Chapter 1. Externalising Behaviour Problems and Parenting Behaviour

Externalising behaviour problems (such as noncompliance, aggression, antisocial behaviour and poor impulse control) in preschool children are a public health concern because of their high prevalence, their high degree of persistence over time, and their association with poor long-term outcomes such as learning and conduct disorders (Hinshaw, 2002; Johnson, Cohen, Kasen, Smailes, & Brook, 2001; Marshall & Watt, 1999; Sanson, Prior, & Smart, 1996). The preschool period is a time of development when children individuate and assert their independence. Parental expectations about compliance are reflected in limit setting and attempts to control children's behaviour (Cunningham & Boyle, 2002). For several decades, research into childhood externalising behaviour problems has focused on the role of parent-child interactions and parental disciplinary practices as causal agents in the development and persistence of externalising problems in preschool children. Thus, parenting behaviour is often identified as a focus for interventions designed to prevent or ameliorate behaviour problems in this age group (Bor et al., 2002; Kazdin, 1997; Neary & Eyberg, 2002; Sanders & Dadds, 1993). Given that the focus of this thesis is the *measurement* of parenting behaviours of preschool-aged children, it is necessary first to obtain an overview of why the particular parenting behaviours examined herein are an important focus for research. The body of literature relevant to externalising behaviour and parenting behaviour is vast and this chapter is designed to provide only a general overview of externalising behaviour problems and parenting behaviours implicated in their development. Chapter 2 will examine the literature relevant to the measurement of these parenting behaviours.

1.1 Externalising Behaviour Problems

The child and adolescent component of the National Survey of Mental Health and Well-being reported that delinquent and aggressive behaviour problems, along with attention and self-regulation difficulties, were some of the most frequently occurring problems in Australian children (Sawyer et al., 2000). Disruptive behaviour disorders are also the main cause of referrals to child and adolescent mental health services (Kazdin, 1995). This section describes contemporary theories about the aetiology of these problems, their prevalence, levels of stability and the financial costs associated with early childhood externalising behaviour problems.

1.1.1 What are externalising behaviour problems?

The ability to engage in self-control (such as the ability to plan, delay gratification, and evaluate and self-regulate problem-solving activities and attention to tasks) is important in young children (Murray & Kochanska, 2002; Normandeau & Guay, 1998). Children with externalising behaviour problems lack the ability to control, inhibit or regulate certain aspects of their behaviour. Externalising problems ("acting out behaviours") exhibited by preschool children include noncompliance, difficulty controlling aggression, antisocial behaviour, overactivity, poor impulse control, tantrums, poor peer relationships and conflict about independence (Hinshaw, 2002). Noncompliance with requests or instructions of caregivers is one of the most common reasons for referral to agencies providing help for children with behaviour disorders (Ducharme, Popynick, Pontes, & Steele, 1996). Whilst externalising behaviours may be annoying or distressing to adults, it is common for preschool children (approximately 20%) to display these behaviours in a transient fashion, as part of their development and individuation, or as a result of age-appropriate conflict and frustration (Campbell, 2002; Kalpidou, Rothbaum, & Rosen, 1998; Prior, Smart, Sanson, Pedlow, & Oberklaid, 1992). Children of preschool age are learning about

their environment and the rules that govern behaviour. For these purposes, high energy levels and the ability to shift attention quickly can be advantageous at this stage of development. However, in some children such behaviours persist or are evident in several settings. Such externalising behaviour problems can be detrimental to children's functioning in that they are associated with impairments in development, learning and social functioning (Normandeau & Guay, 1998; Pavuluri, Luk, Clarkson, & McGee, 1995; Sanson, Smart, Prior, & Oberklaid, 1995).

There are two ways of classifying the degree and frequency of externalising behaviour problems. Classifications can be either dimensional or categorical. In *dimensional* systems, behaviour is viewed as lying along a continuum. Continuous scores are calculated for categories that are derived from rating scales and checklists. Examples of behaviour checklists designed for use with preschool children include the Child Behavior Checklist (1.5-5) (Achenbach & Rescorla, 2000), the Strengths and Difficulties Questionnaire (Goodman, 1997), the Preschool Behaviour Questionnaire (Behar & Stringfield, 1974) and the Preschool Behaviour Checklist, which is designed for use in group settings (McGuire & Richman, 1986). The dimensional approach preserves information about the range and nature of children's problems, as well as having greater statistical power than categorical approaches (Cantwell & Rutter, 1994). Recommended cutoff scores can be used to identify individuals who have levels of behaviour problems similar to those of children attending services for their problems. The statistically derived categories for preschool children include broader externalising (under-controlled behaviours), internalising (over-controlled behaviours including withdrawn and socially avoidant behaviours), and developmental problems factors (Pavuluri & Luk, 1998). At preschool age, there is a high degree of overlap between externalising and internalising behaviours (Pavuluri & Luk, 1998).

Categorical systems classify a disorder as “present” or “absent” based on a set of pre-determined criteria (Hinshaw, 2002; Pavuluri & Luk, 1998). In older children, the diagnosis of mental disorders is usually based on definitions in the Diagnostic and Statistical Manual of Mental Disorders Version IV (DSM-IV; American Psychiatric Association, 1994) or the International Statistical Classification of Diseases and Related Health Problems (ICD-10; World Health Organisation, 1992). There is limited evidence of the validity of DSM-IV criteria for diagnosing disorders in preschoolers (Keenan & Wakschlag, 2000), and studies which have obtained prevalence estimates of externalising behaviour problems in this age group, typically have not relied on formal diagnostic criteria (Campbell, 1995). This is firstly because of the high overlap between symptoms that define particular disorders, the manifestations of transient stress, and developmentally “normative” externalising behaviours (Achenbach & Rescorla, 2000; Campbell, 1995; Fegert, 1996; Keenan & Wakschlag, 2000; Marshall & Watt, 1999), and secondly because the symptoms and signs in the diagnostic manuals have been developed for school-aged children and are not age-appropriate for younger children (e.g., running away from home, truancy, breaking and entry) (Keenan & Wakschlag, 2000; Pavuluri & Luk, 1998).

Campbell (2002) has suggested that the presence of a clinically significant behaviour disorder in preschool children should be based on a clear pattern or constellation of symptoms. These symptoms should persist over time and should not be transient adjustments to stress; they should be pervasive across contexts and caregivers; they should be relatively severe; and they should impact adversely on children’s development and be associated with impairment in children’s functioning.

1.1.2 The aetiology of externalising behaviour problems

There are contrasting theories as to why some children develop externalising problems that are pervasive and persistent. These theories can be roughly divided into those which emphasise a hereditary basis for behaviour (nature), those which suggest that environmental forces primarily shape children's behaviour (nurture), and the prevailing interactional theories which suggest genetic elements of the individual interact with the environment to produce particular constellations of behaviour (Rutter, 2002). It is now generally agreed that there are multiple pathways to childhood externalising behaviour problems (Campbell, 1995; Hinshaw, 2002; Wootton, Frick, Shelton, & Silverthorn, 1997).

Genetically determined (or within-child) characteristics which are thought to influence the development and persistence of childhood behaviour problems include: 1) temperament (characteristic styles of emotional and behavioural responses which are thought to some extent to be genetically determined); and 2) neuropsychological and cognitive function such as mental control and social-cognitive processes (Brannigan, Gemmell, Pevalin, & Wade, 2002; Coy, Speltz, DeKlyen, & Jones, 2001; Nigg, Quamma, Greenberg, & Kusche, 1999; Paterson & Sanson, 1999; Prior, 1992).

Children with abnormally high or low levels of adaptive temperament styles (such as persistence, perspective-taking and effortful control), or very high levels of maladaptive temperament (such as inflexibility, callous traits showing a lack of empathy, and unemotional traits displaying a lack of guilt) have been shown to have higher levels of externalising behaviour problems (Minde, 1992; Murray & Kochanska, 2002; Prior, Smart, Sanson, & Oberklaid, 1993; Wootton et al., 1997). High levels of oppositional (or difficult) temperament, callousness, impulsivity, and low levels of harm avoidance have been incorporated in a concept termed 'antisocial

propensity" which is hypothesised to be influential in the development of early-onset persistent antisocial behaviour (Lahey, Waldman, & McBurnett, 1999). It is suggested that this is because the temperamental qualities of children with antisocial propensity are such that they may fail to understand and manage their emotions; they may fail to inhibit inappropriate behaviours; or they may be desensitised to punishment and thus remain impervious to their parents' attempts at behaviour management (Minde, 1992; Murray & Kochanska, 2002; Wootton et al., 1997).

Children with neuropsychological deficits in the area of social-cognitive processing are thought to engage in more aggressive and antisocial behaviours because of difficulties encoding social information and cues, negative attributional biases about the ambiguous behaviours of others, or the use of less effective problem-solving strategies which are viewed by the children as being effective (Coy et al., 2001; Crick & Dodge, 1994).

Environmental factors which have been associated with the emergence of behaviour problems include parenting practices, family interaction and cohesion, family structure and socioeconomic status, family life events, parental psychopathology, marital conflict and peer rejection (Campbell, Breaux, Ewing, & Szumowski, 1986a; Campbell & Ewing, 1990; Cunningham & Boyle, 2002; Dadds & Powell, 1991; Downey & Coyne, 1990; Fergusson, Horwood, & Shannon, 1984; Gottman & Katz, 1989; Greenwood et al., 1998; Harden & Zoccolillo, 1997; Hinshaw, 2002; Johnson et al., 2001; Larson, Pless, & Miettinen, 1988; Lavigne et al., 1996; Miller-Johnson, Coie, Maumary-Gremaud, & Bierman, 2002; Sawyer et al., 2000). Maternal anxiety in pregnancy, which has been hypothesised to cause changes to the in *utero* environment, may also affect the developing child (O'Connor, Heron, Golding, Beveridge, & Glover, 2002). The strength of the association between different

environmental risk factors and externalising behaviour is unclear, although it is recognised that the occurrence of multiple risk factors, and the interactions between them, are more strongly associated with negative child outcomes than are single specific risk factors (Fonagy, 1999). Environmental family risk factors such as marital conflict, family life events, parental psychopathology and socioeconomic disadvantage are thought to influence the development of children's behaviour both directly, and indirectly via their effects on the abilities of parents to develop and utilise functional parenting techniques (Belsky, 1984; Meyers, 1999). Reciprocal causation models of the development of externalising problems suggest that, in addition to the effects of environment on developing children, children's behaviour and negative affect, in turn, influence aspects of their environment. This may include influencing the parenting behaviour to which they are exposed, causing disruptions to family interactions or family cohesion, prompting marital conflict about ways to manage their behaviour and inducing rejection from their peers (Belsky, 1984; Harvey-Arnold & O'Leary, 1995; Hinshaw, 2002; Meyers, 1999; Miller-Johnson et al., 2002). The role of parenting behaviour in the development of externalising problems will be discussed in more detail in Section 1.3 of this chapter.

Past studies which have investigated children's socialisation have generally treated parenting as the major determinant of socialisation, with other environmental factors (such as peers and neighbourhoods) and heredity considered to be relatively minor influences (Collins, Maccoby, Steinberg, Hetherington, & Bornstein, 2000). More recently, studies of twins and adopted children (Deater-Deckard & Plomin, 1999; van der Valk, van-den-Oord, Verhulst, & Boomsma, 2001) have attempted to provide estimates of the amount of variance in childhood externalising behaviour problems which can be ascribed to factors in children's environments which are both shared (all children in a family are exposed) and unshared (specific to a particular child), and

the variance which can be attributed to genetic inheritance. In this type of research, the effects of environment and heredity are often treated as additive and mutually exclusive, and the effects of environmental factors (such as parenting, peers and school) are generally reported as relatively minor (Collins et al., 2000; Deater-Deckard & Plomin, 1999; van der Valk et al., 2001). However, Maccoby (2000) suggests that environmental influences are largely underestimated in these studies because although similar environments are often shared by individuals who are genetically similar, the variance explained by these shared environments is allocated to heredity rather than environment because of the genetic closeness of the individuals. Recent research has demonstrated that the effects of a shared environment (i.e., family structure) on children's behaviour can actually be quite large, and that this relationship cannot readily be attributed to genetic relatedness because it has been demonstrated to be stronger in families which have greater proportions of unrelated members (O'Connor, Dunn, Jenkins, Pickering, & Rasbash, 2001). It should also be noted that traits that are highly heritable, can still respond to environmental factors and relationships between genetic and environmental influences and are the focus of interactional theories of the development of behavioural maladjustment (Maccoby, 2000).

Interactional theories are receiving more attention as researchers investigate phenomena such as vulnerability and resiliency. These concepts attempt to explain why some children develop behavioural maladjustment whereas others do not, even though they live in the same family or are exposed to similar environmental conditions (Belsky et al., 1997; O'Connor et al., 2001). Genetic characteristics of one child may increase susceptibility to poor environmental conditions whilst another child's genetic composition might protect against the effects of a poor environment (Belsky et al., 1997; Belsky et al., 1998; Rutter, 2002). Research investigating the

development of externalising disorders has focused on the interactions between parental perceptions of child temperament and environmental factors such as family functioning and parenting practices (Brannigan et al., 2002; Paterson & Sanson, 1999; Prior, 1992). Parental perceptions of temperament are more likely to predict later externalising problems in conjunction with other risk factors (e.g., perinatal stress, prematurity, developmental problems and male gender), only when these risk factors co-occur in the presence of environmental adversity (Oberklaid, Sanson, Pedlow, & Prior, 1993; Sanson, Oberklaid, Pedlow, & Prior, 1991). For children experiencing environmental adversity, perceived difficult temperament is a risk factor for behavioural maladjustment, whereas easy temperament operates as a protective factor (Prior et al., 1992). Other examples of interactional theories will be discussed in more detail in Section 1.3.

In addition to the child and family factors implicated in the development of externalising behaviour problems in young children, other significant correlates of these problems have been identified. These correlates include: the general health of the child; childhood chronic illness; early entry into low quality daycare facilities; deviant peer affiliations; and cognitive correlates such as delays in language development, lower IQ and reading difficulties (Hausfather, Toharia, LaRoche, & Engelsmann, 1997; Hinshaw, 2002; Larson et al., 1988; Lavigne et al., 1996; McGee, Partridge, Williams, & Silva, 1991; Richman, Stevenson, & Graham, 1982; Sanson et al., 1996; Thomas, Byrne, Offord, & Boyle, 1991). It is unclear whether these correlates are determinants or outcomes of externalising behaviour problems, or if they are simply associated with the same genetic or environmental factors (Campbell, 1995; Hausfather et al., 1997; Hinshaw, 2002; Sanson et al., 1996).

1.1.3 The prevalence of externalising behaviour problems in preschool children

The prevalence of individual externalising behaviours in young children is relatively high (Campbell, 2002; Chamberlin, 1974; Pavuluri et al., 1995). Parental concern about externalising behaviour varies with the developmental age of the child, with concerns about *discipline* issues peaking at age three years (Chamberlin, 1974; Hickson, Altemeler, & O'Connor, 1983; Taaffe-Young, Davis, Schoen, & Parker, 1998). In non-clinical samples, as children develop the capacity to experience empathy in the early childhood years, an increase in prosocial behaviours occurs concomitantly with a decrease in aggression (Keenan, Shaw, Delliquadri, Giovannelli, & Walsh, 1998; Minde, 1992; O'Leary, Slep, & Reid, 1999). However, it is possible for very active children to display high levels of both kinds of behaviour although the behaviours seem intuitively contradictory (Minde, 1992).

Prevalence estimates of externalising behaviour problems in preschool children vary according to: problem definition, parameters used to establish clinical severity, instruments, informants, sample types and the context/s in which behaviours are measured (Pavuluri et al., 1995; Prior et al., 1992). Using dimensional measures of preschool behaviour problems, almost a quarter of preschool age children have been reported to have clinically significant levels of externalising behaviour problems, and these figures have been consistently found in both community samples and samples of children from low-income families in several countries (Keenan, Shaw, Walsh, Delliquadri, & Giovannelli, 1997; Luk et al., 1991; Pavuluri et al., 1995; Richman, Stevenson, & Graham, 1975). For example, in community samples, 23% of children aged three to four years in Hong Kong, 22% of three-year-old children in England and 22.5% of preschoolers in a New Zealand study were rated as having mild,

moderate or severe behaviour problems (Luk et al., 1991; Pavuluri et al., 1995; Richman et al., 1975). And in a low-income sample of preschoolers in the United States, 26% were found to have scores in the clinical or borderline range on the Externalising Scale of the Child Behavior Checklist (Keenan et al., 1997). These estimates are generally lower when the samples under investigation include older children as well as preschoolers. For example, estimates of the percentage of children above behaviour checklist cutoff scores as reported by parents and/or teachers in community samples of primary school children are approximately 15%-20% (Hofstra, van der Ende, & Verhulst, 2000; Offord et al., 1987; Sawyer et al., 2000; Zubrick et al., 1995). These lower estimates reflect the transient nature of behaviour problems experienced by the large number of preschoolers.

Studies which have assessed the prevalence of externalising behaviour disorders in samples of preschool children using DSM-IIIR criteria have obtained lower estimates, in the order of 15% to 17%, than those using dimensional systems (Keenan et al., 1998; Lavigne et al., 1996; Luk et al., 1991; Pavuluri et al., 1995; Richman et al., 1975). These studies may have reported lower prevalence estimates because of the use of categorical classification systems which include symptoms which are not applicable to preschool-age children, as discussed in Section 1.1.

In preschoolers, gender differences in the prevalence of externalising behaviour problems are generally smaller than in children of primary school age (Prior et al., 1992). Boys and girls of preschool age typically have quite similar levels of behaviour problems, with male:female gender ratios in the order of 1.1:1 (Hofstra et al., 2000; Richman et al., 1975). One exception is a study conducted in Hong Kong which reported a ratio of 1.9:1 (Luk et al., 1991), although cultural differences between this and European samples may explain the gender differences.

1.1.4 The stability of childhood externalising behaviour problems

While it is normal for preschool children to display transient behaviour problems, especially in response to stressful events such as the birth of a sibling or transition to school (Campbell, 2002), longitudinal studies of behavioural outcomes demonstrate that persistent externalising behaviour disorders in preschool children increase the likelihood of developing emotional and conduct disorders in later childhood and adolescence (Feehan, McGee, Williams, & Nada-Raja, 1995; Hofstra et al., 2000; Prior et al., 1992).

The combination of hyperactivity, aggression and noncompliance in young children seems one of the strongest markers for pervasive problems and produces the worst prognosis (Campbell, 1995). In particular, children with high levels of aggressive and antisocial behaviours, when compared with their non-aggressive peers, are at greater risk of delinquency and criminal behaviour in adolescence and of serious psychosocial problems in adolescence and adulthood (Minde, 1992; Verhulst et al., 1993). Other potential outcomes of long-term hyperactivity, aggression and noncompliance include poor social functioning, self-image problems, poor language and cognitive skills, lower levels of reading ability and impaired general academic performance (McGee et al., 1991).

The reported stability (or persistence) of externalising behaviour problems in childhood varies from 18% to 75% (Campbell, 1995; Campbell, Ewing, Breaux, & Szumowski, 1986b; Hofstra et al., 2000; Kingston & Prior, 1995; McGee et al., 1991; McGee, Prior, Williams, Smart, & Sanson, 2002; O'Leary et al., 1999; Richman et al., 1982; Sanson et al., 1991). The high degree of variability in these estimates of stability can be attributed to a number of factors that are outlined below.

1. *The nature of the sample involved*. Generally, higher rates of persistence are found in clinic versus non-clinic samples, and in boys versus girls (Campbell et al., 1986b; Kingston & Prior, 1995; O'Leary et al., 1999; Richman et al., 1982).
2. *The definition and nature of "problem behaviour" at the initial assessment*. Studies which use more inclusive criteria, such as "one standard deviation above the mean" to define significant problem levels of behaviour show lower rates of stability compared with those that use more strict criteria such as clinical diagnosis (McGee et al., 1991; Sanson et al., 1991). Also, specific constellations of behaviour measured in young children, such as aggression, are more likely than others, such as hyperactivity, to be associated with self-reported behaviour problems in adolescence (Hofstra et al., 2000; McGee et al., 2002; Tremblay, Pihl, Vitaro, & Dobkin, 1994).
3. *The definition of "stability"*. Studies which use stricter criteria for stability such as having a significant level of behaviour problems in *at least two* follow-up assessments report lower rates of stability than those studies which only require problems to be present at one follow-up assessment (Lahey et al., 1995).
4. *The age at the initial assessment and the length of the follow-up period.* Children who are toddlers at the initial assessment generally show lower rates of persistence over time than do children who are older at the first assessment. Longer follow-up periods are usually associated with lower estimates of stability than are shorter periods (Campbell, 1995; Hofstra et al., 2000; O'Leary et al., 1999).

In general, preschool children with externalising behaviour problems have a 50% likelihood of early externalising problems persisting into later childhood (Campbell, 1995). Furthermore, of the children who fall into the "problem" category at follow-up assessments, approximately half to two thirds are children who had high levels of

problems at the initial assessment (Campbell & Ewing, 1990; O'Leary et al., 1999; Tremblay et al., 1994).

Although the best predictor of future deviant and antisocial behaviour is the amount and severity of such behaviours in the past, other factors which have been associated with the persistence of early behaviour problems include parenting practices (e.g., physical discipline, parental control via guilt and anxiety-provoking techniques, and inconsistent parenting behaviour), family adversity, hostile sibling interactions, parental perceptions of difficult child temperament, cognitive ability and social competence (Campbell & Ewing, 1990; Feehan et al., 1995; Greenwood et al., 1998; McGee et al., 1991; Pavuluri & Luk, 1998; Sanson et al., 1991).

1.1.5 The cost of pervasive and persistent externalising behaviour problems

The child and adolescent component of the Australian National Survey of Mental Health and Wellbeing found that 13.6% of children aged 4 to 12 years had significant externalising behaviour problems. This corresponds to a population estimate of 165,000 Australian children (Sawyer et al., 2000). The majority of children with clinically significant levels of behaviour problems do not attend professional services for help with their behaviour (Sawyer et al., 2000), and studies have demonstrated evidence of behaviour problems persisting to adolescence in up to 50% of children (Marshall & Watt, 1999). Further, long-term follow-up studies of children who participated in interventions designed to reduce levels of externalising behaviours have reported that treatment groups have better outcomes for children and their parents, particularly in high-risk families, in terms of lower need for special education and lower levels of grade retention, lower school drop-out rates, higher employment

and wages, and decreased levels of delinquent behaviour and criminal arrests, than children who do not receive the interventions (Marshall & Watt, 1999; Ramey et al., 2000; Schweinhart, Barnes, Weikart, Barnett, & Epstein, 1993). Thus, long-term persistent externalising behaviour problems can be associated with substantial costs both to individuals and society. Karoly et al. (1998) categorise the monetary cost of individuals with long-term persistent externalising behaviour problems as costs to the government and costs to the rest of society. Costs to the government include: 1) decreased tax revenue; 2) increased welfare payments; 3) increased expenditures on education, health and other services; and 4) increased criminal justice system costs. Monetary costs to the rest of society include: 1) lower future incomes for people with externalising behaviour problems; and 2) tangible and intangible losses to potential victims of crime. Studies in the United States have shown that it is possible to produce savings of up to $4 for every $1 spent on early interventions designed to prevent externalising problems (Olds et al., 1997).

Mental disorders have been identified by the Australian Federal Government as a National Health Priority Area. Australian expenditure on health services for mental disorders has been estimated at around $3 billion per annum (Australian Institute of Health and Welfare, 2002). However, Australian public sector expenditure on child and adolescent mental health services is only approximately 7% of all expenditure on mental health services (Burgess et al., 2002). This is an issue of concern, given that children and adolescents aged 4-17 years comprise 20% of the population (Australian Bureau of Statistics, 2000), and have levels of clinically significant problems similar to that of adults (Andrews, Henderson, & Hall, 2001). A mid-term review of the Australian Second National Mental Health Plan, an initiative developed under the National Mental Health Strategy, acknowledged the lack of services to meet the needs of children and youth in either the primary care or mental health

sectors, and the adverse outcomes of leaving childhood mental health problems untreated (Thornicroft & Betts, 2001). For example, in addition to the health service delivery costs for children with behaviour problems there are the high costs of youth and adult criminal behaviour associated with long-term persistent externalising problems. The estimated cost of the criminal justice system in Australia was estimated at $6.8 billion in 2000-2001 (Steering Committee for the Review of Commonwealth/State-Service Provision, 2002). This estimate does not include the costs to society of the actual crimes committed, such as earnings lost by the victim, treatment and shelter for victims of violent assault, costs of injury to the victim and monies lost as a result of the crime, these costs are estimated to be in the order of $11 to $13 billion per year (Walker, 1997).

The burden associated with emotional and behavioural problems is not just financial in nature. The child and adolescent component of the National Survey of Mental Health and Well-being reported that higher levels of emotional and behavioural problems were associated with lower quality of life scores for children and adolescents. Self-esteem and participation in activities with friends or at school were particularly affected (Sawyer et al., 2000). Higher behaviour problem scores were also associated with lower levels of family functioning in terms of participation in family activities and family cohesion. Parental quality of life is also affected in that parents of children with higher levels of emotional and behavioural problems reported that they do not have as much time for themselves and had higher levels of parental worry, compared with parents of children with lower levels of problems (Sawyer et al., 2000).

This section of Chapter 1 has provided a brief introduction to externalising behaviour problems in preschool children. This section has highlighted the public health significance of early-onset persistent externalising behaviour problems in terms of their aetiology, relatively high prevalence, long-term stability and associated costs to society and individuals. The next section discusses parenting behaviours that have been associated with the development and persistence of externalising behaviour problems in preschool children.

1.2 Parenting

Different aspects of parenting are important at different developmental stages (Chamberlain & Patterson, 1995; Patterson & Stouthamer-Loeber, 1984; Roberts & Strayer, 1987; Shaw et al., 1998), and the following section briefly explores the literature relevant to parenting behaviours and parenting styles which have been demonstrated to be important in the preschool period. Parenting behaviours are the specific, goal-directed actions which impact directly on the child (e.g., discipline, reasoning) whereas parenting styles incorporate parents' attitudes, beliefs and expectations for child behaviour and thereby influence the context in which parenting occurs (e.g., authoritative styles versus authoritarian styles) (Brenner & Fox, 1999; Darling & Steinberg, 1993; Metsapelto, Pulkkinen, & Poikkeus, 2001; Robinson, Mandleco, Olsen, & Hart, 1995). In this thesis, unless stated otherwise, "parenting" is used to refer to parental discipline and other behaviours that may occur during parent-child interactions. Further, the term "parent", unless otherwise indicated, is used to refer to primary caregivers, with whom the bulk of parenting research has been conducted. Research that examines the relationship between the parenting practices described in this section and children's externalising behaviour are discussed in section 1.3.

1.2.1 What is parenting?

The word 'parent' implies a biological or legal relationship to a child, whereas the term 'parenting' refers to the more dynamic acts of caring, nurturing and protecting children. Parenting involves taking responsibility for the emotional, social and physical growth and development of children (Smith, 1999). Whilst it is possible for a "parent" to have no contact with a child, the act of "parenting" implies an ongoing dynamic bidirectional relationship between the caregiver and child (Smith, 1999). Parenting has also been defined as "anything the parent does or fails to do that may

affect the child" (Kendziora & O'Leary, 1993, p.177). Parenting behaviours include: playing, disciplining, teaching, caring for children's physical needs and establishing a pleasant emotional environment (Kendziora & O'Leary, 1993).

In the scientific field, the definition of parenting has undergone many revisions. In earlier research, parenting was seen as a largely unidirectional process in which children were the recipients of parental actions (Holden, 1983). A more contemporary and holistic view is that parenting is a series of reciprocal interactions between parents and their children, with the behaviour of each participant affecting the behaviour of the other (Chamberlain & Patterson, 1995; Dowdney, Mrazek, Quinton, & Rutter, 1984).

Dimensions of parental behaviour which are important for children's development include: parental discipline in the form of limit-setting and control; instilling moral values, respect and concern for others; and involvement with children as teacher and play partner (Campbell, 1995). Different dimensions of parenting are important at different times of development and in different contexts, depending upon children's needs and upon parental expectations about children's behaviour (Chamberlain & Patterson, 1995; Patterson & Stouthamer-Loeber, 1984; Roberts & Strayer, 1987; Shaw et al., 1998).

Parenting is complex and multiply determined, arising in part from beliefs, attitudes, experiences (both of parenting and being parented) and expectations about both parenting and about children's behaviour (Fox, Platz, & Bentley, 1995; Kendziora & O'Leary, 1993; Kochanska, 1990; Rodriguez & Sutherland, 1999; Rubin, Stewart, & Chen, 1995; Woodworth, Belsky, & Crnic, 1996). These determinants of parenting may in part arise from social or cultural norms (Bradley, Corwyn, Burchinal, McAdoo,

& Coll, 2001). Parental beliefs and behaviour can be affected by a variety of other determinants. Family stresses can undermine parenting, and longitudinal studies show that parents at greatest risk of being ineffective or abusive are young, single parents living in poverty (Greenwood et al., 1998). Context-specific elements such as setting (e.g. at home or in public), presence of other family members (e.g. siblings or partner) and the behaviours exhibited by the child (e.g. stealing versus noncompliance) can have immediate effects on parenting (Dowdney et al., 1984; Johnson, 2001; Meyers, 1999; Mrazek et al., 1982; Pappas-Jones & Adamson, 1987; Socolar, Winsor, Hunter, Catellier, & Kotch, 1999).

Environmental and within-parent determinants, including parental psychopathology, social insularity, marital conflict and disagreements between parents about child-rearing, can also influence parents' abilities to develop or employ effective parenting techniques (Fonagy, 1999; Fox et al., 1995; Gaudin, Polansky, Kilpatrick, & Shilton, 1993; Rubin et al., 1995; Woodworth et al., 1996). It is recognised that these factors in themselves do not occur in isolation and may influence each other as well as having an impact on parenting behaviour. It is the occurrence of multiple risk factors, rather than the operation of a single risk factor, which has the greatest impact on parenting behaviour (Campbell, Shaw, & Gilliom, 2000; Fonagy, 1999; Jackson, Brooks-Gunn, Huang, & Glassman, 2000).

Socioeconomic disadvantage

Socioeconomic disadvantage, as measured by levels of parental unemployment, low household income, family size and composition, frequent residential moves, availability of privacy in the home and low levels of parental education, is thought to affect parenting in a variety of direct and indirect ways (Barkley, 1990; Brenner & Fox, 1999; Fox et al., 1995; Jackson et al., 2000; Johnston, Murray, Hinshaw,

Pelham Jr, & Hoza, 2002; McLoyd, 1990; Pittman, Wright, & Lloyd, 1989; Rothbaum, 1986; Woodworth et al., 1996). Stresses in the family, such as financial strain and the life events associated with economic loss, can leave parents preoccupied, overwhelmed, distressed, frustrated, angry and helpless, with fewer emotional or practical resources to deal with the needs of their children (McLoyd, 1990; Rubin et al., 1995). Frequent residential moves may impact on parenting by disrupting family networks and reducing supportive contacts for parents (Barkley, 1990). Parental educational status has been shown to moderate the relationship between socioeconomic status and parenting behaviour, such that parents who live in lower socioeconomic areas but who have higher education levels display more effective parenting behaviours than parents living in similar areas with lower levels of educational attainment (Fox et al., 1995; Jackson et al., 2000; Johnston et al., 2002).

Parental psychopathology

Antisocial parents are at increased risk of employing ineffective parenting practices compared with non-antisocial parents (Patterson, DeBaryshe, & Ramsey, 1989). This may be because the negative effects of family stressors are amplified, and hence more disruptive, for antisocial parents compared with other parents (Patterson et al., 1989).

Depression is associated with feelings of helplessness and hopelessness. In mothers it is associated with a lack of parental involvement and responsiveness in child-rearing (Johnston et al., 2002; Rubin et al., 1995). Depressed parents view their parenting role less positively and their abilities as inadequate; they may experience negativity towards the demands of parenthood and feelings of rejection and hostility towards their children (Downey & Coyne, 1990). Depressed parents can be more hostile and irritable, put less effort into interacting with their children or issue vague

commands to which children cannot comply, and they may be less sensitive or responsive to their children's needs (Downey & Coyne, 1990; Forehand, Lautenschlager, Faust, & Graziano, 1986; Johnston et al., 2002). Interactions between depressed parents and their children have been described as unanimated and unstimulating, although at times depressed mothers can be intrusive or overstimulating (Field, 1995). Further, the parenting behaviour of depressed parents may be less contingent upon children's behaviour because their depressive symptoms negatively skew their perceptions of, and attributions for, their children's behaviour (Fergusson et al., 1984; Field, 1995; Forehand et al., 1986; Kochanska, 1990; White & Barrowclough, 1998).

Both the stress resulting from major life events and also the cumulative stress from daily hassles have been associated with ineffective parenting practices, possibly by affecting parents' self-esteem, parenting confidence and coping abilities (Belsky, Crnic, & Gable, 1995; Crnic & Greenberg, 1990; Kendziora & O'Leary, 1993; Keown & Woodward, 2002; Koeske & Koeske, 1990; Pittman et al., 1989).

Marital conflict and disagreements about childrearing

Parents involved in marital conflict are often distressed and frustrated, and they may be facing the possibility of marital dissolution (Grych & Fincham, 1990). Parents undergoing marital conflict and distress display styles of parenting that are cold, unresponsive, angry, inconsistent, low in limit setting and even hostile if their children are perceived as causes of the conflict. Parents may attempt to coerce their children into alliances against the other parent (Brook, Zheng, Whiteman, & Brook, 2001; Gottman & Katz, 1989; Grych & Fincham, 1990; Wilson & Gottman, 1995). Parents involved in marital conflict can be absorbed with their own problems and thus become less involved and more inconsistent in their parenting, or the frequent conflict

between partners can reduce the support in childrearing that a spouse provides (Belsky et al., 1995; Wilson & Gottman, 1995).

Social insularity

Parents who are socially insular have few contacts with non-family members, and often have aversive contacts with family and agency representatives (Kendziora & O'Leary, 1993). The ability to parent effectively may be compromised by social isolation because such parents may be exposed to negative social experiences, may have an inability to communicate or problem-solve, and may experience a lack of positive social support (Dadds & McHugh, 1992; Kendziora & O'Leary, 1993). However, the availability of social support may moderate the effects of environmental stressors on parenting behaviour by providing emotional, informational, financial and parenting support, and role modelling (Koeske & Koeske, 1990). These factors can prevent feelings of stress, enhance feelings of competence, strengthen coping abilities and decrease feelings of isolation and helplessness, when parents are confronted by stressful life circumstances (Andresen & Telleen, 1992; Crnic & Greenberg, 1990; Jackson et al., 2000; Koeske & Koeske, 1990; McLoyd, 1990; Rubin et al., 1995).

The following section will examine specific parental behaviours that have been associated with preschool children's behaviour, and will focus on the dimensions of parental control and responsivity.

1.2.2 Parenting behaviour

As described earlier, parenting behaviours are the specific, goal-directed actions by which parents perform their parental duties such as promoting and guiding children's socialisation (Darling & Steinberg, 1993). It has been hypothesised that it is *patterns*

of parental behaviours rather than the specific behaviours *per se* that are important in terms of their direct impact on children's behaviour (Brenner & Fox, 1999; Robinson et al., 1995). This section discusses patterns of parenting behaviours, such as parental control and responsivity, which have been associated with children's behavioural development (Brenner & Fox, 1999; Cunningham & Boyle, 2002; Darling & Steinberg, 1993; Fox et al., 1995; Gardner, 1989; Gardner, Sonuga-Barke, & Sayal, 1999; Granic & Lamey, 2002; Johnston et al., 2002; Keown & Woodward, 2002; Kochanska & Murray, 2000; Rothbaum, 1986; Rubin et al., 1995; Shaw et al., 1998; Stormshak, Bierman, McMahon, Lengua, & Conduct Problems Prevention Research Group, 2000).

Parental control

Parental control incorporates aspects of parental monitoring of children's behaviour, provision of contingent responses to both prosocial and deviant behaviour, handling conflict, problem solving, and limit-setting (Dowdney et al., 1984). It is important for parents to allow children some degree of autonomy to develop a sense of self, but it is necessary also to moderate aggressive behaviours to help children develop appropriate levels of self-control (Dowdney et al., 1984).

Different types of parental control can have different effects on children's behaviour. Parental control which has been shown to lead to more positive child outcomes incorporates limit setting, supervision and monitoring at a level which is developmentally appropriate (Arnold et al., 1993). Contingent and appropriate punishments are also a feature of effective parental control (Chamberlain & Patterson, 1995). Less effective parental control techniques include overreactive parental discipline, power-assertive techniques and permissive or inconsistent

parenting (Arnold et al., 1993; Chamberlain & Patterson, 1995; Cunningham & Boyle, 2002; Patterson & Stouthamer-Loeber, 1984).

"Overreactive" parental discipline is characterised by high levels of parental anger, meanness, irritation and frustration in response to child misbehaviour (Harvey-Arnold & O'Leary, 1995; Harvey-Arnold & O'Leary, 1997), and the use of power-assertive techniques including physical punishment, yelling and threats (Hemphill & Sanson, 2001). While these techniques may be effective in the short-term because they generate a degree of fear and submission to authority, they are less effective in the long-term because they do not foster a generally cooperative attitude in children (Edwards, 1995). Lax parenting (incorporating aspects of permissiveness and inconsistency) occurs when parents allow rules to go unenforced, provide positive consequences for misbehaviour, fail to set limits, and "give in" to children's coercive behaviour (Harvey-Arnold & O'Leary, 1995; Harvey-Arnold & O'Leary, 1997).

Parental responsivity

The concept of parental responsivity refers to the affective dimension of parenting (Rubin et al., 1995). Typically, responsive parenting practices involve positive attention, positive affect, and responding quickly and contingently to children. These practices promote language development, prosocial behaviour and compliance in children through modelling and reinforcement (Downey & Coyne, 1990; Tamis-LeMonda, Bornstein, & Baumwell, 2001). Parental responsivity reflects parents' awareness of children's needs and their sensitivity to children's signals (Dowdney et al., 1984). Responsive parenting is typified by reciprocity, turn-taking, warmth, reasoning and explanation. Hostility and aggression are seldom seen in responsive parenting practices (Dowdney et al., 1984).

Parental warmth is characterised by verbal and physical affection, approval, support and positive closeness between parents and children (Hemphill & Sanson, 2001). Conversely, parenting of a rejecting nature, involves an absence of responsivity and warmth, and can include active denigration of the child. Reasoning and explanation include a sense of openness and non-arbitrariness, and involve providing reasons and explanations for rules or limitations on behaviour, and explaining the consequences of behaviour (Gardner et al., 1999).

1.2.3 Parenting styles

As described earlier, parenting styles are relatively stable constellations of attitudes towards children, which set the context of parent-child relationships (Darling & Steinberg, 1993; Metsapelto et al., 2001). Because parenting styles are thought to act only indirectly on children's behaviour by moderating actual parenting practices, they will not be discussed in great detail in this thesis (Darling & Steinberg, 1993).

Authoritarian, authoritative, permissive and neglectful parenting styles

Baumrind (1967, 1971, 1991) identified four different parenting styles (authoritative, authoritarian, permissive and neglectful), characterised by the values, behaviours and standards expected of children and the way in which these beliefs values and standards are conveyed. These parenting styles have been differentially associated with child outcomes and are described below.

1. *Authoritative parenting* is characterised by high levels of nurturance (characterised by warmth and involvement), encouragement of independence and individuality, communication (characterised by reasoning and soliciting opinions), maturity demands and control (consistency and resisting pressure from the child) (Baumrind, 1971; Baumrind & Black, 1967).

2. *Authoritarian parenting* involves arbitrary and harsh limit-setting (demandingness) in the context of low warmth and responsivity (Baumrind, 1971; Baumrind & Black, 1967).
3. *Permissive parenting* involves tolerance and acceptance of children's impulses, little use of punishment, and few parental demands and expectations of children (Baumrind, 1971; Baumrind & Black, 1967).
4. *Neglectful or indifferent parenting* is characterised both by low levels of parental responsivity and of demandingness. Neglectful parents know very little about their children and try to minimise the time and energy they devote to their children (Baumrind, 1971; Baumrind, 1991).

Research has shown that the authoritative, authoritarian and permissive parenting styles identified by Baumrind correspond well to constellations of *parenting practices* used by parents of young children (Brenner & Fox, 1999; Robinson et al., 1995). As well as finding clusters of parenting behaviours which correspond with Baumrind's authoritative, authoritarian and permissive parenting styles, Brenner and Fox (1999) also identified a large group of "average" parents characterised by low to moderate levels of parental discipline, nurturance and expectations about children's behaviour.

Subsequent work using a revised version of the Parental Authority Questionnaire (PAQ_R) which includes items assessing both parenting beliefs and parenting practices derived directly from Baumrind's descriptions of parenting styles, found moderate levels of support for the authoritarian and permissive parenting styles (Reitman, Rhode, Hupp, & Altobello, 2002). Support was also found for the authoritative parenting style in a Caucasian subsample in this study, but not within two predominantly African-American samples (Reitman et al., 2002). This work highlights the importance of not generalising the results from parenting studies across cultures.

This section has examined the environmental and within-parent determinants of parenting behaviour such as parental psychopathology, marital satisfaction and social support. Both effective and ineffective parenting practices incorporating behaviours along the dimensions of parental control and responsivity were described. Parenting styles that incorporate parental attributions and expectations were also briefly described. The relationships between the parenting behaviours and styles described in this section and children's externalising behaviour problems will be examined in the next section.

1.3 Relationships between Parenting and Behaviour Problems

Although it is not the only influence on children's behaviour, parenting plays a major role in influencing the development and maintenance of aggressive behaviours in childhood. This can become the foundation for poor adjustment later in life (Rubin et al., 1995). Ineffective parenting practices have been described as resulting from a parents' emotions about their children that are too strong, not strong enough, or are poorly matched to their children's needs (Dix, 1991). Lahey et al. (1999) have hypothesised that parents have thresholds for reacting to children's misbehaviour and variations in these thresholds occur both between different parents and across time for the same parent. The experience of psychological stress or environmental adversity may serve to lower such thresholds and this may lead to the use of less effective parenting techniques (Lahey et al., 1999). In turn, ineffective parenting is thought to influence young children's socialisation in a number of ways; for example through modelling, conditioning, or by influencing the security of children's attachment to their parents (Rubin et al., 1995). This section describes research on the relationships between parenting behaviours and children's externalising behaviour. First, simple relationships between the parenting practices described in the previous section and children's externalising behaviour will be discussed, then specific interactional models of influence and the results from parent training programs will be examined.

1.3.1 Relationships between parenting behaviour and children's externalising behaviour

Parental control

Parent-child discipline interactions are important because they teach children behaviours which are appropriate in the context of a particular family or situation (Prushank, 1995). Parental discipline which is lax, permissive and inconsistent, or

overreactive, authoritarian and coercive, contributes to the development and maintenance of serious externalising problems (Brenner & Fox, 1999; Chamberlain & Patterson, 1995; Cunningham & Boyle, 2002; Fox et al., 1995; Gardner, 1989; Granic & Lamey, 2002; Keown & Woodward, 2002; Rubin et al., 1995; Stormshak et al., 2000).

In terms of overreactive and power-assertive discipline, the use of higher levels of verbal and corporal punishment by parents has been associated with higher levels of externalising behaviour problems in preschool children (Brenner & Fox, 1999; Cunningham & Boyle, 2002; Fox et al., 1995). Punitive discipline and spanking have been shown to correlate with all types of children's behaviour problems (including internalising behaviours), and parental physical aggression has been significantly associated with aggressive child behaviours (Fox et al., 1995; Stormshak et al., 2000). Overreactive and punitive discipline may be associated with children's externalising problems because the aggressive or hostile behaviour of parents may act as a model for children, or may serve to prompt similar types of behaviour from children in retaliation (Arnold et al., 1993). Alternatively, the type of behaviours exhibited by the child may influence parental discipline techniques (Gershoff, 2002; Holden & Edwards, 1989; Socolar et al., 1999). For example, a meta-analysis investigating the associations between corporal punishment and child behaviour reported that parents were more likely to use physical punishment for aggressive behaviour rather than noncompliant behaviour, for behaviour which is a threat to themselves or to others, and for escalated disobedience (i.e. if they misbehaved after already being punished) (Gershoff, 2002).

Permissiveness and inconsistency of discipline, both over time and between parents, is associated with hyperactivity, child negative affect, aggressive behaviour and

conduct problems in preschool children (Acker & O'Leary, 1996; Cunningham & Boyle, 2002; Forehand et al., 1986; Gardner, 1989; Granic & Lamey, 2002; Keown & Woodward, 2002; Rubin et al., 1995). Parents who “give in” when their children misbehave, although avoiding conflict in the short term, may reinforce their children’s behaviour (Arnold et al., 1993). Further, parents whose discipline is inconsistent and erratic may contribute to persistence of children’s misbehaviour because when children are uncertain of the consequences of their behaviour such behaviours may become more resistant to extinction (Patterson, 1982). In terms of determining what *type* of inconsistency has the greater significance in terms of influencing children’s behaviour, Gardner (1989) stresses that it is consistency at the level of an entire parent-child discipline interaction rather than at the micro-level of immediate parental reactions to specific child behaviours which is more influential for children’s behavioural outcomes. Further, Acker and O’Leary (1996) suggest that inconsistent parenting which specifically incorporates *positive attention* to children’s misbehaviour is more strongly associated with increases in children’s misbehaviour than is inconsistent parenting that combines ignoring and punishment.

Parental responsivity

It has been hypothesised that positive parenting practices can moderate or “buffer” the effects of negative family risk factors on children’s behaviour problems (Rubin et al., 1995). In situations of high family stress, children with at least one warm and responsive parent have been shown to have lower rates of externalising behaviour problems compared with children without such a parent (Rutter, 1979). Other research has consistently shown that responsivity, parent-child synchrony, warmth and reasoning have a positive association with an array of child functioning variables, including: less aggression and other externalising behaviours; less anxious-inhibited behaviour and psychosomatic problems; moral, helping and altruistic behaviour;

friendliness and social competence; self-esteem; constructive play and intellectual functioning (Johnston et al., 2002; Keown & Woodward, 2002; Kochanska & Murray, 2000; Roberts & Strayer, 1987; Rothbaum, 1986; Shaw et al., 1998; Stormshak et al., 2000). Responsive parents may be modelling and teaching reciprocal and cooperative behaviours (horizontal qualities) in their interactions with their children (Russell, Pettit, & Mize, 1998). Parents who use inductive reasoning, when compared with parents who use more power assertive techniques, have been shown to have children with fewer disruptive playground behaviours, more prosocial behaviours, and higher levels of preference by their peers (Hart, DeWolf, Wozniak, & Burts, 1992).

Hostile and rejecting parenting has been associated with externalising behaviour problems in preschool children (Shaw et al., 1998). Parent-child interactions characterised by *mutual* hostility have been associated with higher levels of comorbid internalising behaviour problems in boys compared with interactions which do not involve such hostility (Granic & Lamey, 2002). Attachment theory posits that infants with experiences of warm, responsive and sensitive care when especially vulnerable and dependent, develop a sense of self-efficacy and trust in the availability and supportiveness of significant others (Shaw & Bell, 1993). Infants with these secure attachments will feel able to trust others and are more likely to develop positive, prosocial relationships with peers, teachers and other adults. On the other hand, infants who experience inconsistent, unresponsive and rejecting parenting may become anxious, irritable, feel themselves unworthy, and perceive their caregivers as unpredictable and untrustworthy, experiencing insecurity in their attachment to their rejecting parents (Granic & Lamey, 2002; Rubin et al., 1995; Shaw & Bell, 1993).

Parenting characterised by low warmth and responsivity and high levels of rejection may not just be associated with children's misbehaviour via the security of

attachment. The children of rejecting parents may misbehave out of frustration, to gain attention, or to make parents withdraw from an interaction (Rubin et al., 1995; Shaw et al., 1998). Further, the lack of warmth and reasoning to which rejected children are exposed may mean that they are unwilling or unable to internalise moral values about the rights of others and thus do not behave prosocially (Shaw et al., 1998).

Authoritarian, authoritative, permissive and neglectful parenting styles

Authoritative parents have been shown to have children who are higher in self-reliance, curiosity, self-control and compliance, cheerful mood, prosocial behaviour and friendliness in peer relationships (Rothbaum, 1986). Authoritarian parenting styles have been associated with lower levels of children's social competence, as reflected in social withdrawal and anxiety, or aggressive and explosive behaviour (Baumrind, 1971; Baumrind & Black, 1967; Reitman et al., 2002). Children of permissive parents have been found to be lacking in the areas of maturity, impulse control, independence and social responsibility (Baumrind, 1971; Baumrind & Black, 1967). Children of neglectful parents have problems with attachment, cognitive development and social and emotional skills (Baumrind, 1971; Baumrind, 1991).

1.3.2 Models of influence

As mentioned in the section describing the aetiology of externalising behaviour problems in children (section 1.1.2), parenting behaviour is not the only influence on the development and persistence of children's externalising problems. Other within-child and environmental factors have been associated both with parenting and children's behaviour (Campbell et al., 1986a; Campbell & Ewing, 1990; Cunningham & Boyle, 2002; Downey & Coyne, 1990; Field, 1995; Gottman & Katz, 1989; Hinshaw, 2002; Larson et al., 1988; Lavigne et al., 1996; Miller-Johnson et al., 2002;

Nigg et al., 1999; O'Connor et al., 2002; Paterson & Sanson, 1999; Prior, 1992). Multiple pathways have been identified between within-child risk factors (e.g., difficult temperament and IQ), family risk factors (e.g., marital conflict, parental psychopathology, socioeconomic disadvantage), parenting behaviours and children's externalising behaviour problems (Belsky, 1984; Hinshaw, 2002; Keown & Woodward, 2002; Patterson et al., 1989; Rubin et al., 1995; Shaw & Bell, 1993). Family risk factors may have a direct impact on children's behaviour through the modelling of negative behaviours, inducing high levels of stress, and by providing opportunities for deviant peer affiliations (Fonagy, 1999; Gottman & Katz, 1989; Hinshaw, 2002; Lahey et al., 1999; Wilson & Gottman, 1995). However, these variables may also indirectly affect children's behaviour because they impact upon parents' abilities to employ effective parenting practices, as discussed in section 1.2.1 (Jackson et al., 2000; Lee & Gotlib, 1991). The role of children's temperament in these dynamic relationships should not be ignored. As described earlier, *interactional* models of parenting suggest that elements of children's environments (such as parenting practices, peer groups and family functioning) influence children's behaviour and development in different ways depending upon inherent child characteristics such as temperament (Paterson & Sanson, 1999; Prior et al., 1992; Rubin et al., 1995; Sanson et al., 1991; Wootton et al., 1997). Further, *transactional* models suggest that these transactions between the environment and within-child characteristics, as well as reciprocal (bidirectional) relationships between parenting practices and children's behaviour, recur across children's development (Belsky, 1984; Belsky et al., 1995; Field, 1995; Harvey-Arnold & O'Leary, 1995; Hinshaw, 2002; Lahey et al., 1999; Patterson et al., 1989; Rubin et al., 1995). Patterson's developmental model of antisocial behaviour (Patterson, 1982; Patterson et al., 1989) and Belsky's ecological model (Belsky, 1984; Belsky et al., 1995) are two transactional models of influence which have attempted to elucidate the pathways to

the development and persistence of children's externalising behaviour problems, and are described below. It is important to note that theories investigating pathways to children's behaviour problems have largely been based on data from male samples, and because socialisation processes differ between boys and girls, it has been recognised that separate models may need to be developed describing the pathways to externalising behaviour problems in girls (Shaw & Bell, 1993).

Patterson's developmental model of antisocial behaviour

Patterson's coercion theory is one of the most widely researched transactional models describing the relationships between parenting and children's behaviour (see Figure 1.1) (Patterson, 1982; Patterson et al., 1989). The model differentiates between "early" and "late" starters in terms of the development of antisocial behaviour. The "early-starter" model employs a social-interactional perspective which suggests early-onset antisocial behaviour arises because in early childhood, children with difficult temperaments elicit ineffective parenting practices (such as poor monitoring, discipline and problem-solving skills) which positively or negatively reinforce coercive child behaviours (Lahey et al., 1999; Patterson et al., 1989; Patterson & Stouthamer-Loeber, 1984).

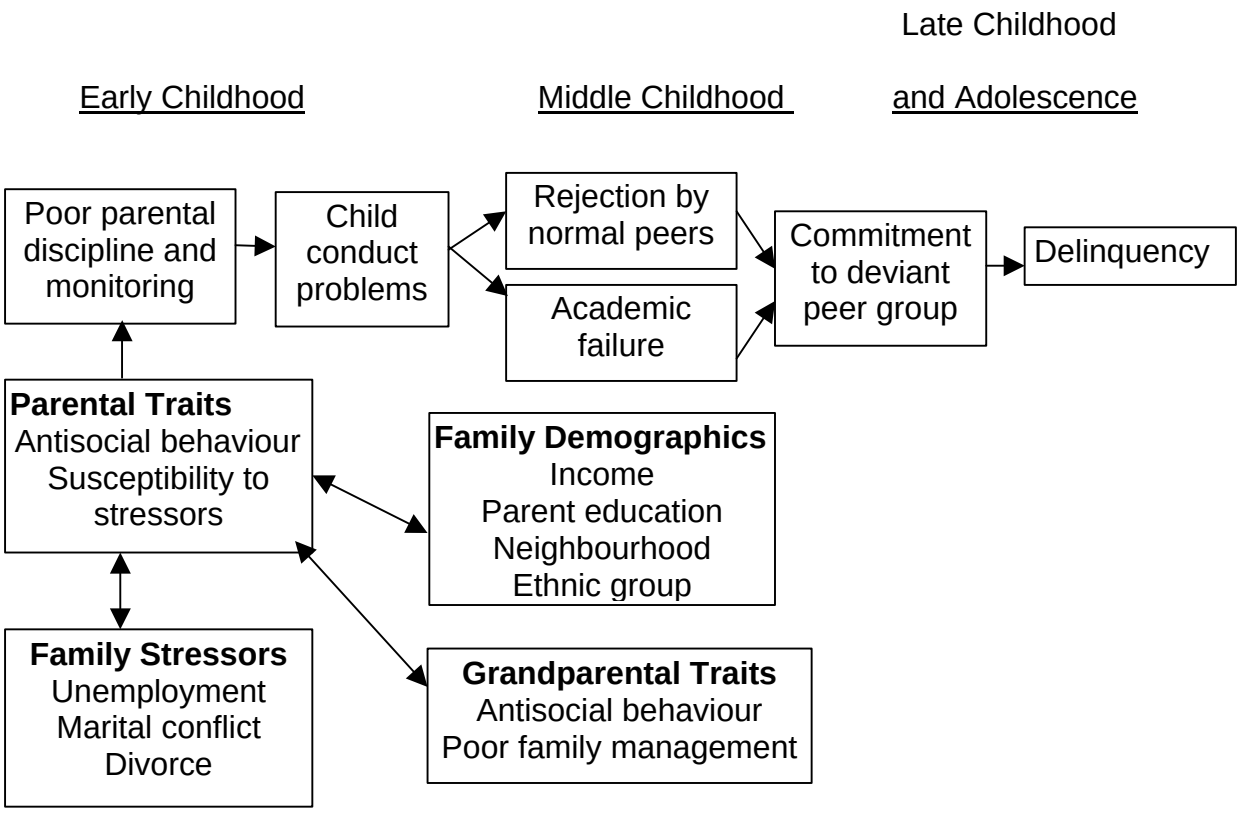

Figure 1.1 Patterson's developmental model of antisocial behaviour (figure adapted from Patterson, 1989)

As the dysfunctional relationships continue, parents and children become enmeshed in a cycle of escalating coercive interactions, eventually terminating when children learn to control their family by their own coercive behaviour (Chamberlain & Patterson, 1995).

The poor quality of the attachment bond which forms between children with difficult temperaments and their parents is thought to set the stage for the coercive interactions which occur later in childhood (Patterson et al., 1989; Shaw & Bell, 1993). The conduct problems which occur as a result of these escalating coercive interactions can lead to academic failure and peer rejection, which in turn are associated with deviant peer affiliations in adolescence (Patterson et al., 1989). In the late-starter model, the fragile balance of families with borderline behaviour

management skills is disrupted by children's move into puberty and adolescence and competing influences from peers. Late-starters are thought to have better long-term prospects than children whose problem behaviours begin earlier in life and persist into adolescence (Campbell et al., 2000; Patterson et al., 1989; Shaw & Bell, 1993).

The model identifies reciprocal transactions between children's behaviour and their social environment as predictable and ongoing, placing children at greater risk for the development of conduct problems (Patterson et al., 1989). In this model, the family stresses and demographic factors that are seen as determinants of parenting behaviour (described in Section 1.2.1) are described as "disruptors". These disruptors are seen to operate on children's behaviour problems indirectly by compromising parents' abilities to parent effectively (Patterson et al., 1989).

Belsky's ecological model

While Patterson's model proposes that interactions between parents and children are the prime determinants of subsequent parent and child behaviour, Belsky's ecological model does not put such emphasis on children's characteristics. This model posits that parenting practices are multiply determined by parents' personal resources, children's characteristics and social contextual factors (e.g., marital relations and social networks). These determinants can interact with each other to influence parenting function, which then influences child behaviour (see Figure 1.2) (Belsky, 1984; Belsky et al., 1995; Meyers, 1999). The model proposes that the influences of these parent, child and contextual factors are unequal: the personal resources of parents (i.e., personality and psychological functioning) have a greater impact on parenting behaviour, whereas child characteristics and contextual forces operate somewhat indirectly upon parenting practices with their influences moderated by

parents' personality characteristics and their susceptibility to stress (Belsky, 1984; Belsky et al., 1995; Campbell et al., 2000).

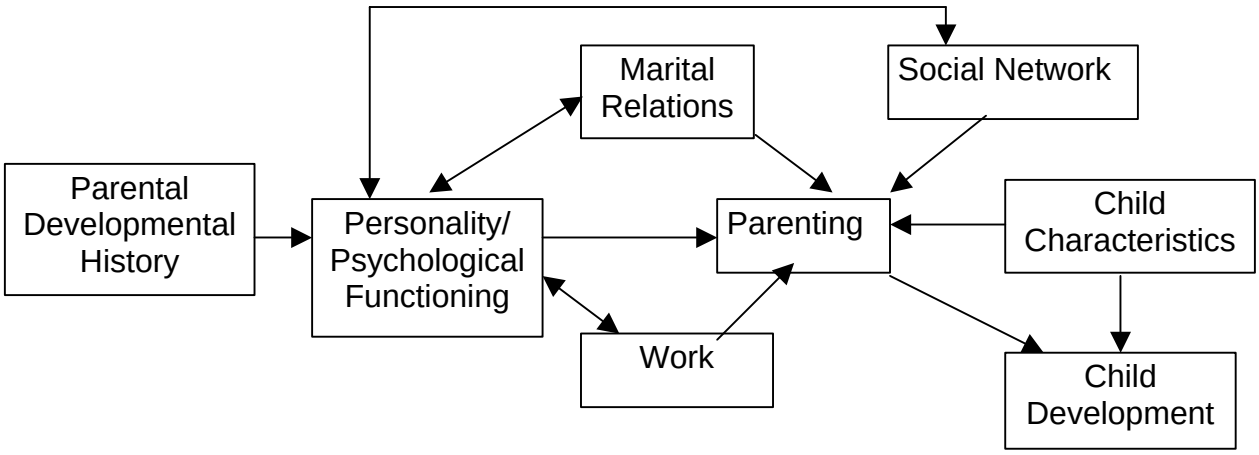

Figure 1.2 Belsky's ecological model of parenting (figure adapted from Belsky, 1984)

Further research involving this model has suggested that the effect of children's characteristics (e.g., temperament characteristics such as infant negativity) on parenting behaviour are limited. This has led to a focus on the influences of perceived child temperament as a moderator of the effects of ineffective parenting practices in the development and persistence of externalising behaviour problems (Belsky et al., 1997; Belsky et al., 1998). The hypothesis put forward in this model is that the perceived negative temperament traits of some children may make them more susceptible to ineffective parenting practices such as low levels of warmth or limit-setting and high levels of punitive or power-assertive discipline practices (Bates, Pettit, Dodge, & Ridge, 1998; Belsky et al., 1997; Belsky et al., 1998; Hemphill & Sanson, 2001; Paterson & Sanson, 1999).

1.3.3 Evidence from parent training programs

The role of parenting in the development and persistence of early-onset externalising behaviour problems is supported by the success of early intervention programs designed to prevent or ameliorate externalising problems by developing and

improving parenting skills (Greenwood et al., 1998; Kazdin, 1997; Marshall & Watt, 1999). In these programs, parenting practices are directly altered and subsequent changes in children's behaviour are compared with control groups in which parenting behaviour has not been changed.

Successful parent management training programs are based on social learning principals and teach parents to promote prosocial behaviours in their children, to identify and monitor children's misbehaviour, and to apply effective and contingent discipline techniques to reduce levels of misbehaviour (Kazdin, 1997; Marshall & Watt, 1999). Kendziora and O'Leary (1993) argue that dysfunctional parenting should be included in diagnostic systems and treated directly to prevent future childhood psychopathology, not just to improve children's existing behaviour problems. It is important to note that because of the multifactorial aetiology of childhood externalising behaviour problems, parenting management training programs alone are generally not sufficient to prevent or treat externalising behaviour problems in all children. However several programs implemented with parents of young children have demonstrated improvements in the behaviour of young children which have been maintained in the longer term (Baum & Forehand, 1981; Emery, Fincham, & Cummings, 1992; Kazdin, 1997; McMahon, 1994; Neary & Eyberg, 2002; Sanders & Markie-Dadds, 1996; Webster-Stratton, 1990). The parenting behaviours which are the foci of specific parent training programs, and the outcomes of these programs are described below.

The Forehand Parent Training Program

Forehand and McMahon (1981) constructed a modified parent training program specifically for parents of *non-compliant* children, which utilises a controlled learning environment in which parents are taught the maladaptive patterns of parent-child interaction. Parenting skills are taught in a clinical setting using didactic instruction, modelling and role play. Parents are expected to practice the skills in the home. The parent training program proceeds in two phases (Baum & Forehand, 1981; Forehand, Furey, & McMahon, 1984). First, parents are taught to only attend to and describe children's appropriate behaviour and ignore misbehaviour, whilst eliminating the use of commands, questions and criticisms. This phase is designed to increase parents' roles as reinforcers by increasing the quality and frequency of social rewards and reducing competing social behaviour. Second, parents are taught to use clear, concise commands, allow time for children's compliance, reward compliance with attention, and institute time-outs for noncompliance. Training proceeds until all phases are mastered, as assessed by observational assessments (Forehand et al., 1984).

In clinical samples, the Forehand Parent Training Program has been associated with improvements in children's externalising behaviour (increased compliance and decreased deviance) and parenting (attending, rewarding and displaying contingent attention), both immediately after treatment completion and at follow-up assessments up to four and a half years later (Baum & Forehand, 1981; McMahon, 1994).

Parent-Child Interaction Therapy

Parent-Child Interaction Therapy (PCIT), uses evidence-based practice to reduce ineffective parent-child interactions in the families of preschool-age children with disruptive behaviour (Neary & Eyberg, 2002). PCIT is based on attachment and social learning theories and focuses on promoting secure attachments between parents and children, and increasing prosocial behaviour, whilst decreasing the incidence of children's misbehaviour (Neary & Eyberg, 2002). The components of PCIT are similar to those used in the Forehand Parent Training Program. The first session of PCIT involves a teaching session with parents during which the skills are explained and demonstrated using modelling and role play. Therapy then progresses in two phases. First, during "Child-Directed Interaction" (CDI) parents are taught to apply praise, reflect appropriate talk, imitate play, describe appropriate play and demonstrate enthusiasm, without issuing commands, questions or criticisms. These techniques are designed to foster secure attachment in children. Second, during "Parent-Directed Interaction" (PDI) parents use the skills gained in CDI and also develop skills to direct children's behaviour such as providing direct, specific, positive, respectful and age-appropriate commands, one at a time as necessary. Parents are taught to give reasons for their instructions to children, to direct their attention contingently to children's responses and to initiate time-out for non-compliance. This training is designed to promote parental consistency and to clarify expectations for children's behaviour (Neary & Eyberg, 2002). Parents are trained in the clinic in weekly one hour sessions and are coached via a bug in the ear device. Parents are expected to practice the skills they obtain in the lab at home, and parents in two-parent families are encouraged to participate in the training together. The program continues until parents have mastered the interaction skills and children's behaviour is within normal limits (Neary & Eyberg, 2002).

Studies have found that children whose parents have undergone PCIT demonstrate behavioural improvements, reaching normative levels of behaviour (Neary & Eyberg, 2002). These results have been shown to persist over time, to have positive effects on family functioning, and to generalise both to siblings of index children and to school settings (Brestan, Eyberg, Boggs, & Algina, 1997; Eyberg & Robinson, 1982; McNeil, Eyberg, Eisenstadt, Newcomb, & Funderburk, 1991). Parents have demonstrated improved interactional styles including decreased criticism and sarcasm and increased physical proximity, compared with controls (Neary & Eyberg, 2002).

The Webster-Stratton Video Modelling Group Discussion Program

This parent training program is based on modelling theory and uses videotape modelling and group discussion to teach parents time-out, logical and natural consequences, monitoring, problem-solving and communication (Webster-Stratton, 1981). Videotapes depicting parent models of different ages, genders and cultural backgrounds are used in 13 therapist-led group sessions. The therapist leads the group in an involved discussion regarding the content of the interactions and the parents' ideas and feelings. The therapist-led video modelling groups have been associated with decreases in children's behaviour problems to normal levels, both immediately after the program and three years later (Scott, Spender, Doolan, Jacobs, & Aspland, 2001; Webster-Stratton, 1990; Webster-Stratton, Hollinsworth, & Kolpacoff, 1989).

Similarly, the PARTNERS program using the video modelling group discussion program described above, provided training to both the parents and teachers of intervention children (Webster-Stratton, 1994; Webster-Stratton et al., 1989). After the intervention, parents were observed to use less critical remarks and commands,

less harsh discipline and were more positive and effective in their discipline. Treatment children were observed to display fewer externalising behaviour problems. These treatment effects for parents and children were maintained at follow-up assessment 12 to 18 months later (Webster-Stratton, 1994; Webster-Stratton & Hancock, 1998).

The Positive Parenting Program (PPP)

The Positive Parenting Program is a multi-level behavioural family intervention in which five different levels of the program are administered to parents based on their level of risk, need and/or preference (Sanders & Markie-Dadds, 1996). The five levels of the intervention are:

Level 1. Self-help information and advice, which includes instructions about dealing with common developmental and behavioural problems;

Level 2. Information plus minimal contact with a therapist about minor behaviour problems;

Level 3. Information and active skills training which involves a brief therapeutic program helping parents to manage particular child behaviour problems;

Level 4. Intensive behavioural parent training, which involves a more intensive therapeutic program focusing on parent-child interactions, and targeting more serious childhood behaviour problems; and

Level 5. Enhanced behavioural family intervention, which involves an intensive intervention program for families of children with severe conduct problems. This level of the intervention focuses on treating behaviour management problems as well as family and marital dysfunction.

Components of the more intensive levels of intervention focus on teaching parents to promote children's competence and development (via techniques such as quality time, physical affection, praise, attention, modelling) and to teach parents to manage

misbehaviour (via strategies such as limit-setting, planned ignoring, logical consequences, and time-out) (Sanders & Markie-Dadds, 1996; Zubrick et al., 2002).

Implemented at various levels in a range of preschool populations in rural and metropolitan Australia, both individual and group administration of PPP have been associated with reductions in the levels of childhood disruptive behaviour and total behaviour problems, ineffective parenting practices (e.g., laxness, verbosity and overreactivity) and family dysfunction (e.g., parental depression, anxiety and marital satisfaction) (Bor et al., 2002; Marshall & Watt, 1999; Sanders & Markie-Dadds, 1996; Zubrick et al., 2002). Improvements in parents' and children's behaviour have been demonstrated to persist over one to two-year follow-up periods in high risk and clinic samples of Australian families (Bor et al., 2002; Zubrick et al., 2002). However, in the study by Zubrick et al. (Zubrick et al., 2002), the effects of the program on family functioning had dissipated by the two-year follow-up assessment.

In summary, two empirically supported transactional models of parenting behaviour were described, in which reciprocal interactions between children's behaviour, parenting and environmental factors occur across the lifespan and are influential for children's development. Evidence of a relationship between parenting and children's behaviour was provided which included the results from parenting interventions in which parenting behaviour was directly manipulated and subsequent improvements in children's behaviour were reported.

Chapter 2. The Measurement of Parenting Behaviour

Given the importance of parenting as a risk factor for the development and persistence of behaviour problems, and the potential of parenting as a target for intervention efforts, there is a need to identify the most valid and reliable measures of parenting behaviour for early childhood (Hemphill & Sanson, 2001; Locke & Prinz, 2002; Reitman et al., 2001). A review by Locke and Prinz (2002) reported that in the past 20 years, 76 self-report questionnaires, 27 interview schedules and 33 observational systems have been developed to measure parental discipline (control) and nurturance (responsiveness) practices. Over this time, there has been a shift from a singular reliance on self-reports (e.g., questionnaires and interviews) to an increasing emphasis on the use of observational measures (Alessi, 1988; Chamberlain & Patterson, 1995; Dowdney et al., 1984; Hemphill & Sanson, 2001; Locke & Prinz, 2002). More recently, studies of parenting behaviour have incorporated both observational and self-report techniques, but because of the high costs of such mixed-method approaches, these studies have often been limited to questionnaire validation studies or to research with small or clinic-based samples (Arnold et al., 1993; Banister et al., 1996; Bor et al., 2002; Crnic & Greenberg, 1990; Cunningham & Boyle, 2002; Denham et al., 2000; Dowdney et al., 1984; Feinberg, Neiderhiser, Howe, & Hetherington, 2001; Harvey-Arnold & O'Leary, 1997; Johnston et al., 2002; Kaplow, Curran, Dodge, & The Conduct Problems Prevention Research Team, 2002; Keown & Woodward, 2002; Roberts & Strayer, 1987; Shaw, Owens, Giovannelli, & Winslow, 2001; Stormshak, Speltz, DeKlyen, & Greenberg, 1997; Strayhorn & Weidman, 1988).

This thesis aims to provide a *systematic* comparison of the methods used to measure current parenting practices at the micro-analytic level in a sample of parents of preschool children. To this end, this chapter examines the existing methods for

measuring parenting behaviour (Section 2.1), and then reviews studies which have directly compared these methods of measuring parenting practices and other phenomena (Section 2.2). The relationships between self-reports and observations of parenting behaviour and factors such as family characteristics and children's externalising behaviour are examined in the final two sections of this chapter.

It should be noted that other methods, such as proxy-reports from partners and children, have been employed to assess parenting practices (Bates & Bayles, 1984; Shelton, Frick, & Wootton, 1996, Brendgen, Vitaro, Tremblay, & Lavoie, 2001; Feinberg et al., 2001; Mattanah, 2001; Patterson & Stouthamer-Loeber, 1984; Rey & Plapp, 1990). However, because these techniques have been used infrequently with parents of preschoolers (Davilla, 1995; Holigrocki, Kaminski, & Frieswyk, 1999; Roberts & Strayer, 1987), this review of parenting measures will not include proxy-report methods. Also, the focus of this thesis is on the measurement of parenting behaviour, and so the large number of self-report instruments designed to assess other aspects of parenting such as parental attitudes, beliefs, sense of competency, satisfaction, coping and parenting stress, are not discussed here (see Baum & Forehand, 1981; Crnic & Greenberg, 1990; Dowdney et al., 1984; Holden & Edwards, 1989; Loyd & Abidin, 1985; Reitman et al., 2002; Sears, 1965; Sonuga-Barke, Daley, Thompson, Laver-Bradbury, & Weeks, 2001).

2.1 Methods Used to Measure Parenting Practices

2.1.1 Self-report measures

Self-report measures, such as questionnaires and interviews, have proliferated over the past twenty years (Locke & Prinz, 2002). This is because highly structured self-report measures are time and cost effective, easy to administer, and enable the comparison of large samples of participants (Arnold et al., 1993; Mrazek et al., 1982).

Further, self-report measures allow researchers to assess a broad range of behaviours (including those which occur infrequently) in a variety of situations, over an extended period of time (Locke & Prinz, 2002; Mrazek et al., 1982). Although the results from some self-report methods (e.g., open-ended interviews) can be more expensive to conduct and analyse, they can allow researchers to explore the sequential and reciprocal nature of parent-child interaction (Mrazek et al., 1982).

The table in Appendix A.1 lists some of the more frequently used self-report measures designed to assess the parenting practices of parents of children aged between two and six years. Because parents display different parenting behaviours depending on the developmental level of their children (Chamberlain & Patterson, 1995; Patterson & Stouthamer-Loeber, 1984; Roberts & Strayer, 1987; Shaw et al., 1998), this table is limited to measures used in early childhood (see Locke and Prinz, 2002 for a review of self-report measures for other age groups).

The self-report measures shown in Appendix A.1 assess aspects of parental control (e.g. consistency, laxness, supervision and limit-setting) (Arnold et al., 1993; Block, 1981; Dwyer, Nicholson, & Battistutta, 2003; Fox, 1994; Hart et al., 1992; Johnston & Behrenz, 1993; Quinton, Rutter, & Liddle, 1984; Sanson, 1996; Shelton et al., 1996; Strayhorn & Weidman, 1988), and/or responsive parenting practices (e.g. warmth, communication, acceptance and sensitivity) (Block, 1981; Dwyer et al., 2003; Fox, 1994; Johnston & Behrenz, 1993; Quinton et al., 1984; Sanson, 1996; Shelton et al., 1996; Strayhorn & Weidman, 1988). Many of these scales were developed in the United States (Arnold et al., 1993; Block, 1981; Fox, 1994; Shelton et al., 1996; Strayhorn & Weidman, 1988), and although they are widely used in other Western cultures, their applicability to populations in other geographic regions is largely unknown. Two of the self-report parenting measures included in Table A.1.1,

although less widely used were included because they were developed in Australia (Dwyer et al., 2003; Sanson, Prior, & Kyrios, 1990).

Frequency ratings on Likert scales are the most commonly used response format for self-report measures of parenting (Locke & Prinz, 2002). Instruments which use this format include the: *Parent Behavior Checklist* (Fox et al., 1995)*, Parent Practices Scale* (Strayhorn & Weidman, 1988), *Child-Rearing Practices Questionnaire* (Sanson, 1996), *Alabama Parenting Scale (Shelton et al., 1996), and Family Risk Factor Checklist-Parent* (Dwyer et al., 2003).

The *Parent Behavior Checklist* (PBC) is a 100-item rating scale with three factors, two of which refer to parenting behaviours (verbal/corporal discipline and nurturing) and one which refers to expectations for the child's development expectations (Fox et al., 1995). Items on the scale are quite context-specific and include: *"I yell at my child for spilling food", "I read to my child at bedtime"* and *"my child should be able to feed him/herself"*. Scores on the scales of the PBC are transformed into T-Scores to account for differences in parenting behaviour for parents of children of different ages. The PBC is psychometrically sound and the parenting practices assessed by the PBC reflect Baumrind's authoritarian, authoritative and permissive parenting styles.

The *Parent Practices Scale* (PPS) includes 34 parenting practices that are frequently targeted in parent training (Strayhorn & Weidman, 1988). Items on the PPS are quite explicit, with concepts such as praise and physical punishment explained to parents, eg *"How often do you praise your child by saying something like 'Good for you!' 'What a nice thing you did!' 'Thank you!' or 'That's good going'?"*; and *"How often do you physically punish your child, for example by spanking?"*. Some items are quite complex and ask parents to rate the fraction of the time that a given event occurs, eg

"What fraction of days does your child get a bath or shower at one particular time, known as his or her bathtime": *Never; Some, but less than a quarter of the days; Between a quarter and half the days; Between half and three quarter of the days; Not all the days but more than three quarters of the days; All the days.* A 27-item adapted version of the PPS, the *Parenting Questionnaire, (PQ*), contains three subscales: Warmth, Consistency and Punitive Discipline (Kaplow et al., 2002; Stormshak et al., 2000). The PPS and the PQ have been used for screening and evaluation purposes and have good psychometric properties (Brannigan et al., 2002; Kaplow et al., 2002; Stormshak et al., 2000; Strayhorn & Weidman, 1988).

The *Child-Rearing Practices Questionnaire* (CRPQ) is a 30-item measure developed in Australia to assess aspects of parental warmth, reasoning, punitive discipline and expectations of child compliance (Hemphill & Sanson, 2001; Paterson & Sanson, 1999; Sanson, 1996). Examples of items on the scale include: *"I express affection by hugging, kissing and holding my child"; "I try to explain to my child why certain things are necessary"; "I slap or hit my child to control his/her behaviour";* and *"I expect my child to give his/her parents unquestioning obedience"*. There is evidence for the validity of the CRPQ in samples of Australian children (Hemphill & Sanson, 2001; Paterson & Sanson, 1999). The CRPQ will be discussed in more detail in the following chapter.

The *Alabama Parenting Questionnaire* (APQ) (Shelton et al., 1996) was developed from a framework by Patterson and colleagues at the Oregon Social Learning Centre. The APQ is designed for parents of children aged five and older, and includes 42 items that are rated for their frequency on 5-point Likert scales. A questionnaire format is used to assess the "typical" frequencies of parenting behaviour and more specific telephone interview probes can be used to ask about behaviour on the items "in the past few days". The APQ assesses elements of both

responsive parenting and parental control with five subscales: Involvement, Positive Parenting, Poor Monitoring/Supervision, Inconsistent Discipline and Corporal Punishment. Examples of items from the different subscales include: *"you have a friendly talk with your child"; "you reward or give something to your child for obeying you or behaving well"; "your child is at home without adult supervision"; you let your child out of a punishment early"*; and *"you spank your child when he/she has done something wrong".* Seven additional items do not load of any of the factors. The APQ has excellent psychometric properties and child-reports can also be obtained on the APQ, but are not recommended for children under nine years of age (Brubaker & Szakowski, 2000; Dwyer et al., 2003; Shelton et al., 1996; Wootton et al., 1997).

The *Family Risk Factor Checklist-Parent (FRFC-P)* is a new screening measure developed to identify risk for persistent child mental health problems (Dwyer et al., 2003). The FRFC-P combines eight parent practices items with another 40 items assessing risk in areas such as: adverse life events and instability; family structure and socioeconomic status; parental verbal conflict and mood problems; and parental antisocial and psychotic behaviour. The Parenting Practices subscale of the FRFC-P contains items related to warmth, involvement, praise, enforcing rules, yelling and physical punishment (Dwyer et al., 2003). The measure has demonstrated adequate psychometric properties in a sample of 1022 Australian parents (Dwyer et al., 2003).

Other response formats for parent-self report instruments include: Q-sorts in which parents sort items into response groups such as often, sometimes or rarely; anchored items in which parents endorse one of two alternatives; and the coding or rating of open-ended responses. The *Child-Rearing Practices Report* (CRPR) (Block, 1981) is a Q-sort instrument in which parents sort 91 items from most to least descriptive. The CRPR is not a pure measure of parenting behaviour as it includes some belief and attitude items. Items include statements such as: "*I teach my child*

that in one way or another punishment will find him when he is bad"; "*I let my child know how disappointed and ashamed I feel when he misbehaves*"; *"I talk it over and reason with my child when he misbehaves"; "I believe in praising a child when he is good and think it gets better results than punishing him when he is bad"*. The CRPR assesses aspects of responsiveness and control such as warmth/nurturance, and restrictiveness/strictness. This measure has been widely used, and demonstrates good internal consistency and validity (Block, 1981; Denham et al., 2000; Locke & Prinz, 2002; Roberts & Strayer, 1987).

The Parenting Scale (PS) (Arnold et al., 1993) was developed from a review of parenting literature and transcripts of interviews with parents. The scale involves parents' endorsements on a continuum anchored with empirically supported ineffective parenting behaviours and their effective counterparts. Because parents rate along a continuum of what they do, rather than how often they do it, parent ratings are frequency-independent. Examples of items include: "*When my child misbehaves I do something about it later*" or "*I do something about it straight away*"; "*When I'm upset and under stress I am picky and on my child's back*" or "*I am no more picky than usual*"; "*When my child misbehaves I give my child a long lecture*" or "*I keep my talks short and to the point*". In the original development of the PS, three factors were identified: Laxness, Overreactivity and Verbosity (Arnold et al., 1993). The robustness of the Overreactivity and Laxness scales has been demonstrated in subsequent factor analyses with a variety of samples, but there has been very limited support for the Verbosity Scale (Collett, Gimpel, Greenson, & Gunderson, 2001; Harvey, Danforth, Ulaszek, & Eberhardt, 2001; Irvine, Biglan, Smolkowski, & Ary, 1999; Reitman et al., 2001). The PS has been widely used and the Overreactivity and Laxness scales have demonstrated good reliability and validity (Bor et al., 2002; Harvey-Arnold & O'Leary, 1995; Harvey-Arnold & O'Leary, 1997; O'Leary et al.,

1999; Zubrick et al., 2002). The PS will be discussed in more detail in the following chapter.

Interview measures of parenting behaviour include the Parental Account of Child's Symptoms (Quinton et al., 1984), Disciplinary Style Interview (Hart et al., 1992), and the Child Management Problem Solving Skills Interview (Johnston & Behrenz, 1993). Each of these measures is designed to assess aspects of parental responsiveness and control. The instruments are context-specific in that they require parents to recall specific parenting experiences (Quinton et al., 1984), or to respond to hypothesised parenting situations (Hart et al., 1992; Johnston & Behrenz, 1993). Examples of prompts in the Disciplinary Style Interview include asking parents what they would do if their child *"picked flowers in the next door neighbours garden", "refused to go to bed at bedtime on a school night", "hurt the feelings of another child by namecalling"* (Hart et al., 1992). Prompts for the Child Management Problem Solving Skills Interview are vignettes such as: "*Your child doesn't like to do her household chores and will argue with you about how much there is to be done. She will continue to argue even after you insist and always tries to get the last word in the argument"* (Johnston & Behrenz, 1993). For all of the interview measures, parents' open-ended responses are transcribed and rated according to predefined coding systems (Hart et al., 1992; Johnston & Behrenz, 1993; Quinton et al., 1984). Intercoder reliabilities are generally high for these measures and all have demonstrated validity (Cunningham & Boyle, 2002; Hart et al., 1992; Johnston & Behrenz, 1993; Keown & Woodward, 2002; Quinton et al., 1984).

2.1.2 Direct Observations

Data collected from behavioural observations are considered more "objective" than self-report data gathered via alternative methods because the systematic observation of behaviour is regarded as a means of depicting reality (Fassnacht, 1982). Observations

have high face validity because they directly measure the actual behaviours of interest. Families are presented with real-life situations, and there is no need for interpretation of questions by parents. Observations also allow investigators to examine sequences of interactions and the fine-grained processes that describe how interactions take place (Mrazek et al., 1982).

The use of direct observations to assess parent-child interactions has been an increasing practice since the 1950s (Locke & Prinz, 2002; Pease & Hawkes, 1960). In the past two decades many diverse approaches towards observation have been developed. Locke and Prinz (2002) identified 31 standardised coding systems designed to assess parental control and responsivity in parents of preschool-aged children. Further examination of the literature revealed an additional five standardised coding systems for this age group (Bor et al., 2002; Dadds & McHugh, 1992; Frankel & Harmon, 1996; Hops, Davis, Leve, & Sheeber, 2003; Kleberg, Westrup, & Stjernqvist, 2000; Whipple et al., 1995; Youngblade & Belsky, 1995). In addition to the formal coding systems, many observational studies have developed their own ad hoc sampling or rating systems for their specific research purposes (Acker & O'Leary, 1996; Arnold et al., 1993; Campbell et al., 1986b; Campbell, March, Pierce, Ewing, & Szumowski, 1991; Crnic & Greenberg, 1990; Denham et al., 2000; Gardner, 1989; Gardner et al., 1999; Hemphill & Sanson, 2001; Mrazek et al., 1982; Russo & Owens Jr, 1982; Strayhorn & Weidman, 1988). The lack of consistency and standardisation between the huge number of coding systems which have proliferated over the past three decades, makes it difficult to make comparisons between observational studies of parenting constructs such as control and responsivity (Locke & Prinz, 2002). Appendix A.2 displays a selection of observational studies that have assessed a range of parenting domains in parents of preschool children. The studies presented in the table are not exhaustive but

were chosen to be representative of the wide variety of observational techniques used to study parenting practices.

The observational coding systems represented in Appendix A.2 reflect a shift in the methods used to code observational data from microanalytic methods which involve calculating absolute or estimated frequencies of behaviours (i.e., event-recording and interval-sampling methods) to an increased use of behaviour ratings (particularly from the mid 1990s onwards). Studies which have used more than one method to code observational data have typically used sequential coding or event sampling to record concrete parent behaviours (eg requests for compliance, physical punishment) and their antecedents, and rating systems to code affective or stylistic qualities such as warmth, anger and authoritarian parenting (Brophy & Dunn, 2002; Campbell et al., 1991; Dowdney et al., 1984; McFadyen-Ketchum, Bates, Dodge, & Pettit, 1996; Roberts & Strayer, 1987). The microanalytic and rating methods used to record the parenting behaviour of parents of preschoolers are briefly examined in this section.

Microanalytic methods

Microanalytic methods produce considerable amounts of data to be coded and analysed (Arnold et al., 1993), and for some coding systems it may take weeks or months for coders to achieve sufficient levels of inter-observer reliability (Belsky et al., 1997; Belsky et al., 1998; Dowdney et al., 1984; Lytton, 1973; Whipple et al., 1995). The main microanalytic coding methods include event recording, interval sampling, sequential systems and narrative recording.

Event recording involves recording a behaviour every time it occurs, and is used to quantify behaviour. This is a comprehensive method of recording data because all behaviours in an observation are recorded, not just a sample of them (Fassnacht, 1982). This continuous coding method relies heavily on the discriminatory power used to describe events and needs definitive criteria about where certain behaviours begin and end (Mrazek et al., 1982). Several standardised event recording systems with good psychometric properties have been used to study parenting behaviour, including the: Parent Child Interaction Code (PACIC) (Lytton, 1973), Dyadic Parent-Child Interaction System (DPCIS) (Robinson & Eyberg, 1981), INTERACT (Dumas & Gibson, 1990), Family Interaction Coding System (FICS) (Kalpidou et al., 1998; Reid, 1978), Early Parenting Coding System (EPCS) (Shaw et al., 2001; Shaw et al., 1998) and Living in Familial Environments Coding System (LIFE) (Hops et al., 2003). The continuous coding in these systems allows the calculation of absolute frequencies and durations of behaviour (Hops et al., 2003; Shaw et al., 2001; Shaw et al., 1998). Typically, in event recording systems several parental behaviours are coded which are then aggregated into patterns of parenting behaviour such as warmth, aversive behaviour, negative control, positive control, inconsistency, critical statements, reinforcing behaviour or aggressive behaviour (Bates et al., 1998; Brophy & Dunn, 2002; Campbell et al., 1986a; Campbell et al., 1986b; Dumas & Gibson, 1990; Gottman & Katz, 1989; Lytton, 1973;

Robinson & Eyberg, 1981). Despite the complex and continuous nature of many of these recording systems, very high levels of inter-observer reliability have been reported (Bates et al., 1998; Brophy & Dunn, 2002; Campbell et al., 1986b; Hops et al., 2003; Kalpidou et al., 1998; Reid, 1978; Robinson & Eyberg, 1981; Shaw et al., 1998). This probably reflects how explicitly behaviours are operationalised in these coding systems, and the extensive training provided to observers.

Interval sampling involves recording behaviour in a series of short time periods and is designed to provide an estimate of the frequencies of the behaviours under observation (Banister et al., 1996). Interval sampling employs checklists upon which operationalised behaviours are marked as present or absent in a series of time intervals (Brandt, 1992). The length of the standard interval should ideally be matched with the 'natural' length of the behaviour so as not to over- or underestimate the relative frequency of behaviour (Fassnacht, 1982). Observational studies of parenting behaviour have typically used interval lengths from ten seconds to one minute. A standardised interval sampling system is the Family Observation System (FOS) and its revised version the FOS-RIII (Bor et al., 2002; Dadds & McHugh, 1992; Dadds, Schwarttz, & Sanders, 1987). The FOS is an intervention-specific measure used to assess correct program implementation and domains of parenting behaviour such as negative parenting and responsivity. Other studies have used interval sampling methods to examine aspects of parental attention, responsivity and negative control (Baum & Forehand, 1981; Campbell et al., 1991; Cunningham & Boyle, 2002; Hemphill & Sanson, 2001; Mrazek et al., 1982).

Sequential coding systems attempt to capture the process of interactions by examining the sequences in which behaviours occur (Dowdney et al., 1984). Sequential systems incorporate information about the reciprocal nature of parent-child interactions and can be used to investigate contingencies between parent-child behaviours. Behavioural sequences can be studied by: a. Investigating which parent

and child behaviours occur in time intervals immediately following or preceding those in which focal behaviours occur (Gardner et al., 1999; Tarver-Behring, Barkley, & Karlsson, 1985); b. Continuously recording behaviour and examining the continual record for particular behavioural patterns (dynamic systems analysis) (Granic & Lamey, 2002); and c. Recording previously defined sequences as they occur (Dowdney et al., 1984; Mrazek et al., 1982). A sequential system used with parents of preschoolers is the Fagot Interactive Behaviour Code which involves a complex context-interaction-recipient-reactor-reaction sequence to record behaviours such as "comment favourably", "instruct", "verbal interaction", "associative and cooperative play", "ignoring child" and "look at child" (Kavanagh, Youngblade, Reid, & Fagot, 1988). Other sequential systems include those employed to study reciprocal aspects of parent-child behaviour such as coercion, agonistic exchanges, and responses to upset (Dowdney et al., 1984; Mrazek et al., 1982; Roberts & Strayer, 1987; Tarver-Behring et al., 1985), and a system designed to study the timing of parental behaviours (Gardner et al., 1999). Sequential systems are used less commonly than other microanalytic systems because of their complexity and the higher costs associated with coding the behaviour of all participants in an interaction.

The least commonly used microanalytic method is narrative records. These include written records of observed behaviours and may also include details of the context of the observation (e.g., day, weather, location), and information about the observer's feelings and impressions during the observation (Banister et al., 1996). Problems of reliability and standardisation occur with narrative records because of the more subjective nature of this type of data collection, and the analysis of narrative information is more difficult than that of quantitative data (Bates et al., 1998; Dowdney et al., 1984; McFadyen-Ketchum et al., 1996; Mrazek et al., 1982).

Rating systems

In addition to, or instead of, recording *what* behaviour occurs and how often, researchers may *rate* behaviour according to a structured scale (Coolican, 1990). Ratings are also known as "judgements" because the rater makes a subjective interpretation about the level of behaviour compared with norms established in training sessions (Banister et al., 1996). Ratings make it easy to quantify and compare behaviours that are otherwise hard to measure (Brandt, 1992). They can also incorporate qualitative and contextual information, enabling comparisons to be made across contexts (Maxwell & Pringle, 1983). The two methods used to rate observed behaviour are global ratings and aggregated ratings.

The global rating method involves making a single rating for a parenting domain (eg "Warmth" or "Punitiveness") for the entire observation period (Arnold et al., 1993; Campbell, 1994; Crnic & Greenberg, 1990; Denham et al., 2000; Dowdney et al., 1984; Kaplow et al., 2002; Olson, Bates, Sandy, & Lanthier, 2000; Olson, Bates, Sandy, & Schilling, 2002; Sears, 1965; Shaw et al., 2001; Stormshak et al., 1997; Strayhorn & Weidman, 1988). This coding method decreases the cost of analysing behavioural data because usually only a single viewing of the observed interaction is required to make the global rating. However, global ratings involve a high degree of abstraction from the actual observed behaviour, and the apparent simplicity of ratings hides the complexity of the observers' judgement process (Dowdney et al., 1984). Global rating systems are often developed for the specific purposes of the study in which they are used (e.g., as a validity check for self-report measures), making it difficult to generalise findings across studies which use global ratings of similar behaviours. No standardised global coding systems have been reported for parents of preschool-aged children, but O'Leary and colleagues have used a global coding system to assess observed overreactivity, laxness

and general levels of ineffective parental discipline in a number of studies (Arnold et al., 1993; Harvey-Arnold & O'Leary, 1995; Harvey-Arnold & O'Leary, 1997).

Aggregated ratings include composites of ratings of different behaviours made for an entire observation period (Bank, Forgatch, Patterson, & Fetrow, 1993; Bates et al., 1998; Caldwell & Bradley, 1984; Conger, Neppl, Kim, & Scaramella, 2003; Frankel & Harmon, 1996; Kaplow et al., 2002; Kleberg et al., 2000; Roberts & Strayer, 1987), or ratings of the same behaviour made at set intervals (Keown & Woodward, 2002; Russo & Owens Jr, 1982; Whipple et al., 1995; Youngblade & Belsky, 1992). Standardised coding systems which make use of aggregated rating systems are often quite large (e.g., 46 rating scales for the Baumrind Rating Scales and 65 items on the Parent Child Early Relational Assessment Scale) (Frankel & Harmon, 1996; Kleberg et al., 2000; Roberts & Strayer, 1987), and require extensive training (e.g., 46 hours training for the Belsky Coding System for Parent-Child Interaction) (Youngblade & Belsky, 1995). However, rating measures such as the Home Observation for Measurement of the Environment (HOME) have been widely used because they employ simple rating scales in which the observer indicates whether items are present or absent on the basis of informal home observations (Bradley, 1993; Caldwell & Bradley, 1984; Jackson et al., 2000; Olson, Bates, & Kaskie, 1992). Inter-rater reliabilities for aggregated ratings are generally lower than those for global ratings because differences between observer ratings are magnified when ratings are added or multiplied (Hutt & Hutt, 1974; Keown & Woodward, 2002; Russo & Owens, 1982; Whipple et al., 1995; Youngblade & Belsky, 1995).

2.2 Direct Relationships between Self-reported and Observed Behaviour

There are many reasons why self-reports and observations of the same parenting behaviours may not yield congruent results. First, perfect agreement between the methods is not expected because responses to self-report measures sample a wider range of behaviours (e.g., over time and in different contexts), compared with the behaviours sampled in the context of a single observation (Dowdney et al., 1984; Gardner, 2000). Other features of self-reports and observations that influence their objectivity and representativeness are briefly discussed here.

Despite their convenience and breadth of assessment, self-report measures of parenting behaviour may not reflect actual behaviour for several reasons (Fassnacht, 1982). Parents may not be used to describing their usual practices and may have distorted recall of their actual parenting behaviours (Holden, 1983; Mrazek et al., 1982). The work of several researchers has suggested that self-reports contain an objective component (a report of actual parenting behaviour) and a subjective component, which may be influenced by parental characteristics such as family structure, education status, family socioeconomic status and parental psychopathology and distress (Alessi, 1988; Bates & Bayles, 1984; Chamberlain & Patterson, 1995; Forehand et al., 1984; Lancaster, Prior, & Adler, 1989; Vitaro, Tremblay, & Gagnon, 1995). These characteristics may influence parents' responses in a systematic but not deliberate fashion (response biases).

Respondents can also be influenced, both consciously and unconsciously by the social desirability of their responses, and their interpretation of the purposes of the study (Rothbaum, 1986). Further, it is unknown to what extent the responses of an individual reflect their *actual* behavioural practices, rather than their *knowledge* of effective parenting practices, particularly when parents have undertaken an

intervention targeting ineffective practices (Aspland & Gardner, 2003; Patterson, 1982).

Even though standardised self-report measures present participants with identical questions, different parents do not always interpret these questions in the same way. This may be because items are ambiguous, leading, not applicable, worded in the third-person, or double-barrelled (i.e. there are two focal statements in a single item) (Holden, 1983). Some scales that are intended to purely measure behaviour are contaminated by items assessing parental attitudes and beliefs which do not necessarily reflect actual behaviour (Holden & Edwards, 1989; Reitman et al., 2001). Different respondents can have different interpretations of the frequency categories provided (e.g., often, sometimes, never), and parents may respond about absolute rather than relative levels of behaviour. Whilst a parent who spends more time in child-rearing tasks might report a higher absolute frequency of parenting behaviours compared with a parent who spends less time with their child, their *relative* frequencies of behaviour could be quite similar. Finally, most self-report questionnaire measures of parenting are not context-specific, in that they ask about parenting practices "in general". In these instances, parents who respond differently to their children's behaviour depending upon the setting, presence of other family members or type of child behaviour, may have difficulty reporting upon their parenting behaviour "in general" (Dowdney et al., 1984; Johnson, 2001; Locke & Prinz, 2002; Mrazek et al., 1982; Pappas-Jones & Adamson, 1987; Socolar et al., 1999).

Because of their high degree of face validity, observational measures of parenting practices are often treated as an objective, gold standard measure of parenting behaviour (Dowdney et al., 1984; Mrazek et al., 1982). However, as with self-report measures of parenting behaviour, observations are also subject to several limitations.

These limitations include context-specificity, participant reactivity (including behaving in socially desirable ways), comparability across families, differences in operationalisation and investigator effects in coding.

Observations are context-specific since data collected are from one point in time and are applicable only to that point in time, in that context (Metsapelto et al., 2001). Parents' and children's behaviour is affected by the presence of other family members, the setting in which the interaction is occurring, and by the state of the participants (e.g., mood, illness, fatigue) (Dowdney et al., 1984; Gardner, 2000; Johnson, 2001; Meyers, 1999; Mrazek et al., 1982; Pappas-Jones & Adamson, 1987; Socolar et al., 1999). The behaviour of participants in one context is not necessarily representative of their behaviour in all contexts, and it is not known to what extent observations of parenting behaviour generalise across contexts (Gardner, 2000).

The use of unstructured, naturalistic observations in which no constraints are placed upon participants' behaviour is considered to provide a more representative picture of participants' "usual" patterns of behaviour (Mrazek et al., 1982). However, it is difficult to make comparisons between families using unstructured observations because the types of activities in which participants engage and the intrusion of extraneous variables is only poorly controlled (Dowdney et al., 1984; Mrazek et al., 1982). Conducting unstructured observations in the home when particular family routines are taking place (e.g. around dinner-time or bedtime) can provide some degree of comparability (Bates et al., 1998; Belsky et al., 1997; Belsky et al., 1998; Brophy & Dunn, 2002; McFadyen-Ketchum et al., 1996; Roberts & Strayer, 1987). Alternatively, standardising or structuring the activities in which participants engage improves the comparability of observations across families. Varying degrees of structure can be employed, from simply supplying toys and a time limit in a free-play situation (Campbell et al., 1986a; Keown & Woodward, 2002; Kleberg et al., 2000;

Russo & Owens Jr, 1982; Strayhorn & Weidman, 1988), to providing explicit requests for the parent to give to their child (Bor et al., 2002; Gomez & Sanson, 1994; Tarver-Behring et al., 1985).

The context in which behaviour is observed includes the presence of an observer which may influence participants' behaviour in ways that cannot be controlled or quantified (Brandt, 1992; Cox, 1975; Metsapelto et al., 2001). This is known as participant reactivity. Participants may alter their behaviour simply because they are being observed, or according to their understanding of the purpose or the focus of the study (Banister et al., 1996; Gardner, 2000). Participant reactivity may manifest itself in such diverse ways as changes in personal grooming or housekeeping practices, changes in voice tone or voice level, gestures and direct references to the camera or changes in the amount of focal behaviour (Kavanagh et al., 1988; Renne, Dowrick, & Wasek, 1983).

The different operationalisation of the same behaviours or constructs across studies can lead to different interpretations of observational data. For example, in some studies the concept of parental "responsiveness" has been operationalised to reflect the timing of parental responses to children's behaviour (Shaw et al., 1998). In other studies "responsiveness" refers to the affect accompanying the parental response, as well as the timing (Johnson, 2001). In the former case it is possible that parental responses which are hostile but occur immediately after children's behaviour could be considered "responsive" (Shaw et al., 1998), whereas this would not be the case in the latter group of studies.

Investigator effects can reduce the reliability of observations because different observers can have different interpretations of the same observed behaviours. This may be due to characteristics of the investigator (e.g., age, race, sex), or to changes within

them (e.g., tiredness, change in attention, inconsistency) (Castorr et al., 1990; Fassnacht, 1982; Metsapelto et al., 2001). Ratings are also influenced by factors such as the time at which the judgement is made (e.g., concurrently versus retrospectively) and the amount of time elapsed since the rater was last trained.

The following review examines the literature that reports direct comparisons of self-reports and observations of parenting behaviour. Because of the paucity of research in this area, studies that examine the direct relationships between observations and reports of other phenomena (e.g. children's temperament and behaviour) are also presented.

2.2.1 Parenting Behaviour

Several mixed-method studies have included self-report and observational measures of the parenting behaviour of parents of preschool-aged children (Arnold et al., 1993; Bor et al., 2002; Crnic & Greenberg, 1990; Cunningham & Boyle, 2002; Denham et al., 2000; Dowdney et al., 1984; Feinberg et al., 2001; Harvey-Arnold & O'Leary, 1997; Johnston et al., 2002; Kaplow et al., 2002; Keown & Woodward, 2002; Roberts & Strayer, 1987; Shaw et al., 2001; Stormshak et al., 1997; Strayhorn & Weidman, 1988). Many of these studies have used the different methods to assess *different domains* of parenting behaviour (Crnic & Greenberg, 1990; Kaplow et al., 2002; Keown & Woodward, 2002; Shaw et al., 2001; Stormshak et al., 1997) or when measuring the *same* dimensions of parenting behaviour have not directly compared self-reports with observations (Bor et al., 2002; Cunningham & Boyle, 2002; Harvey-Arnold & O'Leary, 1997; Roberts & Strayer, 1987).

Researchers have highlighted the need for studies that systematically investigate the associations between self-reports and observations of parenting constructs (Gardner, 2000; Holden & Edwards, 1989; Locke & Prinz, 2002; O'Connor, 2002; Patterson,

1982). Table 2.1 presents a list of studies that have directly compared self-reports and observations of parenting behaviour.

Of the few studies which have directly compared observations and self-reports of parenting behaviour in early childhood, the comparison of the two methods of measurement has generally not been systematic, nor has it been the focal purpose of the study (Arnold et al., 1993; Denham et al., 2000; Dowdney et al., 1984; Strayhorn & Weidman, 1988). Further, the generalisability of the results is often hampered by methodological considerations, such as the limited sample size (Arnold et al., 1993), the generality of the behavioural definitions (Dowdney et al., 1984; Strayhorn & Weidman, 1988), and the fact that all of the studies used only global ratings of observed parenting behaviour (Arnold et al., 1993; Denham et al., 2000; Dowdney et al., 1984; Strayhorn & Weidman, 1988). Ratings of behaviour in general, and global ratings in particular, represent a high degree of abstraction from observed behaviour and are considered more subjective than behaviour counts because they are more susceptible to observer biases than are behaviour frequencies (Aspland & Gardner, 2003; Castorr et al., 1990; Fassnacht, 1982; Maxwell & Pringle, 1983; Metsapelto et al., 2001). Also, in some cases the ratings used to code observed behaviour were not necessarily developed specifically to enable direct comparison with self-report scores (e.g., Denham et al. 2000).

Table 2.1 Studies directly comparing self-reports and observations of parenting behaviour.

Study	Age	N & Sample Type	Parenting Behaviour/s	Self-report Measure/s	Observation	Agreement b/w SR & Obs[a]
Dowdney, Mrazek, Quinton & Rutter, 1984, 1985	2-3.5yo	N=44 Community sample	General parenting quality	Summary ratings from interviews. Parenting quality = "poor" if low warmth and problems in 2 or 3 areas of control, "good" if no problems in any areas and "intermediate" if only some problems. Reliability not stated	Two 2hr observations in the home, not videotaped 1.Unstructured then novel toy presented after 1.5 hours 2.Unstructured then gift given near the end of observation Rating on same 3-point scale used for interviews Inter-observer reliability not stated	73% (32 out of 44) $\chi 2 = 30.99$, df=4, p<.001
Strayhorn & Weidman, 1988	3-4yo	N=125 Children attending Head Start programs	Global domain of parent behaviour	Parent Practices Scale (Strayhorn and Weidman, 1988). 34 items, Variable point Likert scales Total Score α=.79	Two observations, videotaped. Free-play (25min) Global rating (number of points on scale not stated) Inter-observer reliability, N=125(100%) r=.94	r=.33, p=.002
Arnold, O'Leary, Wolff & Acker, 1993	1.5-4yo	N=15 Hard to manage children N=7 Control N=8	Discipline: a) Laxness b) Overreactivity c) Verbosity d) General dysfunctional discipline	Parenting Scale (Arnold et al., 1993) 30 items, 7-point Likert Laxness α=.83 Overreactivity α=.82 Verbosity α=.63 Total Score α=.84 Test-retest reliability (2-week, N=22) Laxness r=.83 Overreactivity r=.82 Verbosity r=.79 Total Score r=.84	Single observation in the home, videotaped. Structured: a) Block sorting task (10min) b) Clean up novel toys (max 10min) c) Telephone call (10 min) Global ratings (7-point scale) Inter-observer reliability, N=15(100%) Laxness rho=.82 Overreactivity rho=.85 Verbosity rho=.88 General discipline rho=.88	Laxness rho=.61, p<.05 Overreactivity rho=.65, p<.01 Verbosity rho=.53, p<.05 General dysfunctional discipline rho=.73, p<.01
Denham, Workman, Cole, Weissbrod, Kendziora & Zahn-Waxler 2000	4-5yo	N=69 Community sample	Self-reported nurturance, Observed supportive presence.	Child Rearing Practices Report (Block 1981). Q-sort, 91 cards Nurturance (18 items), α=.77	Four observations in the lab, not videotaped. Structured: Mother, father and child: a) Tower building (10min) b) Dice game (20min) c) Snack/free time (10min) Mother and child a) Mouse puzzles (10min) b) Story in wordless picture book (10min) c) Emotion Reminiscence (15min) Global ratings on 7-point scale Inter-observer reliability, N=21 (30%) K=.92-.94	r=.43, p<.001

Table 2.1 (continued)

Study	Age	N & Sample Type	Parenting Behaviour/s	Self-report Measure/s	Observation	Agreement b/w SR & Obs[a]
Feinberg, Neiderhiser, Howe & Hetherington, 2001	9-18yo	N=720 Community sample	a) Warmth b) Negativity	Parental warmth: α=.90-.92 Parent Child Relationship Inventory (Hetherington & Clingempeel, 1992) .11 items, 5-point Likert Expression of Affection Inventory (Hetherington & Clingempeel, 1992). 22 items, 7-point Likert Parental negativity: α=.57-.91 Parent Discipline Behaviour Inventory (Hetherington & Clingempeel, 1992): 7-point Likert. Parent-Child Disagreement, Punitiveness and Yielding to Coercion subscales Conflict Tactics Scale (Straus, 1979): Symbolic Aggression Subscale Parent Child Relationship Inventory	Two home observations each for mother-child and father-child, videotaped. Asked to interact for 10 minutes around areas of conflict identified at earlier stage Ratings on 5-point scales composited across five codes for warmth and three codes for parental negativity. Inter-observer reliability, N=>20% Warmth = 73-86%, r=.62-.79 Negativity = 75-80%, r=.78-.81	Paternal: Warmth r=.38 Negativity r=.25 Maternal: Warmth r=.12 Negativity r=.22 (all significant at p=.001)
Johnston, Murray, Hinshaw, Pelham & Hoza, 2002	7-10yo	N=136 Boys with ADHD	Self-reported warmth, involvement and positive-parenting Observed responsiveness	Alabama Parenting Questionnaire (Shelton et al., 1996), 5-point Likert scale. Involvement and Positive Parenting subscales α>.80 Parent-Child Relationship Questionnaire (Burman & Giberson, 1995). 5-point Likert scale: Warmth subscale α>.85	Single observation in the lab, videotaped. Structured: a) free play (4min) b) parent busy (3min) c) paper and pen task (5min) d) clean-up (5min) Ratings every 1min on 7-point scale. Composite responsiveness measure derived from factor analysis. Inter-observer reliability, N= 48 (35%) K=.89-.92	Partial correlations controlling for mother age, education and marital status. r=.01 to -.08

[a]Level of agreement between self-reports and observations of parenting behaviour

For parents of preschool-aged children, the associations between self-reports and observations on the same domains of parenting behaviour range from medium (O'Connor, 2002; Strayhorn & Weidman, 1988) to high (Arnold et al., 1993; Dowdney et al., 1984). The study by Arnold et al. (1993) reported levels of agreement as high as rho=.73 for general dysfunctional discipline, but in a sample of only 15 parents. High levels of agreement were reported by Dowdney et al. (1984) for general parenting quality (73% agreement), but the authors used only three-point scales to rate parent-reports and observations on the same instrument. The comparison of reports on such narrow rating scales increases the likelihood of obtaining agreements between the methods. Studies using larger samples (in the order of 80 or more participants) have reported significant, but more modest, associations between self-reported and observed parenting behaviour (Denham et al., 2000; Strayhorn & Weidman, 1988). The levels of agreement between self-reported and observed behaviour do not appear to be related to factors such as the types of parenting behaviours observed (i.e., parental control and responsivity), the setting in which the observations were conducted (i.e., home or lab), or the number of observations conducted.

Larger studies which have compared methods used to measure the parenting behaviour of parents of older children have found lower degrees of agreement between methods (Feinberg et al., 2001; Johnston et al., 2002). Feinberg et al. (2001) reported associations between self-reports and observations of paternal and maternal warmth and negativity in the range of r=.12 to r=.38. Johnston et al. (2002) reported no associations between observed responsiveness and self-reported warmth, involvement and positive-parenting (r=.01 to -.08) after controlling for maternal age, education and marital status.

These lower correlations between self-reported and observed parenting behaviour for parents of older children may in part be due to the fact that observations which include older children can be less representative of "true" parent-child interactions than observations with younger children (Dowdney et al., 1984; House, 1988; Mrazek et al., 1982). This is because younger children typically display lower levels of reactivity to being observed than do older children. Younger children's behaviour is less likely to be constrained by the observation situation (i.e., they are more likely to behave as they naturally would), and in turn parental responses to children's behaviour are more "naturalistic" in interactions with younger children (Mrazek et al., 1982). Methodological differences between the studies with older and younger children may also explain the varying levels of associations found between these groups of studies. For example, the study of 9- to 18-year-old children used composites of scores from self-report measures, rather than a single scale (Feinberg et al., 2001). Compositing scores across scales can decrease the reliability of the measure (Banister et al., 1996), and this may explain why only small to modest correlations were obtained in this study. Further, the study by Johnson et al. (2002) found no associations between self-reported and observed behaviour after controlling for demographic variables. This is probably due to the fact that these variables (maternal age, education and marital status) are considered determinants of parenting behaviour (Fox et al., 1995; Jackson et al., 2000; Johnston et al., 2002; Rubin et al., 1995), and hence are likely to explain a substantial proportion of the variance in both self-reports and observations of behaviour.

2.2.2 Children's Temperament and Behaviour

When parents report upon their children's characteristics, they are doing so as proxy-reporters, and to some extent as informal observers of these phenomena in their

children. Proxy-reports of behaviour are vulnerable to similar limitations as self-reports of behaviour, particularly with respect to the ways in which parents interpret and respond to the items of a measure, and to the influences of parental characteristics (such as parental depression and anxiety) on perceptions of child characteristics (Bates & Bayles, 1984; Forehand et al., 1984; Vitaro et al., 1995). Studies that have directly examined the relationships between parent-reports of and observations of child temperament or child behaviour are shown in Table 2.2.

Comparisons between parent-reports and observations of children's characteristics have used more detailed and systematic methods than have been used in studies of parenting behaviour (Bates & Bayles, 1984; Bridges, Palmer, Morales, Hurtado, & Tsai, 1993; Campbell, Breaux, Ewing, & Szumowski, 1984; Campbell et al., 1986b; Campbell, Pierce, March, Ewing, & Szumowski, 1994; Campbell, Szumowski, Ewing, Gluck, & Breaux, 1982; Dishion, Duncan, Eddy, Fagot, & Fetrow, 1994; Ducharme et al., 1996; Dunn & Kendrick, 1980; Marcus, 1997; Seifer, Sameroff, Barrett, & Krafchuk, 1994; Stormshak et al., 1997).

For example, several studies have calculated the level of association between parents' and observers' ratings on identical measures, and at times, during the same observation sessions (Bates & Bayles, 1984; Ducharme et al., 1996; Dunn & Kendrick, 1980; Seifer et al., 1994). Conversely, some comparisons between parent-reports and observations of child behaviour have simply examined whether observations of child behaviour differentiate between parent-reported problem and non-problem children (Campbell et al., 1984; Campbell et al., 1986b; Campbell et al., 1982; Dunn & Kendrick, 1980).

Table 2.2 Studies directly comparing parent-reports and observations of children's temperament and behaviour.

Study	Age	N & Sample Type	Child characteristic/s	Parent-report Measure/s	Observation	Agreement b/w PR and Obs[a]
Dunn & Kendrick, 1980	1.5-3yo	N=40 First-born children	Temperament: intensity , negative mood, activity, malleability, approach-withdrawal, persistence and assertiveness Child Behaviour: tearfulness, clinging, demanding, negative behaviour to mother; demanding, naughty, quiet or aggressive during feed; helpful to mother, affectionate to baby, interested in baby	Temperament: Assessment of Temperamental Characteristics Interview Detailed descriptions of how child has behaved in specific situations 37 items rated on 3 point scale Inter-rater agreement for a sample of 25 (63%) of the interviews ranged from 65% to 100% for individual items Child Behaviour: Interviews focussing on the children's reactions after the birth of a sibling. Reliability not reported.	Two 1hr, unstructured observations in the home at two assessment periods, not videotaped. Temperament: Retrospective ratings by observers on selected items from the Assessment of Temperamental Characteristics Interview Inter-observer reliability not assessed Child Behaviour: Event sampled Inter-observer reliability (N=NS) median 88% (range 80-100%)	Temperament Intensity 82%, Mood 81% Activity 57%, Malleability 88%, Withdrawal 85% Persistence 83%, Assertiveness 90% Child Behaviour Except for demandingness, all behaviours reported by the mother to increase after the birth of the sibling, were observed significantly more often post-birth
Campbell et al., 1982, 1984, 1986	2-3yo, 4yo, & 6yo	1982 N=68 (46 problem, 22 control), 1984 N=54 (35 problem, 19 control) 1986 N=59 (32 problem, 27 control)	1982&1984: Activity, attention and impulsivity 1986: Inattentive and uncooperative behaviour	1982 & 1984: Parent-identified problems 1986: Telephone interview to determine if child met DSM-III criteria for ADD/hyperactivity. CBCL (Achenbach & Edelbrock, 1983) SNAP questionnaire (Pelham & Bender, 1982)	1982,1984: 2 videotaped lab observations: a) Child-directed free play (15min) b) Structured tasks Event sampled Inter-observer reliability, ave r=.89 and 86% agreement 1986: Single videotaped lab observation a) 10min free play child alone b) 10 min Lego task with mother c) 30 min structured tasks Event sampled Inter-observer reliability, ave r=.98 & 80-84% agreement	1982:All observed measures discriminated between parent-referred and control children. 1984 & 1986: Behaviour coded during structured tasks (but not free play) differentiated between problem and control groups.
Bates & Bayles, 1984	Data collected at 6,13,24 & 36mths	N ranged from 160 at 6mths to 120 at 3yrs	Difficult/demanding, negative adaptation, noncompliant/ irregular, unexcitable, problem behaviour, language competence, psychomotor incompetence	Factors composited from: Infant Characteristics Questionnaire (Bates et al., 1979), Maternal Perceptions Questionnaire (Olson et al., 1982), Preschool Behavior Questionnaire (Behar, 1977), Minnesota Child Development Inventory (Ireton & Thwing, 1974)	3 hour home observations (2 at 6 & 24 mths, 1 at 13 months) Event sampled and at 6 & 13 mths ratings on the items of the ICQ, at 24mths on CQ and a Post Observation Q'aire (Olson et al. 1982)	Difficult r=.26, negative adaptation r=.24, noncompliant r=.23, unexcitable r=.19 problem behaviour r=.20, language competence r=.47, psychomotor r=.20

Table 2.2 (continued)

Study	Age	N & Sample Type	Child characteristic/s	Parent-report Measure/s	Observation	Agreement b/w PR and Obs[a]
Bridges, Palmer, Morales, Hurtado & Tsai, 1993	6mths	N=71 Sample type not reported	From parent-reports: a) Distress to limitations b) Smiling/laughter From observations: a) Anger b) Struggling c) Facial pleasure d) Vocal pleasure	Infant Behaviour Questionnaire (Rothbart, 1981) 96 items on 7pt scale Relative frequency of child responses during specific activities in the previous week Reliability not reported	1 lab assessment, videotaped. Structured: a) Restraint in seat b) Arm restraint c) Toy retraction d) Puppet game e) Reaction to sound/light display f) Peek-a-boo game Ratings and duration recorded Inter-rater reliability N=71(100%), 88-99%	Distress and anger r=.46 Distress & struggle not significantly correlated (r not reported) Smiling/laughter & facial pleasure r=.30, p<.02 Smiling/laughter & vocal pleasure not significantly correlated (r not reported)
Seifer, Sameroff, Barrett & Krafchuk, 1994	Assess between 16-24 weeks	N=50 First born children	a) Mood b) Activity c) Approach d) Intensity	Mothers rated behaviour for observations on the Temperament Adjective Triad Assessment (TATA; Seifer et al., 1994), Infant Temperament Questionnaire-Revised (Carey & McDevitt, 1978), Infant Characteristics Questionnaire (Bates et al., 1979), Infant Behaviour Questionnaire (Rothbart, 1981) EAS Temperament Survey (Buss & Plomin, 1984)	Weekly home observations for 8 weeks, videotaped. Record at least 10mins of: a) infant with mother (no caretaking) b) infant alone c) mother caretaking with infant Ratings on TATA for each situation Inter-observer reliability, N=50 (100%) r=.80	TATA combined for 8 observations: Mood r=.43-.48, p<.05 Activity r=.25(ns)-.37, p<.05 Approach.23(ns)-.38, p<.05 Intens r=.17 (ns)-.36, p<.05 Total r=.29-.42, p<.05 Other questionnaires ranged from r=.07(ns) (for observed and ITQ intensity) to r=.30 (observed & ITQ approach)
Dishion, Duncan, Eddy, Fagot & Fetrow, 1994	First and fifth grade	N=374 School sample	Parent-report: Aggressive behaviour Observed: Child coercion with parents	Child Behavior Checklist (Achenbach & Edelbrock, 1986) Aggressive Behavior Scale	Observations in the lab Parent-child interaction tasks (not specified) Interval sampled and rated Inter-observer reliability not reported	Interval sampled: r=.24, p<.01 for boys & r=.21, p<.01 for girls Ratings: r=.17, p<.05 for boys and for girls
Ducharme, Popynick, Pontes & Steele, 1996	4-5yo	N=5 Developtal disabilities & oppositional behaviour	Compliance	Parents trained to code compliance during interactions as for the observations (calculate probability of compliance)	Multiple videotaped home observations, parent to request compliance on set tasks Event recording Inter-observer reliability (92% of assessments), ave 96-100%	Average agreement ranged from 83-98%
Stormshak, Speltz, DeKlyen & Greenberg, 1997	4-5yo	N=78 (all boys) Clinic N=44 Control N=34	Parent-reports: Aggressive behaviour Observed: Negative elicitation of father and mother	Father- and mother-reports on the Child Behavior Checklist (Achenbach, 1991) Aggression Scale	10 mins of clinical interview, videotaped Family Intake Coding System Ratings on 4 point scale Inter-observer reliability (22%), ave r=.87 to 1.00	Paternal-reports r=.21 & .25, p<.05 for negative elicitation of mother and father. Maternal-reports r=.19 & .06

[a]Level of agreement between parent-reports and observations of parenting behaviour; ns=not significant

Research which has demonstrated strong associations between parent-reports and observations of children's characteristics have utilised methods which compare data using the same instruments (Ducharme et al., 1996; Dunn & Kendrick, 1980; Seifer et al., 1994). This suggests that parents and observers may be reporting on the same phenomena in these studies, however the associations may be influenced by a number of factors.

The very high levels of agreement between parent-reports and observed temperament reported by Dunn and Kendrick (1980) (with the exception of the Activity dimension all agreements were greater than 80%) may also be the result of the use by both parents and observers three-point scales to rate child temperament. The narrower the range of possible ratings which can be assigned to a behaviour, the greater the likelihood of obtaining agreements between observers. The lower level of agreement for the Activity dimension (57%) was a result of mothers consistently rating their children as more active than did the independent observer (Dunn & Kendrick, 1980). This discrepancy may reflect a tendency for children to be less active in the presence of an observer than they would be at other times, resulting in lower observer ratings.

Seifer et al. (1994) reported small to medium correlations between parent-reports and observations of data obtained on the same measure aggregated over eight assessments. The levels of agreement were not as high as the authors expected given the study design, suggesting parent-reports of child temperament contain a subjective component which yields different reports to that of an independent observer even when coded by parents and observers during the same observation (Seifer et al., 1994). Conversely, high levels of agreement were reported by Ducharme et al. (1996) (83% to 98%) when parents and observers coded child

noncompliance in multiple observations. These levels of agreement may be a result of the very small sample size (N=5), the intensive training parents underwent before coding their child's level of compliance and the fact that parents and observers recorded only the presence or absence of a single behaviour during the interaction, rather than rating multiple child characteristics (Ducharme et al., 1996).

Lower associations between parent-reports and observations of children's characteristics were reported when observations were conducted in a laboratory setting or when comparisons between broad and narrow groups of behaviours were made (Bates & Bayles, 1984; Bridges et al., 1993; Dishion et al., 1994; Seifer et al., 1994; Stormshak et al., 1997). Observations conducted in the lab or clinic may yield less representative pictures of behaviour because they are less naturalistic and may increase children's reactivity to being observed (Banister et al., 1996; Dowdney et al., 1984). This may influence the level of association between what is observed and what parents report as "usual" behaviour. Furthermore, studies which have included parent-reports of broader child characteristics (eg Aggressive Behaviour as measured by the Child Behaviour) and then compared these reports with observations of more narrowly defined behaviours (e.g., child coercion) do not yield as high correlations as studies which compare identical aspects of behaviour (Bates & Bayles, 1984; Bridges et al., 1993; Dishion et al., 1994; Seifer et al., 1994; Stormshak et al., 1997). Similarly, research which calculates the degree of association between parent-reports of child characteristics in general (e.g., laughing and smiling) have found no relationship with behaviours that occur infrequently in the context of a structured observation (Bates & Bayles, 1984; Bridges et al., 1993).

2.2.3 Research Aim One

The importance of parenting behaviour as a risk factor for the development and persistence of externalising behaviour problems, and as a focus for early interventions has led to a desire for the accurate measurement of parenting attitudes and practices. Several methods of measuring parenting behaviour have been employed in theoretical research and program evaluation, including self-report and observational measures. However, the degree of concordance between the methods, and how well either approach measures the constructs of interest are still largely unknown (Gardner, 2000; Holden & Edwards, 1989; O'Connor, 2002).

The research has highlighted a high degree of variability in the results of comparisons between parent-reports and observations of behaviour, and in particular the absence of systematic direct comparisons between methods for measuring parenting behaviour. It is still unknown to what extent self-reports and direct observations of the same parenting behaviours are related (Gardner, 2000). Therefore, the first aim of this thesis is to assess the degree of agreement between self-report questionnaire and direct observation measures of parental responsivity and control. Scores on two frequently-used self-report questionnaires will be directly compared with frequency counts of videotaped parenting behaviours along the domains of parental control and responsivity. To maximise the comparability of the two methods, structured observations and an observational coding system directly developed from the items of two self-report instruments will be used.

2.3 Relationships between Self-reports and Observations of Parenting Behaviour and Family Characteristics

2.3.1 Parenting Behaviour and Family Characteristics

This section will examine relationships between self-reports and observed parenting behaviour indirectly by reporting upon research which has investigated the relationships between parenting behaviour assessed using these methods and measures of family characteristics and parenting determinants highlighted in transactional models of the development of children's behaviour problems (see Sections 1.2 and 1.3) (Belsky, 1984; Patterson et al., 1989). The factors examined include parent gender, child gender, socioeconomic status, marital interaction, parental psychopathology, social support and child temperament. Generally, the literature reported in this section has not provided a direct comparison of the relationships between self-reported and observed parenting and family factors. Because of this, the results from different studies that have examined the relationships between parenting and family variables are compared. This review is limited to research for parents of two- to six-year-old children because of the abundance of research in this area and because of the focus on parenting during this developmental period in this thesis.

Parent Gender

Comparisons of mothers' and fathers' parenting behaviour on various parenting constructs have yielded mixed results. For example, research examining observations of aspects of parental responsivity (warmth, hostility and detached parenting) of mothers and fathers found no differences between parents from the same families (N=131) during storybook interactions with their children (Frosch, Cox,

& Goldman, 2001). Self-reports and observations of the parenting of mothers and fathers within the same families (N=19 couples) have revealed higher levels of overreactive but not lax parenting in mothers compared with fathers (Harvey-Arnold & O'Leary, 1997). However, a larger study which examined the parenting behaviours of primary caregivers in different families found no differences between mothers (N=677) and fathers (=47) on self-reports of parental laxness, overreactivity and total parenting on the Parenting Scale (Collett et al., 2001). The conflicting findings in the two studies that used the Parenting Scale may be due to differences in sample types. The study by Harvey Arnold and O'Leary (1997) included only mothers and fathers of "hard to manage children" whereas Collett et al. (2001) used primary caregivers from a normative sample. In the former study, when levels of depression and the amount of time spent in childrearing was controlled, the differences between mothers and fathers disappeared. This suggests that the higher levels of overreactive parenting of mothers in that sample may be a function of the amount of time spent with their hard to manage children (Harvey-Arnold & O'Leary, 1997).

Child Gender

Studies which have explicitly examined the relationships between the behaviour of parents of boys and girls in early childhood have used self -reports on the Parenting Scale (Collett et al., 2001; O'Leary et al., 1999). No significant differences were found in the levels of overreactive, lax or total ineffective discipline reported by parents of boys and parents of girls in a normative sample (N=768) (Collett et al., 2001). Similarly, no differences in self-reported overreactive discipline were reported by parents of 117 boys and girls at two assessments (one at 18-36 months old and the second 12-42 months later) (O'Leary et al., 1999).

Socioeconomic Status and Parental Education

Both observations and self-reports of general parenting constructs (e.g authoritativeness) have demonstrated significant moderate associations (r=.33-.36) with socioeconomic status (as represented by a composite of income, occupation and education) (Meyers, 1999; Strayhorn & Weidman, 1988). When the elements of socioeconomic status are considered individually, more equivocal results are obtained. For example, two studies using self-reports of overreactivity, laxness and total ineffective discipline report conflicting relationships with parents' education levels (Collett et al., 2001; Reitman et al., 2001). Reitman and colleagues (2001) report no linear relationship between parental years of education and self-reported parenting behaviours (N=193), whereas Collett et al. (2001) found that parents with higher levels of education reported the use of lower levels of ineffective parenting techniques than parents with lower levels of education (N=613). These differences in results may be attributable to the types of samples used in the two studies. Reitman et al. used a sample of parents of children attending a Head Start program, in which the level of parental education was substantially lower than that of the community sample of Collett et al. (only 56% had completed high school in the former sample whereas 67% had at least some college education in the latter study).

Observations of supportive parenting behaviour have shown modest, but significant relationships with education levels, but not income in single-parent families (Jackson et al., 2000). Education has also been found to have a moderating effect on the relationship between income and maternal discipline, with parents with lower incomes but who have higher levels of education displaying lower levels of negative discipline than parents of similar income levels but with lower levels of education (Fox et al., 1995).

Marital Satisfaction

Both self-report and observations of general parenting constructs have been significantly associated with self-reports of marital interaction (Arnold et al., 1993; Gottman & Katz, 1989; Meyers, 1999). Observations of authoritativeness and general parenting behaviour have demonstrated low but significant correlations with marital satisfaction (r=.23-.26) in community samples (Gottman & Katz, 1989; Meyers, 1999). Self-reports of general ineffective discipline and specifically lax, overreactive and verbose parenting behaviours have also shown high negative associations with self-reports of marital satisfaction (r=-.35 to -.53) in a mixed sample of "hard to manage" and control children (Arnold et al., 1993).

Parental Psychopathology

Generally, modest associations between parenting behaviour as assessed by observations and self-reports, and self-reports of depressive symptoms have been found in a number of studies. This has been true for observations and self-reports of responsive parenting (Jackson et al., 2000; Strayhorn & Weidman, 1988), self-reports of overreactive discipline (Arnold et al., 1993; Reitman et al., 2001) and parenting hassles (Crnic & Greenberg, 1990), and observations of commands (Forehand et al., 1986). Surprisingly, self-reports of lax parenting behaviour and observed parental authoritativeness have not been significantly associated with parent-reports of depression (Arnold et al., 1993; Meyers, 1999; Reitman et al., 2001). This may be because laxness is not as influenced by depressive symptoms as behaviours which include more affective components (e.g., overreactivity includes yelling, anger and being picky and on the child's back; responsiveness includes parental warmth and affection) (Arnold et al., 1993; Jackson et al., 2000; Reitman et al., 2001; Strayhorn & Weidman, 1988). Similarly, the maternal "authoritativeness"

construct used in the study by Meyers (1999) was composited from observations of parental control and responsivity, and this may have obscured any significant relationships between specific aspects of authoritative parenting (eg affect) and depression.

Social Support

Perceived social support from family and friends has been found to be significantly associated with both observed and self-reported general parenting constructs reflecting aspects of responsiveness and control (r=.30-.41) (Meyers, 1999; Strayhorn & Weidman, 1988). Conversely, a study investigating instrumental support and observed supportive parenting in single mothers found no relationship (r=-.01). These conflicting results suggest that the type of social support (e.g. emotional versus instrumental support) may be an important determinant of parenting behaviour.

Child Temperament

Observations of maternal derisiveness and maternal intrusiveness have shown only small correlations (r=.02 to .23) with children's inhibited temperament measured concurrently and two years later in a sample of 108 mothers (Rubin, Burgess, & Hastings, 2002). Similarly, Meyers (1999) reported an association of r=.23 between observed maternal authoritativeness and children's "easy" temperament in a sample of 73 mothers and their 5 to 7 year old children. Belsky and colleagues (1998) found a significant relationship (r=.30) between observed "fathering" when children were 2 and 3 years old and child inhibition at 3 years (a composite of observations and parent-reports), but not between "mothering" and child temperament. The prediction of child inhibition at 3 years was much stronger for fathers of children who were

higher in negativity at 2 years than for children low on negativity at that assessment (Belsky et al., 1998).

Relationships between parent-reported temperament characteristics (eg inflexibility, persistence, self-regulation, negativity) and self-reports of parenting behaviour (e.g., punishment, child centeredness) have generally been modest and some studies report non-significant relationships between these variables (Putnam, Sanson, & Rothbart, 2002). It has been suggested that relationships between parenting and temperament may be obscured by the use of broad definitions for temperament and/or parenting variables and the use of larger, heterogeneous samples in which third intervening variables (eg culture, SES, child age, parent gender, child gender, parent personality traits) may moderate any direct associations (Putnam et al., 2002; Rubin et al., 2002). For example, parents of infants who display temperamental difficulties are thought to initially invest more time and effort in parenting, but as children age these parental efforts will decrease (Putnam et al., 2002). Also, it has been suggested that it is parental reports of the child's temperament as "difficult" compared to other children (regardless of their actual scores on measures of temperament dimensions) that are more strongly associated with ineffective parenting practices such as punitiveness and lower responsiveness (Prior, Sanson, Smart, & Oberklaid, 2000; Smart & Sanson, 2001).

2.3.2 Research Aim Two

Given the importance of parenting as a risk factor for the development and persistence of behaviour problems, and the potential of parenting as a target for intervention efforts, there is a need to identify the most valid and reliable measures of parenting behaviour for early childhood (Hemphill & Sanson, 2001; Locke & Prinz, 2002; Reitman et al., 2001). The concurrent validity of the measures of parenting

behaviour will be assessed. Concurrent validity examines how well a test instrument relates to criterion variables (external variables which are expected to show an association with the instrument) measured concurrently. This type of validity can be determined by seeing how well the measure discriminates between groups with known differences relevant to the construct of interest (e.g., parents of boys versus parents of girls), or by correlating the scores on the instrument with measures of related phenomena (e.g., how well do scores on parenting measures correlate with measures of parental psychopathology) (Holden, 1983; Reitman et al., 2001). In this study, concurrent validity was established by examining the relationships between measures of parenting behaviour and parent-reports of determinants of parenting practices as specified by the transactional theories discussed in Chapter 1 (such as child temperament, parental psychopathology, social support and marital satisfaction).

Previous research has not directly compared the associations between self-reported parenting behaviour and family factors with associations between observed parenting behaviour and family characteristics. Also, the results of studies which have studied these relationships separately has often led to mixed results, and differences in the methodologies of these studies (eg sample type, definitions of parenting and family constructs, instruments used) has made interpretation of the results difficult. Therefore, the second aim of this study is to directly compare the relationships between the different measures of parenting behaviour and family characteristics (e.g., parent gender, child gender, socioeconomic status, marital interaction, parental psychological functioning, social support, and child temperament). Parent-reports of these family characteristics collected at the same time as self-reports and observations of parenting behaviour will be used to examine this research aim.

2.4 Relationships between Self-reports and Observations of Parenting Behaviour and Children's Externalising Behaviour

The literature reviewed in Chapter One examined theories of the transactions between parents' and children's behaviour that lead to the development of externalising behaviour problems (Belsky, 1984; Patterson et al., 1989). This section compares research that has utilised self-reports and observations of parenting behaviour to examine relationships with children's externalising behaviour problems.

2.4.1 Parenting Behaviour and Children's Externalising Behaviour

Global Parenting Measures

Self-reports of global parenting techniques (e.g., discipline, overall child-rearing) significantly correlate with parent-reports of externalising behaviour (e.g., total externalising problems, aggression towards peers, hyperactivity and other ADHD symptoms) (Arnold et al., 1993; Brenner & Fox, 1998; Collett et al., 2001; Strayhorn & Weidman, 1988). The size of these correlations ranges from r=.24 to r=.53 with parenting explaining approximately 6% of the variance in children's hyperactive behaviour (Collett et al., 2001; Strayhorn & Weidman, 1988), and up to 28% of general externalising behaviour problems (Arnold et al., 1993; Brenner & Fox, 1998). Because these studies have used parent-reports to assess both parenting and child behaviour, part of these associations may be attributable to common-method variance (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003).

However results from studies which utilise observed and self-reported data suggest that there exist relationships between global parenting constructs and children's externalising behaviour that are not an artefact of common method variance. For example, a meta-analysis by Rothbaum and Weisz (1994) reported significantly stronger associations between parental caregiving and children's externalising

behaviour (noncompliance, aggression and hostility) in non-clinic samples when observations or interviews were used to measure parenting behaviour (r=.28) compared with questionnaires (r=.11) (Rothbaum & Weisz, 1994).

Significant relationships (r=.25 to .67) have also been obtained between self-reported global parenting and observed children's externalising problems (e.g., general child misbehaviour and hyperactivity) (Arnold et al., 1993; Strayhorn & Weidman, 1988). Similarly, observations of supportive parenting on the HOME correlate significantly (r=-.37, p<.01) with parent-reports of children's behaviour problems in a sample of mother-headed families (Jackson et al., 2000). And although observations of general "dysfunctional discipline" have not shown significant relationships with parent-reported externalising behaviour in smaller samples (N=15), correlations in the order of r=.40 have been reported (Arnold et al., 1993).

Parental Control

Both self-reports and observations of negative parental control (including punitiveness, overreactivity, physical discipline, parental anger and aggression) have been found to relate to parent-reports and observations of children's externalising behaviour problems (r=.29-.69) (Arnold et al., 1993; Belsky et al., 1998; O'Leary et al., 1999), aggression (r=.20-.50) (Brook et al., 2001; Hops et al., 2003; Stormshak et al., 1997), and hyperactivity/impulsivity (r=.30-.32) (Collett et al., 2001; Keown & Woodward, 2002). A recent meta-analysis reported a mean effect size of *d*=.44 for the association between the use of corporal punishment and preschool children's aggressive, delinquent and antisocial behaviour (Gershoff, 2002). In general, negative parental control explains between 4% and 22% of the variance in children's externalising behaviour. These associations hold even when demographic and other parenting variables are controlled (Brannigan et al., 2002; Brenner & Fox, 1998; Dwyer et al., 2003; Paterson & Sanson, 1999; Stormshak et al., 2000). The studies

which have reported stronger associations have used the smallest samples (eg N=15 in the study by Arnold et al., 1993).

Self-reports of overreactivity and other negative parental control techniques obtained during the preschool years significantly predict parent-reports of children's externalising problems obtained 12-42 months later (r=.30-.38) (Belsky et al., 1998; O'Leary et al., 1999). Observed negative parental control has also been associated with the onset and stability of behaviour problems in community samples (Kingston & Prior, 1995), but not in clinic samples (Campbell et al., 1986a; Campbell & Ewing, 1990; Campbell et al., 1986b). This may be because levels of child behaviour and negative parenting are more variable in community samples, but are higher in clinic samples and hence parenting does not discriminate between children with and without persistent problems.

Self-reported laxness and inconsistency have been significantly associated with maternal-reports and observations of children's general externalising behaviour problems (r=.41-.62) (Arnold et al., 1993), hyperactivity/impulsivity (r=.12-.19) (Collett et al., 2001; Keown & Woodward, 2002; Stormshak et al., 2000), aggressive behaviour (r=.18-.19) (Brook et al., 2001; Stormshak et al., 2000) and oppositional behaviour (r=.17) (Gardner, 1989; Stormshak et al., 2000). Although the associations between self-reported laxness and the narrow-band externalising problems are significant, they are very modest with parenting behaviour accounting for only 1-4% of the variance in children's hyperactivity, aggression or oppositional behaviour. Longitudinal studies have not reported the prediction of children's behaviour from self-reported laxness or inconsistency.

Observations of parental laxness and inconsistency have not shown such consistent results. For example, observations of lax parenting behaviours were not significantly

associated with parent-reported externalising behaviour in a small sample (N=15) (Arnold et al., 1993), but did differentiate between children at risk of oppositional defiant disorder and control children (Cunningham & Boyle, 2002). Similarly, observations of the use of maternal beta-commands (vague or interrupted commands) were found to be associated with observations of child compliance recorded during the same observation period ($r=-.39$, $p<.05$), but not with parent-reports of child misbehaviour in general ($r=-.09$) (Forehand et al., 1986). It should be noted that the observational system used in this study was frequency dependent: both the number of maternal commands and the frequency of child noncompliance were dependent on the behaviour of the other member of the interaction (ie more beta commands given by parents provide more opportunities for noncompliance and higher levels of child noncompliance prompt more maternal commands). For this reason, frequencies of children's behaviour taken from the same observation as those of parenting behaviour may be more highly associated with parenting behaviour than observations of child behaviour taken on another occasion.

Parental Responsivity

Self-reports and observations of parental warmth, nurturance, affection, rejection and hostility have been significantly associated with parent-reports and observations of children's general externalising behaviour ($r=.13$) (Brenner & Fox, 1998), attention problems ($r=.68$) (Frosch et al., 2001), hyperactivity (Keown & Woodward, 2002), aggression ($r=.14$-$.15$) (Brannigan et al., 2002; Brook et al., 2001; Stormshak et al., 2000), and oppositional behaviour ($r=.17$-$.71$) (Frosch et al., 2001; Shaw et al., 1998; Stormshak et al., 2000). Generally, aspects of parental responsivity explain only 2-3% of the variance in children's externalising behaviour problems (Brenner & Fox, 1998; Frosch et al., 2001; Keown & Woodward, 2002; Shaw et al., 1998; Stormshak et al., 2000). Studies which report stronger relationships ($r=.33$-$.68$) have used

observations of parents' and children's behaviour taken during the same observation periods (Frosch et al., 2001; Shaw et al., 1998). As discussed before, the frequency-dependence of observational measures may mean that correlations between observed behaviours taken from the same observation period are higher than might have been obtained if frequencies of behaviour were obtained from separate observations.

Observational and self-report measures of parental responsivity obtained when children are preschoolers have been found to predict parent-reports of school-age children's externalising behaviour (Denham et al., 2000; Kingston & Prior, 1995; Shaw et al., 2001; Shaw et al., 1998)

2.4.2 Research Aim Three

The literature has demonstrated that both observations and self-reports of parental control and responsivity are significantly related to children's externalising behaviours. However there is limited research about the longer-term prediction of children's behaviour from observations and self-reports of parenting, particularly laxness and inconsistency. Research investigating the relationships between parenting behaviour and the onset and/or persistence of externalising behaviour problems has produced conflicting results and has been limited to negative control and responsivity.

Transactional models of parenting and child behaviour suggest that parenting at one time point will be predictive of children's behaviour problems at a later time. In this study, the predictive validity of observations and self-reports of parenting behaviour will be examined. Predictive validity is the extent to which instrument scores at one time point predict scores on the same or a different measure obtained at a later time point (Holden & Edwards, 1989). In this study, this will be examined by reporting the

associations between the measures of parenting behaviour and parent- and teacher-reports of children's externalising behaviour obtained two years later. Hence, the third aim of this thesis is to examine the relationships between the different measures of parenting behaviour and parent- and teacher-reports of children's externalising behaviour obtained two years later.

Chapter 3. Methodology

This study was conducted in two stages. In the first stage (screening), parents recruited from randomly selected preschools completed a brief questionnaire assessing their children's levels of externalising behaviour problems and parental discipline styles. In the second stage of the study selected participants completed more questionnaires assessing aspects of parenting behaviour, children's temperament, parental agreement about child-rearing, marital satisfaction, social support, parental psychopathology and demographic characteristics. Eighty-one families in the second stage were selected to take part in videotaped observations of parent-child interaction. The 68 families who completed these observations are the focus of this thesis. In a third stage of the study, parent- and teacher reports of children's behaviour were obtained two years after observations were collected.

3.1 Participants

The study design and number of participants at each stage of the study is summarised in Figure 3.1. For the initial screening, families were recruited through 35 preschools randomly selected from a total of 228 preschools in the Adelaide metropolitan area in South Australia, between July and September 2000. At the time of recruitment, 1239 parents of preschool children (70% of parents approached) completed a brief screening questionnaire containing the Externalising Scale of the provisional version of the Child Behaviour Checklist (1.5-5) (Achenbach, 1998), the Parenting Scale (Arnold et al., 1993) and a brief demographic questionnaire. Eight hundred and thirty-one parents indicated their willingness to participate in the second stage of the study (67% of those who completed the questionnaire).

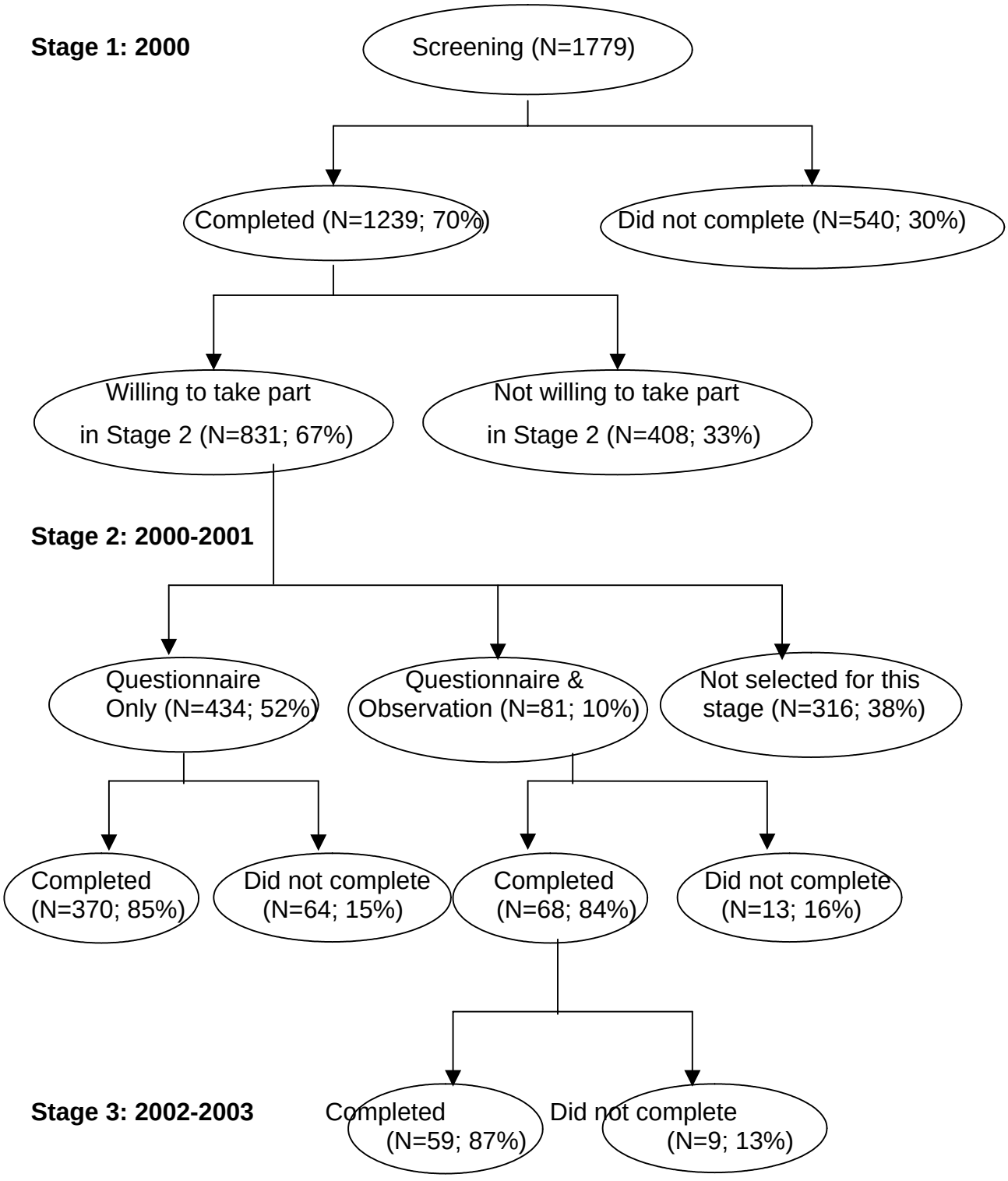

Figure 3.1. Study design, number of participants and response rate at each stage of the study (as a percentage of the previous stage)

Based on scores on the Externalising Scale and the total score of the Parenting Scale, families willing to participate in the second stage of the study were grouped into one of four study groups to be followed in the context of a larger longitudinal study examining the aetiology and persistence of children's externalising behaviour problems. Children were considered to have externalising behaviour problems if their score on the Externalising Scale of the provisional Child Behaviour Checklist (1.5-5) was above the 70th percentile of scores in the screening phase sample. Parents were categorised as utilising dysfunctional parenting techniques if their total score on the Parenting Scale was above the 65th percentile of the scores in the screening sample.

The four groups were thus:

1) children with externalising behaviour problems and parents defined to be using dysfunctional discipline techniques (N=149)
2) children with externalising behaviour problems and parents NOT defined to be using dysfunctional discipline techniques (N=121)
3) children with **no** externalising behaviour problems and parents defined to be using dysfunctional discipline techniques (N=124)
4) children with **no** externalising behaviour problems and parents NOT defined to be using dysfunctional discipline techniques (N=415)

The total number of participants in these four groups was reduced by 22 (N = 809) when siblings of other participants were excluded from the next stage of the study to ensure that each family was included in the study only once. Because of the large number of families in Group 4, if only one sibling in a single family was in Groups 1 to 3, this sibling was selected for the next stage of the study and the brother or sister in Group 4 was excluded. If both siblings fell in any of Groups 1 to 3, or both siblings were in Group 4, only one of the siblings was chosen (at random) to be included in

the next stage. For the purposes of the larger longitudinal study, similar numbers of participants were needed in the four groups. Thus, all of the participants in Groups 1, 2 and 3 were selected to participate in the second stage of the study, and 120 participants from Group 4 were randomly selected using the random sample select facility of the Statistical Package for the Social Sciences (SPSS) software (SPSS Inc, 2000).

In the second stage of the study, 20 families from each of these four groups were randomly selected to take part in additional videotaped observations of parent-child interaction. For the observation component, equal numbers of families were selected from the four groups to ensure that a range of child and parent behaviours could be observed. In this thesis, these participants are treated as a single group for the comparison of self-reports and observations of parenting behaviour. To avoid potential biases and subjectivity in observational coding, the author was kept blind to participants' group membership at all times (Maxwell & Pringle, 1983).

Of the 81 families (including one replacement family) selected to participate in the observation study, 68 families agreed to participate (response rate = 84%). Reasons given by participants (n=13) who refused to participate in the observation component of the study are summarised in Table 3.1. The majority of parents who declined to participate (69%) described themselves as either too busy or unwell.

Table 3.1. Reasons for refusal to participate in the observations

Reason for refusal	**n (%)**
Too busy	6 (46%)
Ill health	3 (23%)
Resent intrusion/questions too personal	2 (15%)
Partner did not want them to do it	1 (8%)
Pregnancy	1 (8%)

To examine the predictive validity of the parenting measures, parent- and teacher-reports were obtained from parents in Stage Three conducted two years after the observations were performed (N=59 for parent-reports and N=54 for teacher-reports). To examine the relationships between parenting and children's behaviour, it was necessary to use reports of children's behaviour obtained at Stage Three, rather than assess concurrent relationships at Stage Two because of the method of sample selection for the observations. That is, selection for the observation component was based upon parents' responses to measures of parenting and externalising behaviour at the screening assessment (ie equal numbers of participants were selected from the four groups described above), and this would have affected any cross-sectional relationships between these variables examined at Stage Two. This aspect of the study is discussed again in Chapter 8.

3.2 Procedure

3.2.1 Stage One: Screening

Preschool directors were first sent a letter explaining the aims of the study and inviting their participation (Appendix B.1). They were then contacted by telephone, and if they were willing to participate, arrangements were made for a research assistant to visit them in person to discuss the study in more detail. Of the 35 preschools approached, only one preschool director declined to participate (because of other research commitments), and this preschool was replaced with another from the same area with a similar number of enrolled children. Ethics approval for this and all subsequent stages of the study was obtained from the Women's and Children's Hospital Human Research Ethics Committee.

A notice about the study aims and procedure was sent to families either in the preschool newsletter or as a flyer sent home with children (Appendix B.2). Following this newsletter/flyer, parents of all children in the preschools were sent a letter about the study, an information sheet and the brief screening questionnaire (Appendices B.3-B.5). As described in the previous section, this questionnaire contained the 26 items of the Externalising Scale of the Provisional Child Behaviour Checklist (1.5-5) (Achenbach, 1998), the 30-item Parenting Scale (Arnold et al., 1993), and some demographic items. The final page of the questionnaire booklet asked for the personal contact details of those families willing to participate in the second, more detailed stage of the study (described below).

Children's primary caregivers were asked to complete the questionnaire and return it to their children's preschool in the envelope provided or alternatively, parents were asked to sign the back of the return envelope and return the blank questionnaire if they did not wish to participate. In this way it was possible to distinguish between

parents who did not wish to participate, and those who may have misplaced the questionnaire or forgotten to complete it. The latter group of parents was then reminded by a letter and/or notice in the preschool newsletters to return the questionnaires.

3.2.2 Stage Two: Detailed questionnaire completion

Parents were sent a letter informing them about their selection for the observation stage of the study (Appendix C.1). This letter briefly described the questionnaire administration and observation procedures and advised participants that the author would contact them to make an appointment to visit their home. The letter was accompanied by an information sheet that explained the purposes and procedures of the study in more detail (Appendix C.2).

The author telephoned the primary caregiver of each child in the study. Telephone contact with parents followed a set protocol (Appendix C.3). Parents were initially asked if they would like to participate in this phase of the study, and then a description of the questionnaire and observation ensued. Parents were instructed that the observation was designed to examine “what families do in real life” and thus to obtain a more naturalistic representation of parent-child interaction, the observations were scheduled to take place in participants’ homes (Banister et al., 1996).

To control for order effects, the order of administering the questionnaire and observation session was determined by random allocation. That is, in half of the families, the questionnaire was administered first, in the other half, the observation was conducted before the questionnaire. Parents were told of the order in which the procedures would be presented at the time of telephone contact. The author made

an appropriate time to visit families, and participants were instructed that the entire visit would last from one and a half to two hours. Two contact telephone numbers to call in the event that home visits needed to be rescheduled, were also provided.

Upon arrival at each participant's home, the author first ensured that the participant had read and understood the Information Sheet describing the study and then asked the parent to sign the consent form. Separate consent to approach the children's teachers for completion of the Teacher-Report Form of the Child Behaviour Checklist for children at preschool was also obtained at this time. (Appendices C.4 and C.5). A "helpful organisations" sheet containing the details of two parenting organisations in South Australia (Appendix C.6), was also given to parents at this time.

For those parents who were scheduled to complete the questionnaire first, the author explained the various sections of the questionnaire and, if necessary, answered questions about questionnaire completion. While parents completed the questionnaire, the author spent time playing with the children (e.g., drawing, colouring, reading, playing games) with the children's own toys. To keep the amount of time with the study toys consistent across all families, the toys to be used in the observation were not produced until it was time for the observation to take place.

3.2.3 Stage Two: Observation component

To provide some standardisation in the observations and thus permit comparisons between families, parents and their children were videotaped engaging in four tasks (Coolican, 1990). Gardner (2000) reports that more structured tasks rather than free play tasks have been shown to discriminate between children with and without hyperactivity, and that concordance between behaviour in naturalistic versus structured observations occurs when parents are either busy or getting the child to

perform a task. Both of these findings may be because structured tasks or discipline situations provide more opportunities to observe those behaviours such as overreactivity and laxness that are associated with children's behaviour problems. Therefore, to enhance the ecological validity of the observations, the observation tasks were designed to replicate activities in which parents and preschool children frequently engage, but would also prompt discipline behaviours on the part of the parents (Coolican, 1990; Mrazek et al., 1982). They were either selected from existing commonly utilised observational situations (i.e., free play and pack-up tasks) (Campbell et al., 1986b; Campbell et al., 1991; Conger et al., 2003; Gardner et al., 1999; Hops et al., 2003; Kavanagh et al., 1988; Lytton, 1973; Minde, 1992; Pappas-Jones & Adamson, 1987; Radke-Yarrow, Mottlemann, Martinez, Fox, & Belmont, 1992; Rothbaum, 1986), or were modified forms of activities commonly used in observational studies of parental discipline and children's behaviour (i.e., drawing and no distraction tasks) (Arnold et al., 1993; Harvey-Arnold & O'Leary, 1997; Pappas-Jones & Adamson, 1987; Sears, 1965). More information about the tasks is provided below.

Activity One. Free Play

The free play task was designed to elicit teaching, turn-taking, problem solving, communication and warmth from participants (Brophy & Dunn, 2002; Crnic & Greenberg, 1990; Kavanagh et al., 1988; Keown & Woodward, 2002; Meyers, 1999; Minde, 1992; Youngblade & Belsky, 1995). The instruction to parents was to "Spend some time playing with your child with the toys provided. Please don't tidy up the toys when I ask you to move to the next activity". The parents and children could play with any of the toys provided in any manner they liked whilst remaining in the same room as the video camera. The free play task was unstructured and non-task-oriented and was designed to allow participants to become accustomed to the presence of the

author and video camera, and to allow for different types of play with a variety of toys (described in the Materials section, below). This first activity lasted approximately ten minutes until parents were instructed to move to Activity Two.

Activity Two. Drawing Task

The drawing task was designed to elicit teaching, turn-taking, communication and warmth from participants, and to see how well parents and children engaged on a cooperative task (Brophy & Dunn, 2002; Rothbaum, 1986). The instruction to parents was to "Draw a picture with your child on the paper provided. Make the picture about something you both enjoy doing together". This drawing task was scheduled to last approximately ten minutes.

Activity Three. Pack-up Task

The pack-up task was designed to reveal non-compliant childhood behaviour in response to parental instructions and to maximise the chance of observing parental discipline in response to children's noncompliance and defiance (Arnold et al., 1993; Brophy & Dunn, 2002; Campbell et al., 1991; Conger et al., 2003; Cunningham & Boyle, 2002; Harvey-Arnold & O'Leary, 1995; Hops et al., 2003; Kalpidou et al., 1998; Kaplow et al., 2002; Meyers, 1999; Shaw et al., 1998; Whipple et al., 1995; Youngblade & Belsky, 1995). It was anticipated that children might misbehave because they were required to pack up the toys having had only ten minutes free play, and because they would be leaving an entertaining task (i.e., the drawing). The instruction to parents was "Together with your child, please pack up the toys you were playing with earlier". The time allocated to this task varied from family to family because of differences in the speed with which they completed the task, however, a maximum of ten minutes was allowed.

Activity Four. No Distraction Task

This task was designed to assess the children's ability to comply with parental instructions and persevere with a nonstimulating task, whilst the parents' attention was focused elsewhere (Acker & O'Leary, 1996; Arnold et al., 1993; Bor et al., 2002; Cunningham & Boyle, 2002; Harvey-Arnold & O'Leary, 1995; Harvey-Arnold & O'Leary, 1997; Johnston et al., 2002). The instruction to parents was to "Please ask your child to sit quietly and play with the building blocks provided while you complete a short questionnaire". A selection of magazines was also provided for parents to read if they completed this "quick questionnaire" before ten minutes had elapsed. This task was also designed to elicit strategies parents used to keep children "on task". Wooden blocks were chosen in anticipation that they would be a less stimulating toy that might not hold the children's attention for the entire ten minutes allocated to this task. In this way, the likelihood of observing children's non-compliance and parental discipline were increased.

Only two family members (the primary caregiver and the index child) were asked to participate in the observations. This was done for two reasons. First, the focus of the study was the interaction between these two family members, particularly parental behaviour toward the index child. Second, the complexity of coding behaviours and difficulties making comparisons between different families increases when additional family members participate in an observation (Dowdney et al., 1984; Fassnacht, 1982).

During the first telephone contact, parents were advised that the observations would be videotaped and a copy of the video would be sent to them during the following year. The observations were videotaped for two main reasons. First, videotaping produced a permanent record of the observation which could then be coded at a later

date and re-viewed as often as necessary to code different aspects of behaviour (Banister et al., 1996). Second, videotaping the observations meant that inter-observer reliability estimates could be obtained at a later date, rather than having the second coder attend the actual observations which would likely increase participant reactivity (Maxwell & Pringle, 1983).

It was requested that the videotaping take place at a time when there would only be the primary caregiver and the index child at home to avoid interruptions from other family members. Parents were also asked to minimise distractions from outside sources (e.g., visitors, phone calls, television) (Banister et al., 1996; Dowdney et al., 1984; Summerfield, 1983). In some cases, parents made alternative arrangements for the care of other children, or the videotaping took place whilst other children in the families were sleeping or at school.

Parents were told that the observation would consist of four activities which were described by the author and descriptions of these activities were also provided on a double-sided laminated A4 sheet which parents used during videotaping to remind themselves of the tasks (Appendix D.1). Parents were told that although they were being instructed about what to do for each activity, the manner in which these activities were to be performed would be left to their discretion. To maximise opportunities for observing parent-child interaction and to decrease reactivity to the recording process, parents were asked to avoid contact with the author and to "pretend that the camera wasn't there" (Renne et al., 1983; Summerfield, 1983). Further attempts were made to decrease participants' reactivity by allowing time for parents and children to play with the study toys whilst the author set up the camera and tripod (Renne et al., 1983). This was designed to familiarise both the parents and the children with the situation, and to the presence of the camera.

Apart from announcing the transitions between tasks (e.g., “please do Activity Two now”), the author refrained from interrupting the observations and engaged in some unrelated type of work whilst the videotaping took place so that parents would not feel that they were under scrutiny and to decrease their reactivity to the process (Dowdney et al., 1984; Mrazek et al., 1982). This technique also provided a distraction for the author so that if the children looked to the author for attention during any of the tasks (and especially the final, “no distraction” task), no engagement between the children and the author occurred. It was found during piloting that children noticed the camera more, and made attempts to engage the author if she was not apparently busy with another task. The author instructed the children that she had some very important work to do now, but their parents would spend some time with them, playing with the toys provided.

After the videotaping, parents and children were congratulated and thanked for their participation and a discussion about the process took place. Parents were able to talk to the author about the observation and any queries were addressed. The children were praised and the author talked about their drawing and, if time permitted, spent time playing with the children. Children were also presented with a “Bugs Bunny Sticker Book” as a token reward for participating in the study.

On the same day as each visit to a participant’s home, the author made global ratings of the parent’s and child’s behaviour. On a separate sheet, the author made notes about any off-screen interaction that took place during the videotaping, and recorded a general summary of the entire visit to the participant’s home. These measures are described in the “Observation Measures” section in 3.3.2, below. The

videotaped observations were coded using a set protocol and coding guide. This procedure is also outlined in detail in the “Observation Measures” section below.

3.2.4 Stage Three: Parent- and teacher-reports of children’s behaviour

Two years after the completion of Stage Two, parents completed measures of their children’s behaviour as part of the larger longitudinal study examining the aetiology and persistence of children’s externalising behaviour problems. Consent to obtain measures of their children’s behaviour from school teachers was also obtained at this time. All but five teachers completed these measures between one and five weeks after consent was obtained.

3.3 Measures and Materials

3.3.1 Parenting questionnaires

The parenting self-report questionnaire measures were administered to parents as part of a booklet of questionnaires. Parents completed two measures of parenting behaviour: the Parenting Scale (Arnold et al., 1993) and the Child-Rearing Practices Questionnaire (Sanson, 1996). These measures were chosen because they assess a range of parenting behaviours along the dimensions of control and responsivity which have demonstrated relationships with externalising behaviour problems in preschool children.

The Parenting Scale (PS; Appendix E.1)

The Parenting Scale, developed in the United States, has been widely used for program evaluation and theoretical research (Bor et al., 2002; Harvey-Arnold & O'Leary, 1995; Harvey-Arnold & O'Leary, 1997; O'Leary et al., 1999; Zubrick et al., 2002). This 30-item scale utilises a response format in which each item includes a description of two types of discipline, anchored at opposite ends of a seven-point scale (Arnold et al., 1993). Parents indicate on the scale which point represents the best description of their usual responses to their children's misbehaviour. For each item, Arnold et al. (1993) designated one of the descriptions of parenting behaviour a "mistake", and the other behaviour the "effective" parental response on the basis of prior empirical research, and of correlations of the items with children's misbehaviour (Arnold et al., 1993). In Appendix E.1, the anchors that have been designated ineffective techniques are underlined. For example, the effective anchor of one item is "*When I want my child to stop doing something I firmly tell my child to stop*", whilst the ineffective anchor is "*When I want my child to stop doing something I coax or beg my child to stop*". In developing this scale the authors sought to create a measure reflecting contemporary empirical knowledge, which assessed the domain of parental

discipline broadly and directly, and which could be easily and inexpensively administered (Arnold et al., 1993).

Parents' responses on the scale are independent of the frequency of children's misbehaviour, because parents indicate which techniques they are more likely to use, rather than how frequently they use them. In the development of the scale, three subscales were identified in addition to the total parenting score: Laxness, Overreactivity and Verbosity. These subscales represent dimensions of behaviour implicated in the development and maintenance of children's externalising behaviour disorders (Arnold et al., 1993). *Laxness* refers to permissive and/or inconsistent limit setting, in which a parent frequently backs down from requests for compliance or gives in to children's inappropriate demands or coercive behaviour. *Overreactivity* reflects parenting characterised by harsh and coercive discipline such as anger, frustration, insults and name calling, and the use of physical punishment. *Verbosity* refers to the use of lengthy or repetitive verbal responses and reprimands and expresses a reliance on talking even when talking is ineffective.

Scores on the total PS score and the subscales can range from 1 to 7 on each scale, with higher scores representing more ineffective parenting styles. The original subscales were derived using principal components analysis of responses from a combined sample of mothers of clinic and non-clinic children (N=168), in which the three factors accounted for 37% of the scale's variance. In separate analyses, the Verbosity factor was supported using principal components analysis in a nonclinic group, but not in a clinic sample (Arnold et al., 1993).

The scales of the PS have demonstrated adequate test-retest reliability over a two week period in a small sample of 22 mothers. Test-retest correlations for the scales

were Laxness = .83, Overreactivity = .82, Verbosity = .79 and Total Score = .84. Internal consistency of the original and modified versions of the Overreactivity and Laxness scales have generally been high, ranging from α=.70 to .87 (Arnold et al., 1993; Harvey et al., 2001; Irvine et al., 1999; Reitman et al., 2001). The internal consistency of the original Verbosity Scale was lower at α=.63 (Arnold et al., 1993).

In a study of only 15 families, scores on the subscales significantly correlated with global ratings on observational measures of parenting and children's behaviour (Arnold et al., 1993). Global ratings of mothers' laxness, overreactivity, verbosity and general dysfunctional discipline and children's misbehaviour were made on seven point-scales by raters familiar with the PS, after a single viewing of each videotape. Spearman rank-order correlations between observations and questionnaire scores on the same dimensions of behaviour were: Laxness = .61; Overreactivity = .65; Verbosity = .53; Total Parenting = .73; and Child Behaviour = .45 (Arnold et al., 1993).

Subsequent factor analyses using PS data from a variety of samples (including parents of adolescents, children attending Headstart programs, and children with ADHD) have reported strong support for the Overreactivity and Laxness scales, but little support for the Verbosity Scale (Collett et al., 2001; Harvey et al., 2001; Irvine et al., 1999; Reitman et al., 2001).

The Laxness and Overreactivity scales have demonstrated discriminant validity, with scores on these scales significantly differentiating mothers of children with behaviour problems (e.g., children attending clinics or with ADHD) from mothers of children without such problems (Arnold et al., 1993; Harvey et al., 2001). The Verbosity Scale has not demonstrated such validity (Arnold et al., 1993). Scores on the Overreactivity

and Laxness scales have also correlated significantly with parent-reports of marital discord, parental social support, psychopathology, stress, family activities and children's behaviour problems (Arnold et al., 1993; Collett et al., 2001; Irvine et al., 1999; Reitman et al., 2001).

Because of the variable results obtained in factor studies using the PS, and because these studies have all used samples of parents from the United States, a new factor analysis was performed in the current study using data from the screening stage (N=1239). The results from this analysis are described in detail in Chapter 5, and these new factors were utilised in the comparison of self-report and observed data in this thesis.

The Child-Rearing Practices Questionnaire (CRPQ; Appendix E.2)

The CRPQ was developed in Australia in the course of the Australian Temperament Project (Sanson, 1996) and has been used with children aged two to nine years of age. This scale was based on the work of Hart et al. (1992), Rubin and colleagues (1995), and Russell et al. (1998). The scale comprises 30 items and asks parents to rate the frequency with which they engage in certain behaviours. The measure has a total score and three subscales rating warmth, inductive reasoning/power assertion and punitiveness. *Warmth* refers to displays of intimacy and responsivity including positive affect, private jokes, laughter, praise and physical intimacy such as hugs and kisses. *Inductive Reasoning/Power Assertion* refers to the use of reasoning and explanation about the need for rules or punishments, while scores describing the use of power assertive techniques to obtain children's compliance load negatively on this subscale. *Punitiveness* includes the use of threats or physical discipline and expectations of unconditional obedience .

The original subscales were derived using a principal components analysis with varimax rotation on a sample of 79 parents of two-year-old children. The three subscales identified relate to dimensions of parental behaviour which, when the levels of the behaviours are lower, are associated with children's behaviour problems (aggression and noncompliance) and, conversely, when the levels of behaviour are higher, with children's positive social adjustment (Hemphill & Sanson, 2001). Individual item and subscale scores on the CRPQ range from one to five, with higher scores indicative of more positive parenting styles. The reliability and validity of the original scales has not been reported, but an adapted version of this scale was used in a study of parenting style, children's temperament and behaviour, which incorporated a fourth parenting factor entitled "Obedience". This subscale reflects the extent to which parents expect unquestioning obedience from their children (Paterson & Sanson, 1999). The internal consistency of the subscales of this adapted version were: Warmth α=.81, Punishment α=.84, Reasoning (Explanation) α=.60, and Obedience α=.69. The CRPQ also demonstrated good predictive validity with scores on the Warmth Scale at 2 years old predicting social skills at 4 years old, and scores on all scales associated with child behaviour problems two years later for children with difficult temperaments (Paterson & Sanson, 1999).

As with the PS, a factor analysis using the items of the CRPQ was repeated in the present study using the data from the sample in the second stage of the study (N=438). The factors resulting from this analysis are discussed in Chapter 5, and form the basis of comparison between self-report and observed parenting behaviour.

3.3.2 Observations of parenting behaviour

Observation Materials

A Panasonic VS3 video camera was used to record the observations. The camera was mounted upon a tripod and the recording was observed on a 2.5 inch colour flip-out screen. The observations were initially recorded onto Panasonic 60 minute Compact Video Cassettes and then transferred to Panasonic VHS Videocassettes for use in standard video cassette recorders.

The toys used in the observation were selected for their novelty, variety and attractiveness to the children, and to promote a range of behaviours on the part of participants (e.g., creative play, cooperative play, problem-solving, and turn-taking). The selection of toys for the observation included:

1. Coloured paper and crayons
2. A plastic toy zoo with lockable doors on the cages, containing animals with different shapes on their bases that can be inserted through the tops of the cages. The doors of the cages could be unlocked with keys of different shapes.
3. Mr Potato Head with 25 parts
4. Fisher Price toy school bus
5. Dancing Bear in the Big Blue House
6. Four small jigsaws of varying degrees of difficulty (shapes, numbers, animals and occupations)
7. Duplo blocks including a small and large hippopotamus in a plastic container in the shape of a hippopotamus.
8. Two small picture books with rhyming text “Crocodile School Bus” and “Penguin Taxi Cab”
9. A small trolley of wooden building blocks

The coloured paper and crayons were provided for the “drawing task” and the wooden blocks were used in the “no distraction” task.

Observation Measures

In the “no distraction” task of the observation, parents completed a questionnaire whilst their children were asked to play with the wooden blocks. The questionnaire contained a mixture of open-ended and yes/no questions designed to gauge parental perceptions of the representativeness of the videotaped observation and their feelings about being observed (Appendix F.1). Specifically, the questionnaire assessed parental perceptions of: 1. The representativeness of the observed interactions; 2. How the observed interaction differed from what would normally occur between the parents and children; 3. If the parents enjoyed the observation and why or why not; 4. What the best things were about being the parent of their children; and 5. Any concerns about being the parents of their children.

During the videotaping any off-screen interaction that took place whilst the video camera was running, was noted on a record sheet (Appendix F.2). This was done to provide the observation coders with contextual information about the parents’ or children’s behaviour which they would otherwise not have (because it was not recorded on the videotape) and this information may be important in the interpretation of the participant’s behaviour (Maxwell & Pringle, 1983; Renne et al., 1983). For example, information about what children were doing when they left the range of the video camera, might assist the observer’s interpretation about whether children were still on-task (e.g., getting a picture to help with the drawing activity) or misbehaving (e.g., going to play outside when they were asked to pack up the toys) (Renne et al., 1983).

On the same record sheet, the author made a general summary of the entire visit to each participant’s home after the visit. This was an unstructured narrative record of

the author's experiences and impressions and included any salient points about the visit (Banister et al., 1996; Brandt, 1992; Renne et al., 1983). These qualitative summaries of the observation visits were made both to provide contextual information when the videotapes were subsequently coded and to reflect the author's experience of the visit. The information in these summaries varied from family to family but in general included information such as: the author's perceptions of the representativeness of each observation; any differences in children's or parents' behaviour between the observed and non-observed periods; and any details about the children's or parents' behaviour that were mentioned in conversation.

Other observational studies have used a wide variety of methods to code observational data (see Appendix A.2), and these different coding methods incorporate different levels of observer subjectivity (Brandt, 1992; Dowdney et al., 1984). For these reasons, observed parenting behaviour was coded in three ways in this study. First, global ratings of parenting behaviour (Laxness, Verbosity, Overreactivity, Punitiveness, Inductive Reasoning/Power Assertion and Warmth) were made on a "global rating sheet" (Appendix F.3), immediately after each home visit. An interval scale from one to seven was used, with one indicating no or very low levels of behaviour and seven indicating very high levels of behaviour (Arnold et al., 1993; Harvey-Arnold & O'Leary, 1995; Harvey-Arnold & O'Leary, 1997). This was based on the author's impressions before, during and after the videotaped period, together with information that arose during conversations with the parents and children. These ratings were made before the factor analyses were performed on the PS and CRPQ and hence were made on the published subscales of these measures.

Second, parent behaviours were directly coded from the videotapes using interval sampling. The coded behaviours can be seen on the coding sheets in Appendix F.4,

and were developed on the basis of congruent items on the PS (Arnold et al., 1993) and the CRPQ (Sanson, 1996). Items on the questionnaires were operationalised to be clearly evident as to what they involved, and these definitions are shown in Appendix F.5. The operationalisation process is described in more detail in Chapter 5. Items were not included in this analysis if they referred to behaviours occurring outside the home and thus could not be seen in the videotapes, or they referred to the attitudes and beliefs of parents, and hence could also not be directly observed. For items that represented reactive discipline (i.e., discipline that occurred in response to children's misbehaviour), children's misbehaviour was defined as non-compliance, defiance, moving off-task and aggression (see Appendix F.5 for operationalisations of these behaviours).

Parental behaviour on the videotapes was recorded on coding sheets which provided spaces for each behavioural item in 30-second intervals, with separate sheets for each activity. The behaviour was recorded using interval sampling in which a box was ticked for a behaviour if it occurred in the 30-second period under scrutiny (Baum & Forehand, 1981; Brandt, 1992; Dadds & McHugh, 1992; Fassnacht, 1982; Forehand et al., 1986). Each behaviour was recorded again if it continued or recurred in the next 30-second period. Behaviour was recorded until the end of the designated activity (e.g., free play) and the time of the end of the activity was marked on the coding sheet. A separate recording sheet was used for each activity. An audiocassette with 30-second intervals marked by the word "time", was used whilst coding observations thus allowing the coder to concentrate fully on the observed interaction without having to consult a clock or video counter. Each videotape was coded from start to finish on a single dimension (e.g., laxness), with a separate coding sheet used for each activity (e.g., free play, drawing etc). In order to reduce familiarity effects with each family, the next participant's entire video was then event

sampled on a different parent behaviour dimension; this process continued until all videotapes had been coded on one behaviour domain each and then the entire procedure was repeated until all domains were coded (Maxwell & Pringle, 1983).

Although the parent behaviours on the coding sheets were grouped according to their published PS and CRPQ subscales (Arnold et al., 1993; Sanson, 1996), they were coded individually and hence could be computed into new behaviour scales according to the results of the factor analyses performed in the current study. Because the observation times were not uniform across families (e.g. some families packed up quicker than others) and because longer observations provide more opportunities for behaviours to occur, the behaviour counts were standardised by dividing the sum of the behaviours by the length of the observation (Cunningham & Boyle, 2002; Kalpidou et al., 1998). That is, subscale and total score behaviour counts were divided by the total time (in minutes) of the observation. This method is described in more detail in Chapter 6.

Third, in addition to the interval sampling, at the end of coding each activity for a particular domain, an overall rating (between one and seven, with seven indicating a higher level of the parenting behaviours on that domain) was given on the coding sheet for the parent's behaviour on that domain during that activity. These ratings were given based on a description of each parenting domain that can be seen in the Parenting Questionnaires section above. This overall rating was intended to be more of a clinical rating, rather than simply a mathematical one (that is, it was intended to reflect the coder's impression of the parents' behaviour rather than the total frequency of behaviours that make up the particular domain). As with the global ratings of behaviour made immediately after the observations, these activity ratings were made before the factor analyses of the parenting questionnaire data were

performed. Hence, they are based on the published questionnaire subscales derived by the instruments' authors (Arnold et al., 1993; Sanson, 1996).

3.3.3 Measures of family risk factors and children's behaviour

In addition to assessments of parenting practices, the questionnaire booklet completed by parents included questionnaires designed to assess children's temperament, parental psychopathology, social support, agreement about child-rearing issues between parents, marital adjustment and demographic factors. Measures of children's behaviour included global ratings of children's externalising behaviour made by the author after the videotaping, and parent- and teacher-reports of children's behaviour obtained at the home visit and two years later. All of these measures are described below.

Child Temperament

Child temperament was assessed with the Short Childhood Temperament Questionnaire - Australian Version (Appendix G.1) (Sanson, Smart, Prior, Oberklaid, & Pedlow, 1994), based on Thomas and Chess' (1977) conception of temperament. This 31-item questionnaire contains subscales measuring aspects of temperament such as inflexibility (negative emotionality and adaptability), persistence (attention and on-task behaviour), approach (sociability and response to new situations), and rhythmicity (regularity and predictability of basic functions such as hunger, excretion and sleep). A global item assesses parents' perceptions of the overall "easiness" or "difficultness" of their children's temperament. The questionnaire has adequate reliability and validity (Sanson et al., 1994). This thesis reports results using only the inflexibility and persistence subscales because of their relevance in the development of externalising problems in children (Prior et al., 1993).

Parental psychopathology

This was measured using the 30-item version of the General Health Questionnaire (GHQ; Goldberg, 1978, Appendix G.2). Parents responded to each item on a four-point Likert response scale that utilised different response labels according to the nature of the different items. Responses to each item were scored according to the level of psychopathology they indicated. For example, item responses which indicated the two lower levels of psychopathology (e.g., not at all or no more than usual) were scored as 0, the next response (e.g., rather more than usual) was scored as 1, and the response indicating the highest level of psychopathology (e.g., much more than usual) was scored as 2. Some items required reversal before they were scored and a total score was calculated by summing all items. Higher total scores indicate higher levels of psychopathology, and whilst the total score can range from 0 to 60, scores of 3 to 4 or higher indicate clinically significant levels of

psychopathology. The GHQ has been widely used to measure non-psychotic psychological impairment among adults in the community and has high levels of test-retest reliability, internal consistency and good validity (Vieweg & Hedlund, 1983).

Social support

Parental social support was measured using the Interview Schedule for Social Interaction – Short Form (ISSI-SF; Unden & Orth-Gomer, 1989, Appendix G.3). This 13-item self-report measure is an abbreviated version of the Interview Schedule for Social Interaction (Henderson, Byrne, & Duncan-Jones, 1981). It assesses the availability and adequacy of social integration (diffuse social relationships), as well as the availability and adequacy of interpersonal attachment (close, confiding and intimate relationships). The ISSI-SF has demonstrated adequate reliability and validity. Only the Availability of Social Integration Scale is reported in the current thesis.

Parental Agreement about Child-rearing

Inter-parental agreement about child-rearing was assessed using the Parent Problem Checklist (PPC; Sanders & Dadds, 1993, Appendix G.4), completed by parents in two-parent families. This scale was developed at the University of Queensland as a measure of inter-parental conflict. Parents indicate which of the 16 parenting problem areas are a source of disagreement between themselves and their partner. Problem areas include, for example, disagreements about rules and discipline for children's misbehaviour, open conflict about child-rearing, and one parent undermining the relationship of the other parent with the children. The total number of parenting problems is summed and ranges from 0 to 16. Scores equal to or greater than 5 indicate levels of parenting disagreement that are clinically significant (Dadds &

Powell, 1991). The PPC has high test-retest reliability and moderate internal consistency and validity (Dadds & Powell, 1991).

Marital Adjustment

Marital satisfaction was measured using the Dyadic Adjustment Scale (DAS; Spanier, 1976, Appendix G.5) which provides an overall measure of the quality of a couple's adjustment in a relationship. This 32-item measure was completed by participants in two-parent families (n=55) and scores range from 0 to 151, with higher scores indicating better marital adjustment. Individual items measure areas of marital consensus, satisfaction, cohesion and affectional expression. The DAS has high internal consistency, and good discriminant, content and construct validity (Fredman & Sherman, 1987).

Demographic Characteristics

Information about the demographic characteristics of participants was collected using a questionnaire developed for this purpose in the Research and Evaluation Unit, Women's and Children's Hospital, South Australia (Appendix G.6). This questionnaire obtains information about the child's age, gender, number of siblings and any illnesses or disorders of the child, the respondent's relationship to the child, the family structure, along with information about the country of birth, educational attainment, the employment status and usual occupations of the child's mother (or maternal figures) and father (or paternal figures).

Children's Externalising Behaviour

Global ratings of children's externalising behaviour were made on the same sheets as ratings of parents' behaviour by the author immediately after each visit to participants' homes (Appendix F.3). These ratings were made on a scale from one to seven, with "one" indicating no or very low levels of behaviour and "seven" indicating

very high levels of behaviour. This was based on the author's impressions from the observation based on the observed behaviour from before, during and after the videotaped period, together with information that arose during conversations with parents and children.

In Stage Three, two years after the observations were completed, childhood emotional and behavioural problems were assessed using the Child Behavior Checklist (6-18) (CBCL(6-18); Achenbach & Rescorla, 2001) and the Teacher Report Form (TRF; Achenbach & Rescorla, 2001) (see Appendices G.7 and G.8). Both of these measures include large sets of items describing a wide range of emotional and behavioural problems. The scores on these items are summed to form raw scores on several narrow-band or syndrome scales, but because in this thesis the focus is on the relationships between parenting behaviour and children's externalising behaviour, only the scores from the Externalising Scale will be reported. The Externalising Scale contains items about behaviours that are aggressive, antisocial and undercontrolled. The CBCL(6-18) and the TRF have adequate test-retest reliability, moderate cross-informant agreement and good discriminant and construct validity (Achenbach & Rescorla, 2001).

Chapter 4. Sample Characteristics

The study was designed to engage a sample of participants with a range of parenting and children's behaviours. The representativeness of the sample is explored in this chapter by comparing their demographic features with those of the rest of the sample from which they were randomly selected, and with a nationally representative sample of families of four- to five-year-old children. This is important to determine whether the results from this study sample are generalisable to other parent-child dyads. The parent-reports of family risk variables and children's behaviour are also described in this chapter.

4.1 Demographic Characteristics

The children in the study sample were aged between three and five years at the time of the observations (1.5% aged three years, 60.3% aged 4 years, and 38.2% aged 5 years). Natural mothers completed the observations and questionnaires in 93% of cases, and natural fathers were the respondents in five families. The first column in Table 4.1 shows the demographic characteristics of the sample (n=68) who participated in the observation component. Almost half of the sample (47%) was male, and 10% of the children had an illness or disability (such as asthma or a speech problem). The majority of children lived in two-parent families, and most children had either one or two siblings. The sample included a predominantly Australian-born group of mothers, 30% of whom had not completed secondary school, and one third of whom were in paid employment.

Table 4.1 Demographic characteristics of the sample

Characteristic	Observation Sample (N=68) (%, 95%CI)	Screening Sample (N=1174) (%, 95%CI)	National Survey (N=564) (%, 95%CI)
Male	47.1 (35.2-59.0)	50.9 (48.0-53.8)	52.8 (48.7-56.9)
Illness or disability	10.3 (3.1-17.5)	not assessed	17.7 (14.6-20.8)
Family Structure			
Two natural parents	75.0 (64.7-85.3)	82.3 (80.1-84.5)	80.1 (76.8-83.4)
Step/blended	4.4 (0.0-9.3)	3.6 (2.5-4.7)	4.4 (2.7-6.1)
Sole parent	20.6 (11.0-30.2)	12.8 (10.9-14.7)	15.5 (12.7-18.3)
Number of siblings			
Only child	10.3 (3.1-17.5)	13.5 (11.5-15.5)	18.3 (15.1-21.5) [a]
One sibling	42.6 (30.8-54.4)	52.1 (49.2-55.0)	50.5 (46.4-54.6)
Two siblings	32.4 (21.3-43.5)	25.1 (22.6-27.6)	23.6 (20.1-27.1)
More than two siblings	14.7 (6.3-23.1)	9.3 (7.6-11.0)	7.6 (5.4-9.8)
Mother born in Australia	86.8 (78.8-94.8)	not assessed	73.0 (69.3-76.7)[a]
Maternal education < Year 12	29.9 (19.0-40.8)	33.5 (30.8-36.2)	not assessed
Mother in paid employment	35.3 (23.9-46.7)	46.2 (43.3-49.1)	42.9 (38.8-47.0)

[a]=significant differences between the current sample and the indicated scores

In order to test if there was a refusal bias for the observation component, chi-square analyses were used to compare the screening data of those who participated in the observation, with those who were selected for observations but declined to participate (n=13). There were no significant differences between these two groups, although the power to detect significant differences with these numbers was limited. The demographic features of the observation sample were also compared with the

larger screening sample from which they were randomly selected (see Table 4.1). Once again, there were no significant differences between the 68 families who participated in the observation component of the study and the larger screening sample (n=1174), although there was a lower proportion of children in the observation sample to have mothers in paid employment (35.3% versus 46.2%, χ^2=3.11, df=1, p=.08) and a higher proportion to come from sole parent families than other family types (20.6% versus 12.8%, χ^2=3.42, df=1, p=.06). These differences between the observation sample and the remainder of the screening sample may reflect characteristics of the sample selection methods. The sampling strategy for the second stage of the study meant that a greater proportion of children with behaviour problems were included in the observation sample than were in the screening sample from which they were drawn. Parent characteristics such as unemployment and sole parent status were associated with children's externalising behaviour in the screening stage . For example, the externalising scores of children who had mothers in paid employment were 10.9±7.6 compared with 12.4±8.8 for children of mothers not in paid employment (t=3.03, df=1203, p=.002). Similarly, in the screening stage the mean externalising behaviour score of children in two-parent families was 11.2±8.0 compared with 13.9±9.4 for children from sole parent families (t=4.02, df=1170, p<.001). Because of these associations, children of mothers who were not in paid employment and children from sole parent families were more likely than other children to be selected for the second stage of the study.

To assess if the demographic characteristics of the current sample were nationally representative, they were compared with data from a national sample of parents of four and five year children who completed a similar demographic questionnaire (see the final column of Table 4.1). It was not possible to perform these comparisons with data from the Australian Bureau of Statistics census because of differences in the

demographic information collected. The data in Table 4.1 were obtained from the child and adolescent component of the National Survey of Mental Health and Wellbeing (Sawyer et al., 2000). Chi-square analyses revealed that in this study sample, the proportions of children who were boys, had an illness or disability, were from different family types and had mothers in paid employment, were not significantly different from the four to five year old children in the national study sample (see Table 4.1). However, in this study sample there were significantly more children from larger families (no or one sibling: current sample = 52.9% versus national sample = 68.8%; two or more siblings: current sample = 47.1% versus national sample = 31.2%; χ^2=4.37, df=3, p=.04). The smaller proportions of children in the current study who were only children or had only one sibling, compared with the Australian sample, seems to be a reflection of the selection of participants for the observation component of the study. There are no significant differences between the remainder of the screening sample and the nationally representative sample in terms of number of siblings.

A significantly greater proportion of mothers in this study sample were born in Australia, compared with mothers in the national study sample (86.8% versus 73.0%; χ^2=6.01, df=1, p=.01). Although no specific provision was made to assist non-English speaking participants in either sample, the national survey was conducted by face-to-face interview which may have overcome some reading and language comprehension barriers. In the current study, participants completed the screening questionnaire without assistance from research staff, which may have excluded people with language difficulties, or people from non-English speaking backgrounds.

Further, there were differences in the sampling strategies employed by the two studies in that the current study recruited participants through preschools, whereas

the national survey sampled participants via households. It is possible that any demographic differences seen between the two samples may reflect a bias between parents whose children attend preschool and those whose children do not. Also, the national sample included non-metropolitan areas whereas the observation sample was recruited from metropolitan preschools only.

In summary, the current study sample was somewhat representative of the screening sample from which it was drawn, however lower proportions of parents were in paid employment and higher proportions of children were from sole parent families than in the larger sample. These differences are largely a result of the sampling strategy for the second stage of the study, in which children from sole parents and/or who had mothers who weren't in paid employment were more likely to be selected for this stage of the study. The current sample resembled a nationally representative sample of parents of four and five-year-old children in a number of demographic areas, with differences in family size, and country of birth. These differences most probably reflect methodological differences between the two studies. The present sample may be biased in the under-representation of families from non-English speaking backgrounds.

4.2 Family Risk Factors

Parents completed questionnaire measures of family risk factors such as children's temperament, parental psychopathology, social support, parental disagreements about child-rearing and marital adjustment. The mean scores from these scales are shown in Table 4.2.

Table 4.2. Mean scores (SD) on the measures of family risk factors

Scale	**Mean (SD) N = 68**	**Possible Range**	**Observed Range**
Temperament Inflexibility	2.82 (.87)	1-6	1.33-4.89
Temperament Persistence	3.42 (.99)	1-6	1.43-5.71
General Health Questionnaire	2.97(4.41)	0-60	0-18
Availability of Social Integration	3.84 (1.62)	1-6	1-6
Parent Problems Checklist	2.76 (3.19)	0-16	0-15
Dyadic Adjustment Scale	115.17(15.81)	0-151	61-141

Child Temperament

Children were reported to demonstrate higher levels of persistence, than difficult, inflexible temperament traits ($t=4.66$, $df=67$, $p<.001$). In addition to completing the 30-item scale, parents were asked to describe their children's temperament compared with the "average child". Parents could choose one of five fixed responses: no parents in this sample described their children as "much more difficult than average"; 14.7% responded that their children were "more difficult than average"; 36.8% that their children's temperament was "average"; 23.5% "easier than average"; 17.6% as "much easier than average"; and five parents (7.4%) did not respond to this question.

The small proportion of children rated as difficult in this study is surprising given that half the sample was selected on the basis of parent-ratings of behavioural difficulties. Although it is consistent with the findings of the Australian Temperament Project which reported that only a small proportion of parents rate their children as “difficult” (Smart & Sanson, 2001). In this study, it may be an indication that parents find behavioural problems and temperament difficulties normative for this child age group.

Parental Psychopathology

Twenty parents (29.4%) in this study reported levels of psychopathology at or above the recommended cut-off score of 4 on the GHQ. The items on which the largest proportion of parents reported “rather more” or “much more trouble than usual” were: feeling constantly under strain (29.4%), experiencing restless nights (27.9%), and losing sleep over worry (22.1%).

Parental Social Support

In general, parents in the study were satisfied with their levels of social support. For example, 82% of parents indicated that the number of people in their daily life was “about right”, with only 4% of parents indicating that they had no-one they could “lean on”.

Parental Agreement about Child-Rearing

On average, participants in two-parent families (N=55) had approximately three areas of disagreement about child-rearing between themselves and their partner (mean = 2.76±3.19). Just over one third (34.5%) of participants identified no areas of disagreement about child-rearing between themselves and their partners. However, one quarter (25.5%) of participants in two-parent families reported disagreement about child-rearing in five or more areas, which is indicative of clinically significant parenting problems (Dadds & Powell, 1991). The most commonly reported parenting

problems in two-parent families were: *one parent is "soft" and one parent is "tough" with children* (34.5%), *disagreements about household rules* (29.1%), and *children preventing parents from being alone* (29.1%). The areas of disagreement least frequently endorsed by participants were: *inability to resolve disagreements about childcare* (1.8%), *disagreements about who should discipline children* (5.5%), and *discussions about childcare turning into arguments* (7.3%).

Marital Adjustment

For participants from two-parent families (N=55), the mean score (SD) on the DAS was similar to other levels of marital satisfaction found in other studies with parents of preschool-aged children in Australia (Dadds & Powell, 1991), New Zealand (Woodward, Taylor, & Dowdney, 1998), and the US (Meyers, 1999). In the current study, the areas of interaction in which the highest levels of disagreement between partners were reported were household tasks (15.1%) and ways of dealing with in-laws (10.3%). The lowest levels of disagreement were found in the areas of religious matters (1.9%), matters of recreation (3.7%), friends (3.7%), and aims and goals (3.7%). The majority (68.5%) of participants in two-parent families in this sample had never discussed or considered divorce or separation. Only 13.5% of participants rated themselves as unhappy (either a little unhappy = 11.5% or fairly unhappy = 1.9%) in their relationship. No participant rated themselves as extremely unhappy in their relationship.

In summary, 15% of parents perceived their children as “more difficult than average”, however no parents described their children as “much more difficult than average”. Twenty-two percent of parents reported levels of psychopathology above the cutoff score on the GHQ, with the most commonly reported problems being strain, restless nights and lack of sleep. In general, parents described their levels of available social support as adequate. One quarter of parents in two-parent families reported clinically significant levels of disagreement with their partners about childrearing. The most frequently reported areas of disagreement were the roles of parents in discipline and disagreements about household rules. The levels of marital satisfaction in this study were similar to those reported by non-clinic samples of parents of preschool children.

Chapter 5. Self-Reported Parenting Behaviour

The Parenting Scale (PS) and the Child-Rearing Practices Questionnaire (CRPQ), were used in the current study because they assess aspects of parental control and responsivity that have been associated with children's externalising behaviour problems. This chapter examines their psychometric properties and reports the summary scores obtained from the measures. Factor analyses using screening data for the PS, and stage two data for the CRPQ are also reported in this chapter. In each section of this chapter, the results are presented separately for the PS and CRPQ because the instruments use different response formats to assess different domains of parenting behaviour. The PS assesses the *types* of parental discipline used (defined by its developers as effective and ineffective), whereas the CRPQ is a measure of the *frequency* with which both negative and positive parenting techniques are used.

5.1 Content Validity of the Self-Report Parenting Questionnaires

5.1.1 Content Validity of the Parenting Scale

Attempts to maximise content validity in the development of the Parenting Scale included a thorough review of the literature on parenting, a review of transcripts of parents discussing discipline practices, and the incorporation of feedback from parents about the clarity of the items (Arnold et al., 1993). As described in Chapter 3, the PS utilises a response format in which each item includes a description of effective versus ineffective discipline techniques, anchored at opposite ends of a seven-point scale, and parents are required to indicate on the scale which point represents the best description of their responses to their children's misbehaviour. This response format enhances the content validity of the PS because parental responses are independent of the frequency of children's misbehaviour as parents indicate which technique they are more likely to use, rather than how frequently they

use it. However, in the use of self-report measures, threats to content validity are common (Holden, 1983), and the PS is no exception. Elements that pose threats to the content validity of the PS are discussed below.

1. Many of the items on the PS do not incorporate information about the specific type of child misbehaviour (e.g., defiance, cheating, stealing) or the context in which it occurs (e.g., at school, in public, at home) (Holden, 1983; Holden & Edwards, 1989). The majority of items on the scale simply state "When my child *misbehaves*, I….", and although a general introductory paragraph provides a definition of child misbehaviour, the behaviours stated are heterogeneous (e.g., not picking up toys, wanting a cookie before dinner and running into the street) which may prompt different parental responses. The measure does not account for flexibility in parenting and the appropriateness of different responses for different behaviours or in different settings. This means that across-parent and even within-parent inconsistencies in responding may occur because parents may be responding to items with different behaviours, contexts and situations in mind. On this measure, parents who describe themselves as using the "effective" parenting technique "always" are regarded as using effective discipline techniques, when they may in fact be responding to their children's behaviour in an inflexible and unfair manner. Parenting which is flexible and situation-specific has been described as adaptive and has been related to positive children's outcomes (Smith, 1999).

2. The PS contains some "double-barrelled items", that is, items which contain two qualitatively different focal phrases (e.g., "*When my child misbehaves I raise my voice <u>or</u> yell*"; "*When my child doesn't do what I ask I often let it go <u>or</u> end up doing it myself*"). In these statements it is unclear as to which part of the question the participant is responding (Holden, 1983).

3. The wording of some items on the PS was described as ambiguous or unclear by participants in the present study. For example, piloting showed that some parents were not sure if the item “When I have to handle a problem I tell my child I am sorry about it” referred to a problem of their own (e.g, having to go in to work on a day off) or that of their children (e.g., children’s noncompliance). Because the Parenting Scale assesses parental control practices, this question was adapted to “*When I have to handle a problem with my child I tell my child I am sorry about it*”, to make the item clearer for parents.

4. In piloting, parents described having difficulty when completing the scale because they were unsure how to use the original response options. Initially, parents were instructed to place a mark in one of the seven tick boxes between the two anchor statements, as an indication of their usual parenting technique. The same response format was employed in this study, and to clarify how the measure should be completed, explicit instructions, a sample item, and labels for the different response choices were included (see Appendix E.1).

In summary, an advantage of the PS is that parental responses to its items are independent of the frequency of children’s misbehaviour. However, the scale does contain some double-barrelled items and its lack of context-specificity may mean that flexibility in parenting is interpreted as dysfunctional. Although in piloting some parents reported confusion in interpreting some items and in understanding the response scale, attempts were made to ameliorate these difficulties before the scale was administered in the current study.

5.1.2 Content Validity of the Child-Rearing Practices Questionnaire

Most of the items on the CRPQ were drawn from other parenting questionnaires with sound psychometric properties (Paterson & Sanson, 1999). The items on the CRPQ have high face validity, and appear to measure the parenting domains of interest (warmth, punitiveness and inductive reasoning/power assertion). In this study, recognised threats to the content validity of this measure which were identified are described below.

1. Because of the response format employed (i.e., never, rarely, sometimes, often, always), items on the CRPQ which relate to the discipline of children's misbehaviour (but not to parental warmth) are frequency-dependent (Holden, 1983; Holden & Edwards, 1989). That is, parental responses regarding how often they employ particular discipline techniques may depend upon the frequency of their children's misbehaviour. Parents whose children do not misbehave very often will only rarely need to employ discipline techniques. They are less likely to report using ineffective techniques "often" or "always" because they are responding to fewer instances of misbehaviour, even though they may use these ineffective parenting techniques every time their children misbehave. Conversely, parents of children who frequently misbehave may report using ineffective techniques more often because they are involved in discipline interactions with their children more often. For these parents, these ineffective parenting techniques may only be one component of their entire behaviour management repertoire. Although frequency-dependent self-report measures are commonly used in parenting research (Crnic & Greenberg, 1990; Fox, 1994; Sanson, 1996; Shelton et al., 1996), their frequency-dependence may mean that spuriously high correlations between ineffective parenting techniques and children's behaviour problems are reported.

2. The CRPQ measure is not a pure measure of parental behaviour as it includes a small number of attitude and belief questions together with behavioural items. Although responses to belief questions can be highly correlated with self-reported frequencies of behaviour (Sears, 1965), it is not always the case that belief predicts behaviour (Reitman et al., 2002; Sonuga-Barke et al., 2001). For example, parents may respond "never" to the question "*I believe physical punishment is the best way to discipline my child*", but respond "sometimes" to the item "*I use physical discipline e.g., smacking for very bad behaviour*".

3. The CRPQ contains "double-barrelled" questions. For example, the item "*I enjoy listening to my child and doing things with him/her*" includes two focal statements that are not necessarily synonymous. As mentioned earlier, the presence of two focal statements in double-barrelled items makes it difficult to interpret respondents' answers because it is not known to which part of the item the participant is responding (Holden, 1983).

4. Although, this measure contains some items which provide information about the type of misbehaviour being disciplined (e.g., "*I use physical punishment for very bad behaviour*"), there is limited contextual information for many of the discipline items on this measure.

In summary, the CRPQ has high face validity. However, like many other self-report parenting measures the instrument contains a small number of double-barrelled items as well as a mixture of belief and behavioural items, and provides limited contextual information for each item. The elements of content validity discussed in this section are relevant to the interpretation of the results obtained using self-reports of parenting behaviour, and will be discussed again in conjunction with these results in Chapters Seven and Eight.

5.2 Construct Validity: Factor Structure and Reliability of the Self-Report Parenting Questionnaires

5.2.1 Factor Structure and Reliability of the Parenting Scale

Factor Structure

As described in Chapter 3, the factor structure of the Parenting Scale has been examined in several studies using a variety of samples (see Appendix H.1 for a summary) (Arnold et al., 1993; Collett et al., 2001; Harvey et al., 2001; Irvine et al., 1999; Reitman et al., 2001). Most of these studies have used relatively small samples of parents from the United States. The decision to perform a principal components analysis using the PS items in the current study, was made because of the mixed results obtained in previous work, the large sample size available in this study, and because the previous factors obtained may not be applicable to the practices of Australian parents.

The data from participants in the screening stage of the current study (N=1239) were used to conduct the principal components analysis of the PS items. The analysis (using varimax rotation) yielded five principal components with eigenvalues greater than one and which made conceptual sense (Kaiser criterion) (Table 5.1). Only the items that had a factor loading of 0.40 or greater (those that share at least 16% of variance with the factor) are shown in the table. These five components accounted for 39.3% of the variance in scale scores, and displayed some similarities to the original factor structure of the scale, but as with other factor analyses using the PS (Harvey et al., 2001; Irvine et al., 1999; Reitman et al., 2001), the presence of a Verbosity factor was not supported.

Table 5.1. Principal components analysis (with varimax rotation) of the Parenting Scale self-report items (N=1239)

Original Domain and Behaviour Item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Laxness					
20. When I give a fair threat or warning, I often don't carry it out	.78				
19. When my child won't do what I ask, I often let it go or end up doing it myself	.68				
30. If my child gets upset, I back down and give in	.65				
7. I threaten to do things I know I won't actually do[a]	.64				
26. When I say my child can't do something, I let my child do it anyway	.64				
21. If saying "no" doesn't work, I offer my child something nice so he/she will behave	.41				
16. When my child does something I don't like, I often let it go		.66			
15. When we're not at home, I let my child get away with a lot more		.61			
8. I let my child do whatever he/she wants		.58			
12. I coax or beg my child to stop		.57			
24. If my child misbehaves and then acts sorry, I let it go at that time		.40			
Overreactivity					
25. I almost always use bad language or curse			.71		
28. I insult my child, say mean things, or call my child names most of the time			.73		
18. I spank, grab, slap, or hit my child most of the time			.65		
17. Things build up and I do things I don't mean to				.45	
22. I get so frustrated or angry that my child can see I'm upset				.50	
3. When I'm upset or under stress, I'm on my child's back				.67	
10. I raise my voice or yell				.58	
9. I give my child a long lecture[b]					.70
6. I usually get into a long argument with my child					.56
14. I often hold a grudge					.55
Verbosity					
23. I make my child tell why he/she did it		.42			
11. If saying "no" doesn't work right away, I keep talking and try to get through to my child					
29. If my child talks back or complains when I handle a problem, I give a talk about not complaining					
2. I give my child several reminders or warnings					
4. I say a lot					
No Factor					
1. I do something about it later		.52			
13. When my child is out of my sight, I often don't know what my child is doing					
27. When I handle a problem, I tell my child I'm sorry about it					
5. I can't ignore my child's pestering					

[a] Loads on the original Laxness and Verbosity Scales
[b] Loads on the original Overreactivity and Verbosity Scales

Items from the Laxness factor (Arnold et al., 1993) loaded onto two components in the present analyses. The first factor involves items related to inconsistent discipline such as backing down and not following through with stated courses of action. The second factor is comprised of permissive parenting behaviours such as “letting misbehaviour go” and not providing discipline in the first place. Items from the Overreactivity Scale (Arnold et al., 1993) loaded onto three separate factors: one involving physical and verbal punishments; the second involving frustration and anger; and the third includes lecturing and arguing with children (Verbal Discipline). The robustness of the factors obtained in the current sample was assessed by conducting additional principal components analyses using different rotations. The results obtained using these alternative rotations were quite consistent with the factor structure displayed in Table 5.1 (see Appendix H.2 for a discussion of the minor differences).

Subscale scores were calculated for each of these factors using the relevant items. For the purpose of this study Factors 3 and 4 were combined into a single “Overreactivity” factor because they were highly correlated (r=.65), and because in the factor analyses using alternative rotations, considerable overlap was found between the items on these two factors (see Appendix H.2). Between the subscales there were significant correlations between Inconsistency scores and both Permissiveness (r=.45) and Overreactivity scores (r=.42), and a non-significant correlation between Permissiveness and Overreactivity scores (r=.24).

The PS was developed using a sample of US parents, and the differences between the original scales and those reported here could reflect the use of different parenting styles in the US and Australia. The different results could also be a result of the different sample types used to conduct the analyses (the US sample included mainly high SES families, whereas the current study included a larger, more representative

sample, n=1239). Other factor analyses using the PS items have recorded Laxness and Overreactivity factors but these have not been obtained consistently across studies, nor do they replicate the original scales (Harvey et al., 2001; Irvine et al., 1999; Reitman et al., 2001). The factor analysis of the PS scores of parents of adolescents by Irvine et al. (1999) reported Laxness and Overreactivity scales similar to the Inconsistency and Overreactivity factors found in the current study. Research including the Verbosity subscale of the PS has found equivocal results, even in the development of the measure (Arnold et al., 1993), and this scale has sometimes been dropped from subsequent research using the PS (Harvey-Arnold & O'Leary, 1995; Harvey-Arnold & O'Leary, 1997; Reitman et al., 2001).

Reliability of the Parenting Scale Factors Derived in the Present Study

Two measures of reliability are reported here for the PS factors derived from the current analysis: internal consistency and temporal stability. Internal consistency is used to assess the consistency of items in an instrument, and essentially examines how well items designed to reflect a single construct yield similar results (Trochim, 2002). In this study the internal consistency of the PS was calculated using Cronbach's alpha with respondents' data in the screening phase (N=1239). The values of alpha for the PS scales were: Inconsistency α=.81, Permissiveness α=.52, Overreactivity α=.68, Verbal Discipline α=.34 and PS Total Score α=.83. Because of the low internal consistency of the fifth factor (Verbal Discipline), this scale was not used in further analyses.

Temporal stability (also known as test-retest reliability) is a measure of the extent to which the same results are obtained on a measure administered to the same sample at two assessments over a relatively brief interval, such as a week or month (Trochim, 2002). The mean PS scores for participants (n=438) who participated in both the screening and second stages of the study are shown in Table 5.2. The test-

retest reliability of the PS was determined in two ways. First, parents' second stage PS scores were correlated with their screening scores obtained between one and three months earlier (Table 5.2). Quite high levels of temporal stability were found for scores on the Inconsistency, Overreactivity and Total PS scales, with a lower correlation obtained between the screening and stage two scores for the Permissiveness Scale.

Table 5.2. T-tests comparing Screening and Stage Two scores on the Parenting Scale (N=438)

	Screening Mean (SD)	**Stage Two Mean (SD)**	**r**	**t**
Inconsistency	2.72 (.86)	2.72 (.80)	.74***	.07 (.95)
Permissiveness	2.85 (.59)	2.95 (.60)	.61***	3.85 (<.001)
Overreactivity	3.11 (.71)	3.04 (.67)	.71***	-2.75 (<.01)
Total Score	**3.10 (.52)**	**3.09 (.53)**	**.78*** **	**-.46 (.65)**

*** p<.001

Although the test-retest correlations obtained in the current study are strong, this simply means that there was a strong linear component of association between participants' at the two assessments (e.g., participants who obtained higher scores at the initial assessment, were likely to have higher scores at the second assessment), but it does not necessarily mean that participants scores' were the same at the screening and stage two assessments. A more conservative measure of temporal stability is the paired samples t-test, which assesses if the difference between individual participants' scores at the two assessments is significantly different from zero. The results of these t-tests for the 438 parents who completed both the screening and stage two assessments are shown in Table 5.2. Parents reported

significantly higher levels of permissiveness, and lower levels of overreactivity, at the second assessment compared with at the screening stage.

These results present mixed evidence for the temporal stability of the PS, but might be explained by the fact that the re-administration of the PS occurred between one and three months after the screening phase. The use of the longer recall period means that parents were unlikely to simply be recalling their previous responses to the questionnaire, but over such a length of time it is possible that the differences in scores may reflect actual changes in parenting practices over time.

In summary, three PS subscales were described in this section: Inconsistency relates to “backing down” and failing to follow through with stated courses of discipline; Permissiveness refers to a failure to discipline or set limits on behaviour; and Overreactivity refers to harsh and punitive discipline practices. The factor analyses and reliability estimates reported in this section were performed with the largest published samples to date. The scales obtained in these analyses may differ from the three original subscales reported by Arnold et al. (1983), and from the Laxness and Overreactivity scales obtained in subsequent analyses (Irvine et al., 1999; Reitman et al., 2001), because of differences in sample size, sample type or because of cultural differences in parenting behaviour. Of the three new subscales reported here, the Inconsistency Scale demonstrated the highest levels of both internal consistency and temporal stability. Results using the Permissiveness Scale should be treated with caution as this factor had only moderate internal consistency and demonstrated moderate levels of temporal stability.

5.2.2 Factor Structure and Reliability of the Child-Rearing Practices Questionnaire

Factor Structure

Parents did not complete the CRPQ in the screening stage of the study, and so principal components analysis with varimax rotation was conducted on the CRPQ items using data from the entire sample who completed the second stage of the study (n=438). The results of this analysis are shown in Table 5.3. Four principal components were obtained which explained 45.3% of the variance in scale scores. Once again, only the items that had a factor loading of 0.40 or greater on any factor are shown in the table.

The principal components analysis yielded similar factors to those obtained in the original questionnaire development, plus a fourth factor relating to parents expectations of obedience from their children. As described in Chapter 3, similar results have been obtained in research with an Australian sample. Paterson and Sanson (1999) incorporated an “Obedience” factor in recent research using the CRPQ on the basis of the results from principal components analyses. As with the items of the PS, additional principal components analyses were performed using different types of rotations on the items of the CRPQ. Each of these analyses yielded the same factor structure as that shown in Table 5.3.

Table 5.3. Principle components analysis (with varimax rotation) of the Child-Rearing Practices Questionnaire self-report items (N=438)

	Factor 1	Factor 2	Factor 3	Factor 4
Warmth				
24. I like to hug and kiss my child	.79			
9. I often hug or hold my child for no particular reason	.73			
6. I express affection by hugging, kissing and holding my child	.69			
30. I joke and play with my child	.69			
17. My child and I have warm, intimate times together	.68			
19. I enjoy listening to my child and doing things with him/her	.61			
14. I tell my child how happy he/she makes me	.60			
3. I give my child comfort and understanding when he/she is scared or upset	.40			
21. I withdraw from my child when he/she displeases me[a]				
11. I prefer going places and doing things without my child[a]				
Inductive Reasoning/Power Assertion				
23. I explain to my child why he/she is being punished or restricted		.77		
20. I give my child reasons why rules should be obeyed		.76		
26. I emphasise the reasons for rules		.74		
10. I explain to my child the consequences of his/her behaviour		.74		
13. I try to explain to my child why certain things are necessary		.72		
29. I talk it over and reason with my child when he/she misbehaves		.61		
1. I expect my child to do what s/he is told without me having to give reasons[a]				.76
7. I expect my child to give his/her parents unquestioning obedience[a]				.76
4. I let my child express his/her feelings about being punished or restricted				
Punitiveness				
22. I use physical punishment, e.g., smacking, for very bad behaviour			.85	
28. I spank when my child is disobedient			.85	
25. I believe that physical punishment is the best way to discipline my child			.84	
2. I think smacking is a good way to make my child behave better			.83	
8. I slap or hit my child to control his/her behaviour			.74	
12. I yell at my child when disciplining him/her				
16. I expect my child to do what he/she is told to do, without stopping to argue about it				.78
No factor				
27. I feel close to my child both when he/she is happy and when he/she is worried	.66			
15. When disciplining my child I send him/her to his/her room with little if any explanation		-.40		
5. When disciplining my child I send him/her to his/her room for five minutes				
18. I take privileges away from my child when he/she misbehaves				

[a]These items are reverse scored on the original factors

Within the scales of the CRPQ, self-reports of Warmth were significantly correlated with self-reported Reasoning (r=.50) and Punitiveness (r=.29), but not with Obedience (r=.07). Self-reports of Reasoning were significantly correlated with Punitiveness (r=.26), as was the Obedience factor (r=.32), and there was a non-significant correlation between Reasoning and Obedience (r=.08).

Reliability of the Child-Rearing Practices Questionnaire Factors Derived in the Present Study

The internal consistency of the CRPQ was calculated using Cronbach's alpha with data from all of the respondents who completed the second stage of the study (N=438). The values of alpha for the CRPQ scales were: Warmth α=.86, Reasoning α=.63, Punitiveness α=.86, Obedience α=.69, and CRPQ Total Score α=.82. These alpha coefficients represent acceptable levels of internal consistency, and are very similar to those reported for slightly modified CRPQ scales in an Australian sample of 74 parents of 5-6-year-old-children (Warmth α=.81; Reasoning α=.60, Punitiveness α=.84 and Obedience α=.69) (Paterson & Sanson, 1999).

It was not possible to perform estimates of temporal stability using the CRPQ scales because parents completed this measure at only one assessment.

In summary, the factor analysis reported here using the self-reports of 438 parents on the CRPQ yielded very similar scales to those published using other Australian samples. Consistent factor structures were obtained across different types of factor analyses. Acceptable levels of internal consistency were found for all of the scales of the CRPQ, although as with previous research, lower levels of consistency were found for the Reasoning Scale.

5.3 Mean Scores for the Self-Report Parenting Questionnaires

5.3.1 Mean scores for self-reports on the Parenting Scale

Mean Scale Scores

The mean scores for participants on the PS are shown in Table 5.4. Parents reported using a range of parenting techniques, including "overreactive" discipline. Independent samples t-tests were used to compare the self-reported parenting practices of parents of boys with parents of girls (Huberty & Morris, 1989; Jaccard & Guilamo-Ramos, 2002a). Consistent with other studies using parent-reports on the PS (Collett et al., 2001; O'Leary et al., 1999), parents of boys did not report significantly different parenting practices on any of the PS scale scores when compared with parents of girls.

Table 5.4. Mean scale scores (SD) on the Parenting Scale

Scale	**All parents N=68**	**Parents of Boys N=32**	**Parents of Girls N=36**	**p-value[a]**
Inconsistency	2.65 (.72)	2.55 (.82)	2.74 (.63)	.28
Permissiveness	2.86 (.58)	2.81 (.66)	2.90 (.52)	.55
Overreactivity	2.99 (.60)	2.98 (.65)	3.00 (.57)	.85
Total Score	**3.02 (.48)**	**2.97 (.53)**	**3.06 (.44)**	**.48**

[a] p-value for t-tests comparing behaviour frequencies between parents of boys and parents of girls

The PS is a widely-used instrument and the mean PS total score of the current study can be compared with those in a variety of samples. The total PS score of the current sample was between the mean total score of samples of clinic and non-clinic preschoolers in the US (Arnold et al., 1993) and New Zealand (Keown & Woodward, 2002), but significantly closer to the scores of the clinic groups in both studies. Studies in Australia and the US which have included only parents of clinic children or

families with lower socioeconomic status have reported significantly higher mean scores on the Total PS scales than were obtained in the current sample (Bor et al., 2002; Harvey-Arnold & O'Leary, 1995; Harvey-Arnold, O'Leary, & Edwards, 1997; Zubrick et al., 2002).

Mean Item Scores

The mean scores for individual items on the PS are shown in Figure 5.1. Each item on the scale has a possible range of one to seven. Parents' responses to most of the items of the PS represent a broad range of possible parenting practices, including responses to items assessing less socially desirable parenting practices. For example, some parents reported the almost exclusive use of physical discipline ("*I spank, slap, grab or hit my child*") to manage children's misbehaviour and some parents reported always "*using bad language and cursing*", whereas other parents reported never or rarely using such techniques. A narrower range of responses was obtained on items that assessed lax parenting behaviours. The most frequently endorsed items were those related to "talking" parenting behaviours (e.g., "*I give my child several reminders", "I make my child tell me why he/she misbehaved", "I say a lot when disciplining my child"*). "*I raise my voice or yell"*, and "*picky and on my child's back"* were two overreactive discipline items that were quite frequently endorsed.

Overreactive subscale items such as "*I insult my child or call my child names", "using bad language or cursing"* and "*I hold a grudge"* were three of the least frequently endorsed discipline styles, as were the lax parenting items "*not reacting to misbehaviour straight away"* and "*coaxing or begging the child to stop misbehaving".*

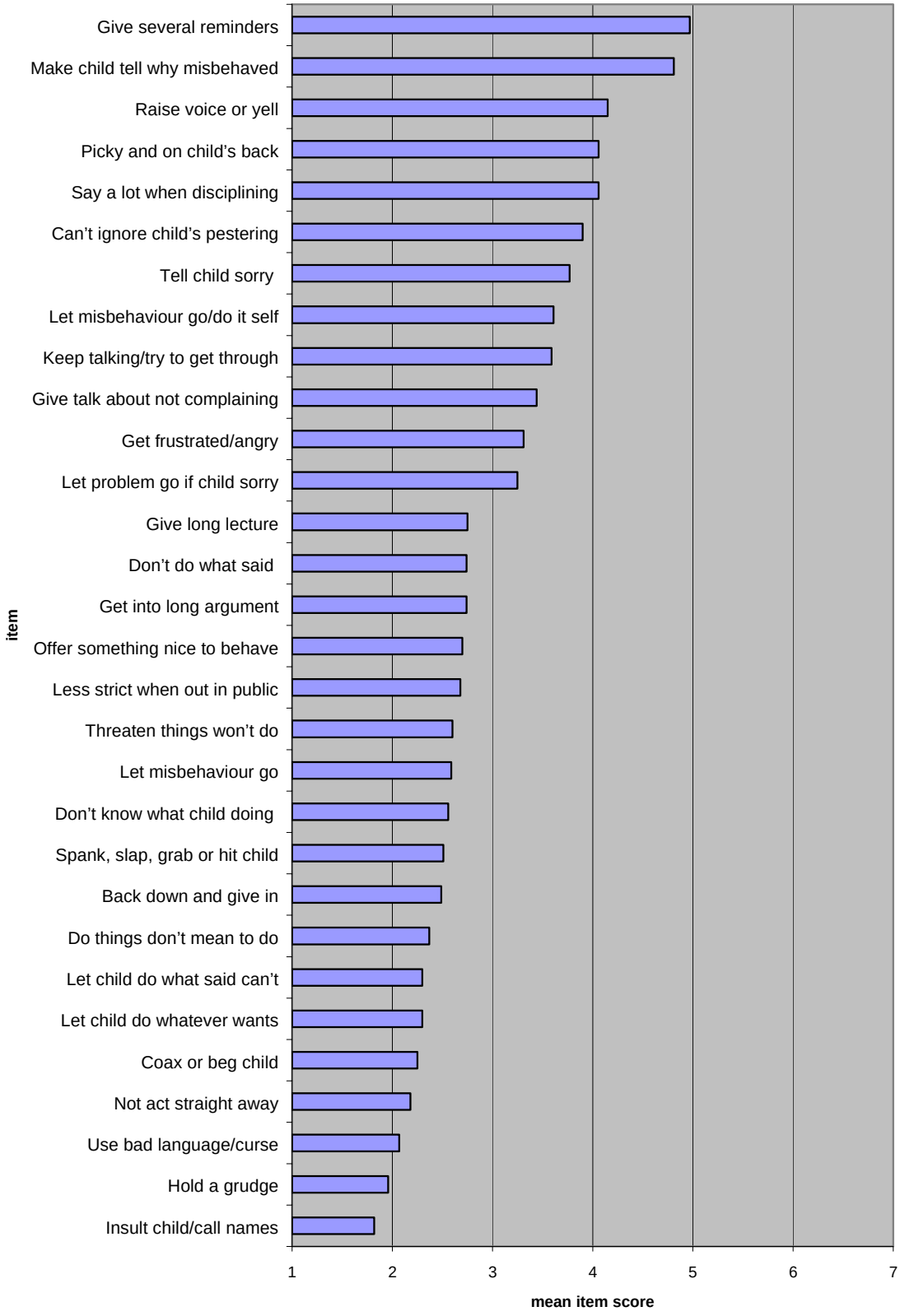

Figure 5.1 Mean item scores for self-reported items on the Parenting Scale

Table 5.5 displays the mean item scores for ten items which were used as part of a modified version of the PS in a sample of 183 parents of three- to four-year-old children enrolled in Head Start programs in the United States (Reitman et al., 2001). In that study, these items were retained in the modified version of the PS as a result of rigorous psychometric testing. The mean scores on these items obtained in the two studies were quite similar, however parents in the current sample had significantly higher mean scores on the item "*picky and on my child's back"* compared with parents in the Head Start sample (t=5.86, df=249, p<.001). This may be a reflection of differences in the perceived meaning of the term "picky and on my child's back" between parents in Australia and those in the US.

Table 5.5. Mean (SD) scores and ranges for items on the Parenting Scale used in both the current study and the study by Reitman et al., 2001.

Self-Reported PS Item	Current Study Mean (SD), N=68	Reitman et al. 2001 Mean (SD), N=183
Raise voice or yell	4.15 (1.21)	3.81 (1.96)
Picky and on child's back	4.06 (1.30)	2.70 (1.74)[a]
Let misbehaviour go/do it self	3.61 (1.34)	3.57 (1.99)
Get into long argument	2.74 (1.07)	3.03 (1.97)
Offer something nice to behave	2.70 (1.09)	3.07 (1.85)
Let misbehaviour go	2.59 (.97)	2.68 (1.72)
Back down and give in	2.49 (1.09)	2.47 (1.73)
Do things don't mean to do	2.37 (1.02)	2.57 (1.70)
Coax or beg child	2.25 (.87)	1.99 (1.46)
Hold a grudge	1.96 (1.23)	2.15 (1.46)

[a]=significant differences between the current sample and the indicated scores

In summary, parents in this study reported using a wide range of parenting practices to manage children's misbehaviour on the Parenting Scale. The most frequently endorsed parenting behaviours largely reflected talking practices and those least frequently endorsed reflected permissive and overreactive behaviours. The mean scores on the Total PS scales reported for this study reflected the nature of the sampling strategy used, in that they were typically intermediate between the scores of parents of clinic and non-clinic children. There was considerable similarity between the mean scores obtained in the current sample and those reported for a sample of Head Start children on ten items retained in a modified version of the PS.

5.3.2. Mean scores for self-reports on the Child-Rearing Practices Questionnaire

Mean Scale Scores

The mean scores for participants on the CRPQ are shown in Table 5.6. It should be noted that scores on the Punitiveness Scale are scored such that higher scores on the scale represent the use of lower levels of punitive behaviour. Parents obtained significantly lower mean scores on the Punitiveness subscale (indicating higher levels of punitive behaviour), than on the Reasoning and Warmth subscales. Further, parents in this sample reported mean levels of Warmth which suggested that they often or always display warm behaviours towards their children. Mean scores on the Reasoning subscale were intermediate between the Warmth and Punitiveness mean scores. As with the self-reports on the PS, no differences were reported in the parenting practices of parents of boys compared with parents of girls.

Table 5.6. Mean scale scores (SD) on the Child-Rearing Practices Questionnaire

Scale	All parents N=68	Parents of Boys N=32	Parents of Girls N=36	p-value[a]
Warmth	4.45 (.49)	4.42 (.54)	4.48 (.44)	.62
Reasoning	4.13 (.47)	4.19 (.49)	4.08 (.45)	.35
Punitiveness	3.78 (.62)	3.78 (.61)	3.77 (.64)	.96
Obedience	2.75 (.65)	2.72 (.73)	2.77 (.57)	.76
Total	**3.90 (.33)**	**3.91 (.37)**	**3.88 (.31)**	**.71**

[a] p-value for t-tests comparing behaviour frequencies between parents of boys and parents of girls

The mean scores on the Warmth, Punitiveness and Obedience scales in the current study were not significantly different to those obtained in a study which used a modified version of the CRPQ with a sample of parents of 74 five- to six-year-old children in Victoria (Paterson & Sanson, 1999). However, parents in the present study reported using higher levels of inductive reasoning than parents in the Victorian sample (4.13±.47 versus 3.87±.42, df=140, t=3.48, p<.001). This may reflect a difference in the nature of the samples in that the Victorian Study consisted of parents of school-aged children, whereas the present study included parents of preschoolers. There may be developmental differences in the use of reasoning for children of different ages. These differences may also be a result of the slight differences in the items included on the Reasoning scales in the two studies.

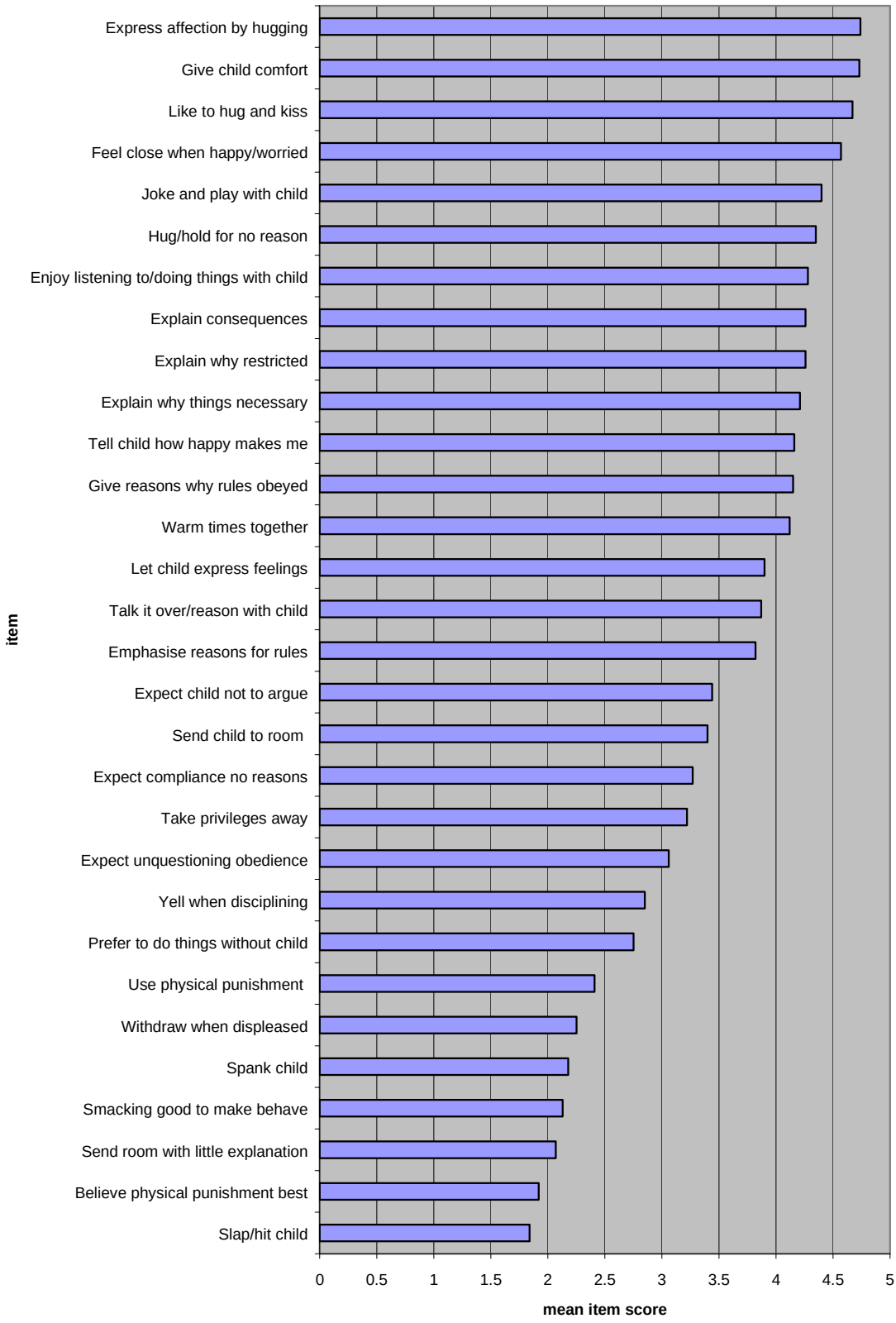

Figure 5.2 Mean item scores for self-reported items on the Child-Rearing Practices Questionnaire

Mean Item Scores

The mean scores and observed ranges for individual items on the CRPQ are shown in Figure 5.2. Each item on the scale has a possible range from one to five. The most frequently endorsed items on this scale were items which comprise the Warmth subscale (for example, "*I express affection by hugging"*, "*I give my child comfort when he/she is upset"*, "*I like to hug and kiss my child"*). These items also showed the narrowest observed range, in that no parents reported using these warm parenting techniques "never" or "rarely".

The least frequently endorsed items included items from the Punitive Scale (e.g., "*I slap or hit my child", "I believe physical punishment is the best way to discipline my child"*), and negatively scored items from the Inductive Reasoning and Warmth scales (e.g., "*I expect my child to do what he/she is told without asking questions", "I prefer to do things without my child, I withdraw from my child when I am displeased"*).

In summary, parents generally reported the use of high levels of warmth on the CRPQ. Mean scores on the subscales of the CRPQ were not significantly different from those reported for a sample of Victorian children on the domains of Warmth, Punitiveness and Obedience. Mean scores in this sample were significantly higher than the Victorian sample on the Reasoning subscale, which may be a result of sampling differences or scale calculation. Consistent with the mean scale scores, individual items which loaded on the Warmth Scale were more frequently endorsed than other items, and the use of punitive behaviours were reported with less frequency.

Chapter 6. Observed Parenting Behaviour

Parents and their preschool-age children were videotaped participating in four semi-structured situations: free play, drawing, pack-up and a no distraction task. Because different coding methods yield different types of information (Dowdney et al., 1984), parenting behaviour was coded in three ways in this study. First, global ratings of parents' behaviour (on a scale from one to seven) were made immediately after the observations on each parenting domain (e.g., Laxness, Overreactivity etc). These ratings represented the author's global impressions of parents' behaviour from the entire observation visit. Second, ratings of behaviour were made from the videotapes for parents in each activity and these were averaged across the four tasks. These ratings were based only on the videotaped behaviour observed during each activity. Third, individual parenting behaviours from the coding sheets in Appendix F.4 were interval-sampled every 30 seconds from the videotapes for each activity (e.g., free play, drawing etc). The occurrence of specific behaviours comprising each corresponding behaviour domain (e.g., Inconsistency, Permissiveness, Warmth) were then summed for the entire observation period. Because families were observed for different lengths of time depending on how quickly they completed the activities, the total behaviour count for each dyad was divided by the length of the observation, giving the average number of behaviours observed per 30-second interval. This enables standardised comparisons of behaviour between families. These relative frequencies of behaviour are less subjective because they do not incorporate observer judgements about the quality of the observed behaviour, and they allow the examination of parenting behaviour at a microanalytic level.

The data obtained using these three coding methods were very highly correlated and showed similar relationships with other variables, such as self-reports of parenting behaviour, children's behaviour and family risk factors. For two reasons, only the

results of analyses performed using the standardised behaviour counts rather than the behaviour ratings, are included in this and the following chapters. First, compared with ratings of behaviour, these standardised behaviour counts represent a lesser degree of abstraction from the observed behaviour (Brandt, 1992). Second, the *ratings* of parents' behaviour were made using the original PS and CRPQ factors, before the factor analyses for the PS and CRPQ scales were performed (see Appendices F.3 and F.4). The results of the factor analyses in the previous section showed that these original scales were not representative of the styles of parenting in the current sample. Thus, the standardised observed behaviour frequencies are used in these chapters because they could be summed in accordance with the modified PS and CRPQ behaviour scales, whereas the global ratings could not. In this chapter the validity and reliability of the observational data is examined and then the descriptive results are presented.

6.1 Validity of the Observed Parenting Behaviour

In this section, the procedure for the operationalisation of observed behaviours is described to enable the reader to understand how attempts were made to maximise the content validity of the definitions for the observed behaviours. Then the validity of the observational coding is examined in two ways. First, in the brief questionnaire completed during the no distraction task, parents were asked if the observed interaction was "normal". Responses to this question give an idea of how valid the observations are in presenting a sample of the dyad's "usual" behaviour. Second, the author composed narrative descriptions after each observation visit and recorded her perceptions of the representativeness of the observations.

6.1.1 Operationalisation of the observed behaviours

As described in Chapter 3, the behaviours coded during the viewing of the videotaped observations were operationalised directly from items on the PS and the

CRPQ. These descriptions of behaviour were originally developed and then, if necessary, they were modified in consultation with the author's supervisors and Professor Margot Prior from the University of Melbourne. The procedure for operationalisation of items was as follows:

1. Items which could not be observed in the context of the videotaping, or which related to parents emotions or feelings were excluded from the list of behaviours (e.g., "*when we are not at home, I let my child get away with a lot more*"; "*when disciplining my child I send him/her to his/her room for five minutes*"; "*I <u>feel</u> close to my child when he/she is happy or worried*");
2. The remaining items were operationalised verbatim if possible (e.g., "*I hug or kiss my child*");
3. Items which could not be operationalised verbatim or required interpretation, were described with stringent parameters (See Appendix F.5 for detailed operationalisations); and
4. Definitions for observed children's misbehaviour were developed because the parental behaviours that referred to discipline were reactive; that is, they occurred in response to children's misbehaviour (see Appendix F.5).

6.1.2. Parent reports of the representativeness of the observations

The first item of the questionnaire completed by parents during the no distraction task of the observation, asked "Was this a normal interaction with your child today?" (Appendix F.1). Just over half of the parents (54%) reported that this was a normal interaction, 35% said that it was not, and 10% said "yes and no". Parents gave 57 comments about how the interactions might have been different from a "normal" interaction. Reasons given included: being engaged in different activities than normal (28 comments); other children usually being present (11 comments); the children being more cooperative than usual (9 comments); and parents not usually being able

to spend this much time one-on-one with their children (8 comments). A small proportion of parents (5 comments) also reported that their parenting would have been different (e.g., less forced, or they would have tolerated less misbehaviour). Some verbatim examples of comments from parents for these questions are given in Appendix I.1. These parental perceptions of the representativeness of the interactions are examined again in Chapter 7, when the relationships between self-reported and observed behaviour are reported.

6.1.3 Author's perceptions of the representativeness of the observations

The information described in this section was collected from the author's narrative descriptions of the observations which were written immediately after each home visit (Banister et al., 1996; Brandt, 1992; Renne et al., 1983). Some examples of these summaries are shown in Appendix I.2. Because the information recorded was not standardised across participants, it is only an anecdotal summary of the perceived representativeness of the observations.

On the whole, parents' and children's behaviour during the observation was seen by the author as representative of the behaviour seen outside the recorded part of the visit. Only seldom did parental behaviour seemed constrained or forced whilst being videotaped, but this reactivity dissipated as the observation progressed. In a couple of instances, parents seemed more warm and interested in what their children were doing whilst being filmed, than at other times. The presence of other children in the room before or after the videotaping enabled the observer to see how parents interacted differently with different children. The presence of other children often changed both the behaviour of parents and that of index children. Typically, parenting seemed to be more stressful and demanding at these times, and the behaviour of index children generally was more disruptive and noncompliant. Although, on one or

two occasions the presence of a sibling was associated with a decrease in children's misbehaviour because it provided a distraction or relief from boredom for children.

Children who displayed externalising behaviours during the observation generally also displayed these behaviours outside the videotaped session. In some instances, children who were relatively well-behaved during the taping would demonstrate externalising behaviour before or after the taping, and the reverse was also true. In some cases, the children demonstrated no reactivity to the videotaping because they did not realise that their behaviour was being recorded (as evidenced by comments such as "when is she going to take a picture of me" whilst the video was running). However, on a couple of occasions the behaviour of the children altered slightly when they noticed the video camera following their movement about the room. These children often reacted by putting on a performance for the camera, but it was possible to decrease this reactivity, by closing the viewing screen of the video-camera. The novelty of having their parents' undivided attention for almost 40 minutes meant that some children were more compliant (e.g., packing up quickly, sitting still in the no distraction task) during the observation than they were in the non-recorded part of the observation visit, when more "normal" interactions were resumed.

Although, in general, the interactions between parents and children were perceived by the parents and the author as representative of the dyad's behaviour outside the recorded period, some discrepancies were noted. It is possible that the videotapes contain fewer instances of children's misbehaviour (and hence parental reactive discipline) than would "naturally" occur because of the novelty of the situation, the presence of the observer, the absence of other family members, and the amount of parental attention children received.

6.2 Inter-Observer Reliability of the Observed Parenting Behaviour

Inter-observer reliability (IOR) is a measurement of the extent to which two coders agree upon the presence or absence of observed behaviours (Maxwell & Pringle, 1983). In this study, approximately thirty percent (N=20) of the tapes were coded by both the author and a second coder to measure the reliability of ratings for the behavioural categories. Both coders were blind to participants' scores on any of the screening and baseline questionnaires (including the parenting measures) to avoid potential biases in the coding of behaviour.

6.2.1 Training

The procedure for training the second coder was that specified by Castorr et al. (1990). A videotape with examples of the behaviours was provided for training. The behaviour examples were selected by the author as representative of the types of behaviours being coded and were taken from the observations that were not being used for IOR assessments. Three pilot observations were used as practice coding tapes and were coded until 90% agreement on the three tapes was obtained (Fassnacht, 1982). The entire training process took approximately thirty hours.

6.2.2. Procedure

The second coder was an Honours Psychology graduate working as a research assistant in the Research and Evaluation Unit, Women's and Children's Hospital, South Australia. Video tapes were coded by both coders in the manner described in section 3.3.2. The second coder was provided with a coding protocol that summarised the coding procedures and could be referred to throughout the inter-coding process (Appendix I.3). Information about interruptions and off-screen activities during the observations were given to the second coder to make her aware of what was occurring during off-screen gaps in the observation, and she coded a

maximum of four tapes per day to avoid boredom and fatigue (Maxwell & Pringle, 1983).

6.2.3. Inter-Observer Reliability

Observational studies have calculated IOR in a number of ways (Castorr et al., 1990), and this study has used two of the most commonly reported methods: correlations and percentage agreements (see Table 6.1). These two techniques were used because they examine different aspects of the data. Because of the non-normal distributions of the behaviour frequencies in each parenting domain, rank-order correlations were used to examine if there was an association between the total behaviour frequencies recorded by the two coders. That is, were the rank orders of participants similar across the two coders? Percentage agreements were calculated to examine the precise level of agreement between the two coders about the presence of specific parenting behaviours (Castorr et al., 1990; Uebersax, 1987). The equation for this calculation is:

$$\text{Level of agreement} = \frac{\text{number of agreements about the presence of behaviour}}{\text{number of agreements + disagreements}}$$

This calculation is a more conservative estimate than those which also include agreements about the absence of behaviour (Dadds et al., 1987; Kavanagh et al., 1988), which can result in inflated percentage agreements for behaviours which occur relatively infrequently (such as harsh and punitive parenting) (Uebersax, 1987).

Table 6.1. Rank-order correlations and percentage agreements between behaviour counts of the two coders (N=20)

Observed Parenting	**Correlations**	**% agreement**

Domain		(95% CI)
PS		
Inconsistency (5 items)	.77**	76.7 (66.7-86.7)
Permissiveness (6 items)	.81***	74.8 (64.5-85.1)
Overreactivity (6 items)	.89***	72.4 (61.8-83.0)
PS Total Score[a]	**.96*** **	**75.9 (65.7-86.1)**
CRPQ		
Warmth (6 items)	.61**	66.1 (54.8-77.4)
Reasoning (6 items)	.57**	64.6 (53.2-76.0)
Punitiveness (5 items)	.85***	69.0 (58.0-80.0)
Obedience (2 items)	.61**	69.7 (58.8-80.6)
CRPQ Total Score[a]	**.67*** **	**66.1 (54.8-77.4)**

*=p<.05, **=p<.01, ***=p<.001. [a]Includes all of the behaviours derived from the PS or CRPQ questionnaires

The two methods returned similar results with higher reliability estimates for the domains of negative and ineffective discipline (i.e., Inconsistency, Permissiveness, Overreactivity and Punitiveness), and lower estimates for the positive parenting domains measured by the CRPQ (Warmth and Reasoning). Examination of the data showed that the discrepancies were the result of one observer not recording as many instances of positive behaviour, rather than the two coders coding behaviours differently. Harsh and punitive behaviours are thought to be more reliably observed because their relatively low frequency and negative nature means they are more salient to the observer and are less likely to be missed in coding (Fassnacht, 1982; Shaw et al., 1998). Conversely, the responsive parenting practices coded by the two observers occurred with higher relative frequencies than most of the negative parenting behaviours (with the exception of parental laxness) and thus provided more opportunities for disagreements between the coders (Fassnacht, 1982). Warm and

positive behaviours may also require more subjective interpretation on the part of the coders than do ineffective discipline practices. That is, the coders could have different interpretations of what constitutes “warmth”, “enjoyment” and “intimacy” in an interaction, whereas “spanking”, “insulting” and “failure to discipline” were more stringently operationalised and required less interpretation. Other studies that have used micro-analytic coding methods to measure more subjective parenting behaviours such as unresponsiveness, displeasure, warmth and reasoning have reported lower inter-observer reliability rates for these parenting domains compared with more objective or salient parenting techniques such as physical punishment and negative control (Brophy & Dunn, 2002; Gottman & Katz, 1989; Lytton, 1973; Shaw et al., 2001; Shaw et al., 1998).

Consistent inter-observer reliability estimates were obtained for the domains of parenting behaviour using correlations and percentage agreements. Parenting practices that reflected inconsistency, permissiveness, overreactivity and punitiveness were coded more reliably than were practices that involved parental warmth and reasoning. This may be a reflection of the salience of these behaviours or of the level of subjectivity involved in the interpretation of the different types of behaviour. These reliability estimates will be discussed again when the results using observational data are considered in the following chapters.

6.3 Mean Scores for the Observed Parenting Behaviours

6.3.1 Mean observed scale frequencies

The mean(SD) standardised frequencies for the observed PS and CRPQ scales for the entire observation period are shown in Table 6.2. Both frequencies for the sample as a whole and separate frequencies for parents of boys and girls are displayed. On average, parents displayed significantly higher levels of warmth than any other behaviour, followed by permissive parenting behaviours. Behaviours that made up the Punitiveness, Inconsistency and Obedience scales were observed with much lower frequencies.

No significant differences were found when the frequencies of Warmth and Reasoning were compared between parents of boys and girls (Table 6.2). However, parents of boys when compared with parents of girls demonstrated higher mean levels of Laxness, Overreactivity, Punitiveness and Obedience. These gender differences were not found when self-reported parenting on the corresponding domains of the PS and CRPQ were examined (Chapter 5), nor in other studies which have used parent self-reports on the PS (Collett et al., 2001; O'Leary et al., 1999).

Table 6.2 Mean frequency of behaviours per 30 second interval

Behaviour Domain	All parents N=68	Parents of Boys N=32	Parents of Girls N=36	p-value[a]
PS				
Inconsistency	.03 (.05)	.05 (.07)	.02 (.04)	.11
Permissiveness	.38 (.29)	.49 (.31)	.29 (.23)	<.01
Overreactivity	.15 (.14)	.19 (.16)	.11 (.11)	.01
PS Total Score	**.69 (.44)**	**.88 (.47)**	**.51 (.33)**	**<.001**
CRPQ				
Warmth	.78 (.31)	.72 (.32)	.83 (.31)	.17
Reasoning	.21 (.12)	.23 (.13)	.19 (.11)	.18
Punitiveness	.02 (.03)	.03 (.04)	.01 (.01)	<.01
Obedience	.07 (.06)	.09 (.07)	.05 (.05)	<.01
CRPQ Total Score	**.95 (.37)**	**.88 (.39)**	**1.00 (.36)**	**.18**

[a] p-value for t-tests comparing behaviour frequencies between parents of boys and parents of girls

In this study, gender differences in observed parenting behaviour may reflect different levels of observed children's externalising behaviours during the observations (as rated by the author) (Harvey-Arnold & O'Leary, 1997). Boys misbehaved with greater frequency during the observed interactions compared with girls (boys: 3.63±1.58; girls: 2.33±1.37, t=3.61, df=67, p=.001), and thus interactions with boys provided more opportunities to observe parental discipline. However, when this possibility was examined using general linear modelling to predict frequencies of observed parenting behaviour, interactions between child gender and their levels of externalising behaviour during the observations were not significant for any analyses. Therefore, these observations may have revealed true differences in the parenting practices of parents of boys compared with those of parents of girls, which are not related to levels of children's misbehaviour. Parents may be more likely to let boys'

misbehaviour go undisciplined or to respond with harsh discipline than they are for girls' misbehaviour.

Differences across the four activities

As described in Chapter Three, the different demand characteristics of the four observation activities were expected to yield different frequencies of parenting behaviours (i.e., higher levels of observed ineffective parental discipline and lower levels of responsive parenting during the pack-up and no distraction tasks, and the reverse for the free play and drawing tasks) (Brophy & Dunn, 2002; Kavanagh et al., 1988; Metsapelto et al., 2001). To determine if this was the case, the mean PS and CRPQ behaviour counts across the four observation activities were examined and are displayed in Figure 6.1 (the results for individual participants are shown in Appendix I.4). Repeated measures analyses of variance were performed (Huberty & Morris, 1989; Jaccard & Guilamo-Ramos, 2002a), and significant within- and between-subjects effects were found for all of these domains of parenting behaviour. This indicates that for both individual participants and the group as a whole the mean observed behaviour frequencies differed between the activities. Post-hoc comparisons were performed to identify activities for which participants demonstrated different behavioural frequencies. Significantly higher levels of Warmth were demonstrated per 30-second interval in the free play and drawing tasks compared with the final two tasks, with the lowest levels of Warmth in the no distraction task. This latter result is not surprising given that little opportunity for displays of Warmth were provided during the no distraction task because parents and children were engaged in separate tasks (Cunningham & Boyle, 2002; Johnston et al., 2002; Summerfield, 1983).

Participants were also observed to demonstrate higher levels of Inconsistency, Overreactivity, Punitiveness and Total Behaviours per interval in the pack-up task than in the free-play and drawing tasks. However, individual participants did not display significantly higher levels of Overreactivity or Punitiveness in the no distraction task compared with the free play and drawing tasks. These unexpected findings, may reflect the nature of the no distraction activity. Although this activity was one in which children were expected to misbehave (due to the boredom of playing with a set of blocks suitable for children of a younger developmental age), it was the only activity in which parents' attention was not directed towards their children (parents were completing a questionnaire). Because parents' concentration was directed elsewhere, they were more likely to fail to notice if their children had not complied or had moved off-task (display permissiveness), rather than to display reactive discipline (e.g., overreactivity, punitiveness and inconsistency). This was reflected in the higher levels of observed Permissiveness in the no distraction task than in either the free play or drawing tasks, and the higher frequencies of Permissiveness compared with Overreactivity, Punitiveness and Inconsistency in this final task.

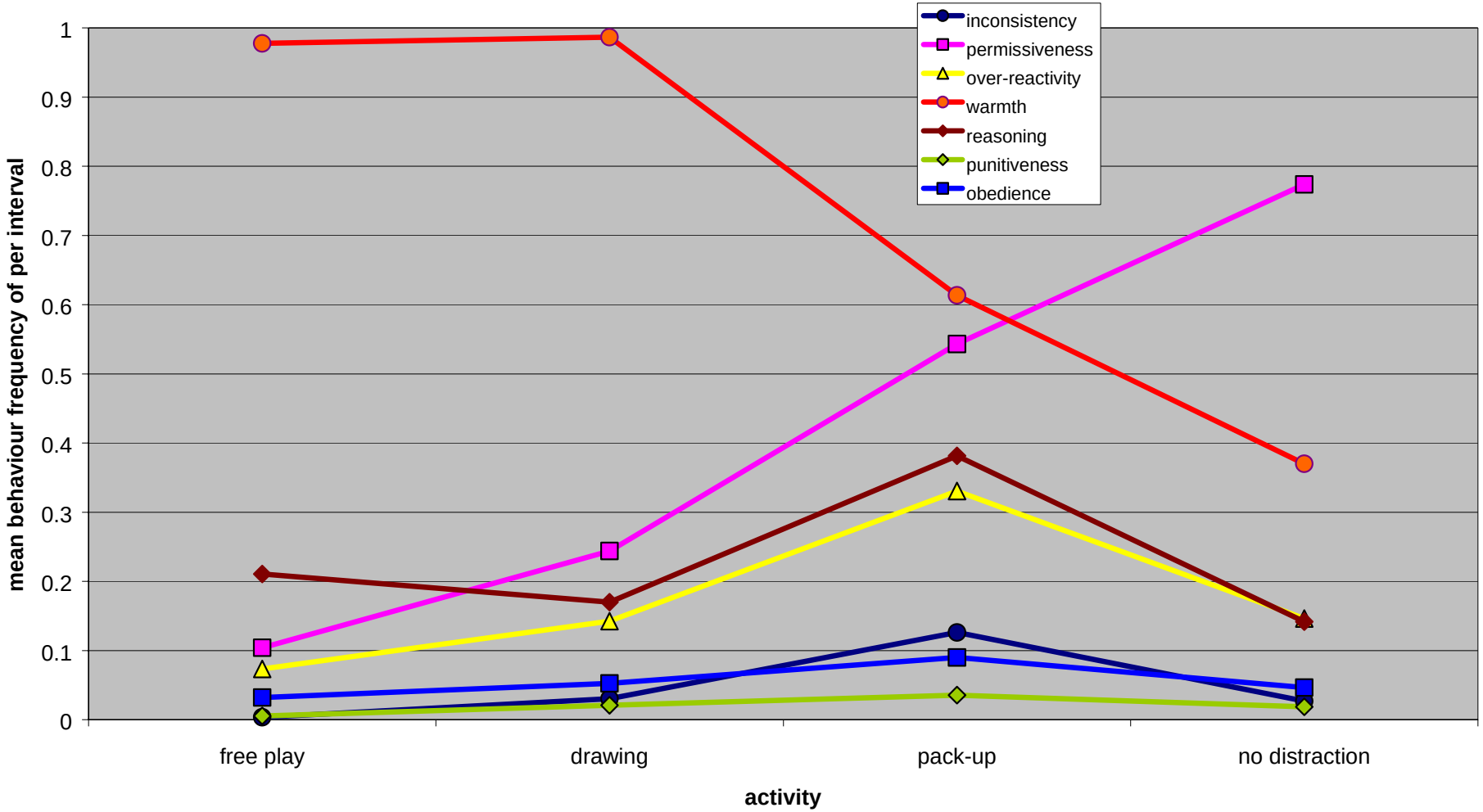

Figure 6.1 Standardised frequencies of parenting behaviour per 30 second interval in the four activities

In summary, Warmth and Permissiveness were observed more frequently than were other domains of behaviour. Parents of boys had higher mean frequencies of ineffective parental discipline styles compared with parents of girls, which in part reflects the higher levels of misbehaviour displayed by boys during the observation. No differences in the frequencies of observed Warmth and Reasoning were found between parents of boys and girls, but these are generally spontaneous rather than reactive parenting practices and hence do not rely upon the frequency of children's misbehaviour to be observed.

The demand characteristics of the four activities served to elicit different levels of specific parenting behaviours. The pack-up task served to elicit higher levels of reactive parental discipline, whereas the cooperative tasks such as free play and drawing elicited higher levels of Warmth. The no distraction task did not facilitate the observation of harsh and punitive discipline, but was useful for observing permissive practices. This failure to notice or react to misbehaviour reflects the parents' focus on a different task to that of their children during this final activity.

6.3.2 Mean observed item frequencies

Parenting Scale

To examine which behaviours were most and least frequently observed, the mean scores and ranges of observed behaviours which correspond to individual items on the PS are presented in Table 6.3.

Table 6.3. Mean number of 30 second intervals (SD) in which individual behaviours corresponding to the Parenting Scale were observed

Observed PS Behaviour	**Mean (SD)**	**Range**
Let child do whatever he/she wants	12.76 (10.53)	0-40
Not react to misbehaviour immediately	9.71 (8.60)	0-42
Keep talking when child noncompliant	5.32 (5.03)	0-25
Misbehaviour go unpunished	3.93 (4.10)	0-15
Can't ignore child's pestering	3.53 (3.03)	0-11
Argues with child	1.72 (1.88)	0-8
Give several reminders/warnings	1.50 (2.43)	0-11
Raise voice or yell	1.46 (2.39)	0-13
Picky and on child's back	1.38 (2.82)	0-15
End up doing what asked child to do	1.26 (3.01)	0-19
Say a lot when making a request	1.03 (1.48)	0-7
Frustrated or angry	.91 (2.17)	0-14
Insult child or says mean things	.79 (1.31)	0-5
Not know what child doing if out of sight	.75 (1.96)	0-11
Offer reward to obtain compliance	.53 (.85)	0-3
Spank, slap, grab or hit child	.49 (1.00)	0-5
Make child explain misbehaviour	.43 (.80)	0-4
Let child do something when said no	.40 (.96)	0-4
Back down/give in if child upset at no	.18 (.57)	0-3
Coax or beg child to stop	.18 (.54)	0-3
Give long lecture when misbehave	.13 (.75)	0-6
Threats not carried out	.13 (.49)	0-3
Hold a grudge	.04 (.21)	0-1
Give talk about not complaining	.03 (.24)	0-2
Don't discipline if child sorry	.01 (.12)	0-1
Use of bad language	.00 (.00)	0-0
Do things don't mean to do	.00 (.00)	0-0
Apologise to child for discipline	.00 (.00)	0-0

There were large differences in the frequencies of specific behaviours. Reflecting the mean behaviour scores reported in the previous section, the PS behaviours which were observed most frequently were largely those associated with permissive discipline items (such as "*letting the child do whatever he/she wants"*, "*not reacting to misbehaviour immediately"*, and "*letting misbehaviour go unpunished")*. As described earlier, the "no distraction" task provided an opportunity for observing behaviours from the Permissiveness Scale of the PS. Other frequently observed PS behaviours included items originally from the Verbosity Scale such as "*keep talking trying to get through to the child"*, "*get into an argument"* with the child and "*giving the child several reminders or warnings"*.

The observational techniques used in this study did not permit the observation of certain types of parenting behaviours. Some behaviours which were not observed during any of the 68 videotaped observations included "*using bad language or cursing"*, "*doing things did not mean to do"* and "*apologising to their children for discipline"*. Also, "*letting a problem go if their child was sorry"*, "*giving their child a talk about not complaining"*, and "*holding a grudge against their child"*, were very infrequently observed parenting practices. These behaviours were less frequently observed for a number of reasons. First, some of these behaviours are those which occur with a naturally low frequency (Bates & Bayles, 1984; Bates et al., 1998; Summerfield, 1983). This is supported by comparing these results with those in Chapter Five in which parents endorsed these behaviours on the PS questionnaire with a lower frequency. Second, the context and length of the observation may not have provided an opportunity to observe some of these behaviours, in particular those practices which occur in reaction to specific children's behaviours such as complaining or appearing sorry after misbehaviour (Bates et al., 1998). Finally, some of these less frequently observed items, such as using bad language or holding a

grudge, may have been less frequently observed because they are less socially desirable (Bates et al., 1998; Renne et al., 1983; Janssens, 2004).

Child-Rearing Practices Questionnaire

The mean scores and observed ranges for observed behaviours that correspond to individual items on the CRPQ are presented in Table 6.4. As with the PS behaviours, the frequencies of these behaviours ranged widely both within and between families. The observational system used in this study yielded more frequent observations of behaviours which loaded on the instrument's Warmth Scale such as: "*Enjoys listening to and doing things with the child"*, "*warmth and intimacy",* and "*jokes and plays with the child"*. On average, parents displayed behaviours representing enjoyment in interactions with the child more than three times as frequently as any other CRPQ behaviour. Behaviours from the Reasoning Scale were also amongst the more frequently observed behaviours (e.g., "*explaining why things are necessary"* and "*talking things over and reasoning with the child"*).

As with the PS, some of the operationalised behaviours were not observed in any of the 68 observations, or were observed only once or twice. These behaviours represented all three of the parenting domains assessed and included behaviours such as "*smacking, slapping* or *spanking their child*", "*telling their child how happy he/she makes them*", "*emphasising reasons for rules or punishment*" and "*hugging and kissing their child*".

Table 6.4. Mean number of 30 second intervals (SD) in which individual behaviours corresponding to the Child-Rearing Practices Questionnaire were observed

Observed CRPQ Behaviour	Mean (SD)	Range
Enjoy listening to/doing things with child	37.32 (14.43)	3-71
Have warmth and intimacy	11.59 (10.15)	0-44
Explain why things necessary	8.18 (5.41)	0-23
Joke and play with child	5.46 (4.82)	0-20
Talk it over/reason with child	4.87 (4.28)	0-16
Expect child to do what told immediately	3.18 (3.84)	0-17
Prefer to do things without child	1.90 (5.25)	0-30
Explain consequences of behaviour	1.90 (2.34)	0-30
Let child talk about punishment	1.38 (1.75)	0-7
Hugs or holds child	1.03 (1.81)	0-8
Give child comfort and empathy	.97 (1.63)	0-8
Yell at child	.68 (1.62)	0-9
Other forms of physical punishment	.44 (.94)	0-4
Expect child to do what told/no questions	.19 (.55)	0-3
Give reasons why rules to be followed	.13 (.45)	0-3
Withdraw from child when displeased	.12 (.53)	0-3
Hug and kiss child	.07 (.43)	0-3
Explain why punished/restricted	.07 (.61)	0-5
Emphasise reasons for rules	.04 (.27)	0-2
Slap or hit child	.03 (.17)	0-1
Smack child	.01 (.12)	0-1
Tell child how happy makes me	.00 (.00)	0-0
Spank child	.00 (.00)	0-0

As with the observed items on the PS, some of these behaviours, such as those relating to physical discipline practices, may have been infrequently observed because they occur with lower natural frequencies or because they are not socially desirable (Bates et al., 1998). However, in these observations some of the infrequently *observed* behaviours (e.g., "*hugging and kissing the child"* and "*telling the child how happy he/she made the parent"*) were actually some of the most frequently endorsed self-report items (see Chapter Five). In fact, parents self-reported that, on average, they told their children how happy they made them "often", however this behaviour was not observed during any of the videotaped observations.

These discrepancies may have occurred for a number of reasons. First, the set tasks in the observations meant that parents and children were concentrating on the activities at hand and may have shown different behaviours to those which would normally be seen in an unstructured observation (Mrazek et al., 1982). Further, the presence of an observer may have created participant reactivity and hindered the expression of affection between parents and children, because of shyness or nervousness in front of a stranger (Banister et al., 1996; Brandt, 1992; Kavanagh et al., 1988). It could also be that parents are responding to the self-report items about physical affection in a socially desirable manner, and their reports are not supported by the observed frequencies of such behaviour (Rothbaum, 1986).

In summary, particular parenting behaviours adapted from the Parenting Scale and the Child Rearing Practices Questionnaire, such as permissive and warm practices, were observed with higher frequencies whilst others, such as physical discipline, were observed seldom or not at all. It is suggested that this observational coding system is better for observing behaviours which: a) occur with greater natural frequency; b) are not susceptible to social desirability biases or participant reactivity; c) do not depend upon the occurrence of specific infrequently occurring child behaviours (such as crying or complaining) to be observed; and d) are likely to be observed in the context of the semi-structured activities.

Chapter 7. Level of Agreement Between Self-Reported and Observed Parenting Behaviour

The literature examining direct comparisons between self-reports and observations of parenting behaviour was examined in Chapter 2. Generally, studies which have directly compared the two methods in parents of preschoolers have found modest to high levels of agreement (Arnold et al., 1993; Denham et al., 2000; Dowdney et al., 1984; Strayhorn & Weidman, 1988). But the generalisability of the results from these studies is hampered by methodological considerations, such as the limited sample size (Arnold et al., 1993), the generality of the behavioural definitions (Dowdney et al., 1984; Strayhorn & Weidman, 1988), and the fact that all of the studies used only global ratings of observed parenting behaviour (Arnold et al., 1993; Denham et al., 2000; Dowdney et al., 1984; Strayhorn & Weidman, 1988). Studies of parents of older children which have used larger samples and composite ratings or behaviour counts have found more modest levels of agreement between the two methods (Feinberg et al., 2001; Johnston et al., 2002).

This chapter presents the results of data analyses which examine the first study aim: to assess the degree of agreement between self-report questionnaire and direct observation measures of parental responsivity and control. Specifically, it explores the agreement between the self-reported and observed parenting behaviours in several domains of parenting and for individual behaviours. These relationships were not investigated at the level of total parenting behaviour scores on the PS and CRPQ, because for the observations these total scores were mostly comprised of behaviours on the Permissiveness Scale for the PS, and the Warmth Scale for the CRPQ. The distributions of the variables used in the analyses for Chapters 7 and 8 are shown in Appendix J.1. Where variables were not normally distributed, appropriate transformations were made, and these transformed variables are used in subsequent

analyses. Where an appropriate transformation could not be identified (observation totals for the punitiveness scale and scores on the General Health Questionnaire), the variables were dichotomised for use in further analysis.

7.1 Correlations Between Self-reported and Observed Parenting Domains

The levels of agreement between self-reported and observed domains of parenting behaviour were examined using correlations (Table 7.1). Correlations determine the strength and direction of a linear relationship between two variables (Pallant, 2004). That is, a correlation can assess the extent to which parents who reported higher levels of parenting behaviours were also observed to use these techniques more frequently and to which parents who reported using the techniques less often were also observed to do so infrequently. In this way, it is not possible to explore non-linear (e.g. curvilinear) relationships between the variables using correlations. However, the relationships between the results obtained from the two methods were not expected to be non-linear, and examination of the scatterplots between the observed and self-reported domains of behaviour (Appendix J.2) did not support the use of a non-linear approach.

Table 7.1 shows that the only significant correlation was between transformed self-reported Warmth scores and observed frequencies of Warmth ($r=.36$, $p<.01$). The size of this association is similar to correlations obtained in other studies measuring the relationship between observations and self-reports of responsive parenting practices (Denham et al., 2000; Feinberg et al., 2001; Strayhorn & Weidman, 1988). One exception is the study by Johnston et al. (2002) which reported no association between self-report scores of warmth, involvement and positive parenting and composite ratings of observed responsiveness, however in this study the analyses

controlled for maternal age, education and marital status which would have influenced the size of the correlation.

Table 7.1 Correlations between corresponding behaviour domains on the self-report and observed measures

Parenting Domain	**All parents N=68**
PS	
Inconsistency	.01
Permissiveness	.05
Overreactivity	.11
CRPQ	
Warmth	.36**
Reasoning	.12
Obedience	.04

** = p <.01

The correlations for the other domains ranged from r=.01 to r=.12, indicating almost no linear relationship between self-reported and observed parenting behaviour for these particular practices (Table 7.1). In addition, the observed Punitiveness scores were dichotomised into parents who displayed punitiveness during the observation and those who did not. The self-reported scores on the Punitiveness scale were not significantly different for these two groups of parents. The relationships between observations and self-reports on corresponding parenting domains were not significantly different for parents of boys compared with parents of girls. For the entire sample, the relationships between observations and self-reports of parental discipline (i.e., Inconsistency, Permissiveness, Overreactivity and Punitiveness) obtained in this

study do not mirror the significant associations reported in other studies (Feinberg et al., 2001; O'Leary et al., 1999). Specifically, previous work using the PS has reported quite strong correlations between self-reported and observed parenting behaviour (Arnold et al., 1993). In the original scale development, the Spearman's rank order correlations between corresponding self-report and observational scales were Laxness = .61; and Overreactivity = .65 (Arnold et al., 1993). Although, the study by Arnold et al. differs from the current study in both sample size and type (Arnold et al. used a sample of only 15 parents with relatively high socioeconomic status), several other factors might explain the differences between the two studies. Five factors that might explain the lack of correlation between self-reported and observed parental discipline in this study are examined in the next section.

7.2 Factors which might Influence the Relationships Between Self-reported and Observed Parenting Domains

7.2.1. Use of behaviour counts versus ratings

The results in Table 7.1 used interval sampled behaviour counts as the method for coding the observational data. Other studies which have directly compared the degree of concurrence between self-reported and observational data have used global ratings of parenting behaviour or style (Arnold et al., 1993; Denham et al., 2000; Dowdney et al., 1984; Feinberg et al., 2001; Rothbaum, 1986; Strayhorn & Weidman, 1988). This study was also unique in that the behaviour counts obtained were directly operationalised from the parenting self-report measures, rather than being a global rating of parenting style. It may be that relationships between observed and self-reported parenting behaviour that are evident when a global definition of parenting is used are not evident at a more micro-analytic level.

Global ratings of parenting behaviour were made in the current study, immediately after the observations were completed, but as described in Chapter Six these ratings were based on the original scale definitions and therefore could not be compared directly with the modified self-report scales presented in this section. However, Table 7.2 shows the relationships between these global ratings when the self-report scores for participants were recalculated using the original scale scores. Because of the non-normal distribution of the ratings of behaviour, rank-order correlations were performed. Once again, the correlations between self-reports and observations of the ineffective parenting domains remained low and non-significant, whereas that between self-reports and observations of parental Warmth were significant. One difference with these results was that the correlation for parental Reasoning which was significant at the $p<.05$ level (rho=.25). The reliability of the microanalytic coding

(interval sampling) on the Reasoning subscale was the lowest of the seven scales. This might explain the differences in the results obtained using behaviour counts versus behaviour ratings.

Table 7.2 Rank-order correlations between global ratings and self-reports of parenting behaviour on the original PS domains

	rho
PS	
Laxness	.05
Overreactivity	.02
Verbosity	.06
CRPQ	
Warmth	.26*
Punitiveness	-.09
Reasoning	.25*

* = p <.05

7.2.2. Differences between the observations and self-reports of parental discipline in their dependence on the level of children's misbehaviour

Parental responses to the items of the PS questionnaire are not dependent on the frequency of children's misbehaviour, because parents indicate their typical responses to misbehaviour, rather than how frequently they use a particular technique (Arnold et al., 1993). Conversely, in the observations of parenting behaviour, the frequencies with which parents use discipline techniques are dependent on how often children misbehaved during recorded periods. Parents who might typically use more harsh or punitive techniques, may be recorded during the observation as using none of these behaviours because their children have not

misbehaved and therefore have provided no opportunity to observe such parental behaviour (Janssens, 2004). This frequency dependence of the observed behaviours could be seen when the levels of PS behaviours were compared for children who were rated by the author as demonstrating different levels of misbehaviour during the observations (Appendix J.3). Parents of children who were rated as displaying higher levels of externalising behaviour demonstrated significantly greater frequencies of ineffective parental discipline behaviours such as permissiveness and overreactivity.

Self-reports of parental discipline (but not warmth) on the CRPQ are also frequency dependent because the response scale is one of frequency (parents respond on a five-point scale from “never” to “always”). Thus, the frequency with which children misbehave influences the frequency with which parental discipline will be reported. Given the self-reports and observations of the CRPQ *discipline* scales were both frequency-dependent, it might be expected that the correlations between self-reports and observations on these scales would be higher than on the PS scales. However, higher correlations were obtained only between observations and self-reports on the Warmth Scale of the CRPQ. This may be because both observations and self-reports of warmth are less likely to be frequency-dependent because warmth can be a spontaneous behaviour as well as a reaction to children’s behaviour. As with the PS, to demonstrate the frequency-dependence of some of the scales of the observed CRPQ behaviours, their frequencies were compared for children who demonstrated higher levels of misbehaviour during the observation with the rest of the sample (Appendix J.2). The levels of Obedience were significantly higher in families with children with higher ratings of externalising behaviour, and there was a trend for Reasoning to be higher in this group. There was no significant difference in the levels of Warmth between the two groups. There was a trend for parents who had children

with higher levels of behaviour problems to demonstrate punitive behaviours, but this trend was non-significant.

To determine if the correlations between self-reports and observations would be stronger if the level of children's misbehaviour during the observation were taken into account, the correlations between observations and self-reports of the behaviour domains were repeated whilst adjusting for the global ratings of children's behaviour made by the author at the time of the observation using partial correlation coefficients (Table 7.3). Adjusting for the level of children's misbehaviour did not make a difference to the size of the correlations between observed and self-reported behaviour. Also, and analysis of variance was conducted to compare the levels of self-reported punitiveness between parents who did and did not display punitive behaviours during the observations, whilst controlling for children's observed behaviour ratings. This analysis returned a non-significant F-value of 0.38.

7.2.3. Removal of non-observed items from the questionnaire scores

Discrepancies between the self-reports and observations may have occurred because, as described in Chapter 3, a small number of items that formed the self-report scales could not be observed. The self-report scales were recalculated with these items removed, and these revised scores were then correlated with the observed data (see the final column of Table 7.3). These correlations show that the lack of agreement between self-reported and observed parenting behaviour was not a result of differences in scale composition, as once again the only significant correlation was between observed and self-reported Warmth. The other correlations remained small and non-significant. Consistent with this, parents who displayed behaviours from the Punitiveness scale during the observations did not have

significantly higher self-reports on the Punitiveness scale when non-observed items were removed.

Table 7.3 Correlations between corresponding behaviour domains on the self-report and observed measures (N=68)

Parenting Domain	Unadjusted correlations for the sample	Correlations adjusted for children's observed behaviour[a]	Correlation with items removed from questionnaire
PS			
Inconsistency	.01	-.01	-.01
Permissiveness	.05	.04	.06
Overreactivity	.11	.14	.04
CRPQ			
Warmth	.36**	.36**	.39**
Reasoning	.12	.08	.12
Obedience	.04	.00	.06

* = p <.05, ** = p <.01

[a] Correlation adjusted for the square root of the rating of children's externalising behaviour during the observation

7.2.4. Representative versus non-representative interactions

A fourth explanation for the lack of correlation between some of the parenting domains is that a 40 minute observation may not be representative of the interactions that normally occur between parents and children (Mrazek et al., 1982). This could occur for a number of reasons. First, the observation may not provide enough

opportunities to observe parenting behaviours such as ineffective discipline which occur with a lower natural frequency than behaviours such as warmth (Fassnacht, 1982). However, it should be noted that low correlations between observations and self-reports were obtained for permissive parenting behaviour practices, which were observed with relatively high frequencies. Observations might also not be representative of the "usual" levels of behaviours such as ineffective discipline because such behaviours are subject to social desirability biases. That is, parents who might normally use such techniques might not display them in the presence of an unfamiliar observer (Kavanagh et al., 1988).

To examine the degree to which representativeness of the observations influenced the correlation between observed and self-reported behaviour, the results were re-examined using parental perceptions of the representativeness of each recorded interaction as reported on the questionnaire completed during the no distraction task. It would be expected that the correlations between observed and self-reported behaviour would be stronger for those families who reported the interaction as "normal" compared with those who did not. Figure 7.1 displays the correlations between self-reported and observed parenting behaviour, repeated separately for parents who described the interaction as representative and those who did not.

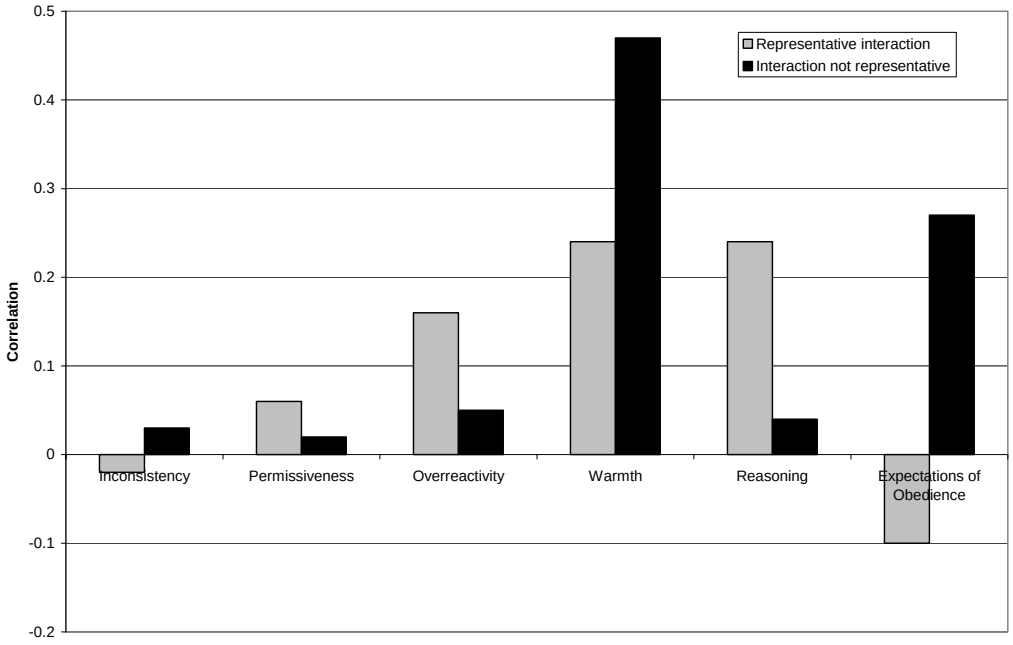

Figure 7.1 Correlations between observed and self-reported parenting behaviour for parents who described the interaction as normal and those who did not

Fisher z transformations were performed on the correlation coefficients and these were used to assess if the associations for parents who described the interaction as representative were significantly different from those who did not. Surprisingly, for Warmth and Obedience the correlations were stronger for families who described the interaction as not representative of what would normally occur, although these differences were not significant. The fact that parents said that the observation did not represent a normal interaction between themselves and their children does not necessarily mean that their parenting behaviour would have been different. In fact, differences in parenting were often not cited as reasons for the interaction not being "normal". Rather, the presence of other children or differences in the children's behaviour were more often reported as sources of difference. A two-way between-groups ANOVA was performed to determine if there was a significant difference between self-reported Punitievness scores for parents who displayed punitive behaviour during the observation and those who did not, taking into consideration the

representativeness of the observation. The result was not significant at the $p<.05$ level.

7.2.5 Reliability of the observed and self-reported behaviour

The lack of association between some observed and self-reported parenting domains might be expected for those domains that were less reliably reported or observed. Observations or self-reports that are not reliable also demonstrate low validity, and therefore are not objective measurements of the phenomena of interest. This lack of objectivity means these measurements are unlikely to correspond with more objective measures of the same phenomenon (Alessi, 1988).

This explanation does not necessarily account for the results in this study because, not all of the domains which demonstrated low correlations between observed and self-reported behaviour had lower levels of reliability. For example, observations and self-reports of Inconsistency, Overreactivity and Punitiveness all demonstrated high levels of reliability, but were not strongly associated across the two methods. Conversely, observations and self-reports of Warmth were significantly correlated even though the level of inter-observer reliability for Warmth was lower than the levels for the ineffective discipline styles.

In summary, when the relationships between corresponding behaviour domains on the self-report and observational measures were examined, the only significant correlation was for the Warmth Scale. No significant correlations were obtained for the ineffective parental discipline techniques. These results were unexpected given that other studies using observations and self-reports on PS domains have reported substantial correlations. Several factors that could have influenced these relationships were examined. These included the use of frequency counts rather than ratings of parents' behaviour, the influence of the level of children's misbehaviour during the observations, the removal of unobserved items from the self-report scales, the degree of representativeness of the observations, and the reliabilities of the observed and self-reported behaviours. However, analyses suggested that none of these factors explained the poor agreement. This extensive investigation suggests that the questionnaires of parenting behaviour and the behaviours coded from a 40-minute parent-child interaction may not be measuring the same phenomena. This may be due to social desirability biases, differences in natural frequencies of behaviour, poor operationalisation of the observed behaviours, or because some parents do not respond objectively about their own behaviour.

Chapter 8. Relationships between Parenting Behaviour, Family Risk Variables and Children's Behaviour

The results presented in the previous chapter showed that in this study self-reports and observations of ineffective parenting behaviours were not very strongly associated. This chapter examines the second and third aims of the study by presenting the results of comparisons between the two methods of measuring parenting behaviour and family and child variables. First, the concurrent validity of the measures of parenting behaviour will be presented. Concurrent validity assesses how well a test instrument relates to criterion variables (external variables which are expected to show an association with the instrument) measured concurrently. This type of validity can be determined by seeing how well the measure discriminates between groups with known differences relevant to the construct of interest (e.g., parents of boys versus parents of girls), or by correlating the scores on the instrument with measures of related phenomena (e.g., how well do scores on parenting measures correlate with measures of parental psychopathology) (Holden, 1983; Reitman et al., 2001). In this first section, concurrent validity was established by examining the relationships between measures of parenting behaviour and parent-reports of determinants of parenting practices as specified by the transactional theories discussed in Chapter 1 (such as child temperament, parental psychopathology, social support and marital satisfaction).

Second, the predictive validity of the measures of parenting behaviour will be described. Predictive validity is the extent to which instrument scores at one time point predict scores on the same or a different measure obtained at a later time point (Holden & Edwards, 1989). In this study, this was examined by reporting the associations between the measures of parenting behaviour and parent- and teacher-

reports of children's externalising behaviour obtained two years later. In these analyses, only results using the broad-band externalising scales from the Child Behaviour Checklist and Teacher Report Form were used. Narrow-band scale scores for attention problems and aggressive behaviours were also calculated but these did not show any differing relationships with the parenting variables, and hence parent-reports on the broader Externalising Scales are used. The analyses in Section 8.2 will also examine the incremental validity of the self-report and observational data collected in this study. As with the results in Chapter 7, the analyses were performed with transformed variables as appropriate (see Appendix J.1).

A large number of analyses have been performed in this section which increases the possibility of making a Type I error. Instead of correcting the probability values to reflect the number of analyses the results have been presented as originally conducted (Jaccard & Guilamo-Ramos, 2002b). This has been done to highlight interesting findings and consistent trends in the data, and to reduce the likelihood of making a Type II error (Jaccard & Guilamo-Ramos, 2002b), but the reader is advised to interpret the findings with this caution in mind.

In the following analyses which utilise self-reports from the same informant, it is possible that significant results between two self-report measures could be due to common method variance. To assess the extent to which common method variance may be a concern, Podsakoff (2003) suggests placing all scores from the same source (i.e., parents) in an unrotated factor analysis. If a single factor (or a first factor that accounts for a large proportion of the covariance between the measures) emerges, this is taken as evidence of high levels of common method variance. It should be noted that this technique does not statistically control for common method variance, and it is possible that the emergence of a single factor reflects causal

relationships between the variables rather than indicating methodological biases. This method was used to assess the degree to which relationships between the self-report measures in this study may be vulnerable to common method variance. The results of the unrotated factor solution can be seen in Appendix K.1. The first factor in the unrotated factor solution only accounts for 27.9% of the variance in the scores. This does not suggest strong common method variance effects (Podsakoff, 2003).

8.1 Relationships between Parenting Behaviour and Family Characteristics

The second aim of this study was to directly compare the relationships between the different measures of parenting behaviour and family characteristics (e.g., socioeconomic status, marital satisfaction, parental psychological functioning and social support, and child temperament). Differences in the parenting styles of parents of boys and parents of girls were examined in Chapters 5 and 6. The comparisons between mothers and fathers (all of which were non-significant) are not displayed in this chapter because of the very small number of fathers in the study (see Appendix K.2).

8.1.1 Parental education and employment

Education

There were no significant differences between parents with high school education (N=31) versus those with tertiary or trade qualifications (N=37) on any of the self-reported and observed parenting scores. This is consistent with research using self-reports on the Parenting Scale to assess parental discipline practices in a sample of children attending the Head Start program in the US (Reitman et al., 2001), but not in a community sample (Collett et al., 2001). In the latter study, parents with higher levels of education reported lower Overreactivity and Laxness scores. The differing

results may be explained by the fact that although all of these studies used parent-reports on the Parenting Scale, the actual items comprising the subscales differed in all three studies (see Appendix H.1). Also, the composition of the samples in terms of education levels of parents in the three studies differed, with parents in the current study intermediate to parents in the other two studies. The fact that the frequencies of observed parenting behaviours were not significantly different for parents with higher and lower levels of education in the present study adds support to the results obtained using the self-reports of parenting behaviour.

Paid employment

Table 8.1 displays the self-reported parenting scores and frequencies of observed behaviour for parents in paid employment and those not in paid employment. Compared with employed parents, those not in paid employment displayed significantly higher frequencies of permissive parenting behaviours and reported higher mean scores on the scales assessing harsh parenting techniques, Overreactivity and Punitiveness. Also, parents not in paid employment had lower mean frequencies of observed Warmth (which was not reflected in parents' self-reported scores), greater observed Overreactivity and higher self-report scores on the Inconsistency subscale, but these differences were not significant at the $p<.05$ level. A Chi-square analysis did not reveal a significant association between parents' paid employment status and their demonstration of punitive behaviours in the observation.

Table 8.1. Self-reported parenting scores and frequencies of parenting behaviour for parents in paid employment and parents not in paid employment

Parenting Domain	Paid Employment (N=25) Mean(SD)	Not in Paid Employment (N=42) Mean(SD)	p[a]
PS			
Inconsistency			
Self-report	2.43 (.76)	2.78 (.68)	.06
Observed (square root)	.11 (.11)	.16 (.13)	.17
Permissiveness			
Self-report	2.86 (.57)	2.85 (.60)	.91
Observed	.27 (.20)	.45 (.32)	.01
Overreactivity			
Self-report	2.77 (.58)	3.12 (.59)	.02
Observed (square root)	.30 (.13)	.37 (.19)	.12
CRPQ			
Warmth			
Self-report (reflect/inverse)	.72 (.18)	.69 (.18)	.50
Observed	.87 (.29)	.73 (.32)	.07
Reasoning			
Self-report	4.19 (.49)	4.11 (.47)	.50
Observed	.22 (.11)	.21 (.13)	.62
Punitiveness[b]			
Self-report	4.10 (.47)	3.57 (.61)	<.001
Obedience			
Self-report	2.84 (.66)	2.69 (.65)	.37
Observed (square root)	.22 (.12)	.24 (.14)	.49

[a] p-value for t-tests comparing scores of parents in paid employment and parents not in paid employment

[b] lower scores on the Punitiveness Scale indicate higher levels of punitive behaviours

Two factors might explain the relationships between self-reported harsh discipline and employment status. First, parents who are not in paid employment generally spend more time with their children and thus may be exposed to more instances of children's misbehaviour and coercive discipline interactions. Second, parental stress may be higher for parents who are not in paid employment due to financial strain, which may be reflected in higher levels of overreactive and harsh parenting (Jackson et al., 2000). In both of these instances, parents would report higher levels of harsh discipline that may not be reflected in the context of a 40-minute observation.

However, parents who were not in paid employment were observed to display higher frequencies of permissive parenting behaviours than employed parents, and this was not reflected in self-report scores on this scale. Parents who spend more time with their children may "choose their battles", letting some misbehaviour go undisciplined depending on its type (e.g., defiance compared with moving off-task during a boring activity) rather than reacting to every instance of child misbehaviour (Gershoff, 2002). Such parents would be unlikely to report this as permissive parenting, even though in observed interactions it would be regarded as such due to the strict operationalisation of parenting behaviours.

8.1.2 Relationship adjustment and parental disagreements about child-rearing

Only the 55 participants from two-parent families completed questionnaires assessing marital satisfaction and parental disagreements about child-rearing. Correlations between self-reported and observed parenting behaviour scores and marital satisfaction as measured using scores on the Dyadic Adjustment Scale (Spanier, 1976) are shown in Table 8.2.

Table 8.2 Correlations between measures of parenting behaviour and parent-reports on the Dyadic Adjustment Scale and the Parent Problems Checklist (N=55)

Parenting Domain	Dyadic Adjustment Scale r	Parent Problems Checklist r
PS		
Inconsistency		
Self-report	-.29*	.34*
Observed	.07	.28*
Permissiveness		
Self-report	-.21	.16
Observed	-.02	.38**
Overreactivity		
Self-report	-.16	.19
Observed	-.21	.25
CRPQ		
Warmth		
Self-report	.15	-.06
Observed	.20	-.13
Reasoning		
Self-report	.32*	-.20
Observed	-.07	-.09
Punitiveness		
Self-report	.06	-.15
Obedience		
Self-report	-.05	-.19
Observed	-.07	-.03

* p<.05, ** p<.01

Only self-reports on the Inconsistency (r=-.29, p<.05) and Reasoning (r=.32, p<.05) scales were significantly correlated with Dyadic Adjustment Scale scores. This is consistent with previous studies and theory which suggest that parents involved in marital conflict can be inconsistent in their parenting because they are preoccupied by their own relationship problems (Arnold et al., 1993. However, in the current study this was not reflected in the relationships between Dyadic Adjustment Scale scores and the frequencies of observed Inconsistency or Permissiveness. T-tests comparing parents who demonstrated punitive behaviours during the observations with those who did not, revealed no significant differences in Dyadic Adjustment Scale or Parent Problem checklist scores.

Table 8.2 also displays the correlations between parenting behaviour scores and parents' levels of disagreements about childrearing as measured using the Parent Problems Checklist (Dadds & Powell, 1991). Self-reported Inconsistency scores and observations of inconsistent parenting behaviours were significantly correlated with parent disagreements about child-rearing. Observed Permissiveness was also significantly related to scores on the parent problems checklist (r=.38). This might suggest that disagreements between parents specifically about childrearing are more strongly related with the frequencies of behaviour displayed during the context of a 40-minute observation than with parents' perceptions of their own behaviour. More general aspects of marital functioning, such as those assessed using the Dyadic Adjustment Scale, do not show these relationships. It may also be that the observed parenting behaviours (i.e., Inconsistency and Permissiveness) which were significantly correlated with PPC scores are those behaviours which are the focus of disagreements between parents about childrearing (e.g., as assessed by items such as "one parent is soft, one parent is tough with children" and "inconsistency between parents") (Dadds & Powell, 1991).

8.1.3 Parental psychological functioning

The mean frequencies for the observed and self-reported parenting domains for parents who scored above and below the cutoff on the General Health Questionnaire (Goldberg, 1978) are shown in Table 8.3.

Table 8.3 Mean (SD) frequencies of self-reported and observed parenting behaviour for parents who scored above or below the cutoff on the GHQ.

Parenting Domain	Below cutoff on the GHQ (N=48)	Above cutoff on the GHQ (N=20)	p[a]
PS			
Inconsistency			
Self-report	2.52 (.68)	2.99 (.73)	.02
Observed (square root)	.16 (.11)	.08 (.12)	.01
Permissiveness			
Self-report	2.81 (.60)	2.99 (.53)	.24
Observed	.37 (.28)	.41 (.32)	.65
Overreactivity			
Self-report	2.96 (.53)	3.07 (.77)	.49
Observed (square root)	.32 (.17)	.40 (.16)	.09
CRPQ			
Warmth			
Self-report (reflect/inverse)	.69 (.17)	.72 (.22)	.47
Observed	.78 (.33)	.78 (.28)	.97
Reasoning			
Self-report	4.13 (.44)	4.15 (.56)	.86
Observed	.21 (.11)	.21(.14)	.84
Punitiveness			
Self-report	3.85 (.50)	3.59 (.83)	.20
Obedience			
Self-report	2.71 (.63)	2.85 (.69)	.42
Observed (square root)	.23 (.12)	.24 (.16)	.84

[a] p-value for t-tests comparing scores of parents above and below the cutoff on the GHQ

The only significant differences between these two groups of parents were seen with both self-reports and observations on the Inconsistency scale, however the results were in opposite directions. Parents who were above the cutoff on the GHQ scale reported using more inconsistent parenting techniques than parents below the cutoff, but were actually observed to do so less often. This contradiction in findings may be the result of a response bias reflected in the self-reports of people with higher GHQ scores. That is, a tendency to perceive themselves as using higher levels inconsistent discipline, even though in relative terms they are not observed to do so.

In addition, A chi-square analysis revealed no association between the use of punitive discipline during the observation period and being above or below the cutoff on the GHQ. The results contradict research that has reported significant correlations between self-reported depressive symptoms and both observations and self-reports of responsive parenting and overreactive parental discipline (Jackson et al., 2000; Strayhorn & Weidman, 1988;. However, in the current study a measure was used that broadly assesses parental psychopathology (the GHQ-30) (Goldberg, 1978), compared with the specific measures of depression used in other studies. It is possible that specific depressive symptoms, rather than general psychopathology, are related to specific aspects of parenting behaviour.

8.1.4 Availability of social support

Transformed scores on the Availability of Social Integration subscale of the Interview Schedule for Social Interaction – Short Form (Unden & Orth-Gomer, 1989) were significantly correlated with the frequency of observed Warmth ($r=-.40$, $p<.01$) and self-reported Punitiveness scores ($r=-.28$, $p<.05$) (Table 8.4). Please note that because the transformed social support variable used the inverse square root of the

original scores, the relationships are in the opposite direction to what would be expected.

Table 8.4 Correlations (r) between scores on the Availability of Social Integration and self-reported and observed parenting behaviour scores (N=68)

Parenting Domain	Availability of Social Integration[a]
PS	
Inconsistency	
Self-report	.01
Observed (square root)	.10
Permissiveness	
Self-report	-.18
Observed	.17
Overreactivity	
Self-report	.13
Observed (square root)	.22
CRPQ	
Warmth	
Self-report (reflect/inverse)	-.14
Observed	-.40**
Reasoning	
Self-report	-.15
Observed	.11
Punitiveness[b]	
Self-report	-.28*
Obedience	
Self-report	-.20
Observed (square root)	.22

* p<.05, ** p<.01, [a] Analyses use the reflected square root of the Availability of Integration scale

[b] lower scores on the Punitiveness Scale indicate higher levels of punitive behaviours

The relationships between observed Warmth and self-reports of the availability of social integration are consistent with results using observed responsiveness (Meyers, 1999), but not with self-reported Warmth (Strayhorn & Weidman, 1988). Parents

generally reported demonstrating high warmth towards their children with high frequency. It is possible that the context of a structured interaction allows for a finer differentiation between the different levels of warmth demonstrated by parents. Conversely, t-tests revealed that observations of punitive parenting behaviour were not significantly correlated with parent-reports of social support whereas self-reports on the Punitiveness Scale of the CRPQ were. This may be because observations do not provide a representative sample of punitive behaviours (because of their lack of social desirability and their relatively low natural frequency) and hence relationships between these variables may be obscured.

The relationships between the availability of social integration and the parenting behaviours examined may be bidirectional. Parents who are warmer and happier may have more social contacts because of their disposition, whereas parents who display antisocial or punitive behaviours (both socially and in parent-child interactions) may find they have few social contacts. Conversely, the availability of social integration may enhance feelings of parenting competence and opportunities to observe or discuss effective parenting strategies with other parents, and these factors would generally promote displays of warm parenting and discourage the use of punitive techniques (Andresen & Telleen, 1992; Crnic & Greenberg, 1990; Jackson et al., 2000; Koeske & Koeske, 1990; McLoyd, 1990; Rubin et al., 1995).

8.1.5 Child temperament

The correlations between parent-reports of children's inflexibility and persistence obtained using the Short Childhood Temperament Questionnaire - Australian Version (Sanson et al., 1994) and parents' self-reported parenting scores and observation frequencies are shown in Table 8.5.

Table 8.5 Correlations (r) between parent-reported child temperament and self-reported and observed parenting behaviour scores (N=68)

Parenting Domain	Inflexibility	Persistence	Global temperament rating[a]
PS			
Inconsistency			
Self-report	.14	.00	-.19
Observed (square root)	.00	.25*	-.26*
Permissiveness			
Self-report	.13	.06	.02
Observed	.17	.19	-.36**
Overreactivity			
Self-report	.18	.13	-.20
Observed (square root)	.30*	.23	-.38**
CRPQ			
Warmth			
Self-report (reflect/inverse)	.04	-.23	-.11
Observed	-.14	-.23	.24
Reasoning			
Self-report	.12	-.01	-.26*
Observed	.20	-.02	-.08
Punitiveness			
Self-report	-.13	-.15	.11
Obedience			
Self-report	.08	-.11	.02
Observed (square root)	.17	.06	-.28*

* p<.05, ** p<.01; [a] Uses the log(10) of the global temperament score

None of the self-reported parenting scores were significantly correlated with Inflexibility or Persistence scores. Children's Inflexibility scores were significantly correlated with observed frequencies of Overreactivity (r=.30, p<.05), and parent-reports of children's persistence were significantly correlated with the frequency of observed parental inconsistency (r=.25, p<.05). Parents of children with persistent temperament traits may be "worn down" by their children's persistence and thus may be more likely to back down and give in to their child. In this way, parents may be inadvertently rewarding their child's persistence and thus increasing the likelihood of persistent behaviour in future interactions. Correlations of Persistence scores and both self-reported Warmth scores and observed frequencies of Warmth were of a similar size to the correlations with Inconsistency (r=-.23 for both self-reported and observed Warmth), but these were not significant. The magnitude of the correlations between observed parenting and the narrow-band domains of child temperament are similar to those reported in other observational studies that have explored broader dimensions of parenting behaviour (e.g. authoritative or negative parenting) and child temperament (Bates et al., 1998; Meyers, 1999; Rubin et al., 2002).

Also in Table 8.5 are the associations between the measures of parenting behaviour and parents' global ratings of their children's temperament. These ratings have been conceptualised in previous research as indicating the "goodness of fit" between parents' expectations for behaviour and their children's temperament. In this study, parents' global ratings of their children's temperament were significantly associated with standardised frequencies of observed Inconsistency (r=-.26, p<.05), Permissiveness (r=-.36, p<.01), Overreactivity (r=-.38, , p<.01) and Obedience (r=-.28, , p<.05). The direction of the correlations indicates that parents who rated their children as more difficult than average, demonstrated higher levels of these parenting

behaviours. It may be that parents who perceive their children as being temperamentally “easier” than the average child may use more effective parenting techniques. However, its also possible that parents who reported having more temperamentally difficult children were provided with more opportunities to use parental discipline during the observations, and hence these parents demonstrated higher levels of discipline in general, and ineffective discipline in particular.

For self-reported parenting behaviours, only Reasoning was significantly associated with the global ratings of child temperament. Surprisingly this correlation was negative ($r=-.26$, $p<.05$) indicating that parents who reported using reasoning behaviours more frequently also reported their child as being more difficult than the average child. Reasoning behaviour may be one of the techniques used by parents to deal with the behaviour of their children who they perceive as temperamentally “difficult”. If this type of parenting behaviour is incompatible with the child’s temperament, it would increase the mismatch between parents’ expectations and their children’s behaviour (Sanson & Rothbart, 1995; Smart & Sanson, 2001).

T-tests comparing the mean scores on the narrow-band and global temperament scales for parents who did and did not demonstrate punitive behaviours in the observations did not reveal any significant differences between these two groups.

In summary, both parent self-report scores and frequencies of observed parenting practices were generally not associated with family characteristics such as parent gender, parental level of education, and parent-reported scores on the General Health Questionnaire. These results replicate previous work to some extent, but may also be the result of the small numbers in some groups (i.e., n=5 fathers), or the breadth of the assessment (i.e., parental psychopathology instead of parental depression).

Parents who were not in paid employment reported significantly higher scores than employed parents on the harsh parenting scales and were observed to demonstrate higher frequencies of permissive parenting behaviours. These results may be a function of the longer amounts of time parents who are not employed spend with their children compared with employed parents.

Self-reported scores on the Inconsistency and Reasoning scales were significantly correlated with parent reports of marital adjustment. Self-reports and observations on the Inconsistency Scale and observations on the Permissiveness Scale were correlated with the levels of marital disagreements specifically about child-rearing as measured by the Parent Problems Checklist. These relationships may be explained by the fact that some of the specific areas for disagreements measured using this instrument include inconsistency between parents and parents being “soft” with children.

Observed parental Warmth frequencies and self-reported Punitiveness scores were significantly related to levels of available social support. These results may reflect relationships between the availability of social supports and specific parenting

behaviours that are not simply unidirectional (e.g., parents might have antisocial personality traits that affect both parenting style and the ability to obtain social support).

None of the self-reported parenting scores showed significant associations with specific aspects of child temperament, and only parent reports of reasoning behaviour were associated with parents' global ratings of their children's temperament. Observed parenting behaviour frequencies of overreactivity were associated with parent-reports of children's inflexibility and observed Inconsistency scores were positively correlated with children's Persistence scores. Similarly, observed Overreactivity, Obedience and Permissiveness were associated with parents' ratings of their children's overall temperament as "easy" or "difficult". Parents who rated their children as difficult were also more likely to report higher scores on the Reasoning Scale. These associations might reflect parenting in response to particular temperamental traits, but it is also possible that there exists a bidirectional effect between child temperament and parenting behaviour. The cross-sectional nature of this analysis means this cannot be explored further in this thesis.

8.2 Relationships between Parenting Behaviour and Children's Externalising Behaviour

The third aim of this thesis was to examine the relationships between the different measures of parenting behaviour and parent- and teacher-reports of children's externalising behaviour obtained two years later. Comparisons between parenting measures and child behaviour measures that were obtained concurrently could not be made because of the method of sample selection. The sample for this study was selected on the basis of parent-reports of parenting and children's externalising behaviour in the screening stage (see Chapter 3). Equal numbers of participants were selected from the four groups at the end of the screening stage. These four groups were determined on the basis of higher and lower parenting scores, and higher and lower externalising behaviour scores. In this way, any associations between these variables that might have existed in the larger screening sample would not have been found in the smaller observation sample.

8.2.1 Unadjusted relationships between children's externalising behaviour and measures of parenting behaviour

Preschoolers' externalising behaviour scores

The effect of the method used for sample selection can be seen in the correlations between the parenting behaviour measures and externalising behaviour assessed concurrently shown in Appendix K.3. Parents' scores on the self-reported variables were generally not significantly associated with either parent- or teacher-reports on the Externalising Scale of the Child Behavior Checklist (1.5-5) or Caregiver-Teacher Report Form. Conversely, parent-reports of preschoolers' externalising behaviour problems were significantly correlated with frequencies of observed parenting behaviours, specifically parental Permissiveness ($r=.38$, $p<.01$) and Overreactivity ($r=.36$, $p<.01$). Teacher-reports were significantly correlated with almost all of the

observed behaviour frequencies: Inconsistency r=.25, p<.05; Permissiveness r=.42, p<.01; Overreactivity r=.38, p<.01; and Warmth r=-.28, p<.05. These results may be due to the frequency-dependence of the observations, wherein children had to display misbehaviour during the observation to enable the observation of parental discipline. For parents who have children who display very low levels of misbehaviour, it would be more unlikely to observe parent-child discipline interactions than for children with higher levels of problems. T-tests revealed no association between observed punitiveness and either parent- or teacher-reports of preschoolers' externalising behaviour.

School-aged children's externalising behaviour scores

Table 8.6 displays the unadjusted correlations between the parenting variables and parent- and teacher-reports of children's externalising behaviour problems obtained two years later. Parents' self-reported scores on the Overreactivity (r=.27), Punitiveness (r=-.31) and Obedience (r=.28) scales obtained when their children were at preschool were significantly correlated (all p<.05) with parent-reports on the Externalising Scale of the CBCL(6-18) obtained two years later. Parent self-reports of Overreactivity (r=.30) and Reasoning (r=-.28) also correlated significantly with teacher-reports of children's externalising behaviour at school. These results are consistent with other studies that have found significant correlations between self-reported parenting scores for negative control (such as harsh and overreactive parenting) and children's behaviour assessed prospectively (Belsky et al., 1998; O'Leary et al., 1999).

Table 8.6 Correlations between parenting behaviour and externalising behaviour reported two years later

Parenting Domain		Parent-reports on CBCL (6-18)[a] N=59	Teacher-reports on TRF[a] N=54
PS			
Inconsistency			
	Self-Report	.13	.27
	Observed	.33*	.30*
Permissiveness			
	Self-Report	.11	.19
	Observed	.45**	.47**
Overreactivity			
	Self-Report	.27*	.30*
	Observed	.37**	.41**
CRPQ			
Warmth			
	Self-Report	-.13	-.21
	Observed	-.30*	-.40**
Reasoning			
	Self-Report	-.10	-.28*
	Observed	-.11	.17
Punitiveness[b]			
	Self-Report	-.31*	-.23
Obedience			
	Self-Report	-.28*	-.17
	Observed	.14	.40**

* = p <.05, ** = p <.01. [a] Analyses use the square root of the Externalising scales

[b] lower scores on the Punitiveness Scale indicate higher levels of punitive behaviours

As with the measures of children's behaviour two years earlier, most of the observed parenting frequencies were significantly correlated with parent- and teacher-reports of children's externalising behaviour at school-age. Observations of punitive behaviour during the observations were also related to parent-reports, but not teacher-reports of externalising behaviour (mean (SD) CBCL score for parents who

used punitive techniques = 3.15 (1.34); for those who did not = 2.31 (1.37); t=2.31, df =57, p=.02) .The frequency of observed Reasoning was not related to either parent- or teacher-reports of children's behaviour two years later. These results replicate findings obtained with community samples (Denham et al., 2000; Kingston & Prior, 1995; Shaw et al., 2001; Shaw et al., 1998), but not in clinically referred samples (Campbell et al., 1986a; Campbell & Ewing, 1990; Campbell et al., 1986b)

8.2.2 Predicting externalising behaviour problem scores using initial problem levels and parenting behaviour

Predicting school-aged children's behaviour problem scores from preschool behaviour and parenting behaviour

The two questions examined in this section are: 1. Do self-report parenting scores explain any additional variance in children's externalising behaviour scores after controlling for levels of externalising behaviour obtained two years earlier? Previous research has shown that the strongest predictor of children's externalising behaviour is previous levels of behaviour (Campbell, 2002). In studies of child psychopathology, there would be no need to obtain self-reports of parenting practices if they did not contribute uniquely to the explanation of children's externalising behaviour. 2. Do frequencies of parenting behaviour obtained from videotaped observations enhance the prediction of children's externalising behaviour scores over and above what is predicted using self-reported parenting behaviour scores? That is, do they display incremental validity? If this is the case, it would highlight a need for observations of parenting behaviour to be conducted instead of, or in addition to, obtaining self-reports of parenting practices. If it is not the case, self-reports of parenting behaviour may be adequate to assess parenting practices.

To address these questions, a series of hierarchical linear regressions were performed to predict parent- and teacher-reports of children's externalising behaviour

obtained two years after the initial assessment. Children's behaviour at school-age was the dependent variable. In the first step of the regression, parent- or teacher-reports of preschoolers' externalising behaviour were entered to determine what proportion of current child behaviour is explained by previous levels of behaviour. Second, the relevant parenting self-report score was entered into the equation to see if self-reports of parenting behaviour on the various parenting domains explained any additional variance in children's behaviour scores. Finally, frequencies of observed parenting behaviours were entered into the equations to see if they contributed any additional predictive power to the explanation of children's externalising behaviour scores over and above that predicted by self-reports of parenting. The results from the second and third steps of these analyses for the prediction of parent- and teacher-reports of externalising behaviour scores are shown in Tables 8.7 and 8.8, respectively. The analyses were performed using each parenting domain (e.g., Inconsistency, Permissiveness, Warmth etc) as the independent variables in separate regressions.

Parent-reports of children's externalising behaviour problems

For each of the analyses shown in Table 8.7, in the first step of the equations, initial parent-reports on the CBCL(1.5-5) accounted for 31.6% (R=.562) of the variance in parent-reports of children's externalising behaviour two years later. In the second step of the equations, only parents' self-report scores on the Obedience scale predicted unique variance (5%) in children's later externalising behaviour scores.

Table 8.7 The prediction of parent-reports of externalising behaviour scores on the Child Behaviour Checklist (N=59)

Independent Variables	**ΔR^2**	**F change**	**Standardised β**
PS			
Inconsistency			

Self-report	.00	.19	.05
Observed (square root)	.08	6.87*	.28
Permissiveness			
Self-report	.00	.30	.06
Observed	.07	5.78*	.28
Overreactivity			
Self-report	.04	3.00	.18
Observed (square root)	.03	2.32	.18
CRPQ			
Warmth			
Self-report (reflect/inverse)	.01	.64	-.03
Observed	.03	2.07	-.17
Reasoning			
Self-report	.03	2.43	-.16
Observed	.01	1.04	-.11
Punitiveness			
Self-report	.04	3.40	-.19
Observed (categorical)	.04	3.99	.21
Obedience			
Self-report	.05	4.60*	-.23
Observed (square root)	.00	.18	.05

* p<.05

In the current study, observed frequencies of Expectations of Obedience did not explain any additional variance in parent-reports on the CBCL Externalising Scale, after initial children's problem scores and self-reports of parenting had been controlled. Parent-reports on this parenting domain may provide a broader picture of parent's overreactivity and expectations of obedience that is relevant to the prediction of parent-reports of children's behaviour problems at later stages of life. The inclusion of observed frequencies of behaviour on the same dimension did not explain levels of children's behaviour problems that are not already accounted for by parent self-reports of parenting behaviour or previous levels of children's externalising problems. It is possible that the information provided in a structured context-specific observation is only related to concurrent levels of behaviour problems because of the frequency-

dependence of the behaviours coded during these observations. When this is taken into account, observed parenting Expectations for Obedience does not account for any additional variance in children's externalising behaviour assessed two years later. It is also possible that the observations did not provide sufficient counts of this behaviour to predict children's behaviour. The results also support the findings of O'Leary et al. (1999) who found that parental Overreactivity self-reports at Time 1 did not significantly predict children's externalising behaviour problems at Time 2, after controlling for the effects of children's behaviour at Time 1.

The frequencies of observed Inconsistency and Permissiveness significantly improved the prediction of parent-reports of children's externalising behaviour obtained two years later, after self-reported parenting behaviour scores and initial levels of children's problems were controlled. Observations of behaviours on these dimensions of laxness, provide information that is relevant to the prediction of children's externalising behaviour problems that is not provided by parent self-reports on these dimensions. This may be because parents and observers have different concepts of what permissive and inconsistent parenting behaviours are. The observation of strictly operationalised parenting behaviours may provide information about parenting behaviour that is relevant to the prediction of parent-reports of later children's behaviour problems, whereas self-reports that are more subjective do not.

Teacher-reports of children's externalising behaviour problems

For each of the regressions shown in Table 8.8, teacher-reports of preschoolers' behaviour on the C-TRF accounted for 21.0% (R=.459) of the variance in teacher-reports of children's externalising behaviour two years later. Self-reported scores on the Overreactivity scale of the Parenting Scale predicted unique levels of variance (6%) in teacher-reports of children's externalising behaviour. The observed frequencies of parent behaviours on this domain did not explain any unique variance

in teacher-reports of children's externalising behaviour problems beyond that predicted by the self-reports and teacher-reports of behaviour at preschool.

Observation frequencies on the Permissiveness, Warmth and Obedience scales explained unique levels of variance (9%, 9% and 11%, respectively) in teacher-reports of externalising behaviour even after the initial levels of children's behaviour and self-reports on these parenting domains were controlled. Self-reports on these parenting domains did not account for any variation in teacher-reports that was not already accounted for by previous levels of behaviour as reported by preschool teachers. These results are interesting, given that self-reports and observations on the Warmth Scale were significantly correlated ($r=39$, $p<.01$), but they do not show the same predictive relationships with respect to teacher-reports of children's behaviour. As described earlier, parents in general reported high Warmth scores on the CRPQ, but there was a wider spread of scores on the observed frequencies of Warmth. The microanalytic nature of the observational coding scheme and the structured nature of the tasks might allow for a finer differentiation between the different levels of warmth demonstrated by parents in the observations. This greater sensitivity may enhance the predictive validity of observed Warmth compared with self-reported Warmth scores.

Table 8.8 The prediction of teacher-reports of externalising behaviour scores on the Teacher Report Form (N=54)

Independent Variable	ΔR^2	F change	Standardised β
PS			
Inconsistency			
Self-report	.04	2.49	.20
Observed (square root)	.04	2.73	.21
Permissiveness			

Self-report	.02	1.57	.15
Observed	.09	6.53*	.33
Overreactivity			
Self-report	.06	4.36*	.24
Observed (square root)	.05	3.97	.25
CRPQ			
Warmth			
Self-report (reflect/inverse)	.02	1.54	-.07
Observed	.06	3.98*	-.27
Reasoning			
Self-report	.03	2.21	-.21
Observed	.03	2.28	.19
Punitiveness			
Self-report	.03	1.91	-.17
Observed (categorical)	.00	.07	.03
Obedience			
Self-report	.01	.83	-.13
Observed (square root)	.12	9.32**	.36

* p<.05, ** p<.01

8.2.3 Examining the relationships between parenting practices and the stability of externalising behaviour problem scores

For children on whom follow-up data were available (N=59), the sample was divided into four groups on the basis of the stability of their parent-reported externalising behaviour scores. These children were divided into four groups as shown in Table 8.9. Nine out of the 21 children who were above the cutoff on the CBCL at preschool were still above the cutoff two years later. This level of stability (43%) is comparable with levels obtained in other studies (Campbell, 1995; Campbell et al., 1986b; Hofstra et al., 2000; Kingston & Prior, 1995; McGee et al., 1991; McGee et al., 2002; O'Leary et al., 1999; Richman et al., 1982; Sanson et al., 1991). Further, 69% (9 out of 13) of the children above the clinical cutoff at the second assessment had been above the cutoff at the preschool-age assessment.

Table 8.9. Number of children above and below the CBCL cutoff at the two assessments

		Preschool CBCL(1.5-5)	
		Below clinical cutoff	Above clinical cutoff
School CBCL (6-18)	Below clinical cutoff	34	12
	Above clinical cutoff	4	9

Analyses of variance were then performed to assess whether there were significant differences in the self-reported and observed parenting scores for these four groups of participants. Only two measures of parenting behaviour produced significant results in these analyses: observed Permissiveness frequencies ($F=2.74$, $p=.05$) and the square root of observed Overreactivity ($F=3.99$, $p=.01$). Post hoc analyses revealed that parents of children with stable, high levels of externalising problems displayed significantly higher frequencies of permissive ($.62 \pm .43$ versus $.33 \pm .22$) and

overreactive (.45±.10 versus .28±.16) parenting behaviours than parents of children who were below the clinical cutoff at both assessments .

Because of the small numbers of participants in some of the groups, and the number of analyses performed, these results should be interpreted with caution. However, these results to some extent replicate those reported by Kingston and Prior (1995) who found that parent-reports of physical punishment and control through guilt and anxiety distinguished groups of children with persistent behaviour problems from children without such problems.

In summary, parent self-reported Overreactivity scores obtained when children were preschoolers were significantly associated with parent- and teacher-reported externalising behaviour scores assessed two years later. For teacher reports, these relationships still held when the initial levels of children's behaviour problems were controlled. Observed frequencies of parenting behaviour on most scales (Reasoning excepted) were significantly associated with parent- and teacher-reports of children's externalising behaviour obtained two years after the observations were conducted. However, after controlling for the initial levels of children's misbehaviour and self-reports of parenting behaviour, only observed Inconsistency and Permissiveness scores were still associated with parent-reports of children's externalising behaviour, and frequencies of Permissiveness, Warmth and Expectations of Obedience were assoicated with teacher-reported externalising scores. These results suggest that self-reports and observations yield different types of information that are important for the prediction of children's externalising behaviour problems. Parent-reports may be more valid when obtaining information about overreactive parenting behaviours, whereas observations may be more valid for measuring inconsistent, permissive and responsive parenting practices.

Chapter 9. Summary, Study Limitations and Conclusion

This is the first study to systematically compare self-reports and interval-sampled observations of parenting, using observational measures derived directly from self-report questionnaires. Sixty-eight parents recruited from preschools in the Adelaide metropolitan area, completed two self-report questionnaires (the Parenting Scale and the Child-Rearing Practices Questionnaire), and participated in videotaped semi-structured home observations with their children. The study used a microanalytic approach in which parenting behaviour was interval-sampled during four activities. The parenting measures were compared directly, and their relationships with family characteristics and parent and teacher-reports of children's behaviour were calculated. These analyses involved family, parent and child variables highlighted in the transactional models of Patterson (1982, 1989) and Belsky (1984) as being determinants of parenting behaviour. In addition, the incremental validity of the parenting measures was assessed in multivariate analyses after controlling for levels of children's behaviour when they were preschoolers. The study findings have been discussed in detail in Chapters 4 to 8, and this chapter will provide an overall summary of the research and recommendations for each domain of parenting behaviour. The second part of this chapter will examine the limitations of the study, and provide recommendations for future research.

9.1 General Summary and Recommendations

9.1.1. Relationships between Self-reported and Observed Behaviour

The results of this study support a mixed-method approach to measure parenting behaviour. Self-reports and observations appear to provide types of information that are differentially associated with children's externalising behaviour problem scores. The general absence of any relationship between parenting self-reports and

observations of the same behaviour, suggests that the two methods are actually measuring different things. This may be occurring for a number of reasons. First, perfect agreement between the two methods cannot be expected, because self-report questionnaires assess parenting behaviour "in general" whereas the observations provided frequencies of behaviour in structured tasks over a 40-minute period with a stranger present (Dowdney et al., 1984; Gardner, 2000; Johnson et al., 2001; Locke & Prinz, 2002; Metsapelto et al., 2001; Pappas-Jones & Adamson, 1987). Second, because observational data is considered more objective than self-report data, a lack of agreement between these measures may reflect parents' lack of awareness about their behaviour (e.g., parents are under-reporting certain behaviours) or their different interpretations of the behavioural items on the questionnaires (Holden, 1983; Mrazek et al., 1982). The way in which the parenting and child behaviours were operationalised for the observational coding system does not necessarily correspond to how parents interpret the same behavioural items (Holden, 1983; Johnson et al., 2001; Mrazek et al., 1982; Shaw & Bell, 1993). Parents could also be giving misleading reports about certain aspects of behaviour because of intentional or unintentional response biases (Alessi, 1988; Bates & Bayles, 1984; Holden & Edwards, 1989; Lancaster et al., 1989; Podsakoff et al., 2003). Third, it is possible that due to participants' reactivity to the observations, they may not contain accurate representations of parent-child interactions for some families (Banister et al., 1996; Brandt, 1992; Cox, Puckering, Pound, & Mills, 1987; Gardner, 2000; Kavanagh et al., 1988; Metsapelto et al., 2001; Renne et al., 1983). For example, some parents reported that their children were much better behaved than normal during the observations. As a result, the videotaped interactions contained fewer instances of children's misbehaviour, and hence opportunities for parental reactive discipline, than would normally occur (Brandt, 1992).

Other authors have reported a lack of relationship between observations and self-reports of parenting behaviour with parents of older children (Ten Haaf et al., 1994; Feinberg et al., 2001; Johnston et al., 2002). Research with parents of preschool children that has reported significant associations between the two methods, has been limited by the use of ratings of parenting behaviour, broad definitions of parenting behaviour, and small sample sizes (Arnold et al., 1993; Denham et al., 2000; Dowdney et al., 1984; Strayhorn & Weidman, 1988). In this study, a variety of factors could have been responsible for the lack of associations between observed and self-reported parenting behaviour on the various parenting domains. These factors included the use of frequency counts rather than ratings of parents' behaviour (Dowdney et al., 1984), the influence of the level of children's misbehaviour during the observations, the removal of unobserved items from the self-report scales, the degree of representativeness of the observations (Banister et al., 1996; Brandt, 1992; Dowdney et al., 1984), and the reliabilities of the observed parenting behaviours (Alessi, 1988; Maxwell & Pringle, 1983). None of these factors were found to have an effect on the level of association between the two methods.

The results of the reliability and validity assessments for the seven specific parenting domains examined in this study are discussed below.

9.1.2. Inconsistency

Inconsistency was the first factor derived from the principal components analysis of the Parenting Scale items. It includes items from the original Laxness factor (Arnold et al., 1993), and refers to parental behaviour which involves "backing down" and failing to follow through with stated courses of action. Observation and self-report scores on the parental Inconsistency Scale demonstrated acceptable levels of reliability (Pallant, 2004). The highest levels of Inconsistency were observed in the

pack-up task, which has been used in previous research to assess aspects of lax parenting behaviour (Harvey-Arnold & O'Leary, 1995; Harvey-Arnold & O'Leary, 1997).

The concurrent validity of parent's self-reports on the Inconsistency Scale was supported by a negative correlation with parent reports of marital satisfaction and a positive correlation with parental disagreements about child-rearing. This finding supports previous research examining inconsistent and lax parenting behaviour (Arnold et al., 1993; Gottman & Katz, 1989; Meyers, 1999). Observation scores on the Inconsistency Scale were also significantly correlated with parents' ratings of couples' disagreements about child-rearing, which may reflect issues about inconsistency between parents and one parent being "soft" with children (Dadds et al., 1987). Both self-reports and observed frequencies of Inconsistency were significantly different between parents who scored above the cutoff on the General Health Questionnaire and those who scored below. However, the relationships were in opposite directions with parents above the cutoff reporting higher levels of inconsistency but demonstrating less in the observations than parents below the cutoff. Observed, but not self-reported, Inconsistency Scale scores were also positively correlated with parent-reports of children's temperament on the Persistence scale and with parents' global ratings of temperament. As described earlier, this relationship might be due to the intermittent reinforcement parental inconsistency provides to persistent children (Arnold et al., 1993; Patterson, 1982), or the wearing down of parents by their persistent children (Chamberlain & Patterson, 1995; Cunningham & Boyle, 2002; Patterson & Stouthamer-Loeber, 1984). However, it is surprising that the self-report questionnaire scores do not reflect this relationship. This may be an indication that parents are unaware that they are behaving in a manner that is inconsistent, and thus do not report such behaviour as occurring. This

suggestion is also supported by the lack of association between self-reports and observations on the Inconsistency Scale in the current study.

The predictive validity of observations on the Inconsistency Scale was indicated by significant associations with parent- and teacher-reports of children's externalising behaviour obtained two years after the observations were conducted. This finding replicates previous research findings using observational measures of inconsistent parenting (Denham et al., 2000; Kingston & Prior, 1995; Shaw et al., 2001; Shaw et al., 1998). In the current study, this relationship was not found with the self-reported scores. Further, the incremental validity of observed Inconsistency was demonstrated. After controlling for the initial levels of children's misbehaviour and self-reports of parenting behaviour, observed Inconsistency scores explained a significant degree of additional variance in parent-reports of children's externalising behaviour two years later. These relationships were not replicated using teacher-reports of the children's behaviour, suggesting that parental inconsistency may be more important for the development of behaviour problems at home, rather than at school. This might be particularly true if there is not only inconsistency in the way an individual parent manages their child's behaviour, but if there is also inconsistency in the way parents and teachers manage the same child's behaviour (e.g., if the child's teacher follows through with discipline, but the parent does not) (Gagnon, Vitaro, & Tremblay, 1992; Rubin et al., 1995).

These results would suggest that observations of parental inconsistency as described in this study are reliable and valid, and may be particularly useful in the prediction of children's behaviour problems at home. Parent-reports of inconsistent parenting behaviours while reliable, do not demonstrate the same levels of validity as observations. This may be because what an observer considers inconsistent

parenting would not be rated by the parents as such (Holden, 1983; Mrazek et al., 1982). Evaluation studies that include interventions designed to improve parental consistency should use observations to more accurately measure these parenting behaviours. The lack of agreement between the two methods for assessing inconsistent parenting practices and the poor concurrent and predictive validity of the self-reports, suggest that parent-reports of inconsistency are less useful than observations for informing clinical practice. The highest frequencies of inconsistent parenting behaviour were observed in the pack-up task, which suggests this activity may be a quick and useful way for clinicians to gain a valid and reliable measure of parental inconsistency. It should be noted that for some parenting domains the highest levels of observed parental discipline occurred during transitions to tasks. This was particularly true for the transition from the drawing activity to the pack-up task, when children showed resistance at leaving an enjoyable activity to pack up the toys.

9.1.3. Permissiveness

The second factor extracted from the principal components analysis of the PS item was Permissiveness. This scale also contained items from the original Laxness factor (Arnold et al., 1993), and it refers to a failure to discipline or set limits on behaviour. Self-reported scores on the Permissiveness Scale had moderate levels of reliability as assessed via internal consistency and temporal stability (Pallant, 2004). Good inter-observer reliability for the Permissiveness Scale was obtained. Permissive parenting practices were one of the most frequently observed behaviours during the videotaped observations. In particular, the no distraction task was useful for observing permissive practices, as the children were more likely to move off-task whilst parents were concentrating on another task. Hence, there were more opportunities to observe permissive parenting during this activity. Other studies have

also used this type of activity to assess features of lax and permissive parenting behaviour (Harvey-Arnold & O'Leary, 1995; Harvey-Arnold & O'Leary, 1997).

Parents who were not in paid employment demonstrated higher frequencies of permissive parenting behaviours than did parents who were in paid employment. In part, this may be a function of the longer amounts of time parents who are not employed spend with their children compared with employed parents. Parents who spend more time with their children can be exposed to more frequent instances of child misbehaviour, and thus might “choose their battles” and let some misbehaviour go (Gershoff, 2002). Observations on the Permissiveness Scale were correlated with parent-reports on the Parent Problems Checklist. As with the Inconsistency scores, it is possible that permissive parenting practices are one of the areas of disagreement between parents about childrearing (Dadds et al., 1987). Parent-reports of the difficultness of their children’s temperaments were significantly correlated with observed parental Permissiveness Scale scores. Parents who reported their child as “more difficult than average” demonstrated more permissive parenting practices. Previous studies have used this temperament item as an indicator “goodness of fit”, i.e. the extent to which parents’ expectations for their child’s temperament fit with the child’s characteristics (Sanson & Rothbart, 1995; Smart & Sanson, 2001; Thomas & Chess, 1977). For some parents a perception of their child as more difficult than average may reflect attributions for the child’s behaviour which are internal, stable and global (Thomas & Chess, 1977). These parents may be less likely to respond to instances of child misbehaviour if they believe that discipline will have no effect on the child’s behaviour.

Observed frequencies on the Permissiveness Scale were significantly associated with parent- and teacher-reports of children’s externalising behaviour obtained two

years after the observations were conducted. As with observations on the Inconsistency Scale, further evidence for the validity of the Permissiveness Scale was demonstrated when this scale accounted for unique variance in parent-reports of children's behaviour after controlling for the initial levels of children's misbehaviour and self-reports of parenting behaviour. This was also true for predicting teacher-reports of children's externalising behaviour. Self-reports on the Permissiveness Scale did not demonstrate this predictive validity with either parent- or teacher-reports of externalising behaviour after controlling for the initial levels of children's behaviour.

These results are very similar to those obtained with the Inconsistency Scale. The lack of concordance between observations and self-reports on the Permissiveness Scale may reflect parents' lack of awareness of their permissive behaviour. This was supported by comments parents made to the observer after the no distraction task, when children were engaged with something quiet, despite being off-task, some parents commented that they did not consider this misbehaviour and so did not respond. Observed frequencies of this type of behaviour (which was categorised as permissive) were both concurrently and predictively associated with a number of child and parent variables, as described above. As with the self-reported scores on the Inconsistency Scale, self-reported Permissiveness Scale scores did not display concurrent or predictive validity, and were less reliable than other scales on the Parenting Scale. The recommendation from these results is that when the focus of a study or intervention is on permissive parenting practices, observations of parent-child interaction should be used to obtain valid and reliable measures of permissive parenting practices. The no distraction task could be used for this purpose, as high levels of permissive parenting were observed during this activity.

9.1.4. Overreactivity

The Overreactivity Scale is the final factor derived from the factor analysis of the PS items. It refers to physical discipline practices and anger when dealing with misbehaviour. This scale includes most of the items from the original Overreactivity factor (Arnold et al., 1993). Scores on the self-reported Overreactivity Scale had lower levels of internal consistency than scores on the Inconsistency and Permissiveness scales, but the reliability of this scale was comparable with estimates from previous studies (Arnold et al., 1993; Harvey-Arnold & O'Leary, 1995; Harvey-Arnold & O'Leary, 1997). Observed parenting practices on the Overreactivity Scale had the highest levels of inter-observer reliability, reflecting the salience of these parenting behaviours and the lack of subjectivity about their coding (Fassnacht, 1982; Shaw et al., 2001; Shaw et al., 1998). The pack-up task served to elicit higher levels of overreactive parenting behaviour, with higher levels of confrontation between parents and children occurring during this activity. The pack-up activity has been used in several studies to facilitate the observation of overreactive and harsh parenting practices (Arnold et al., 1993; Campbell, 1994; Campbell et al., 1991; Conger et al., 2003; Harvey-Arnold & O'Leary, 1995; Harvey-Arnold & O'Leary, 1997; Hops et al., 2003; Kaplow et al., 2002; Shaw et al., 2001; Shaw et al., 1998).

Parents not in paid employment reported higher scores on the Overreactivity Scale. As with the Permissiveness Scale, these results may be a function of the longer amounts of time that parents who are not employed spend with their children compared with employed parents, or they may reflect a relationship between financial stress and the use of harsh parenting techniques (Jackson et al., 2000). This relationship was not found for observations on this behaviour domain. Observed parenting behaviour frequencies on the Overreactivity Scale were associated with parent-reported children's Inflexibility scores and global ratings of temperament,

which could reflect parents' frustration with an inflexible child (Prior et al., 2000; Sanson & Rothbart, 1995; Smart & Sanson, 2001; Thomas & Chess, 1977).

Of all the self-report measures, only Overreactivity scores obtained when children were preschoolers were significantly associated with parent- and teacher-reported externalising behaviour scores assessed two years later. These relationships still held for teacher-reports but not parent-reports when the analyses controlled for the levels of children's behaviour problems reported at the preschool assessment. Observations of parental overreactivity, while significantly associated with parent- and teacher-reports of children's externalising behaviour obtained two years after the observations were conducted, did not account for any additional variance in the children's behaviour after self-reports on this domain and children's initial levels of behaviour were controlled. This may be a function of the frequency-dependence of the observations. Children with higher levels of externalising behaviour problems at the preschool assessment would have provided more opportunities for the observation of overreactive parenting techniques. When the levels of these problems were controlled in the multivariate analyses, the relationship between observed Overreactivity Scale scores and children's later behaviour did not remain. Self-reports on the Overreactivity Scale were not frequency-dependent and their relationship with children's externalising behaviour problems remained even after controlling for preschool-age behaviour.

Self-reports on the Parenting Scale are a reliable means of assessing overreactive parenting practices. They demonstrate some concurrent and predictive validity, and are able to assess the use of these behaviours over a long period of time (Arnold et al., 1993). Conversely, observations on the Overreactivity Scale while demonstrating good reliability, only demonstrate some concurrent validity and do not show

incremental validity. The most likely reason for the lack of agreement between parent-reports and observations on this behaviour domain is the generally low frequency of overreactive behaviours obtained in the observation session (Fassnacht, 1982; Shaw et al., 2001; Shaw et al., 1998). Those behaviours demonstrated by parents in the context of a 40 minute observation with an observer present are unlikely to be representative because of the naturally low frequencies of overreactive parenting behaviours, and participant reactivity to being observed which could decrease the levels of these socially undesirable behaviours (Bates et al., 1998).

The recommendation for researchers and clinicians wishing to measure levels of overreactive parenting techniques is to utilise self-report measures of this domain. Repeated observational measures could overcome the limitations of a single observation, by providing more opportunities to observe overreactive parenting behaviour and by decreasing participant reactivity, but repeated observations are both time-consuming and costly. Further, self-reports of Overreactivity provide a measure that has value in predicting children's behaviour problems both at home and at school.

9.1.5. Punitiveness

The Punitiveness Scale from the Child-Rearing Practices Questionnaire refers to harsh parenting behaviours including the use of physical discipline (Sanson, 1996). Self-reports on the Punitiveness Scale demonstrated acceptable internal consistency, at a level consistent with previous research (Paterson & Sanson, 1999). Inter-observer reliability on the Punitiveness Scale was good. As with the observations on the Overreactivity Scale, the pack-up task served to elicit relatively higher levels of punitive parental discipline, however, across tasks the absolute frequencies of these

behaviours was very low. As with observations on the Overreactivity Scale, this reflects participants' reactivity to the observation process and the fact that a 40-minute observation is unlikely to yield very many instances of punitive parenting behaviour (Fassnacht, 1982). Therefore, for subsequent analyses, observations on the Punitiveness scale were dichotomised to represent parents who demonstrated any behaviours from that scale during the observation versus those who did not.

The relationships found for parents' self-reports on the Punitiveness Scale and family risk factors were similar to those obtained with the self-reports on the Overreactivity Scale. For example, self-reports on the Punitiveness Scale differentiated parents who were not in paid employment from employed parents, whereas observed frequencies did not. In addition, self-reported Punitiveness scores were significantly negatively correlated with parent reports of levels of available social support. These results could reflect a relationship in which parents with few social supports do not have social opportunities to learn alternatives to punitive parenting practices, or another factor (e.g. parental mental health or personality) could affect both parenting behaviour and the availability of social supports (Andresen & Telleen, 1992; Crnic & Greenberg, 1990; Jackson et al., 2000). Neither observed Punitiveness scores or self-reports showed predictive validity with parent- and teacher reports of children's externalising behaviour problems.

The results for the measurement of punitive parenting practices are mixed. Self-reports on this domain demonstrated some concurrent validity, but not predictive validity, whereas observations demonstrated predictive validity to some extent. The absolute levels of punitive parenting behaviour that could be observed in the context of a brief observation were very low. Given that there is overlap in some of the items on the Punitiveness and Overreactivity scales, the Overreactivity Scale of the PS

appears a more valid and reliable measure of harsh parenting practices.

9.1.6. Reasoning

The Reasoning Scale of the CRPQ refers to the use of reasoning and explanation about the need for rules or punishment (Sanson, 1996). Interestingly, some of the items on this scale which are seen as effective parenting practices (e.g., I talk it over and reason with my child when he/she misbehaves), are considered ineffective anchors on the Parenting Scale (Arnold et al., 1993). Self-reports and observations on the Reasoning Scale demonstrated only modest reliability, but the internal consistency of the self-reports was consistent with that reported by Paterson and Sanson (1999). The moderate inter-observer reliability may reflect the lack of salience of these behaviours, with many parents providing reasons for rules and consequences of misbehaviour indirectly as part of imaginary play during the observations. There was no association between parent-reports and observations on this measure. While the observations provided many opportunities to observe Reasoning behaviour items, these did not correlate with parents' reports on this scale of the CRPQ. This may be because parents' may be interpreting the items on the self-report Reasoning Scale in a different way to which they were operationalised for the observational coding system. As described above, there was lower reliability for this scale than for other parenting domains. If two trained observers return different counts of reasoning behaviour, it would not be surprising if parents' and observers' ratings do not provide agreement (Castorr et al., 1990).

There was very limited evidence for the validity of observations and self-reports on the Reasoning Scale. The observed Reasoning Scale did not display any of the three forms of validity assessed (concurrent, predictive and incremental). Self-reports on the Reasoning Scale demonstrated some concurrent validity, with significant

associations with parent reports of marital satisfaction, and with the global item assessing children's easiness or difficultness. Parents who reported the use of higher levels of Reasoning reported greater relationship adjustment, but surprisingly also reported their children as being more difficult than average. Parents of more difficult children may have more opportunities to use reasoning as they respond to their children's difficult behaviour (Putnam et al., 2002; Thomas & Chess, 1977). Self-reports on the Reasoning scale were negatively associated with teacher-reports of externalising behaviour obtained two years later, but this relationship did not hold after controlling for the initial levels of children's externalising behaviour.

There has been only limited evidence for the importance of parental reasoning behaviour in the development and maintenance of children's behaviour problems (Hart et al., 1992; Paterson & Sanson, 1999). There are discrepancies between reasoning being seen as an effective parenting strategy, whilst a reliance on such talking techniques is seen as ineffective in other research (Arnold et al., 1993; Sanson, 1996). Perhaps talking might be effective with some children (e.g., older children), or in certain situations (e.g., where the importance of household rules need to be explained) but ineffective in other circumstances. Without a context-specific measure, it is difficult to assess this. Further investigation about the utility of studying parental reasoning is needed. Self-reports of this behaviour may suit this purpose, but they appear to have only limited reliability and validity.

9.1.7. Warmth

The Warmth Scale of the CRPQ refers to displays of intimacy and responsivity including physical affection, humour and praise (Sanson, 1996). Self-reports on the Warmth Scale had high internal consistency, whereas observations of the same items showed only modest inter-observer reliability. While the Warmth Scale did

contain some behaviours that were easily operationalised (e.g. hugs or holds child), other behaviour items required more subjective judgements for their coding (e.g., enjoys listening to and doing things with child). This may have influenced the level of inter-observer agreement (Fassnacht, 1982). Also, parenting behaviours which involve warmth and affection, because of their higher natural frequency, may not be as salient as more punitive parenting behaviours (Brandt, 1992; Fassnacht, 1982). High frequencies of Warmth were observed during the free play and drawing tasks during the observations. The Warmth Scale was the only scale that demonstrated a significant association between observed and self-reported behaviour. The size of this correlation was similar to correlations obtained in other studies measuring the relationship between observations and self-reports of responsive parenting practices (Denham et al., 2000; Feinberg et al., 2001; Strayhorn & Weidman, 1988).

There was limited evidence for the concurrent validity of the Warmth Scale. Self-reports of Warmth did not correlate with any of the family risk variables, and observed levels of Warmth were only significantly related to levels of available social support. This association might reflect the influence of parental personality or mental health on both parental warmth and the ability to find social support (Andresen & Telleen, 1992; Crnic & Greenberg, 1990; Jackson et al., 2000).

Observed frequencies on the Warmth Scale were significantly negatively associated with parent- and teacher-reports of children's externalising behaviour obtained two years after the observations were conducted. However, after controlling for the initial levels of children's externalising behaviour problems and self-reports of parenting behaviour, observed Warmth was still negatively associated with teacher- but not parent-reports of children's externalising behaviour. Observations of parental warmth can then add to the explanation of variance in children's externalising behaviour

problems above and beyond that accounted for by self-reported parenting behaviour, and children's initial level of problems.

Self-reports of parental warmth demonstrated high levels of reliability, but did not demonstrate strong validity as assessed in this study. Conversely, observations of parental warmth, while demonstrating lower levels of reliability than other domains of parenting behaviour, did demonstrate concurrent, predictive and incremental validity. A brief free play activity with a variety of toys is a quick way of obtaining an observational measure of parental warmth. It may be necessary to refine the operationalisations of the observed behaviours on the Warmth Scale to increase the degree of reliability between observers.

9.1.6. Obedience

The Obedience Scale of the CRPQ is a measure of parental expectations of obedience (Paterson & Sanson, 1999). It is not so much a measure of parenting behaviour, but more a measure of parental attitudes and expectations. Observations of the items on the Obedience Scale required some subjective judgement on the part of the observer. Reflecting this, the inter-observer reliability on this measure was lower than for the more salient and less subjective measures of parenting. Self-reports on the Obedience Scale demonstrated a level of internal consistency that was consistent with that reported by Paterson and Sanson (1999). Frequencies of observed behaviours on the Obedience Scale were quite low, and this largely reflected that parents had to verbalise their expectations of obedience for the behaviour to be coded, which occurred infrequently. These observed expectations were greatest in the pack-up task, but these frequencies were only slightly higher than in other tasks.

Observed parenting behaviour frequencies for Obedience were associated with parent-reports of children's overall easiness/difficultness. Parents who reported having greater expectations of unconditional obedience tended to rate their children as being more difficult. It might be possible that for some parents, the perception of their child as "difficult" reflects a mismatch between their child's inflexible temperament and parents' expectations for obedience (i.e. poorness of fit) (Prior et al., 2000; Putnam et al., 2002; Smart & Sanson, 2001; Thomas & Chess, 1977).

Observed frequencies of parents' expectations of obedience were significantly associated with teacher-reports of children's externalising behaviour obtained two years later. Observed Obedience accounted for unique variance in teacher-reports of children's behaviour problems, even after controlling for the initial levels of children's misbehaviour and self-reports of on the Obedience Scale. Surprisingly, self-reports of expectations of Obedience were negatively correlated with parent-reports of children's behaviour problems two years later. It is hard to interpret these results given parental expectations of obedience themselves, are unlikely to directly affect children's behaviour (Holden & Edwards, 1989). It is how these expectations are conveyed and enforced (i.e. the behaviour that accompanies these expectations) that will affect how children behave.

The Obedience Scale of the CRPQ has shown limited reliability and validity when measured using observational methods and when using self-reports. It is more useful to use measures of the parenting styles through which expectations for obedience conveyed. Such behaviours might include overreactivity and punitiveness. Observations of behaviours on the Obedience Scale during a pack-up task may be useful for clinicians and researchers wishing to examine goodness of fit relationships between parental expectations and children's temperament. It should be noted that

the inter-observer reliability of the Obedience Scale used in this study was not high, and the frequency of the observed behaviour was quite low. Further research might need to refine the operationalisation of the behaviours on the Obedience Scale or involve participants in an activity in which parents might demonstrate more expectations for their children's behaviour (e.g. a trip to the supermarket).

9.2 Study Limitations and Future Research

This section describes the way in which the results from this study were limited by the features of using self-report and observational methods, and other factors such as sample selection. This section describes ways in which these limitations may affect the interpretation of the results described above, attempts in this study to overcome these limitations, and ways in which they could be addressed in future research. These methods include extending this research to larger samples and different populations and using different methodologies to study parenting behaviour from different perspectives and in different settings.

9.2.1 Sample Selection, Representativeness and Size

The major limitation of this study is the way in which the observation sample was selected. Embedding the project in a larger study of child behaviour, permitted the recruitment of parents and children who displayed a wide range of behaviours. This may not have been possible if a random sample of parents from the screening stage had been included in the project. Despite this advantage, the study design has limited the cross-sectional comparison of child and parent behaviour, because the observation sample was selected on the basis of their screening parenting and child behaviour scores. This may also have influenced the representativeness of the sample in terms of demographic characteristics. In particular, a lower proportion of parents were in paid employment and a higher proportion of children were from sole parent families in the observational component than were in the larger screening sample.

Although the method of recruitment through the preschools was successful, largely in part due to the very enthusiastic assistance of preschool directors and staff, it would be valuable to study the relationships between observed and self-reported parenting

practices in those who refused to participate in either the initial screening stage or the observation component. In studies which involve voluntary participation, the groups who consent to participate often have different characteristics to those who do not, may mean that the results are not generalisable to families with more severe problems (Banister et al., 1996; Minde, 1992). This is supported in the current study by the fact that parents who were selected to participate in the observations but did not participate had significantly higher screening scores on the Laxness and Overreactivity scales of the PS, than parents who did take part in the observations. Unfortunately it was not possible to obtain parenting behaviour data on those who refused to participate in even the screening stage, but it is reasonable to expect that parents who did not complete the screening stage would include parents with more dysfunctional parenting and behaviour problem children (Minde, 1992). However, it should be noted that in the current sample, a wide range of behaviours were observed, including those which may be considered more severe (e.g., physical discipline and yelling), which suggests that the current recruitment strategy was sufficient to yield a diverse sample of parents utilising a variety of parenting techniques. Additionally, the self-report nature of the questionnaire measures meant that parents from non-English-speaking backgrounds were under-represented. To yield a more representative picture, future research could employ methods to recruit families who ordinarily would not participate in parenting research conducted in this way (Minde, 1992). This may include taking a more individualised approach in recruiting families, translating the measures into different languages, or providing incentives for participation in studies of parenting behaviour (such as free parenting programs or child care) (Minde, 1992).

The sample size in this study was limited due to the time-consuming nature of coding the videotapes several times. The smaller sample size hindered the ability to use

more complicated statistical methods and to make comparisons across different observation situations (this is discussed in more detail, below). Increasing the sample size, or using targeted recruitment strategies, would also make it possible to explore the relationships between the methods of measuring parental behaviour for a number of populations. For example, in the current study, only primary caregivers (mainly mothers) were recruited, different study methods could be used to examine the measurement of parenting behaviour separately for mothers and fathers. Mothers and fathers in the same families have been observed to use different parental discipline techniques in similar parenting situations (Harvey-Arnold & O'Leary, 1997; Robinson & Eyberg, 1981), and to provide divergent self-reports of parenting behaviour (Harvey-Arnold & O'Leary, 1997). Future research would benefit from comparisons between self-reports and observations of the parenting behaviour of mothers and fathers in the same family (Denham et al., 2000; Johnson, 2001). Increasing the size of the sample would also allow a more detailed examination of the parenting differences reported between parents of boys and girls highlighted in the current study, and for examining differences in parenting behaviours in the different parenting activities.

Different parenting behaviours are important at different stages of child development (Chamberlain & Patterson, 1995; Patterson & Stouthamer-Loeber, 1984; Roberts & Strayer, 1987; Shaw et al., 1998), and it would be valuable to determine if the same relationships between these different dimensions of parenting behaviour are found in samples of parents of older and younger children.

Large-scale, longitudinal studies of child development which have incorporated, or have the potential to include, mixed-method measures of parenting behaviour provide excellent opportunities to further examine the relationships between self-reports and

observations of parenting behaviour in different populations of interest (e.g., parents from different cultural backgrounds, fathers and mothers, parents of children of different ages). Such studies include the Longitudinal Study of Australian Children (LSAC) in Australia (Sanson et al., 2002), the Early Childhood Longitudinal Survey in the US (U.S. Department of Education. National Centre for Education Statistics, 1997), and the Millennium Cohort Study and the Avon Longitudinal Study of Parents and Children in the UK (Golding, 1996; Smith & Joshi, 2002). Because these studies are longitudinal, the information from these studies could also be used to examine the predictive validity and stability of parenting behaviour measures.

9.2.2 Methods used to Measure Parenting Behaviour

Parents shared a lot of information about their parenting techniques anecdotally during the informal parts of the observation home visits. This information was incorporated in the study through the use of global ratings and qualitative reports of parenting and child behaviour made after the observation visit. Standardised interviews would be a more reliable way to obtain this information (Banister et al., 1996). Future research incorporating standardised interview measures to gather more detailed self-reports of parenting behaviour could then be compared with observational data and questionnaire self-report data. A standardised semi-structured format would help to explore the cross-contextual stability of parenting behaviour and the dynamic relationships between parent and child behaviour (Mrazek et al., 1982). There is also great potential in examining the relationships between such interview data and data from sequential observational coding methods which capture the dynamic and bidirectional nature of parent-child interaction, such as dynamic systems analysis (Granic & Lamey, 2002).

The way in which the parenting and child behaviours were operationalised for the observational coding system does not necessarily correspond to how parents interpret the behavioural items on the questionnaires. For example, the "Inconsistency" and "Permissiveness" ratings used in a coded observation may not be what parents would consider "inconsistent" or "permissive" parenting (Holden, 1983; Mrazek et al., 1982). For example, parents' failure to discipline their children in the no distraction task when he/she moved off-task to look at a book or do some other quiet activity, may not have been considered permissive by parents because their child was being quiet and not requiring parental intervention. Future research which uses parents' self-reports and observations of parenting behaviour before, during and after a parenting program could compare the agreement between methods as parents' awareness and monitoring of ineffective parenting practices increases (Scott et al., 2001). In this way, parents and observers would have increasingly similar perceptions and operationalisations of the parenting behaviours of interest.

9.2.3 Reports of Parenting Behaviour and Family Risk Factors

The reports of family risk factors were all obtained from the primary caregiver, who also provided the parent self-reports. Some of the relationships between the self-reports of parenting behaviour and parent-reports of family risk factors may be explained by common method variance (Podsakoff et al., 2003). Although, the factor analysis performed in Chapter 8 to assess for common method variance suggests that this was not a significant issue in the current study, future studies might benefit from using multiple informants for obtaining reports of parenting behaviour. This would also allow the examination of subjective and objective components of parent self-reports by comparing the reports of parenting behaviour from different informants

(such as the parent, partner, child/ren, and independent observer) (Wade & Kendler, 2000).

To date, the vast majority of research which has utilised child-reports of parenting behaviour has been conducted with samples of children aged six years or older (e.g., Brendgen et al., 2001; Feinberg et al., 2001; Mattanah, 2001; Patterson & Stouthamer-Loeber, 1984; Rey & Plapp, 1990). Shelton et al. (1996) reported that the use of questionnaires and standardised interview techniques to obtain reports of parenting behaviour from younger children (under the age of nine years), do not produce valid data, particularly because young children tend to use extreme response sets when answering questionnaire items. More developmentally appropriate methods have been utilised to collect valid and reliable reports of parent-child interactions from preschool-aged children (Davilla, 1995; Holigrocki et al., 1999; Roberts & Strayer, 1987). These methods include the use of toys and vignettes to prompt children's responses to imaginary scenarios (such as a trip to the zoo) (Holigrocki et al., 1999; Roberts & Strayer, 1987), and the combination of interviewing techniques in conjunction with analysis of children's play and drawings (Davilla, 1995). Also, methods used to collect information from young children about their own behaviour and self-perceptions (including responding to pictorial representations of behaviour and completing stories using teddy bear characters), could be adapted in future research to collect child-reports of parenting behaviour (Mueller, 1996; Valla, Bergeron, & Smolla, 2000).

9.2.4. Context-Specificity of the Observations

Any observation is context-specific, in that the behaviour observed is the behaviour of an individual in a particular situation on a specific day (Brandt, 1992; Metsapelto et al., 2001). In this study, the focus of the study was the relationships between the

behaviour of the primary caregiver and that of the index child, and thus observations examined the behaviour of parents with only those two family members present. This was done to facilitate the coding of the parent-child interactions, because as the number of participants involved in an interaction increases, so does the complexity of coding the interactions (Banister et al., 1996). Contextual differences (such as parenting in different settings, situations, or in the presence of other family members) changes the dynamics of parent-child interaction (Dowdney et al., 1984; Johnson, 2001; Meyers, 1999; Mrazek et al., 1982; Pappas-Jones & Adamson, 1987; Socolar et al., 1999). Therefore the types or intensities of the parenting behaviours observed may be different depending upon the number of people participating in the interaction, their relationships with each other, and the setting of the observation (Johnson, 2001; Metsapelto et al., 2001; Pappas-Jones & Adamson, 1987). Many parents reported the absence of other family members during the videotaped interactions as a reason for why the observations may not have been representative of 'normal' interactions between themselves and their children. A more realistic picture of parent-child interaction (and one that could have corresponded more strongly with parent self-reports of behaviour) may have been obtained using naturalistic observations in which the observation context is minimally structured and in which different family members participate (Bank et al., 1993; Belsky et al., 1995; Brandt, 1992; Dumas & Gibson, 1990; Johnson, 2001; Kalpidou et al., 1998). However, naturalistic observations may mean that it is harder to make comparisons between families, because they may be observed during different activities, and the behaviours of interest may not occur during a naturalistic observation (Dowdney et al., 1984; Mrazek et al., 1982).

This study involved a single observation, but future research would benefit from the use of multiple observations. It is believed that more representative measurements of

observed behaviour are obtained when participants are visited on more than one occasion (Seifer et al., 1994). This is because multiple visits can increase the rapport between the observer and the participants and reduce participant reactivity (Belsky et al., 1997; Belsky et al., 1998; Mrazek et al., 1982; Gardner, 2000; Kalpidou et al., 1998; Youngblade & Belsky, 1995), and the observer can gain some idea of the consistency or stability of parental behaviour across time and in different contexts (Gardner, 2000; Seifer et al., 1994). In this way, measures of test-retest reliability for the observational coding system can be calculated, and there is an increased opportunity to observe less frequently occurring behaviours (Gardner, 2000).

Although, many observational studies have often sampled parenting behaviour in only one situation (Campbell et al., 1986a; Campbell et al., 1986b; Gardner et al., 1999; Shaw et al., 1998; Strayhorn & Weidman, 1988), the present study attempted to overcome the limitation of a single observational visit, by sampling behaviours using four semi-structured activities (Arnold et al., 1993; Kalpidou et al., 1998; Metsapelto et al., 2001; Rothbaum, 1986; Whipple et al., 1995; Youngblade & Belsky, 1995). The activities in which parents and children took part were selected to facilitate the observation of particular behaviours (e.g., warmth in the free-play and drawing activities, and laxness and punitiveness in the pack-up and no distraction activities) and to represent activities that parents and children frequently engage in. As reported in Chapter 5, examination of the frequencies of parenting behaviour during the observations, indicate that parents did behave differently during these different activities. Although imposing structure during the observations promotes comparability across participants (because they are performing similar tasks) and can increase reliability (because varying situational influences are decreased), there is little evidence that what participants do in a more structured situation is representative of what occurs in a more naturalistic setting (Gardner, 2000). It would

be useful to see if the levels of behaviours in the four activities are similarly related to self-reports of behaviour, and if they differentially predict children's behaviour problems. This information would have implications for future observational research designs. For example, is parenting behaviour during the less constrained and more "naturalistic" free play task more strongly related to self-reports of parenting behaviour, than is the behaviour observed in the more structured no distraction task? It was not possible to examine this in this thesis because more participants are required to provide sufficient variability in the frequencies of behaviours across the activities.

In this study, parents were asked to rate their parenting "in general", while the observations examined behaviour in four tasks in a 40 minute period with a stranger present. In light of this, perfect agreement between self-reports and observations would not be expected (Dowdney et al., 1984; Gardner, 2000). One way to improve the comparability of the two methods would be to ask parents to report on their parenting behaviour during the observation period only using a context-specific questionnaire. That is, one that taps into parenting behaviour during the observation period only. A study design could be used that is similar to that developed by Seifer et al. (1994) to study the levels of agreement between parent-reports and observations of child temperament. These authors used repeated observations and parent-reports on both general and context-specific measures of child temperament. They reported larger correlations between the context specific mothers' reports and observations of temperament, than those obtained between the observations and mothers' reports on four widely used questionnaires that assess temperament in general (Seifer et al., 1994).

9.2.5. Frequency-Dependence of the Observations

The coding of the observations of parental disciplinary practices (i.e., laxness, overreactivity, and punitiveness) was dependent on the frequency of children's misbehaviour. That is, children had to misbehave during the observation in order to observe parental responses to this misbehaviour. These observational frequencies were compared with the non-frequency-dependent parental responses to the PS, on which parents indicated the techniques they were more likely to use to handle child misbehaviour, rather than indicating how often they used the techniques. These differences in degree of frequency-dependence may in part account for the lack of concordance between the parent self-reports and observations on the PS items. Using non-frequency-dependent behaviour counts (such as the proportion of discipline exchanges that included the behaviours of interest) or behaviour ratings might overcome this limitation. However, these are not useful if the children do not misbehave at all during the observation period, because the parents of these children would be seen as not using ineffective practices, when there has not been an opportunity to use any practices at all (Janssens, 2004). The structured tasks in the current study were designed to provide these opportunities, but it may be necessary to use repeated observations to see more instances for parental discipline (Bates et al., 1998).

In this study, an attempt was made to statistically overcome this limitation by repeating the analyses in Chapter 7, whilst controlling for the level of child misbehaviour during the observation. This did not affect the size of the correlations between the self-reported and observed parenting behaviour.

9.2.6. Participant Reactivity

As with any observational study, it is very difficult to gauge the participants' reactivity to the process of being observed (Brandt, 1992; Cox, 1975; Gardner, 2000). Varying levels of participant reactivity will decrease the reliability and validity of the results and the veracity of conclusions made from the observed data. In the current study, several steps were taken to reduce levels of participant reactivity and to assess how the behaviour of participants differed from their "normal" behaviour. These included things such as using a non-threatening and non-judgemental approach, and the observations taking place in the context of a larger child behaviour study in which parents consented to participate to further knowledge about the development of behaviour problems (Brandt, 1992; Lytton, 1973). Further, the author engaged in unrelated work whilst the observations were being taped, to reduce the feeling of "being watched" (Dowdney et al., 1984; Mrazek et al., 1982). Many parents commented to the author that after a while they forgot she was there. Children also did not attempt to engage the author as she appeared occupied with another task. The structured nature of the observation also allowed parents and children to concentrate on the activities at hand, rather than focus on the fact that they were being videotaped.

The author recorded how participants reacted to the videotaping component of the observations, and these qualitative reports generally suggested low levels of participant reactivity to this part of the home visit. That is, the author reported that by and large participants' behaviour during the observations seemed representative of their behaviour observed outside the videotaped session. Further, as part of the questionnaire completed during the no distraction task, parents described how, if at all, the observed interactions were different from "normal" interactions between themselves and their children. Very few parents stated the presence of the observer

and video-camera as a reason for a non-normal interaction, but rather reported this was more due to the absence of other family members, the novelty of the situation or the longer amount of time than normal that parents were able to interact with their children. As described in Chapter 6, the analyses of agreement between self-reported and observational data were repeated separately for those parents who reported the interaction as “normal” and for those who did not, with no significant differences being found between the two groups.

9.3 Conclusion

In general, the decision about which research tool is most appropriate for a study is determined by considerations such as the constructs being measured, the instrument's availability and cost, ease of administration, and psychometric properties. In the absence of a "gold standard" measure of parenting behaviour, research that compares the available measurement methods is crucial. The results of this project suggest that a mixed-method approach to studying parental discipline techniques may be yield the best quality data. Self-report methods should be used to measure parental discipline that involves harsh, punitive or overreactive behaviours that are less likely to be observed because they occur infrequently or are subject to social desirability biases. Whereas the measurement of behaviours that reflect permissive and inconsistent parenting techniques yields more valid information when it involves the observations in the context of a pack up or no-distraction task. Future empirical research is needed to determine if these results hold in different contexts, and if the different results yielded by observational and self-report methods is a result of mis- or under-reporting by parents or if the two methods are actually measuring different phenomena. Further, where different researchers are measuring similar constructs, they should employ methods which are not only valid and reliable, but which are consistent across studies to permit the accurate comparison of results obtained in different studies.