

CHAPTER FIVE

METHODOLOGY

“To hone in on characteristics that are universal among and specific to individuals with a particular syndrome or disability, researchers often compare the performance of a target group to that of one or more groups with other disabilities and/or to a group of children who are developing normally” (Mervis & Klein-Tasman, 2004, p. 7).

The previous three chapters looked at the communication characteristics (Chapter two), cognitive processing characteristics (Chapter three) and theory of mind difficulties (Chapter four) seen in PDD and discussed how these different aspects may be related. Furthermore, a model of language processing, a model of cognitive processing and a model of theory of mind, that were felt to be of use in understanding these characteristics, were presented. These chapters provide the theoretical framework underlying the current study. In this chapter the methodology of the study is outlined, with the aims, research design, participants, preliminary investigations, assessment battery used, procedures, analysis of data and inter-rater agreement for rated data being described.

5.1 AIMS OF THE STUDY

The broad aim of this study was to explore the underlying basis of the communication impairments in children with high functioning PDD (HFPDD), compared to children with specific language impairment (SLI) and children with no history of developmental difficulties (NDD). The study had the following specific aims:

- i) To determine whether group differences existed on a number of variables and whether the three groups obtained different profiles on the following groups of variables:
 - Communication - Based on the communication characteristics previously described in PDD, it was hypothesized that the HFPDD group would experience the most difficulty with the semantics, discourse and pragmatic aspects of the

assessment, while the SLI group would experience the most difficulty with the structural aspects of language.

- Cognitive processing - Based on the cognitive processing characteristics of PDD reported in the literature, it was hypothesized that the HFPDD group would experience more difficulty with more central cognitive processes such as planning and attention and may perform well on lower level processes such as successive and simultaneous processing. In contrast it was expected that the SLI group would experience more difficulty with lower level perceptual processes such as successive and simultaneous processing.
 - Theory of mind - Based on the theory of mind literature, it was hypothesized that the HFPDD group would experience marked difficulty on all the theory of mind tasks. While it was expected that the SLI group would experience some difficulty on the theory of mind tasks, particularly those tasks involving more language, it was hypothesized that their difficulties would not be as marked as in the HFPDD group.
- ii) To determine whether particular communication deficits are linked to particular cognitive processing difficulties and whether these are linked to particular theory of mind deficits. It was hypothesized that a relationship may be seen between the semantic, pragmatic and discourse aspects of communication and the planning and attention aspects of cognitive processing, which in turn would also be expected to have a close relationship with theory of mind.
- iii) To determine which measures from the assessment battery best differentiated the three groups. This will have important implications for clinical assessment.

5.2 **RESEARCH DESIGN**

A standard-group comparison design was utilized. In this design, groups formed on the basis of one dependent variable are compared according to the same or other dependent variables (Hegde, 1987). Standard group comparisons do not involve manipulation of an independent variable, but rather just involve measurement of the dependent variable (Hegde, 1987).

5.3 **PARTICIPANTS**

Children between 5.0 and 7.11 years of age were included in the study. It was decided to use younger rather than older children in order to minimize the effect of developmental maturation and the effects of having undergone different interventions. These effects have previously been highlighted to affect results in studies in children with PDD (Lord & McGee et al., 2001). Children from 5 years of age were chosen, as this has been found to be the youngest age at which a variety of cognitive behaviours can be reliably assessed (Das et al., 1994). Certain aspects, such as planning abilities (Das et al., 1994) and certain theory of mind abilities (Baron-Cohen & Swettenham, 1997), have not fully developed by the age of 4 years. Furthermore, the results of a study by Wetherby et al. (1998) suggest that deviances in verbal pragmatics become most evident after 4 years of age in the PDD population. In a study by Schmaman (1998), 5 years was regarded by a number of professionals working in the field of PDD as being a reliable age for making a diagnosis of Asperger's syndrome.

5.3.1 Selection criteria for the three groups

As mentioned above three groups of participants were included in the study. Twenty-six participants were included in each group. The groups were as follows:

5.3.1.1 Children with high functioning PDD (HFPDD group)

It has been suggested that it is useful to study high functioning individuals with PDD as the effects of other confounding variables such as mental retardation are ruled out (Goodman, 1989; Rutter, 1996). By studying high functioning individuals one is able to study the nature of autism in a purer form (Goodman, 1989; Rutter, 1996).

Children for the HFPDD group were included in the study if they had previously been diagnosed with one of the following:

- High Functioning Autism (HFA)
- Asperger's syndrome
- Pervasive developmental disorder not otherwise specified (PDD-NOS), although a general diagnosis of PDD was also accepted, if the child was felt to fall on the high end of the spectrum.

- Semantic-pragmatic disorder (SPD) with PDD features

The children diagnosed with autism (HFA) needed to have been identified as being on the high end of the spectrum and needed to be verbal. The diagnoses of Asperger's syndrome, HFA and PDD-NOS (or PDD on the higher end of the spectrum) ought to have been made by a neurodevelopmental paediatrician, psychiatrist, psychologist or a team of professionals working in the area of PDD. In addition, children diagnosed as having SPD were recruited through speech therapists experienced in working with children with PDD. To be included in this group the child's primary communication difficulty needed to be a semantic-pragmatic one, although other communication impairments could be present. Furthermore, these children needed to present with features suggestive of PDD (i.e. reported social impairment and ritualistic behaviour). A study by Schmaman (1998) indicated that within the South African context paediatricians, psychiatrists, psychologists and speech therapists working in the area of PDD all play an important role in making a diagnosis. According to a study by Mahoney et al. (1998, as cited by Towbin, 2005) experienced clinicians are able to recognise PDD reliably from other disorders (showing 91% agreement).

A short description of each of the groups included on the high functioning end of the PDD spectrum for the purposes of this study follows:

- High functioning autism (HFA) – These children are taken to meet the diagnostic criteria for autism but to have an IQ of 70 or above (Baron-Cohen et al., 2005). Diagnostic criteria for autistic disorder according to *DSM-IV* (American Psychiatric Association, 1994) include impairments in three different areas: impaired social interaction, impaired communication, and restricted, repetitive and stereotyped behaviours, interests and activities. Impairments in social interaction include poor use of non-verbal behaviours (such as eye contact, facial expression and gesture) to regulate social interaction, failure to develop appropriate peer relationships, lack of sharing interest or enjoyment with others and lack of social and emotional reciprocity (Smith & Damico, 1996). Impairments in social communication include delays in spoken language development, impaired ability to initiate and sustain conversation, stereotypical and repetitive use of language and poor development of pretend and social imitative play (Smith & Damico, 1996). Restricted interests include pre-occupation with stereotyped and restricted interests, adherence to non-functional routines and rituals, stereotyped, repetitive motor mannerisms and preoccupation with

parts of objects (Smith & Damico, 1996). Children with HFA are generally verbal but have difficulties with the content, grammar, conversational and non-verbal aspects of communication (Smith & Damico, 1996). Only verbal children were included in the current study.

- ii) Asperger's syndrome - Children with Asperger's syndrome do not have a significant delay in receptive or expressive language or cognitive development (Baron-Cohen et al., 1995). According to the criteria for the diagnosis of Asperger's syndrome, individuals with this disorder should have acquired words by 2 years and phrases by 3 years (Klin et al., 2005). This would mean that a child could have a mild delay in language milestones and still be diagnosed as having Asperger's syndrome. Motor milestones may be delayed with motor clumsiness being common (Klin et al., 2005). These children often have isolated special skills related to their pre-occupations (Klin et al., 2005). Qualitative impairments of social interaction occur (Klin et al., 2005). The child demonstrates restricted, repetitive, stereotyped patterns of behaviour or activities or has an unusually intense, circumscribed interest (Klin et al., 2005). Pragmatic deficits are severe, while semantic and syntactic deficits may or may not occur (Smith & Damico, 1996).
- iii) Pervasive developmental disorder not otherwise specified (PDD-NOS) - These children show similar features to children with autism but fail to meet the full criteria for autism. When compared to children with autism, these children show more differentiated cognitive development, social relatedness and communication skills. Their eye contact and affective engagement is usually of a higher order than children with autism. A restricted range of interests and levels of attachment may or may not be present. Delays in both receptive and expressive language usually occur, as do over-sensitivities to sensory stimuli. Deficits in visual and spatial processing may occur. Rote capacities such as reciting numbers may be well developed and hyperlexia may be present (Smith & Damico, 1996).
- iv) Semantic-pragmatic disorder (SPD) with features of PDD – SPD is characterized by receptive language difficulties, generally related to semantic language problems, with particular difficulties with non-literal language being noted and expressive deficits related particularly to the semantic and pragmatic aspects of language (Gagnon et al.,

1997). Although not their primary difficulty, some problems in the areas of syntax and phonology may be present (Firth, 1999). Their conversational skills are poor and they often use language in an inappropriate manner (Gagnon et al., 1997). Language is often used in a stereotypical and repetitive manner (Gagnon et al., 1997). Language milestones may be delayed (Gagnon et al., 1997). In addition to their communication difficulties, poor social skills, and difficulty in the area of imaginative play and with the non-verbal aspects of communication have been noted (Shields et al., 1996, as cited by Gagnon et al., 1997). In addition, some restricted repetitive and stereotypical patterns of behaviour may occur and there may be mild tendencies to ritualistic and obsessional behaviour (Gagnon et al., 1997).

Children with HFA and SPD perform similarly on various neuropsychological tests (Gagnon et al., 1997), as well as on theory of mind tasks (Botting, 1998). A diagnosis of SPD is sometimes given by some professionals who are reluctant to give a diagnosis of autism in mildly affected individuals, due to the connotations of the term autism (Gagnon et al., 1997). SPD is more frequently found in PDD than any other condition (Rapin & Allen, 1998). According to Rapin (1995, as cited by Gagnon et al., 1997) it is usually seen in individuals on the high functioning end of the PDD spectrum and is seldom seen in other conditions. Some see SPD as a description synonymous with PDD-NOS (Towbin, 2005). SPD's inclusion into the category of PDD is, however, controversial and, more recently, Bishop (2000) suggested that rather than SPD the term pragmatic language impairment should be used. Bishop's (2000) description of pragmatic language impairment appears to be broader than the description of SPD and according to Bishop (2000) pragmatic language impairment can co-occur with PDD or language impairment. In the current study, in order to ensure that the children with SPD who were included fell on the PDD spectrum rather than the language impaired spectrum, the narrower description of SPD rather than the broad description of pragmatic language impairment was followed. Furthermore, features of PDD, including the triad of impairments, needed to have been reported in the SPD subjects.

Selection criteria for participation in the study were kept broad as in the South African context children do not generally need to receive a diagnosis of PDD to receive access to services. Professionals, therefore, appear to be more reluctant to attach a firm label to a child

and often rather discuss a child in qualitative terms or give a broad label to a child, for example saying that the child presents with a pervasive developmental disorder and appears to fall on the higher end of the spectrum. Bishop (1989) showed that children with HFPDD often received a different diagnosis depending on the professional who made the diagnosis. It was, therefore, decided that it would be too limiting to include children with only one diagnosis (for example, Asperger's syndrome) and that all the different diagnoses included on the high end of the spectrum should be included. One of the limitations of recruiting participants in this way was that the HFPDD group might not have been homogenous and as a result this could affect the extent to which the results can be generalised to the PDD population as a whole. Also it was not possible to confirm the validity of the diagnoses made across a range of professionals. One of the benefits of recruiting participants in this way, however, may be that differences in how children on the HFPDD spectrum process information may be more readily seen if group differences exist. Furthermore, by not having strict diagnostic criteria, it would seem that children on the very high end of the spectrum could be included (who may have been excluded if stringent diagnostic criteria had been followed). It was hoped that this would assist in capturing PDD in its purer form. Twenty- six children met these criteria and were included in the HFPDD group. The different diagnoses that had been given to participants in the HFPDD group are included in table 5.1.

Table 5.1: Previous diagnoses given to participants in the HFPDD group

Diagnosis/Description	n
HFA/Autistic disorder on the high end of the spectrum	6
Asperger's syndrome	7
Asperger's syndrome/HFA	1
PDD-NOS/PDD on the high end of the spectrum	3
SPD with PDD features	9

5.3.1.2 Children with specific language impairment (SLI group)

These children were chosen as a comparison group as they have selective impairments in mastering language but are developing normally in other respects (Bishop, 2000). These children often appear to have disproportionate difficulties with language structure (Bishop, 2000; Kamhi, 1996). It is useful to compare children who experience difficulty with the structural aspects of language (i.e. phonological and syntactical deficits) with children who experience semantic and pragmatic difficulties as they are contrasted well in terms of their

general communication profiles (Shields et al., 1996a, 1996b). As their communication profiles differ, it was hypothesized that their underlying cognitive processing and theory of mind profiles would differ too.

Children from the SLI group were recruited from speech-language therapists experienced in working with children with different types of language impairment. Speech-language therapists were given the following selection criteria as guidelines:

- A significant delay in language skills in the presence of normal hearing and relatively normal performance in other areas tapping general cognitive ability, such as non-verbal ability and play skills (Rescorla & Carlson Lee, 2000).
- No obvious signs of neurological damage (Rescorla & Carlson Lee, 2000).
- The child's main difficulty should be with the structure or form of language, including difficulty with phonology, morphology, grammar and vocabulary acquisition (Kamhi, 1996; Owens, 1999), particularly with grammatical morphology (Kamhi, 1996). Other communication impairments could be present as structural aspects may also influence pragmatic ability (Craig, 1995). However, in relation to their grammatical language difficulties, the children with SLI's conversational abilities should basically be intact and they should be purposeful and responsive communicators (Kamhi, 1996), i.e. any pragmatic language difficulties should be seen as secondary to their difficulty with the structure of language and not as their primary difficulty (Owens, 1999).
- The absence of significant social impairment and the absence of ritualistic patterns of behaviour, to rule out the presence of PDD (Rescorla & Carlson Lee, 2000).

5.3.1.3 Children with no history of developmental difficulties (NDD group)

These children were mainly recruited from teachers at mainstream schools. They needed to have met the following criteria:

- A parent and teacher report of no significant history of developmental difficulties.
- A teacher report that the child was coping well in all aspects of development, particularly in the areas of communication, social skills and play.

5.3.2 Matching

Children from the three groups were individually matched to each other according to chronological age. Shields (1991) and Shields et al. (1996a, 1996b) demonstrated that when comparing children with SLI, semantic-pragmatic communication impairments and children with normal development, chronological age was a useful matching device. Developmental research has indicated that during the pre-school years children acquire new skills within the space of a few months (Ames, Gillespie, & Haines, 1980). In an attempt to control for this, participants were matched by age within three months of each other. Each matched triad's age, therefore, had to fall within a three-month interval. The mean ages and standard deviations for the three groups are included in table 5.2.

Table 5.2: Mean ages and standard deviations for the three groups

Age	PDD group	SLI group	NDD group
Mean	6.2 years (74.7 months)	6.2 years (74.7 months)	6.2 years (74.7 months)
SD	9.0 months	7.9 months	8.3 months

Participants were also matched according to gender, as the majority of children with HFPDD are boys, with a ratio greater than 4:1 in this group being predicted (Bryson, 1997). In this study in each group there were 21 boys and 5 girls, giving a ratio of 4.2:1. Groups were not matched on language ability or cognitive ability, as these were dependent variables in the study. It has been suggested that it is not useful to match groups on these variables when delineating developmental profiles based on these variables (Burack, Iarocci, Flanagan, & Bowler, 2004).

5.3.3 Additional selection criteria

Additional subject criteria included the following:

- i) An attempt was made to include children with normal intelligence within the study. IQ tests were not conducted on the children in order to reduce the amount of testing and as IQ assessments may not be reflective of a child's intelligence in such young children with communicative impairments. Furthermore, traditional IQ measures have been shown not to be an accurate representation of overall cognitive ability,

particularly when difficulties in the areas of planning and attention are anticipated (Das et al., 1994). Children were mainly chosen from settings for the education of children of normal intelligence. Some children were taken from settings for children with a range of intelligence. This mainly applied to the PDD group. These children were taken from the upper classes within these settings. See case history information in appendix A2, A3 and A4 for further information. While intelligence could not be closely controlled for, the children's full scale Cognitive Assessment System (Naglieri & Das, 1997) results gave an indication of their overall intelligence. All the children in the study achieved a full scale standard score of above 70. According to Baron-Cohen et al. (2005) children require an IQ of greater than 70 to be included on the high functioning end of the PDD spectrum.

- ii) All the participants included in the study required English as a first language or English needed to be one of the primary languages spoken at home. For all the participants included in the study, English needed to be the medium of instruction at school.
- iii) The children in the HFPDD and SLI groups needed a report of normal hearing, either from a hearing screening or formal audiological assessment. This was difficult to control for in the NDD group, as a number of these children had never previously undergone a hearing evaluation. A report from their parents and teacher that they had never previously been concerned about the child's hearing was, however, required.

As inclusion of children into the study was based on a more general than on more specific diagnoses, as well as being based on descriptions of the children, particularly in the case of the children with SPD and some features of PDD, detailed case history information was obtained to describe the children included in the study more fully. This was also done in order to be aware of other variables such as birth history, developmental history, medical history, etc. Although these variables were not controlled for, they were documented in order to facilitate interpretation of the data.

Parents were asked to complete a case history form and where necessary this was followed up with a short interview if further information was required. In certain cases parents did not complete the case history form but gave the researcher consent to access this information

from the child's school/clinic file. For a limited number of participants, parents did not complete a case history form and no case history information was available in the child's school file. This mainly occurred for a small number of the NDD participants. An example of the case history form, as well as case history information for the three groups, is presented in appendix A1.3, A2, A3 and A4. Please see this for further information and description of participants.

5.4 **PRELIMINARY INVESTIGATIONS OF THE RESEARCH BATTERY**

A preliminary assessment battery was devised and administered to one child with HFPDD (with a diagnosis of Asperger's syndrome), one child with SLI and one child with NDD. All three children were male, were 5.11 years at the time of the assessment and had normal hearing. The aim was to determine whether the proposed assessment battery was able to distinguish between a child with HFPDD, a child with SLI and a child with NDD. A further aim was to revise the assessment battery based on the usefulness of the measures. Results of this pilot study were presented at the South African Speech, Language and Hearing Association Conference held in June 1999 (Shaw-Ridley, Penn, & Rosenthal, 1999). A summary of these results follows.

5.4.1 Assessment battery

The assessment battery consisted of the following:

5.4.1.1 Communication assessment

This consisted of the following measures:

- Receptive language - The Test of Auditory Comprehension of Language - Revised (TACL-R) (Carrow-Woolfolk, 1985) and the Linguistic Concepts sub-test from the Clinical Evaluation of Language Functions – Revised (CELF-R) (Semel et al., 1987) were used to assess receptive language. In addition tasks assessing the child's understanding of deixis were carried out. These included tasks looking at the child's comprehension of 'here' versus 'there' (adapted from Charney, 1979), 'this' versus 'that' (adapted from Webb & Abrahamson, 1976) and 'I' versus 'you' (adapted from Loveland, 1984).

- Expressive semantics - The Word Finding Vocabulary Scale (Renfrew, 1988), the Oral Vocabulary sub-test from The Test of Language Development – Primary (TOLD – P) (Newcomer & Hammill, 1988) and the Test of Problem Solving (TOPS) (Zachman, Jorgensen, Huisinigh, & Barrett, 1984) were administered to assess expressive semantics.
- Expressive grammar - Renfrew's (1991) procedure for scoring sentence length and procedure for scoring subordinate clauses were followed to provide an indication of grammatical complexity.
- Narrative discourse - Narrative discourse was assessed by getting each child to sequence and relate a mechanical, behavioural and intentional story, following Baron-Cohen et al.'s (1986) procedure and analysis. Stories were scored according to whether they were sequenced correctly and whether the child's narration fell into one of the following of Baron-Cohen et al.'s (1986) categories: descriptive, causal or mental state.
- Conversational discourse/pragmatics – A sample of conversational discourse of approximately 35 minutes was obtained using the sampling procedure suggested by Creaghead (1981), Joffe (1990) and Lucas (1980), which consisted of both communicative temptations and discussion topics. Both a macro- and micro-analysis was carried out on the results. The macro-analysis consisted of scoring the child's pragmatic behaviours according to The Pragmatic Profile (Prutting & Kirchner, 1987) using the 5-point rating scale suggested by Penn (1988) for assessing pragmatic behaviours. The micro-analysis consisted of Bishop and Adam's (1989) judgement of inappropriacy analysis. The percent of inappropriate utterances was recorded. Inappropriate utterances were then also categorised into various different categories of inappropriacy, according to Bishop and Adam's (1989) guidelines.

5.4.1.2 Cognitive processing assessment

The Cognitive Assessment System (CAS) (Naglieri & Das, 1997) was used to assess the children's planning, simultaneous processing, attention and successive processing.

5.4.1.3 Theory of mind assessment

This consisted of the following:

- Understanding perception/perspective taking – This included Dawson and Fernald's (1987) visual perceptual role taking and conceptual role taking tasks.
- Understanding the mental significance of the eyes – This included Baron-Cohen and Goodhart's (1994) seeing leads to knowing task, Baron-Cohen et al.'s (1995) eye pointing task and Baron-Cohen et al.'s (1995) thought detection task.
- Unexpected belief – This consisted of Baron-Cohen et al.'s (1985) unexpected location task/the Sally-Anne experiment and Gopnik and Astington's (1988) unexpected identity task.
- Deception – This consisted of two tasks, the first based on Sodian and Frith's (1992) deception versus sabotage game and the second consisting of Baron-Cohen's (1992) penny hiding game.
- Imagination and pretence – Each child's teacher was asked to rate the child's imagination, pretence and interaction with other children on a visual analogue scale.

5.4.2 Findings

The child with HFPDD was noted to present with significant semantic and pragmatic language difficulties. The grammatical language difficulties that were noted appeared related to difficulty with the conceptual aspects of grammar, rather than due to a structural grammatical sequencing difficulty. The child with SLI's greatest area of difficulty appeared to be in the area of expressive grammar, although some semantic and pragmatic difficulties were also evident. The quality of these was, however, different when compared to the child with HFPDD. The control participant coped well with all the communication tasks. The child with HFPDD presented with successive and simultaneous processing skills in the superior range (above high average) but with planning and attention abilities in the significantly below average range. Both the child with SLI and the child with NDD presented with more even cognitive profiles. The child with SLI's successive processing and planning fell in the low average range, while his simultaneous processing and attention fell in the average range. The participant without a history of developmental difficulty's processing across all four areas fell within the high average or average range. Both the HFPDD child and the child with SLI experienced some difficulty on the theory of mind tasks. The child with HFPDD,

however, experienced greater difficulty and the quality of the difficulties that were experienced was noted to be different. The participant with NDD coped well with all the theory of mind tasks. It appeared that the child with HFPDD's difficulties with planning and attention, markedly uneven cognitive profile, as well as his greater difficulty on the theory of mind tasks could explain a number of his communication difficulties.

Based on the preliminary investigation it was felt that a number of adaptations to the assessment battery needed to be made, based on the following:

- While the TACL-R was felt to be a useful measure to administer, it was found to take long to administer in relation to the information that it provided. Furthermore, the vocabulary sub-test was not felt to give a comprehensive enough idea of receptive vocabulary.
- The Linguistic Concepts sub-test was found to be a useful measure but to take long to administer in relation to the information that it provided.
- The receptive measures of deixis were found to be cumbersome to administer and the children were sometimes unclear in their responses.
- All the expressive semantic measures used were found to be useful. In addition, it was felt that a measure of expressive deixis or pronoun shifting would provide valuable information. Furthermore, while the TOPS was felt to be a useful measure, it was felt that some adaptations to its administration needed to be made. This will be discussed later under the revised assessment battery used.
- The grammatical measures used were found to be useful but were felt to be superficial measures of grammar and it was felt that a more comprehensive analysis of grammar was needed.
- Analysing story telling by sequencing and narration types (mechanical, causal or mentalistic) was found to be useful but was not felt to provide enough of an in-depth analysis of narratives. It was felt that in particular, the coherence of the children's narratives needed to be examined.
- The macro-analysis of pragmatics was found to be useful. However, it was felt that for the purposes of this study, some adaptations to Prutting and Kirchner's (1987) scale should be made. This mainly included combining certain of the items and adding certain items. This will be discussed in more depth later. The micro-analysis of pragmatics was felt to be too time consuming for a group study and it was, therefore, decided not to carry this out. Furthermore, the three different participants

performed differently on the macro-analysis of pragmatics, so that it was believed that carrying out a micro-analysis in addition to this was not necessary.

- From the pragmatic analysis, it was felt that a 35-minute conversational sample was not required and that a 15-minute sample was adequate. It was decided that it was not necessary to carry out such structured tasks as communication temptations (as suggested by Creaghead, 1981; Joffe, 1990 and Lucas, 1980) but that discussing pictures about topics familiar to children of this age was adequate in obtaining a sample of the children's pragmatic abilities.
- The Cognitive Assessment System (Naglieri & Das, 1997) was found to be a useful measure of cognitive ability with different profiles being obtained for the three different participants. This was, therefore, kept in the revised battery.
- The perceptual visual role taking task of understanding perception was found to be useful. However, it was felt that the conceptual task of understanding perception did not add that much extra information to the other theory of mind tasks and it was, therefore, decided not to include this task in the revised battery.
- All the mental significance of the eyes tasks were found to be useful. It was, however, felt that some changes to these tasks needed to be made. These included increasing the number of items for each task, increasing the field of pictures presented on one page in the eye pointing task to six on a page and changing the pictures used on the thought detection task to actual photographs.
- The unexpected belief and unexpected location tasks were found to be useful and both these tasks were included in the revised battery. Two trials (rather than one) of each of these tasks were, however, included in the revised battery. As both the SLI and NDD participants passed both these belief tasks without difficulty it was felt that a more difficult belief task should also be included in the battery, particularly as children older than 5.11 years were going to be included in the group study and in order to show the effect of development. It was, therefore, decided also to include a second-order false belief task in the revised battery.
- As Sodian and Frith's (1992) deception task required a verbal response, it was decided to change the deception task to Sodian's (1991) deception task. This task also assessed deception in a similar manner to Sodian and Frith's (1992) deception task but instead of a verbal response only a pointing response was required. It was felt that the children's responses in the pilot study on the penny hiding game (Baron-Cohen, 1992) were not always clear and were sometimes ambiguous. Furthermore, it did not

seem to add additional information to that provided by the initial deception task. It was, therefore, decided to exclude this task from the revised battery.

- It was felt that a more comprehensive evaluation of pretence was required than just an evaluation of the child's general pretence on a rating scale by the child's teacher.

5.5 ASSESSMENT BATTERY USED

The revised assessment, which was used in the group study, consisted of three batteries of measures:

5.5.1 Assessment of communication skills

This consisted of both measures of receptive language and expressive language.

5.5.1.1 Receptive language

a) Receptive semantics

- **The British Picture Vocabulary Scale: Second edition (Dunn, Dunn, & Whetton, 1997)** - This measure aims to assess receptive (hearing) vocabulary in English (Dunn et al., 1997). All the items included in this measure for this age group appeared to be words that are commonly used in South African English.
- **The Basic Concepts sub-test from The Clinical Evaluation of Language Fundamentals – Pre-school (Wiig, Secord, & Semel, 1992)** – This measure is designed to assess more abstract word meanings (Wiig et al., 1992). All the words included in this measure appeared to be words commonly used by South African speakers of English.

b) Receptive grammar

The Grammatical Understanding sub-test from The Test of Language Development – Primary : Third edition (TOLD-P:3) (Newcomer & Hammill, 1997) - This measure assesses a child's ability to comprehend sentences (Newcomer & Hammill, 1997). The word "pitch" (part of item 17 on the Grammatical

Understanding sub-test) was changed to “throw” as “pitch” is not a word commonly used in South African English to indicate “throw”. The adapted item 17, therefore, read “He is going to throw”.

c) Receptive pragmatics

Understanding conversation - The pragmatic assessment included two items that evaluated the child’s comprehension skills within context. These were combined to make up the understanding conversation measure. Please see a description of this in appendix B4 under I: Comprehension.

5.5.1.2 Expressive language

a) Expressive semantics

The expressive semantic assessment was designed to follow a path of greater semantic complexity or depths of meaning. Norris and Hoffman (1993) suggest graded variety of levels of meaning and their levels were used as a guide.

- Labelling

The Word Finding Vocabulary Test: Fourth edition from The Renfrew Language Scales (Renfrew, 1995) – This is a measure of expressive vocabulary and word retrieval (Renfrew, 1995). All the items consisted of words used in South African English.

- Pronoun alternation task

An assessment of first versus second person pronoun use was included in the study, as children with PDD have been shown to experience significant difficulty with first (e.g. “I”) versus second (e.g. “you”) person pronouns (Jordan, 1989; Lee et al., 1994). An adaptation of the pronoun alternation task used by Tanz (1980) was carried out. In this experiment the child was required to ask a third person (a puppet) a question which required him/her to alter a pronoun from either a first, second or third person pronoun to a first, second or third person pronoun. For example, “Ask Roger what his favourite colour is?” needed to be changed into something similar to “Roger, what is your favourite colour?”. Only pronoun shifting and not the grammar the child used

was scored. See appendix B1 for further information on administration and scoring of this task.

- Description

The Oral Vocabulary sub-test from the TOLD:P-3 (Newcomer & Hammill, 1997) – This measure assesses a child’s ability to give oral definitions to common English words (Newcomer & Hammill, 1997). All the words used for this age group consisted of words used in South African English.

- Interpretation, inference and evaluation

Modified Test of Problem Solving (TOPS-M) - The Test of Problem Solving (TOPS) (Zachman et al., 1984) was administered. However, results from the pilot study indicated that for the TOPS participants tended to require more encouragement than was allowed in the administration procedure. This included saying to the children “Tell me more”, “Think carefully”, pointing at relevant parts of the pictures and rephrasing questions, when it was apparent that the child had not understood what had been asked. The norms presented in the test manual could, therefore, not be used, as the TOPS had been administered in a slightly different manner. The scoring guidelines were, however, used. Due to the changes in the administration of the TOPS, the TOPS will now be referred to as the modified TOPS (TOPS-M).

b) Expressive grammar

A fifteen-minute sample of language was obtained from a conversational interaction between the participant and researcher, using picture stimuli of photos of everyday situations of interest to young children, such as a girl and her mother making chocolate pudding, a man next to a broken-down car and children at a birthday party. Following the procedure used by Adams and Bishop (1989), the conversational samples involved discussion of each situation, with the aim of eliciting discussion of the participant’s own personal experiences of the pictured or related situations. The samples were recorded on audiotape, transcribed orthographically and then analysed.

The samples were analysed using the following measures:

- The number of terminable units (T-units) – A T-unit consists of one independent clause and all the dependent clauses that modify it (Hunt, 1965, as cited by Scott, 1988). Main clauses beginning with coordinating conjunctions such as “and”, “but”, “or” etc. were taken to indicate a new T-unit unless co-referential subject deletion occurred in the second clause (Scott, 1988). The sentence “Jane went to the store and bought some Coke” would, therefore, be one T-unit, while the sentence “Jane went to the store and she bought some Coke” would be regarded as two T-units (examples taken from Scott, 1988, p.55).
- Number of dependent clauses – These are clauses that are dependent on the independent clause. An independent clause is made up of the noun phrase and verb phrase and represents the main clause. Dependent clauses include subordinate clauses and non-finite clauses (Chapman, Levin, Wanek, Weyrauch, & Kufera, 1998).
- Mean length of utterance (MLU) – This was taken as the number of morphemes per T-unit. This has been suggested by Paul et al. (1996) to be a useful measure of grammar. A morpheme has been defined as “a minimal meaningful unit of a language: for example, *dog* or plural-*s*” (Chapman, 1981, p. 24). The guidelines provided by Chapman (1981) for counting morphemes were followed. These included:
 - 1) Repetitions due to non-fluency were counted as one morpheme. However, repetitions of a word for emphasis, e.g. “The lady said no, no, no”, were counted separately.
 - 2) Fillers such as “mm” or “oh” were not counted but words such as “no”, “yes” and “hi” were.
 - 3) Compound words (for example, “birthday”), proper names (for example, “John Black”) and ritualized reduplications (for example, “choo-choo”) were counted as single words, as there is no evidence that the child sees these as separate morphemes.
 - 4) Irregular past tense verbs (for example, “got”, “did” and “saw”) were counted as one morpheme, as no evidence exists that the child relates these to the present form.
 - 5) Diminutive forms (for example, “doggie”) were counted as one morpheme as the child does not use the suffix productively here.

- 6) Auxiliaries (for example, “is”, “can” and “must”) were counted as separate morphemes.
 - 7) Catenatives (for example, “gonna”, “wanna” and “gotta”) were counted as single morphemes, as no evidence exists that they function as anything differently for the child.
 - 8) Inflections, for example, the possessive “s”, the plural “s”, the regular past tense “-ed” and present progressive “-ing”, were each counted as separate morphemes.
- Complex grammatical clauses used - Each dependent clause was also analyzed according to the categories at stage V and stage VI of the Language Assessment, Remediation and Screening Procedure (LARSP) (Crystal, 1991). Dependent clauses were categorised into one of the following categories:
 - 1) **Coord 1** – This was reserved for coordinating clauses where co-referential subject deletion occurred in the second clause (Scott, 1988). In main clauses beginning with coordinating conjunctions where co-referential subject deletion did not occur in the second clause, these clauses were taken to indicate a new T-unit, according to Scott’s (1988) definition of a T-unit. For example in the utterance, “I get all the food out and chop the potatoes”, “chop the potatoes” would be regarded as the dependent clause. If the child had said “I get all the food out and I chop the potatoes”, then “and I chop the potatoes” would be regarded as a new T-unit and not as a dependent clause.
 - 2) **Coord 1+** - This was reserved for coordinating clauses where co-referential subject deletion occurred in the third clause (Scott, 1988), as described above for Coord 1, for example, “I get all the food out and chop the potatoes and cook them”.
 - 3) **Subord A 1** – Here the dependent clause consisted of an adverbial clause. For example “He arrived when it was dark” with “when it was dark” being the adverbial clause (example taken from Crystal, 1991, p. 34).
 - 4) **Subord A 1+** - This occurred when two dependent adverbial clauses occurred in the same T-unit. For example, “He arrived when it was dark and when it was raining” with “when it was dark” and “when it was raining” being the two adverbial dependent clauses (example taken from Crystal, 1991, p. 34).
 - 5) **Subord S** – Here the dependent clause formed the subject of the sentence. For example, “What I said was important” with “What I said” being the subject clause (example taken from Crystal, 1991, p. 35).

- 6) **Subord C** – Here the dependent clause formed the complement of the sentence. For example, “That is what I wanted” with “what I wanted” being the complement clause (example taken from Crystal, 1991, p. 35). Comparative and more advanced complement and “how/what” clauses at stage VI of the LARSP were included under the category Subord C as they both consisted of dependent complement clauses.
 - 7) **Subord O** – Here the dependent clause formed the object of the sentence. For example, “He knew what I wanted” with “what I wanted” forming the object clause (example taken from Crystal, 1991, p. 35).
 - 8) **Postmod clause 1** – This consisted of a clause introduced as part of the noun-phrase structure. For example, “The man who saw me is outside” with “who saw me” forming the postmodified clause (example taken from Crystal, 1991, p. 36).
 - 9) **Postmod clause 1+** – This consisted of two clauses introduced as part of the noun-phrase structure. For example, “There’s the car which you drove and which was bumped” with “which you drove” and “which was bumped” forming the postmodified clauses (example taken from Crystal, 1991, p. 36).
- Minor, major and phrasal utterances - These were categorised according to the definition of the LARSP analysis (Crystal, Fletcher, & Garman, 1988; Crystal, 1991) as the following:
 - 1) **Minor utterances** – Minor utterances are utterances whose elements are not able to combine with other elements according to the language’s grammatical rules, for example “yes”, “no”, “oh” or using a name as a signal (e.g. “John” to get the person’s attention) (Crystal, 1991).
 - 2) **Major utterances** – These consist of single words that could be combined with other elements according to the language’s grammatical rules, for example “boy” (noun), “running” (verb) etc., but which were not combined (Crystal, 1991).
 - 3) **Phrasal utterances** – These consist of a string of more than single words, which are not a clause, i.e. they cannot be classified as a T-unit, for example “in the box”, “three dogs”. Only phrasal utterances used, that were not included as part of a T-unit, were included under phrasal utterances. Such utterances were not analysed further but were recoded as a phrasal utterance.

The categories of minor, major and phrasal utterances used in the research gave an idea of the utterances that were not T-units used by the child. They were, however, not analysed further than this.

All samples were analysed by the researcher, who has experience in analysing children's grammar. Where it was uncertain how an utterance should be analysed, this utterance was discussed with other therapists experienced in analysing children's grammar until a consensus on how that utterance should be classified was reached. An example of the scoring sheet used for the grammatical analysis is found in appendix B2.

c) Narrative assessment

The wordless picture book "Frog On His Own" (Mayer, 1973) was used to elicit a narrative, following a similar procedure to that used by Tager-Flusberg and Sullivan (1995). This was felt to be an appropriate book as it has been shown to elicit numerous emotion and cognitive terms in spontaneous narratives (Tager-Flusberg & Sullivan, 1995).

The story "Frog On His Own" (Mayer, 1973) is about a frog that goes to the park with his owner, a young boy, together with the boy's other pets. The frog escapes and goes off exploring. The story consists of four episodes, each about the frog's adventures with various minor characters (Tager-Flusberg & Sullivan, 1995).

The researcher went through the pictures in the book once with the child. Then the researcher went back to the beginning of the book and asked the child to tell the story. Unlike in the study by Tager-Flusberg and Sullivan (1995) the child was not required to tell the story to a second experimenter, due to logistical reasons. See appendix B3.1 for a further description of this procedure.

Narrative rating

Narratives were audiotaped, orthographically transcribed and then rated according to:

- Sinoff's (1993) coherence analysis of narratives
- Ulatowska, Freedman-Stern, Weiss Doyel, & Macaluso-Haynes's (1983) clarity analysis of narratives

Coherence analysis of narratives

Sinoff (1993) devised a coherence analysis of narratives where narratives were rated on a five-point scale according to the following parameters:

- 1) Temporal organisation – This refers to how well sequenced the events in a story are presented, i.e. ranging from logical and sequential to arbitrary presentation (Sinoff, 1993).
- 2) Relevance – This refers to the appropriateness of the information and its relationship to the story as a whole (Sinoff, 1993).
- 3) Development of character(s) – This refers to the inclusion of one or more characters at the centre of the story, the extension of these characters and how this contributes to the overall effectiveness of the story (Sinoff, 1993).
- 4) Supporting description – This refers to the setting of the story, included to orientate the reader and enhance the story (Sinoff, 1993).
- 5) Ending – This refers to the outcome or resolution of the story (Sinoff, 1993).

Sinoff's (1993) coherence analysis of narratives was felt to be an appropriate narrative analysis for HFPDD as all these parameters have been shown to be difficult for children in this population. A coherence analysis of narratives such as this may better differentiate children with HFPDD from children with SLI, than previous narrative analyses reported in the research literature have been able to. See appendix B3.2 for further information on the coherence analysis and instructions to the raters.

Clarity rating of narratives

Ulatowska et al.'s (1983) clarity rating of narratives was added to the above coherence analysis of narratives. The clarity rating consisted of only one item, with clarity being rated in response to the question "How comprehensible is the language of the story?" (Ulatowska et al., 1983). This was felt to be a broad measure of cohesion (Henshilwood & Ogilvy, 1999). Cohesion was seen as being expressed through vocabulary and grammar (Henshilwood & Ogilvy, 1999). This aimed to evaluate the language of the story at a sentence level and to evaluate the use of linguistic devices within the story. See appendix B3.2 for a further description of this.

d) Pragmatic assessment

The fifteen-minute conversational sample (elicitation described under the assessment of expressive grammar) was videotaped and rated according to a pragmatic rating scale devised from aspects of Adams and Bishop's (1989), Bishop and Adams's (1989), Penn's (1983, 1988), Prutting and Kirchner's (1983, 1987) and Sonnenberg's

(1995) pragmatic analyses. The pragmatic analysis designed, however, mainly followed Prutting and Kirchner's (1987) Pragmatic Profile. Prutting and Kirchner (1987) devised a comprehensive pragmatic assessment consisting of thirty parameters. This tool was designed to be used with children from 5 years of age, as the developmental literature has indicated that by 5 years children show some form of the parameters evaluated on this protocol (Prutting & Kirchner, 1987). It was decided that this was an appropriate scale to use for the present study due to both its comprehensiveness, as well as it being developmentally appropriate for the age group being evaluated.

In the Pragmatic Protocol three aspects of pragmatics are evaluated. They are:

- Verbal aspects - This includes parameters such as speech acts, topic, turn taking, lexical selection and stylistic variations (Prutting & Kirchner, 1987).
- Paralinguistic aspects – This includes aspects such as intelligibility, vocal intensity, vocal quality, prosody, and fluency (Prutting & Kirchner, 1987).
- Non-verbal aspects – This includes aspects such as physical proximity, physical contacts, body posture, foot/leg and hand/arm movements, gestures, facial expression and eye gaze (Prutting & Kirchner, 1987).

On the adapted scale verbal aspects, paralinguistic aspects and non-verbal aspects were also evaluated. However, some of the sub-parameters within these scales were combined to allow for greater ease of rating. Furthermore, some additional parameters were added based on the pragmatic analyses of Adams and Bishop (1989), Bishop and Adams (1989), Penn (1983, 1988) and Sonnenberg (1995). Prutting and Kirchner's (1983, 1987) original scale consisted of a two-point scale. It was, however, felt that this might not differentiate adequately enough between the HFPDD and SLI groups and in the adapted scale a five-point scale was used, ranging from inappropriate to appropriate behaviour. A five-point scale has previously been shown to be effective in rating pragmatic behaviours (Penn, 1988; Sonnenberg, 1995).

The following behaviours were rated on the adapted scale:

- a) Comprehension/understanding conversation
 - i) Comprehension of literal meaning
 - ii) Use of context in comprehension

- b) Verbal aspects
 - i) Speech acts
 - ii) Topic skills
 - Selection
 - Introduction, maintenance, change
 - Coherence
 - iii) Turn taking
 - Exchange structure
 - Repair/revision
 - Pause time/interruption/overlap
 - Quantity/Conciseness
 - iv) Lexical selection
 - Specificity/accuracy
 - Cohesion
 - v) Stylistic variations
 - Communicative sensitivity
- c) Paralinguistic aspects
 - i) Prosody
 - ii) Voice
 - iii) Speech rate
 - iv) Fluency
 - v) Intelligibility
- d) Non-verbal aspects
 - i) Body posture
 - ii) Eye contact
 - iii) Facial expression
 - iv) Body movements

See appendix B4 for a further description of the devised scale and a description of these behaviours, as well as the guidelines for rating given to the raters.

5.5.2 Assessment of cognitive processing

The Cognitive Assessment System (CAS) (Naglieri & Das, 1997) was administered. This is based on the PASS model. It has measures of planning, attention, simultaneous and

successive processing and can be administered to children from 5.0 years of age. This battery is based on measures developed by Das et al. (Das, 1992; Das et al., 1979, 1994, 1996; Das, Mensink, & Mishra, 1990; Naglieri & Das, 1988, 1990) over a number of years.

The CAS consists of four scales, each consisting of a number of sub-tests. Guidelines for administration and scoring of the CAS as outlined in the CAS manual (Naglieri & Das, 1997) were followed. Only the sub-tests relevant for the 5.0 to 7.11 year age group were administered. These consisted of the following:

5.5.2.1 Assessment of planning

The planning sub-tests require the child to create a plan of action, apply the plan, monitor whether the action taken is in line with the original goal and if necessary to modify the plan (Naglieri & Das, 1997). The sub-tests involve tasks that require the individual to make a decision or decisions about solving novel tasks (Naglieri & Das, 1997). They consist of the following:

- a) Matching numbers – This consists of two items, each item consisting of eight rows of numbers, with each row containing six numbers. Two numbers in each row are the same and the child is asked to underline the matching numbers in each row. Numbers gradually increase in digit length across the rows, with there being four rows for each digit length. Each item is timed and has a time limit. The rows of numbers were developed to maximize the benefits of strategy use in identifying correct matches (Naglieri & Das, 1997). Previous research found this sub-test to be related to other planning tests (Naglieri & Das, 1997).
- b) Planned codes – This sub-test consists of two items each containing their own set of codes. At the top of each page is a legend that shows a correspondence of letters to specific codes. Below the legend are seven rows and eight columns of letters that do not contain the codes. The child is required to fill in the codes in the empty boxes under the letters, corresponding to the codes presented in the legend. The correspondence of letters to codes is different for each item. The position of the letters is also different so that on the first item the letters are configured vertically on the page, while on the second item the letters are configured diagonally. A time limit

for each item is given. Planned Codes is similar to other planning codes measures described in the literature (Naglieri & Das, 1997).

- c) Planned connections – This sub-test consists of five items. Each item requires the child to connect numbers in sequential order. Each item is timed, with the test score being the total amount of time in seconds used to complete the items. This sub-test is similar to the Trail Making procedure and has been found to correlate with other planning tests in several studies (Naglieri & Das, 1997).

5.5.2.2 Assessment of simultaneous processing

The simultaneous processing sub-tests require separate elements to be synthesized into an interrelated group. They involve strong visual-spatial and logico-grammatical components (Naglieri & Das, 1997). Two of these sub-tests use only visual content, while one of these sub-tests uses both verbal and visual content.

- a) Non-verbal matrices – This sub-test consists of thirty-three items presented in multiple choice format, where the child is required to choose the best option of six options to complete a diagram. Each item uses a combination of shapes and geometric elements. The child is required to decode the relationships among the parts of the item in order to make his/her choice. Items progress from relatively simple to complex. Items are scored as either correct or incorrect and the administration of the sub-test is terminated when the child fails four items in a row. This sub-test is similar to other matrix tasks and matrix tasks have been used as a measure of simultaneous processing in a number of studies (Naglieri & Das, 1997).
- b) Verbal-spatial relations – This sub-test consists of twenty-seven items requiring the comprehension of logical and grammatical descriptions of spatial relationships. Each item consists of six drawings, involving objects and shapes arranged in different spatial configurations. The child is asked a question and is asked to select the drawing that best matches the verbal description. Each item is presented for a maximum of thirty seconds and each item is scored as either correct or incorrect. The sub-test is discontinued when the child fails four consecutive items (Naglieri & Das, 1997).

- c) Figure memory – This sub-test consists of twenty-seven paper-and-pencil items. The child is shown a page with a two- or three-dimensional geometric figure for five seconds. This is then removed and the child is given a response page, containing the original figure in a larger, more complex pattern. The child is required to identify the original figure in the larger figure by tracing over it. This sub-test is discontinued when the child fails four consecutive items. Drawing from memory and copying designs have previously been used as measures of simultaneous processing (Naglieri & Das, 1997).

5.5.2.3 Assessment of attention

The attention sub-tests require attention to be focused, selective, sustained and effortful. They require a particular stimulus to be detected while responses to irrelevant competing stimuli are inhibited (Naglieri & Das, 1997).

- a) Expressive attention – This sub-test measures selectivity and shifting attention. It involves an interference condition after the administration of items without interference. The child is presented with pictures of common animals and is asked whether the animals depicted are big or small. The child is meant to respond based on the size of the animals in real life and to ignore their relative size on the page. In the first item the animals presented are the same size. In the second item the animals are sized appropriately (i.e. big animals are drawn larger than small animals). In the third item the animals' realistic and relative sizes depicted in the item differ (i.e. an elephant drawn small in relation to a smaller animal). This last item is a measure of selective attention. A ratio score is calculated from a combination of the child's accuracy and the time taken to complete each item. This sub-test has been found in various studies to load on an attention factor (Naglieri & Das, 1997).
- b) Number detection – This sub-test is designed to measure selectivity, shifting attention and resistance to distraction. The child is given a page of numbers (consisting of eighteen rows of ten numbers containing forty-five targets) and is asked to underline the specific numbers that appear at the top of the page. In the first item the child is required to underline the numbers 1, 2, and 3, while on the second item he/she is required to underline the numbers 4, 5, and 6. The child is required to complete the

page by working from left to right and top to bottom and on completion may not go back to check the page. A ratio score is obtained based on the child's accuracy and time taken to complete each item (Naglieri & Das, 1997).

- c) Receptive attention – This sub-test consists of four pages of pictures arranged in pairs. Two conditions are presented. In the first condition the child is required to underline pairs of pictures that are identical. In the second condition the child is required to underline pairs of pictures that are not identical but have the same name (for example, two different trees). Four items are presented, each consisting of fifty pairs of drawings per page with approximately 25% targets per page.

5.5.2.4 Assessment of successive processing

The measures of successive processing assess the individual's ability to integrate stimuli into a specific serial order, with successive processing having strong serial and syntactic components (Naglieri & Das, 1997). These sub-tests require perception and reproduction of the serial nature of stimuli (Naglieri & Das, 1997).

- a) Word series – This sub-test consists of nine single-syllable high-frequency words. The first item consists of a sequence of two words and items gradually increase in length. The words are read at the rate of one word per second and the child is required to repeat the words in the same order that they are presented. A point is given for each correct sequence of words repeated and testing is discontinued when the child fails four consecutive items. Previous studies have accepted the repetition of words and digits as a marker for successive processing (Naglieri & Das, 1997).
- b) Sentence repetition – This sub-test consists of sentences composed of colour words (for example, "The blue is yellowing"). These are read to the child one sentence at a time and the child is required to repeat the sentence exactly as it was presented. Colour words are used to reduce the influence of simultaneous processing by allowing the sentences to contain little meaning. To complete this task successfully an appreciation of the sentence syntax is required. A point is given for each sentence that is repeated correctly and testing is discontinued after the child fails four consecutive

items. This sub-test aims to assess syntactic structure based on the serial relationship between words (Naglieri & Das, 1997).

- c) Speech rate – This sub-test consists of eight items, which are timed. A three word series is read to the child and the child is required to repeat the series until he/she is told to stop. The child is required to repeat each series ten times in a row. Speech rate has been used to measure successive processing in a number of previous studies (Naglieri & Das, 1997).

5.5.3 Assessment of theory of mind abilities

A battery of theory of mind tasks was devised. It assessed important theory of mind skills that should have developed by 5.0 years or which develop during the age range 5.0 to 7.11 years. Tasks were chosen both according to which aspects of theory of mind they assessed, as well as the age they are usually mastered in children without developmental difficulties. Background on these various aspects of theory of mind is provided in Chapter four. The assessment of theory of mind consisted of the following:

5.5.3.1 Visual perceptual role taking

The perceptual role taking tasks used by Dawson & Fernald (1987) with a group of autistic children and by Flavell et al. (1975, as cited by Dawson & Fernald, 1987) and Zahn-Waxler et al., (1977, as cited by Dawson & Fernald, 1987) with groups of normal children were carried out. These tasks assess a child's ability to take another person's perspective from a perceptual point of view, for example understanding that if they are sitting across the table from someone and they have to show that person a picture of someone upside down (while keeping the picture flat on the table) then the picture must be the right way up for them. This consisted of five tasks:

- Upside down person – When seated across the table from the experimenter the child was asked to show the researcher a picture of a boy so that the researcher could see the boy standing on his head (Dawson & Fernald, 1987).

- Face - When seated across the table from the researcher the child was asked to show the researcher a picture of a boy so that the boy in the picture was looking at the researcher (Dawson & Fernald, 1987).
- Spontaneous – The child was asked to choose one of two drawings placed crosswise between the child and researcher and to show it to the researcher (Dawson & Fernald, 1987).
- Cubes – The child and researcher each had an identical cube. The researcher rotated the cube looking at a different picture on each of the four trials. The child was asked to turn his/her cube so that he/she could see the same thing on his/her cube that the researcher was looking at (Dawson & Fernald, 1987).
- Single cubes – One cube was used, with the researcher rotating the cube so that a different picture was presented to the child on each of the four trials. On each trial the child had to name both the picture that he/she could see as well as the picture that the researcher could see (Dawson & Fernald, 1987).

The total score that could be obtained over these 5 tasks was 15. The administration and scoring of each of these is described in appendix B5.

5.5.3.2 Mental significance of the eyes

This consisted of three tasks:

a) Seeing leads to knowing

Baron-Cohen and Goodhart's (1994) 'seeing-leads-to-knowing' task was carried out. Two dolls were used, one who looked in a box and one who did not, and the child was then asked "Who knows what is in the box?". This was carried out five times with five different boxes, giving a total score of 5 points.

b) Eye pointing

These tasks assessed whether the child was able to use eye direction information in determining someone's desires and thoughts. This consisted of an adaptation of Baron-Cohen et al.'s (1995) desire, goal and refer conditions and Lee et al.'s (1998) eye direction condition. In each of these conditions a boy named Larry looked at one of six surrounding pictures. The different conditions were as follows:

Desire condition – Here the child was asked "What does Larry want?".

Goal condition – Here the child was asked “Which one is he going to take?”.

Refer condition - Here the child was asked “‘There’s a beb/don/sen’. Which one does Larry say is the beb/don/sen?”.

Eye direction condition – Here the child was asked “Where is Larry looking?”.

All three conditions were tested using Lee et al.’s (1998) procedure where a possible six answers were provided for each item, to reduce the possibility of scoring correctly by chance. Three different items for each condition were presented, giving a total score of 12 across the four conditions.

The above three procedures, as well as their scoring are described in more detail in appendix B6.

c) Thought detection

An adaptation of Baron-Cohen et al.’s (1995) thought detection task was used. Baron-Cohen et al. (1995) used sixteen pairs of photos and six pairs of line drawings. The current study used eight photo pairs. One of the people in each photo pair adopt a thinking posture. For each photo pair the child was asked “Which one is thinking?”. All the photos where the person adopted a thinking posture had previously been judged by two independent raters to depict ‘thinking’. A total score of 8 could be obtained for this task.

5.5.3.3 Belief

This consisted of two first-order false belief tasks and one second-order false belief task.

a) First-order false belief tasks

In first-order false belief tasks a person’s ability to think about another person’s thoughts about an event is assessed. Unexpected identity tasks similar to those used by Davis and Pratt (1995), Gordon and Olson (1998), Slaughter (1998), Symons, McLaughlin, and Moore (1997) and Zelazo et al. (1996) were used. An unexpected location task (the Sally-Anne experiment), similar to that used by Baron-Cohen et al. (1985), Charman and Campbell (1997) and Scott and Baron-Cohen (1996), was used. The tasks were as follows:

- Unexpected identity task – Here the child was shown buttons in a smartie box and golf balls in an egg box. He/she was then asked what a puppet who had not seen

inside the boxes would think was in the smartie box and was in the egg box. This was scored out of 2.

- Unexpected location task – In the Sally-Anne experiment, a story involving two dolls was acted out to the child. In order to understand the story and answer the questions correctly, the child has to be able to understand the story from both characters' perspective and needs to understand that if a character did not see an object being moved to a new position he/she will still think that it is in the position where he/she originally saw it being placed. This was scored out of 2.

These procedures, as well as their scoring are described in more detail in appendix B7.

b) Second-order false belief tasks

In second order false belief tasks a person's ability to think about what another person is thinking about a third person is assessed. Baron-Cohen's (1989b) "John thinks that Mary thinks that ..." experiment was carried out. Here a story was acted out using model people and a model village. Test questions consisted of a belief question testing the child's second-order belief and a justification question, where the child was required to explain his/her belief. The justification question was scored according to whether the child's answer reflected understanding of the initial location (1 point), the character's first-order belief (2 points) or the character's second-order belief (3 points). A total score of 2 for the belief question and 6 for the justification question could be obtained. The procedure and scoring for this task are described in more detail in appendix B7.

5.5.3.4 Deception

An adaptation of Sodian's (1991) deception task was carried out. Here the child hid a star in one of two boxes. A good king puppet and bad robber puppet were produced and on each trial one of the puppets asked the child where the star was. If the child told the good king puppet where the star was he/she got another star. If the child told the bad robber puppet where the star was the robber stole the star. If the child told him the incorrect location (i.e. deceived him) then he went away with nothing and the child kept the star. A total score of 4 could be obtained on this measure. The administration and scoring of this are described in more detail in appendix B8.

5.5.3.5 Emotions

After the participants had completed telling the story “Frog On His Own” by Mayer (1973) in the narrative task, the experimenter turned to four different pages in the story and asked the child to relate how each character in that picture felt and to explain why they felt that way. This procedure followed that of Tager-Flusberg and Sullivan (1995) in their narrative study using the same book. The first two pictures chosen were the same as those used by Tager-Flusberg and Sullivan (1995). An additional two pictures were also chosen, as they were felt to reflect a number of additional emotions. Pictures used consisted of those depicted on pages 13 (picture 13), 18 (picture 18), 23 (picture 22) and 28-29 (picture 26) of the book “Frog On His Own” (Mayer, 1973). Five speech-language therapists working in child language were asked to write out the emotion for each character shown in the four pictures and were asked to relate why the character felt that way. These five therapists’ responses were used to create a template to score the children’s responses. This template, as well as the administration and scoring of the emotion task is described in more detail in appendix B9.

5.5.3.6 Pretence

Pretence was assessed following an adaptation of Jarrold et al.’s (1996) generating pretend acts measure. The child was required to sit on a chair in the middle of a square marked out by white cardboard. In the square were a number of props. The child was engaged in a game where he/she had to generate how many things he/she could pretend to do in the time given (six minutes). The props could be used but pretend acts without the use of the props were also accepted. Each child was videotaped and scoring was conducted afterwards from the videos. Both the number of instances of true pretence and the number of attempts at pretending were scored. The administration, scoring and instructions to raters for this task are described in more detail in appendix B10.

5.6 **PROCEDURES**

5.6.1 Ethical clearance

Ethical clearance for the research was obtained from the Committee for Research on Human Subjects (Humanities), University of the Witwatersrand, Johannesburg.

5.6.2 Parental consent

All parents/guardians were given an information sheet regarding the study and were asked to complete a consent form before their child was included in the study. Please see examples of these in appendix A1. The researcher made herself available to discuss any concerns the parents had.

5.6.3 Data collection

The aim was to get the best possible results from each child, while still remaining as objective as possible. It was, therefore, not always possible to administer tests in exactly the same way to each individual child and some adjustments sometimes had to be made. These included the following:

- Tests were not always administered in exactly the same order to each child, although the researcher attempted to follow a general format.
- The time of testing sessions was not the same for all participants. The length of testing sessions depended on whether the individual participant appeared to be tiring or not and breaks were given according to each individual child's needs. As a result a different number of testing sessions were conducted with different participants.
- Different levels of reinforcement were given to different children. For example, some of the children from the HFPDD group required tangible reinforcement for every few items they completed in order to keep them going. Children from the NDD group generally required little reinforcement, apart from social reinforcement.
- Practice items were given according to each child's needs. Olswang, Bain, and, Johnson (1992) indicate that younger children generally require more prompts than older children to learn a task. While practice items were given to all participants, they were repeated when participants did not appear to understand the task. If they still did not appear to understand the task, the researcher attempted to teach them the task through a game format, but without using the actual test stimuli. This was done so each child's performance would represent their ability in that area rather than whether or not they had understood what was expected of them. Children were excluded from the study if they repeatedly did not appear to understand the format of activities as well as practice items and example tasks. This resulted in some children with HFPDD

being excluded. It was, however, felt that this resulted in the HFPDD group representing children on the high functioning end of the spectrum.

The researcher carried out the assessment of all the children in the study. The reason for this was the time taken to conduct the testing and that both clinical experience with the PDD population as well as in-depth understanding and competence in administering the battery of measures were required. Unfortunately, the researcher was not blind to each participant's diagnostic group. However, the majority of the measures were scored on pre-determined criteria. In addition to the researcher, a percentage of the data that required rating was rated by two independent raters. These two independent raters were blind to the diagnostic group of the children they rated.

The aim of the assessment with each child was to obtain a profile of that child's strengths and difficulties across the different areas. Testing was conducted to get the profile most representative of each child's abilities. In order to do this testing had to be conducted in a clinical framework. While testing was generally static in nature, some elements of dynamic assessment were included. This included aiming at eliciting the child's best response and, where necessary, teaching the child the nature of a task before it was administered. In order to obtain the children's best responses, the researcher attempted to make the testing as much fun as possible for the children.

5.6.4 Recording of data

The conversational sample was video recorded so that an analysis of the child's pragmatics could be undertaken later. The child's responses on all the measures of expressive semantics, as well as on the measure of expressive narrative ability, were audio taped so that they could later be transcribed. In addition to being video recorded the conversational sample was also audio taped so that the child's utterances could later be transcribed for the grammatical analysis. The theory of mind tasks requiring a verbal response, were both audio and video taped so that the child's responses could later be analysed. The pretence task was video recorded so that instances of pretence and attempts at pretence could later be recorded.

5.6.5 Transcription and scoring of data

Data for the expressive semantic, conversational discourse sample, narrative assessments and data for the theory of mind tasks requiring a verbal response were orthographically transcribed from the audio and video tapes. Scoring of this data, including the grammatical analysis and narrative analysis, then took place from the orthographically transcribed data. Some data could be scored on line, i.e. the receptive language measures and CAS (Naglieri & Das, 1997) measures.

All measures administered were scored in one of two ways:

- Formal language and cognitive tests that were used were scored according to the manuals' scoring for that particular test. Raw scores were converted into all the different scores possible for that particular test (i.e. age scores, standard scores, standardized scores, percentile ranks).
- Clinician constructed measures and measures based on previous research articles were scored according to the scoring outlined in appendix B.

5.6.6 Inter-observer agreement of rating data

Twenty-three percent (six triads for each measure rated) of the data that required subjective rating of behaviours using rating scales was assessed for reliability by two independent raters blind to the diagnostic status of the participants. Rating was carried out on the narrative task, pragmatic task and pretence task.

Data to be rated was randomly selected by randomly selecting six triads. A triad consisted of a HFPDD subject and his or her matched SLI and matched NDD subject, so that each of these measures was rated for eighteen participants (i.e. six participants from each group, who were matched across groups). This random selection included children from a variety of age ranges, spread across the groups. The order of the presentation of the subjects was random, so that there was no specific order as to how the HFPDD, SLI and NDD subjects' data was presented.

The raters used (two for the pragmatic analysis and two different raters for the narrative analysis and pretence task) were experienced in working with children of this age group. Furthermore, these raters were experienced with working in the area of early language difficulties and PDD.

Instructions and guidelines to follow for the rating of each of these tasks were given to the raters. These are presented in appendix B3.2 for the narrative task, B4.2 for the pragmatic task and B10.2 for the pretence task. For each of these tasks training of raters took place on data from a child with HFPDD, SLI and NDD, whose data had not been included in the data to be rated. After discussing the instructions and guidelines together, each rater rated this practice data. Differences obtained on the ratings for the three raters were then discussed and any misunderstandings regarding the variables to be rated were clarified. The remaining data was then rated independently.

5.7 **ANALYSIS OF DATA**

Further analysis of the data to assist with answering the research questions, was carried out. This consisted of the following:

In order to establish whether the three groups differed on the three batteries of measures and whether different profiles could be obtained the following statistical procedures were used:

- Descriptive statistics – Means, standard deviations and skewness measures were obtained for all the different communication, theory of mind and cognitive measures that were administered.
- Analysis of variance – Analysis of variance measures were run on all the measures that did not involve the rating of data on an ordinal scale, in order to determine group differences. These included: the BPVS, the Basic Concepts sub-test, the Grammatical Understanding sub-test, the Word Finding Vocabulary Test, the pronoun alternation measure, the Oral Vocabulary sub-test, the TOPS-M, measures of expressive grammar, the CAS measures and the theory of mind measures. The aim of the analysis of variance is to determine whether the difference observed between the means of the groups is greater than what would be expected from chance alone (Hopkins & Glass, 1978).

- Kruskal-Wallis Test – In order to test for group differences on measures that involved behaviours rated on an ordinal scale, the Kruskal-Wallis Test was used. This was carried out on the narrative (including the parameters of temporal organisation, relevance, developing character(s), supporting description, ending and clarity) and pragmatic (including the parameters of understanding conversation, verbal aspects, paralinguistic aspects and non-verbal aspects) analyses. The Kruskal-Wallis Test is a non-parametric one-way analysis of variance that helps to determine whether the differences in ordinal data observed between groups are due to genuine group differences (Siegel & Castellan, 1988).
- Bonferroni t tests – Pairwise comparisons using the Bonferroni t test were carried out on all the different communication measures, the different parameters tested on the CAS (i.e. planning, attention simultaneous and successive processing) and the theory of mind measures, in order to determine whether differences between pairs of groups existed (i.e. between the HFPDD group and the SLI group, or between the HFPDD group and the NDD group).
- Frequency distributions – Frequency distributions were carried out on the number of different types of complex grammatical clauses used and the different parameters rated as part of the narrative analysis. This was done as Bonferroni t tests could not be run on these parameters as a t test cannot be run on one item of a measure alone. The aim of the frequency distribution is to provide information on the shape of the distribution for the different groups (Hopkins & Glass, 1978).

In order to determine whether particular communication deficits were linked to particular cognitive processing deficits and whether these were linked to particular theory of mind deficits, Pearson correlation coefficients were undertaken. Correlations were taken of the following: the receptive language score, the expressive semantics score, the number of dependent clauses, the MLU, the total narrative score, the receptive pragmatics (understanding conversation) score, the expressive pragmatic score (comprising verbal, paralinguistic and non-verbal aspects), the total pragmatic score, the theory of mind score, the planning standard score, the simultaneous processing standard score, the attention standard score and the successive processing standard score. The aim of the correlation is to determine the degree of the relationship between two variables (Hopkins & Glass, 1978). Correlations were run on the three groups combined (i.e. the total sample). This was done as

the more homogenous a group is, the lower the correlation, and the more difficult it is to see the correlation between variables (Hopkins & Glass, 1978).

In order to determine which measures best differentiated between the three groups discriminant function analyses were carried out. The aim of discriminant function analysis is to determine the fewest dimensions required to distinguish between groups and to determine which dimensions discriminate better between groups, to classify cases into groups based on the discriminant selected variables and to test the underlying theory of the measures by determining whether cases are classified as predicted (Garson, 2003). This was carried out on each of the following: all the different communication sub-tests, the four parameters making up the CAS (i.e. planning, attention, simultaneous and successive processing), all the different theory of mind sub-tests and on the summary scores for the communication, CAS and theory of mind batteries. The summary scores consisted of the following:

- Receptive language score (consisting of a combination of the BPVS raw score, Basic Concepts raw score and Understanding Grammar raw score)
- Expressive semantics score (consisting of a combination of the Word Finding Vocabulary Test raw score, the pronoun alternation score, the Oral Vocabulary raw score and the TOPS-M raw score)
- Number of dependent clauses
- MLU (number of morphemes per T-unit)
- Total narrative score (consisting of a combination of the temporal organisation, relevance, developing character(s), supporting description, ending and clarity scores)
- Total pragmatic score (consisting of a combination of the understanding conversation, verbal aspect, paralinguistic aspect and non-verbal aspect scores)
- Planning standard score
- Simultaneous standard score
- Attention standard score
- Successive standard score
- Theory of mind score (consisting of a combination of the visual perceptual role taking, mental significance of the eyes, belief, deception, emotion and pretence true scores).

In order to determine inter-rater agreement the following statistical procedures were used:

- Kappa coefficients on rated data - Kappa coefficients were run on the results of the ratings that involved rating using an ordinal scale, i.e. the narrative and pragmatic measures. The Kappa is a coefficient of agreement for nominally scaled data (Siegal & Castellan, 1988), the aim here being to determine the extent that the independent raters' ratings agree with the researcher's ratings, in order to provide an indication of the accuracy of the researcher's ratings, as the researcher rated the additional 77% of the data.
- Analysis of variance measures on rated data – For the overall narrative coherence score, overall narrative score, pragmatic comprehension score, pragmatic verbal score, pragmatic paralinguistic score, pragmatic non-verbal score, instances of true pretence and number of attempts at pretence, an analysis of variance was carried out. This was done in order to determine whether significant differences existed between the three raters.

5.8 INTER-RATER AGREEMENT FOR RATED DATA

Inter-rater agreement was evaluated by carrying out Cohen's (1960) kappa analyses on the ratings obtained for the narrative and pragmatic assessments. Rater A was always the researcher. Raters B and C were the two independent raters, who were blind to the diagnosis of the children. Raters B and C differed for the narrative and pragmatic tasks. Results for each pair of raters and for the narrative and pragmatic ratings are presented in table 5.3.

Table 5.3: Summary of kappa results for raters' ratings of narrative and pragmatic measures

Assessments	Raters	HFPDD	SLI	NDD	All
Narrative Analysis	A-B1	0.78	0.33	0.61	0.57
	B1-C1	0.50	0.28	0.50	0.43
	C1-A	0.56	0.44	0.53	0.51
Pragmatic Analysis	A-B2	0.42	0.71	0.99	0.71
	B2-C2	0.52	0.64	1.00	0.72
	C2-A	0.48	0.68	0.99	0.72

According to Landis and Koch (1977) kappa values greater than 0.75 represent excellent agreement, values between 0.40 and 0.70 represent good agreement and values less than 0.40 represent poor agreement.

For the narrative measure, excellent or good agreement was obtained for the HFPDD group for all the rater pairs. Acceptable agreement was obtained for raters A-C1 but not for raters A-B1 and raters B1-C1 on the narrative measure for the SLI group. While the agreement was poor, raters A-B1 achieved better agreement than raters B1-C1. For the NDD group good agreement was obtained for all rater pairs for the narrative measure. Good overall agreement was obtained for all the rater pairs for the narrative measure. The poorer agreement for the SLI group may have been due to this group's coherence scores generally falling between appropriate and inappropriate and as a result being more difficult to establish agreement on a rating. For example, one rater may have given a child a rating of 3 on a measure and another rater may have given a child a rating of 4. While these reflect similar ratings, the kappa procedure would still record them as not agreeing.

For the pragmatic measure, results indicate good agreement for all rater pairs for the HFPDD and SLI groups, with there being better agreement for the SLI group than the HFPDD group. For the NDD group excellent agreement was obtained for all rater pairs. For the pragmatic measure overall good agreement was obtained for all the rater pairs. Qualitative analysis of the results appeared to indicate that poorer agreement occurred for the HFPDD group as raters generally identified that pragmatic behaviours were inappropriate but did not always give the same exact rating. For example, one rater may have given a rating of 1 and another rater a rating of 2. Again while these reflect similar ratings, the kappa procedure would still record them as not agreeing. The overall results of the narrative and pragmatic ratings, would appear to suggest that it may be easier for raters to rate inappropriate versus appropriate pragmatic behaviours than measures of coherence and clarity on the narrative measure.

Analysis of variance measures were carried out for the overall narrative coherence score, overall narrative score, pragmatic comprehension score, pragmatic verbal score, pragmatic paralinguistic score, pragmatic non-verbal score and pretence assessment (number of instances of pretence and number of attempts at pretence) in order to determine whether significant differences existed between the three raters. Results are presented in table 5.4.

Table 5.4: Analysis of variance results for rated data

Measure		SS	df	MS	F	p
Total coherence	Between	22.26	2	11.13	0.23	0.8
	Within	2455.22	51	48.14		
Total narrative	Between	28.26	2	14.13	0.22	0.8
	Within	3277	51	64.26		
Pragmatic comprehension	Between	0.15	2	0.07	0.01	0.99
	Within	314.22	51	6.16		
Pragmatics verbal	Between	9.59	2	4.79	0.02	0.98
	Within	12902.78	51	253.0		
Pragmatics paralinguistics	Between	3	2	1.5	0.12	0.88
	Within	620.5	51	12.17		
Pragmatics non-verbal	Between	0.15	2	0.07	0	1.0
	Within	1287.5	51	25.25		
Pretence: Correct	Between	2.70	2	1.35	0.04	0.97
	Within	1967.44	51	38.58		
Pretence: Attempts	Between	4.33	2	2.17	0.04	0.96
	Within	2727	51	53.47		

No significant difference/effect was found, indicating that the three raters did not differ significantly in their ratings of the data for any of these measures.

The results from the ratings suggest that, although overall good inter-rater reliability was obtained, inter-rater reliability appeared better on the pragmatic and pretence measures than on the narrative measure. However, overall inter-rater reliability was found to be acceptable for all three of the assessments rated.

This chapter has included a description of the methodology of the study. The following chapter includes a description of the results of the study and a discussion of these.