. Results of the studies presented herein suggest that cognitive testing in the auditory domain – even when the focus is on existing well established abilities – holds significant potential for furthering our understanding of intellectual functioning. It should be stressed that this thesis represents only a preliminary step forward. Previous studies of the auditory modality and its relation to human intelligence have not been long lived. Thus, it is important that differential psychologists make use of advances in technology and both continue and encourage the study of the expression of intellectual abilities in all of our modalities; olfaction, gustation and tactition included.

Future studies in the area of auditory abilities and indeed all of our other modalities must ground themselves adequately in existing theory. This will be fundamental to the success of this movement. Herein I have used the CHC taxonomy as a basis for understanding the processes underpinning auditory performance, and for guiding the development of auditory tests. Elsewhere, however, as already highlighted, researchers have generally failed to consider their findings in relation to existing data regarding performance in the auditory modality (i.e., temporal discrimination and auditory IT studies). This only serves to hamper advances in this area because conclusions regarding the nature of the processes measured by these tasks are often erroneous.

As Carroll (1993) has indicated elsewhere, the taxonomy of intellectual abilities is far from complete. It is lacking in its understanding of abilities in modalities other than visual, as well as how the existing broad cognitive abilities are expressed within each of these. Well designed and detailed research programmes are necessary to overcome these problems. Only once this is achieved will we arrive at a truly balanced understanding of human intelligence (Horn, 1968).

References

Acton, G. S., & Schroeder, D. H. (2001). Sensory discrimination as related to general intelligence. *Intelligence, 29*, 263-271.

ASVAB. (2000). *ASVAB : Armed Services Vocational Aptitude Battery* (2nd ed.). New York: Learning Express.

Barrett, P. (2007). Structural equation modelling: Adjudging model fit. *Personality and Individual Differences, 42*, 815-824.

Bates, T. C. (2005). Auditory inspection time and intelligence. *Personality and Individual Differences, 38*, 115-127.

Bentley, A. (1966). *Measures of musical abilities*. New York: October House Inc.

Binet, A., & Henri, V. (1896). La psychologie individuelle. *L'Annee Psychologique, 2*(411-465).

Binet, A., & Simon, T. (1905). Methodes nouvelles pour le diagnostic du niveau intellectuel des anomaux. *L'Annke psychologique, 11*, 191-244.

Bors, D. A., & Stokes, T. L. (1998). Raven's Advanced Progressive Matrices: norms for first-year university students and the development of a short from. *Educational and Psychological Measurement, 58*, 382-399.

Brand, C., & Deary, I. J. (1982). Intelligence and 'inspection time'. In E. H.J. (Ed.), *A model for intelligence* (pp. 133-148). Berlin: Springer-Verlag.

Brebner, J. T., & Welford, A. T. (1980). Introduction: an historical background sketch. In A. T. Welford (Ed.), *Reaction times* (pp. 1-23). New York: Academic Press.

Breitmeyer, B. G. (2007). Visual masking: past accomplishments, present status, future developments. *Advances in Cognitive Psychology, 3*, 9-20.

Breitmeyer, B. G., & Ogmen, H. (2000). Recent models and findings in visual backward masking: A comparison, review, and update. *Perception & Psychophysics, 62*, 1572-1595.

Brody, N. (2000). History of Theories and Measurements of Intelligence. In R. J. Sternberg (Ed.), *Handbook of Intelligence* (pp. 16-33). Cambridge; New York: Cambridge University Press.

Burns, N. R. (1998). *Inspection time and cognitive abilities: An event-related potential study.* Unpublished PhD Thesis, University of Adelaide, South Australia (Unpublished PhD thesis).

Burns, N. R., & Nettelbeck, T. (2003). Inspection time in the structure of cognitive abilities: Where does IT fit? *Intelligence, 31*, 237-255.

Burns, N. R., & Nettelbeck, T. (2005). Inspection time and speed of processing: Sex differences on perceptual speed but not IT. *Personality and Individual Differences, 39*(2), 439-446.

Burns, N. R., & Nettelbeck, T. (2008, December 11-13). *Inspection time in the structure of cognitive abilities: more on where IT might fit.* Paper presented at the 9th Annual International Society for Intelligence Research Conference, Decatur, GA, USA, December 11-13.

Burns, N. R., Nettelbeck, T., & Cooper, C. (2000). Event related potential correlates of some human cognitive ability constructs. *Personality and Individual Differences, 29*, 157-168.

Burns, N. R., Nettelbeck, T., & Cooper, C. J. (1999). Inspection time correlates with general speed of processing but not fluid ability. *Intelligence, 27*, 37-44.

Burns, N. R., Nettelbeck, T., & McPherson, J. (2009). Attention and intelligence: A factor analytic study. *Journal of Individual Differences, 30*, 44-57.

Burns, N. R., Nettelbeck, T., McPherson, J., & Stankov, L. (2007). Perceptual learing on inspection time and motion perception. *Journal of General Psychology, 134*(1), 83-100.

Burns, N. R., Nettelbeck, T., & White, M. (1998). Testing the interpretation of inspection time as a measure of speed of sensory processing. *Personality and Individual Differences, 24*, 25-39.

Burt, C. (1909/1910). Experimental tests of general intelligence *British Journal of Psychology, 3*, 94-177.

Buttsworth, L., Fogarty, G., & Rorke, P. (1993). Predicting aural performance in a tertiary music training programme. *Psychology of Music, 21*, 114-126.

Campbell, L. J. (1995). *Intelligence, arousal, inspection time and flicker: Evidence for a basic processing speed factor?* Honours Thesis, Aberdeen University, Aberdeen.

Carroll, J. B. (1992). Cognitive Abilities: The State of the Art. *Psychological Science, 3*(5), 266-270.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge; New York: Cambridge University Press.

Cattell, R. B. (1941). Some theoretical issues in adult intelligence testng. *Psychological Bulletin, 38*, 592.

Cattell, R. B. (1943). The measurement of adult intelligence. *Psychological Bulletin, 40*, 153-193.

Cattell, R. B. (1961). *Culture free intelligence test, scale 3*. Champaign, IL: Institute for Personality and Ability Testing.

Cattell, R. B. (1963). Theory for fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology, 54*, 1-22.

Cooper, C. (1999). *Intelligence and abilities*. London: Routledge.

Crawford, J., & Stankov, L. (1983). Fluid and crystallized intelligence and primacy/recency componenets of short-term memory. *Intelligence, 7*, 227-252.

Danthiir, V. (2005). *Speed on simple and complex tasks: Structure and validity within a framework of cognitive abilities.* PhD, Humboldt-University, Berlin.

Danthiir, V., Roberts, R. D., Pallier, G., & Stankov, L. (2001). What the nose knows Olfaction and cognitive abilities. *Intelligence, 29*, 337-361.

Danthiir, V., Roberts, R. D., Schulze, R., & Wilhelm, O. (2005). Mental speed: On frameworks, paradigms, and a platform for the future. In O. Wilhelm & R. W. Engle (Eds.), *Handbook of understanding and measuring intelligence*. Thousand Oaks California: Sage.

Danthiir, V., Wilhelm, O., Schulze, R., & Roberts, R. D. (2005). Factor structure and validity of paper-and-pencil measures of mental speed: Evidence for a higher-order model? *Intelligence, 33*, 491-514.

Deary, I. J. (1992). *Auditory inspection time and intelligence.* doctoral dissertation, Univeristy of Edinburgh.

Deary, I. J. (1994). Intelligence and auditory discrimination: separating processing speed and fidelity of stimulus representation. *Intelligence, 18*, 189-213.

Deary, I. J. (2000a). *Looking down on human intelligence; from psychometrics to the brain*. New York: Oxford University Press.

Deary, I. J. (2000b). Simple information processing and intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 267-284). Cambride; New York: Cambridge University Press.

Deary, I. J. (2003). Reaction time and psychometric intelligence: Jensen's contributions. In H. Nyborg (Ed.), *The scientific study of general intelligence: tribute to Arthur R. Jensen* (pp. 53-76). Amsterdam; Boston: Permagon.

Deary, I. J., Bell, P. J., Bell, A. J., Campbell, M. L., & Fazal, N. D. (2004). Sensory discrimination and intelligence: Testing Spearman's other hypothesis. *The American Journal of Psychology, 117*, 1-18.

Deary, I. J., Caryl, P. G., Egan, V., & Wight, D. (1989). Visual and auditory inspection time: Their interrelationship and correlations with IQ in high ability subjects. *Personality and Individual Differences, 10*, 525-533.

Deary, I. J., Head, B., & Egan, V. (1989). Auditory inspection time, intelligence and pitch discrimination. *Intelligence, 13*, 135-147.

Downing, G. (1965). The preparation of teachers for schools in culturally deprived neighborhoods (The Bridge Project): Final Report.

Drake, R. M. (1933a). Four new tests of musical talent. *Journal of Applied Psychology, 17*, 136-147.

Drake, R. M. (1933b). The validity and reliability of tests of musical talent. *Journal of Applied Psychology, 17*, 447-458.

Duncan, C. C., Kosmidis, M. H., & Mirsky, A. F. (2005). Closed head injury-related information processing deficits: An event-related potential analysis. *International Journal of Psychophysiology, 58*(2-3), 133-157.

Ebbinghaus, H. (1897). Uber eine neue Methode zur Prufung geistiger Fahigkeiten und ihre Anwendung bei Schulkindern. *Zeitschrift fur angewandte Psychologie, 13*(401-459).

Ekstrom, R. B., French, J. W., Harman, H. H., & Derman, D. (1976). *Kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service.

Elbert, T., Ulrich, R., Rockstroth, B., & Lutzenberger, W. (1991). The processing of temporal intervals reflected by CNV-loke brain intervals. *Psychophysiology, 28*, 648-655.

Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions in the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102-134). Cambridge: Cambridge University Press.

Evans, G., & Nettelbeck, T. (1993). Inspection time: A flash mask to reduce apparent movement effects. *Personality and Individual Differences, 15*, 91-94.

Fogarty, G., & Stankov, L. (1988). Abilities involved in performance on competing tasks. *Personality and Individual Differences, 9*, 35-49.

Gallacher, J. (2004). Hearing, cognitive impairment and aging: a critical review. *Reviews in Clinical Gerontology, 14*, 199-209.

Galton, F. (1865). Hereditary talent and character. *Macmillan's Magazine, 12*, 157-166.

Galton, F. (1892). *Hereditary genius: an inquiry into its laws and consequences* (2nd ed.). London: Macmillan and Co.

Gibbons, H., Brandler, S., & Rammsayer, T. H. (2002). Dissociating aspects of temporal and frequency processing: A functional ERP in humans. *Cortex, 39*, 947-965.

Gordon, E. E. (1965). *Musical Aptitude Profile*. Boston: Houghton-Mifflin.

Gordon, E. E. (1989). *Advanced measures of music audiation*. Chicago: GIA publications.

Gregory, T., Nettelbeck, T., Burns, N. R., Danthiir, V., Wilson, C., & Wittert, G. (2011). Sex differences in inspection time with age. *Personality and Individual Differences, 50*, 593-596.

Gregory, T., Nettelbeck, T., & Wilson, C. (2009). Inspection time and everyday functioning: A longitudinal study. *Personality and Individual Differences, 47*, 999-1002.

Grudnik, J. L., & Kranzler, J. H. (2001). Meta-analysis of the relationship between inspection intelligence and inspection time. *Intelligence, 29*, 523-535.

Guttman, R. (1974). Genetic analysis of analytical spatial ability: Raven's progressive matrices. *Behavior Genetics, 4*, 273-284.

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis*. New Jersey: Prentice-Hall.

Harrington, D. L., Haaland, K. Y., & Knight, R. T. (1998). Cortical networks underlying mechanisms of time perception. *Journal of Neuroscience, 18*, 1085-1095.

Helmbold, N., & Rammsayer, T. H. (2006). Timing performance as a predictor of psychometric intelligence as measured by speed and power tests. *Journal of Individual Differences, 27*, 20-37.

Helmbold, N., Troche, S. J., & Rammsayer, T. H. (2006). Temporal Information Processing and Pitch Discrimination as Predictors of General Intelligence. *Canadian Journal of Experimental Psychology, 60*, 294-306.

Helmbold, N., Troche, S. J., & Rammsayer, T. H. (2007). Processing of temporal and nontemporal information as predictors of psychometric intelligence: A structural-equation-modeling approach. *Journal of Personality, 75*, 985-1005.

Hendrickson, A. E. (1982). The Biological Basis of Intelligence Part I: Theory. In H. J. Eysenck (Ed.), *A Model for Intelligence* (pp. 151-196). Berlin: Springer-Verlag.

Hendrickson, D. E. (1982). The Biological Basis of Intelligence Part II: Measurement. In H. J. Eysenck (Ed.), *A Model for Intelligence* (pp. 197-230). Berlin: Springer-Verlag.

Hick, W. E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology, 4*, 11-26.

Holmes, J. (1954). A substrata analysis of spelling ability for elements of auditory images. *Journal of Experimental Education, 22*, 329-349.

Holmstrom, L. G. (1969). Intelligence v. progress in musical education. *Journal of Research in Music Education, 17*, 76-81.

Horn, J. L. (1968). Organization of abilities and the development of intelligence. *Psychological Review, 75*, 242-259.

Horn, J. L. (1988). Thinking about human abilities. In J. R. Nesselroade & R. B. Cattell (Eds.), *Handbook of multivariate experimental psychology* (Second edition ed., pp. 645-685). New York: Plenum.

Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized intelligences. *Journal of Educational Psychology, 57*, 253-270.

Horn, J. L., & Noll, J. (1994). A system for understanding cognitive capabilities: A theory and the evidence on which it is based. In D. K. Detterman (Ed.), *Current topics in human intelligence: Theories of intelligence* (4 ed., pp. 151-203). Norwood NJ: Ablex.

Horn, J. L., & Stankov, L. (1982). Auditory and visual factors of intelligence. *Intelligence, 6*(2), 165-185.

Horn, W. (1983). *Leistungsprufsystem*. Göttingen, Germany: Hogrefe.

Irwin, R. J. (1984). Inspection time and its relation to intelligence. *Intelligence, 8*, 47-65.

Jaaskelainen, I. P., Hautamake, M., Naatanen, R., & Ilmoniemi, R. J. (1999). Temporal span of human echoic memory and mismatch negativity: revisited. *Neuroreport, 10*, 1305-1308.

Jäger, A. O., Süβ, H.-M., & Beauducel, A. (1997). *Berliner intelligenzstruktur test form 4*. Göttingen, Germany Hogrefe.

Jausovec, N., & Jausovec, K. (2000). Correlations between ERP parameters and intelligence: a reconsideration. *Biological Psychology, 50*, 137-154.

Jensen, A. R. (1982). Reaction time and psycyometric g. In H. J. Eysenck (Ed.), *A model for intelligence* (pp. 93-132). Berlin: Springer.

Jensen, A. R. (1998). Jensen on "Jensinism". *Intelligence, 26*, 181-208.

Jensen, A. R. (2005). Mental chronometry and the unification of differential psychology. In R. J. Sternberg & J. E. Pretz (Eds.), *Cognition and Intelligence: Identifying the mechanisms of the mind* (pp. 26-50). Cambridge: Cambridge University Press.

Jensen, A. R. (2006). *Clocking the mind: Mental chronometry and individual differences*. Amsterdam: Elsevier.

Kidd, R. G., Watson, C. S., & Gygi, B. (2007). Individual differences in auditory abilities. *Acoustical Society of America, 122*, 418-435.

Kranzler, J. H., & Jensen, A. R. (1989). Inspection time and intelligence: a meta-analysis. *Intelligence, 13*, 329-347.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence, 14*, 389-433.

Langsford, P. B., Mackenzie, B. D., & Maher, D. P. (1994). Auditory inspection time, sustained attention, and the fundamentality of mental speed. *Personality and Individual Differences, 16*, 487-497.

Lohman, D. F. (2000). Complex information processing and intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 285-340). Cambride; New York: Cambridge University Press.

Macpherson, E. A., & Middlebrooks, J. C. (2000). Localization of brief sounds: effects of level and background noise. *Journal of the Acoustical Society of America, 108*, 1834-1849.

McCrory, C., & Cooper, C. (2005). The relationship between three auditory inspection time tasks and general intelligence. *Personality and Individual Differences, 38*, 1835-1845.

McGrew, K. S. (2005). The Cattell–Horn–Carroll theory of cognitive abilities. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (2 ed., pp. 136-181). New York: Guilford Press.

McGrew, K. S., & Evans, J. (2004). Carroll Human Cognitive Abilities Project: Research Report No. 2. *Internal and external factorial extensions to the Cattell-Horn-Carroll (CHC) theory of cognitive abilities: A review of factor analytic research since Carroll's seminal 1993 treatise*. St. Could, MN: Institute for Applied Psychometrics.

McLeish, J. (1950). The validation of Seashore's measures of musical talent by factorial methods. *British Journal of Psychology (Statistical Section), 3*, 129-140.

McPherson, J., & Burns, N. R. (2005). A speeded coding task using a computer-based mouse response. *Behavior Research Methods, 37*(3), 538-544.

Michon, J. A. (1990). Implicit and explicit representations of time. In R. A. Block (Ed.), *Cognitive models of psychological time* (pp. 37-58). Hillsdale, NJ: Lawrence Erlbaum Associates.

Muthen, L. K., & Muthen, B. O. (1998). *Mplus. The comprehensive modelling program for applied research. User's guide.* Los Angeles: Muthen and Muthen.

Muthen, L. K., & Muthen, B. O. (2007). *Mplus User's guide.* (5th ed.). Los Angeles, CA: Muthen and Muthen.

Nettelbeck, T. (1987). Inspection time and intelligence. In P. A. Vernon (Ed.), *Speed of information processing and intelligence* (pp. 295-346). Norwood, NJ: Ablex.

Nettelbeck, T. (2001). Correlation between inspection time and psychometric abilities; A personal interpretation. *Intelligence, 29*, 459-474.

Nettelbeck, T. (2003). Inspection Time and g. In H. Nyborg (Ed.), *The scientific study of general intelligence: tribute to Arthur R. Jensen* (pp. 77-92). Amsterdam; Boston: Permagon.

Nettelbeck, T., Edwards, C., & Vreugdenhil, A. (1986). Inspection time and IQ: evidence for a mental speed-ability association. *Personality and Individual Differences, 7*, 633-641.

Neubauer, A. C., & Fink, A. (2005). Basic information processing and the psychophysiology of intelligence. In R. J. Sternberg & J. E. Pretz (Eds.), *Cognition and Intelligence: Identifying the mechanisms of the mind* (pp. 68-87). Cambridge: Cambridge University Press.

Norman, G. R., & Streiner, D. L. (2000). *Biostatistics: The bare essentials* (2nd ed.). Hamilton, London: B.C. Decker.

O'Connor, T. A., & Burns, N. R. (2003). Inspection time and general speed of processing. *Personality and Individual Differences, 35*, 713-724.

Olsson, H., Bjorkman, C., Haag, K., & Juslin, P. (1998). Auditory inspection time: On the importance of selecting the appropriate continuum. *Personality and Individual Differences, 25*, 627-634.

Parker, D. M., Crawford, J. R., & Stephen, E. (1999). Auditory Inspection Time and Intelligence: A New Spatial Localisation Task. *Intelligence, 27*, 131-139.

Pelletier, H. W. (1963). *An investigation of the relation between training in instrumental music and selected aspects of language growth in third grade children.* Ed.D., Arizona State University.

Petrill, S. A., Luo, D., Thompson, L. A., & Detterman, D. K. (2001). Inspection time and the relationship among elementary cognitive tasks, general intelligence, and specific cognitive abilities. *Intelligence, 29*, 487-496.

Pfefferbaum, A., Ford, J. M., White, P. M., & Roth, W. T. (1989). P3 in Schizophrenia is affected by stimulus modality, response requirements, medication status, and negative symptoms. *Archives of General Psychiatry, 46*, 1035-1044.

Poon, P. W. F., Yu, W. Y., & Chan, J. W. C. (1986). Correlation between auditory reaction time and intelligence. *Perceptual and Motor Skills, 63*, 375-378.

Rammsayer, T. H., & Brandler, S. (2002). On the relationship between general fluid intelligence and psychophysical indicators of temporal resolution in the brain. *Journal of Research in Personality, 36*, 507-530.

Rammsayer, T. H., & Brandler, S. (2004). Aspects of temporal information processing: A dimensional analysis. *Psychological Research, 69*, 115-123.

Rammsayer, T. H., & Brandler, S. (2007). Performance on temporal information processing as an index of general intelligence. *Intelligence, 35*, 123-139.

Raven, J. C., Raven, J. E., & Court, J. H. (1998). *Progressive matrices*. Oxford: Oxford Psychologists Press.

Raz, N., Willerman, L., & Yama, M. (1987). On sense and senses: intelligence and auditory information processing. *Personality and Individual Differences, 8*, 201-210.

Reed, T., & Jensen, A. (1993). Choice reaction time and visual pathway nerve conduction velocity both correlate with intelligence but appear not to correlate with each other: Implications for information processing. *Intelligence, 17*, 191-203.

Roberts, R. D., Pallier, G., & Goff, G. N. (1999). Sensory processes within the structure of human cognitive abilities. In P. L. Ackerman, P. C. Kyllonen & R. D. Roberts (Eds.), *Learning and individual differences: process, trait, and content determinants*. Washington: American Psychological Association.

Roberts, R. D., & Stankov, L. (1999). Individual differences in speed of mental processing and human cognitive abilities: Toward a taxonomic model. *Learning and Individual Differences, 11*, 1-120.

Roberts, R. D., Stankov, L., Pallier, G., & Bradley, D. (1997). Charting the cognitive sphere: Tactile-kinesthetic performance within the structure of intelligence. *intelligence, 25*, 111-148.

Roederer, J. G. (1973). *Introduction to the physics and psychophysics of music*. London: English Universities Press; Springer-Verlag.

Rust, J., & Golombok, S. (2009). *Modern psychometrics : the science of psychological assessment* (3rd ed.). Hove, East Sussex ; New York: Routledge.

Seashore, C. E. (1919). *The psychology of musical talent*. New York: Silver Burdett.

Seashore, C. E., Lewis, D., & Saetveit, J. C. (1960). *Manual of instruction and interpretations for the Seashore Measures of Musical Talents* (2nd ed.). New York: The Psychological Cooperation.

Seidel, K. (2005). *The auditory abilities: A neglected domain in intelligence models and tests*. Paper presented at the International Society for the Study of Individual Differences, Adelaide, South Australia, Adelaide.

Seidel, K. (2008). *Assesment of social and auditory intelligence: New persepctives and approaches*. Lengerich: Pabst Science Publishers.

Sen, A., Jensen, A. R., Sen, A. K., & Arora, I. (1983). Correlation between reaction time and intelligence in psychometrically similar groups in america and india. *Applied Research in Mental Retardation, 4*, 139-152.

Sheppard, L. D., & Vernon, P. A. (2008). Intelligence and speed of information-processing: A review of 50 years of research. *Intelligence, 44*, 535-551.

Shuter-Dyson, R. (1968). *The psychology of musical ability*. London: Methuen.

Sloutsky, V. M., & Napolitano, A. C. (2003). Is a picture worth a thousand words? Preference for auditory modality in young children. *Child Development, 74*, 822-833.

Sparrow, W. A., Bradshaw, E. J., Lamoureux, E., & Tirosh, O. (2002). Ageing effects on the attention demands of walking. *Human Movement Science, 21*, 961-972.

Spearman, C. (1904). "General Intelligence," objectively determined and measured. *American Journal of Psychology, 15*, 201-293.

Spearman, C. (1923). *The nature of intelligence and the principles of cognition*. London: Macmillan.

Spearman, C. (1927). *The abilities of man: Their nature and measurement*. London: Macmillan.

Stålnacke, B.-M., Elgh, E., & Sojka, P. (2007). One-year follow-up of mild traumatic brain injury: Cognition, disability and life satisfaction of patients seekeing consultation. *Journal of Rehabilitation Medicine, 39*, 405-411.

Stankov, L. (1971). *Hierarchical structure of auditory abilities and relationship between auditory and visual modalities.* doctorol dissertation, University of Denver.

Stankov, L. (1978). Fluid and crystallized intelligence and broad perceptual factors among 11 to 12 year olds. *Journal of Educational Psychology, 70*, 324-334.

Stankov, L. (1983). *The role of competition in human abilities revealed through auditory tests*. Sydney, Australia: Society of Multivariate Experimental Psychology.

Stankov, L. (1994a). Auditory abilities. In R. J. Sternberg (Ed.), *Encyclopedia of human intelligence* (pp. 157-162). New York: Macmillan.

Stankov, L. (1994b). The complexity effect phenomenon is an epiphenomenon of age-related fluid intelligence decline. *Personality and Individual Differences, 16*(2), 265-288.

Stankov, L. (2005). Reductionism versus charting. In R. J. Sternberg & J. E. Pretz (Eds.), *Cognition and Intelligence: Identifying the mechanisms of the mind* (pp. 51-67). Cambridge: Cambridge University Press.

Stankov, L., & Horn, J. L. (1980). Human Abilities Revealed Through Auditory Tests. *Journal of Educational Psychology, 72*(1), 21-44.

Stankov, L., Horn, J. L., & Roy, T. (1980). On the relationship between Gf/Gc theory and Jensen's Level I/Level II theory. *Journal of Educational Psychology, 72*, 796-809.

Stankov, L., Seizova-Cajic, T., & Roberts, R. D. (2001). Tactile and kinesthetic perceptual processes within the taxonomy of human cognitive abilities. *Intelligence, 29*, 1-29.

Sternberg, R. J. (2000). The Concept of Intelligence. In R. J. Sternberg (Ed.), *Handbook of Intelligence* (pp. 3-15). Cambridge; New York: Cambridge University Press.

Suprenant, A. M., & Watson, C. S. (2001). Individual differences in the processing of speech and nonspech sounds by normal hearing listeners. *The Journal of the Acoustical Society of America, 110*, 2085-2095.

Surwillo, W. W. (1968). Timing of Behaviour in senescence and the role of the central nervous system. In T. G.A. (Ed.), *Human aging and behavior* (pp. 1-35). New York: Academic Press.

Swani, P. (1995). *A study of Projective and Chronometric Correlates of Fluid and Crstallized Intelligence.* PhD Doctoral Dissertation, Guru Nanak Dev University, Amritsar, Punjab.

Teng, S., & Whitney, D. (2008). Position discrimination of auditory stimuli in early visual cortex. *Journal of Vision, 8*. doi: 10.1167/8.6.166

Thurstone, L. L. (1931). Multiple Factor Analysis. *Psychological Review, 38*, 406-427.

Thurstone, L. L. (1936). The factorial isolation of primary abilities. *Psychometrika, 1*(3), 175-182.

Thurstone, L. L. (1938). *Primary mental abilities*. Chicago, IL: University of Chicago Press.

Troche, S. J., & Rammsayer, T. H. (2009a). The influence of temporal resolution power and working memory capacity on psychometric intelligence. *Intelligence, 37*, 479-486.

Troche, S. J., & Rammsayer, T. H. (2009b). Temporal and non-temporal sensory discrimination and their predictions of capacity- and speed-related aspects of psychometric intelligence. *Personality and Individual Differences, 47*, 52-57.

Vernon, M. (2005). Fifty years of research on the intelligence of deaf and hard-of-hearing children: A review of literature and discussion of implications. *Journal of deaf studies and deaf education, 10*, 225-231.

Vernon, P. A. (1993). Intelligence and Neural Efficiency. In D. K. Detterman (Ed.), *Current Topics in Human Intelligence* (Vol. 3, pp. 171-188). New Jersey: Ablex Publishing Corporation.

Vickers, D., Nettelbeck, T., & Willson, R. J. (1972). Perceptual indices of performance: The measurement of 'inspection time' and 'noise' in the visual system. *Perception, 1*, 263-295.

Walhovd, K., & Fjell, A. M. (2001). Two- and three-stimuli auditory oddball ERP tasks and neuropsychological measures in aging. *Neuroreport, 12*, 3149-3153.

Watson, C. S., Johnson, D. M., Lehman, J. R., Kelly, W. J., & Jensen, J. K. (1982). An auditory discrimination test battery. *The Journal of the Acoustical Society of America, Suppl., 171*, 73.

Watson, C. S., & Miller, T. K. (1993). Auditory perception, phonological processing, and reading ability / disability. *Journal of Speech and Hearing Research, 36*, 850-863.

Weiss, R. H. (1971). *Grundintelligenztest CFT 3 skala 3.* Braunschweig, Germany: Westermann.

Wetherill, G. B., & Levitt, H. (1965). Sequential estimation of points on a psychometric function. *British Journal of Mathematical and Statistical Psychology, 18*, 1-10.

White, M. (1996). Interpreting inspection time as a measure of the speed of sensory processing. *Personality and Individual Differences, 20*, 351-363.

Whyte, J., Polansky, M., Fleming, M., Coslett, H. B., & Cavallucci, C. (1995). Sustained arousal and attention after traumatic brain injury. *Neuropsychologia, 33*(7), 797-813.

Wing, H. D. (1948). Tests of musical ability and appreciation. *British Journal of Psychology Monograph Supplement, 27*, 88.

Wing, H. D. (1962). A revision of the Wing Musical Aptitude Test. *Journal of research in Musical Education, 10*, 743-791.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III*. Itasca IL: Riverside Publishing.

Yerkes, R. M. (1921). Psychological examining in the United States Army. *Memoirs, National Academy of Sciences, 15*.

Zajac, I. T., & Burns, N. R. (2007). Measuring auditory inspection time in primary school children. *Journal of Individual Differences, 28*, 45-52.

Zajac, I. T., & Burns, N. R. (2009). *Do auditory temporal discrimination tasks measure temporal resolution of the CNS?* Paper presented at the International Society for the Study of Individual Differences. Evanston, Illinois, USA. July 18-22, Evanston, Illinois, USA.

Zajac, I. T., & Burns, N. R. (2011). Relationships between three auditory inspection time tasks and processing speed. *Australian Journal of Psychology*, 63, 163-172. doi: 10.1111/j.1742-9536.2011.00020.x

Zajac, I. T., Burns, N. R., Danthiir, V., & Nettelbeck, T. (2009). *Can auditory tests measure general speediness?* Paper presented at the International Society for the Study of Individual Differences. Evanston, Illinois, USA. July 18-22, Evanston, Illinois, USA.

Appendices

Appendix 1

Frequency distribution of stimuli in the stimuli set presented within the Dissociation Paradigm

	Dimension 1	Stimulus Duration							
Dimension 2		125 ms	150 ms	175 ms	200 ms (*S*)	225 ms	250 ms	275 ms	∑
Auditory (Pitch)	964 Hz	1	1	1	3	1	1	1	*9*
	976 Hz	1	1	1	3	1	1	1	*9*
	988 Hz	1	1	1	3	1	1	1	*9*
	1000 Hz (*S*)	3	3	3	9	3	3	3	*27*
	1012 Hz	1	1	1	3	1	1	1	*9*
	1024 Hz	1	1	1	3	1	1	1	*9*
	1036 Hz	1	1	1	3	1	1	1	*9*
	∑	*9*	*9*	*9*	*27*	*9*	*9*	*9*	*81*
Visual (Line Length)	7 cm	1	1	1	3	1	1	1	*9*
	8 cm	1	1	1	3	1	1	1	*9*
	9 cm	1	1	1	3	1	1	1	*9*
	10 cm (*S*)	3	3	3	9	3	3	3	*27*
	11 cm	1	1	1	3	1	1	1	*9*
	12 cm	1	1	1	3	1	1	1	*9*
	13 cm	1	1	1	3	1	1	1	*9*
	∑	*9*	*9*	*9*	*27*	*9*	*9*	*9*	*81*

S = Standard

Paper Reprints

Zajac, I.T. & Burns, N.R. (2011) Do auditory temporal discrimination tasks measure temporal resolution of the CNS?
Psychology, v. 2 (7), pp. 743-753

NOTE:
This publication is included in the print copy
of the thesis held in the University of Adelaide Library.

It is also available online to authorised users at:

http://dx.doi.org/10.4236/psych.2011.27114

Zajac, I.T. & Nicholas R. Burns, N.R, (2011) Relationships between three auditory inspection time tasks and processing speed.
Australian Journal of Psychology, v. 63 (3), pp. 163-172