CHAPTER 4

METHODOLOGY

4.1   INTRODUCTION

The research question was formulated in Chapter 1, namely: *"Does the measurement instrument measure the communication-related behaviours that it claims to measure and does the package of play activities facilitate the development of communication-related behaviours that it claims to facilitate?"* This question refers to validation. The methodology implemented to answer the research question is described in this chapter.

The methodology is a description of the process by which the play package (the measurement instrument and package of play activities) was validated and it included a description of the research carried out in three domains. In the conceptual domain concepts and relations were considered in abstract form while in the methodological domain instruments and techniques for obtaining observations and for relating and comparing sets of observations were included. In the substantive domain "states and processes in the real world (events) and sets of relations among events (phenomena)" were included (Brinberg & Kidder, 1982: 6). In this chapter a description of the method employed in the pre-experimental and experimental phases reflects the conceptual and methodological domains. The description of the data analysis procedures is indicative of the way in which validation will be carried our in the substantive domain.

Based on knowledge of the normal development of communication and communication-related skills, severe communicative disability, and scientific investigative methods, specific aims for the investigation into the validity of the proposed play package, were formulated.

4.2   MAIN AIM

The aim of this research is to validate a play package for the facilitation of communication-related skills by:

---

a)   the development of a play package consisting of specifically selected activities and presentation methods based on existing theories on childhood development and facilitation principles and procedures; and

b)   experimentation to establish a cause-effect relationship between the activities and the changes in single gradually acquired behaviours of children with intellectual impairments.

As this aim represents a complex problem incorporating a variety of facets to be investigated in different phases, it was necessary to formulate a number of sub-aims through which each main aim could be realised before, during and after the application of the play package.

It is evident that there are two distinctive phases in this research, and they will be discussed separately in a pre-experimental as well as an experimental phase.

4.3   SUB-AIMS

4.3.1   Pre-experimental phase

4.3.1.1 To develop the play package:
-   by evaluating each activity of the play package used in previous research (Uys, 1997) in terms of its value in facilitating communication-related behaviours. This refinement is necessary in order to limit the number of activities, while covering all the necessary underlying behaviours;

-   by establishing presentation methods for each activity, which will ensure the facilitation of communication-related behaviours in children with intellectual impairments, incorporating theories on childhood development and facilitation principles and procedures.

4.3.1.2 To select and develop assessment tools for the identification of behaviours related to play activities:
-       by selecting appropriate pre- and post-intervention assessment tools;
-       by identifying behavioural indicators, which will be facilitated and measured during each play session with the child; and

- by establishing face, and content validity of the daily multiple measurement instrument (DMMI) to be used daily, involving the evaluation by experts in the field.

4.3.2 Experimental phase

4.3.2.1 To establish the reliability of the daily multiple measurement instrument by determining intra-rater and inter-rater reliability

4.3.2.2 To establish outcome validity of the play package by determining and comparing the levels of performance of the participants on the sensorimotor, cognitive, communication, and social-emotional activities during the pre-intervention phase, intervention phase, post-intervention phase and post-withdrawal phase, using four authentic assessment tools.

4.3.2.3 To establish construct validity of the play package by determining the changes in behavioural indicators reflecting the sensorimotor, cognitive, social –emotional and communication constructs.

4.3.2.4 To establish convergent validity of the daily multiple measurement instrument by comparing the performance of the participants on this instrument with their performance as measured by three other authentic measurement instruments, using two as comparable assessment tools and one as a countermeasure.

4.4 RESEACH DESIGN

A small group pretest-posttest design was selected to meet the aims of this research. A quasi-experimental design was used, as a comparison (control) group was omitted (Brink, 1999; Campbell & Stanley, 1966). The research was carried out over an eight-week period.

4.4.1 Schematic presentation of the research design
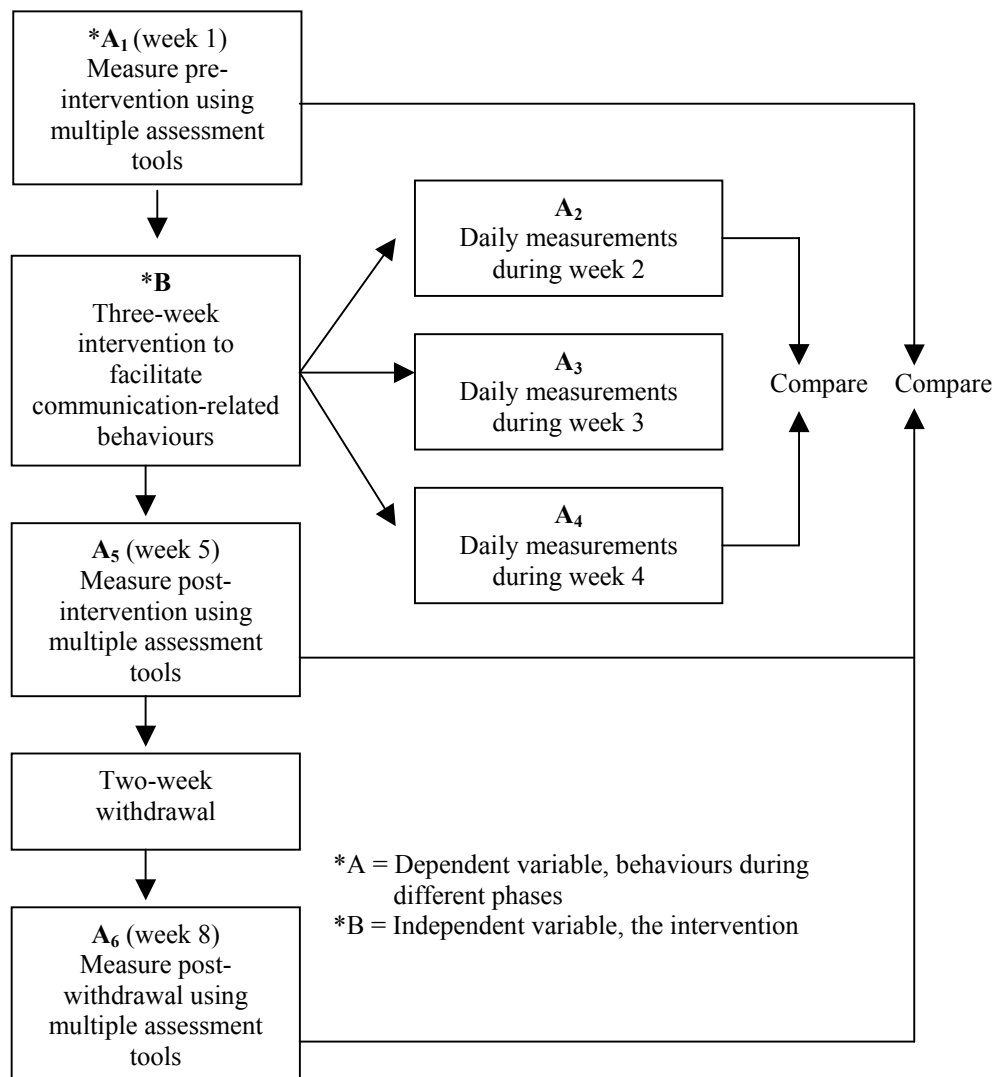
Figure 4.1 One-group pretest-posttest design

This design was used to determine how the independent variable affects the dependent variables namely behaviours of the selected group. When a decision is made about the effectiveness of intervention the rule of a quasi-experimental design is that it is necessary to consider the context in which the research took place and the particular pattern of results obtained (Robson, 1994). By measuring on a daily basis during intervention, a trend towards a particular pattern can be obtained as the interpretation process is enhanced. Change in behaviour from the pre-intervention phase, through the intervention phase, to the post-intervention phase of each subject on multiple measurements, is obtained and can be compared. This method allows for inter-method cross validation where the more formal testing in the pre- and post-intervention phases is correlated with the daily evaluations made

during intervention (Mouton & Marais, 1985; Waltz *et al.*, 1991), thus improving the reliability of the testing procedure implemented in the research.

Although this is, by necessity (because of the heterogeneity of the population) a quasi-experimental design, certain measures of validity and reliability are required by this design.

4.4.2   Validity measures

Specific measures of estimating validity had to be introduced prior to the determination of the validity of the play package – the aim of this study. Full detail of this process is discussed under 4.5.1. Incorporated in the initial establishment of validity is the determination of the validity of the DMMI (See 4.5.2.2 (d)). The final validity of the play package, however, can only be accepted or rejected once the findings have been analysed, evaluated, and interpreted.

Internal validity is an important part of validity measures. It has to do with the ways in which the research design ensures that the introduction of the independent variable (play package) can be identified as the sole cause of change in the dependent variable (behaviours) (Williams *et al.*, 1995). Certain measures were taken to ensure internal validity in this study.

a)   Maturation

Since the subjects used in this study are young, the design could not be stretched over a long period as normal development or maturation could influence the results.

b)   Testing

When selecting pre- and post-intervention tests or assessments, a criterion should be that performance on the post-test should not be influenced by memory traces of the pre-test.

c)   Instrument error

In the selection and development of measuring instruments it should be noted that the instruments are valid and reliable. The administration of the tests should also be valid and reliable in order to limit misinterpretation by external observers. For the administration of the tests see 4.6.3.

4.4.3   Reliability measures

During the quantitative evaluation phase of the development of the DMMI certain reliability measures were introduced (See 4.5.2.2 (c)).

4.5   PRE-EXPERIMENTAL PHASE OF THE RESEARCH

This research consisted of two phases namely pre-experimental phase and experimental phase. Each phase will be discussed separately after the schematic presentation of the two phases in Figure 4.2.

The pre-experimental phase aim was threefold. Firstly, to develop intervention material and procedures to be utilised in the experimental phase. Secondly, to select and develop appropriate measurement instruments to be used during the experimental phase. Lastly, the developed play package and assessment instruments were submitted to a pilot study.
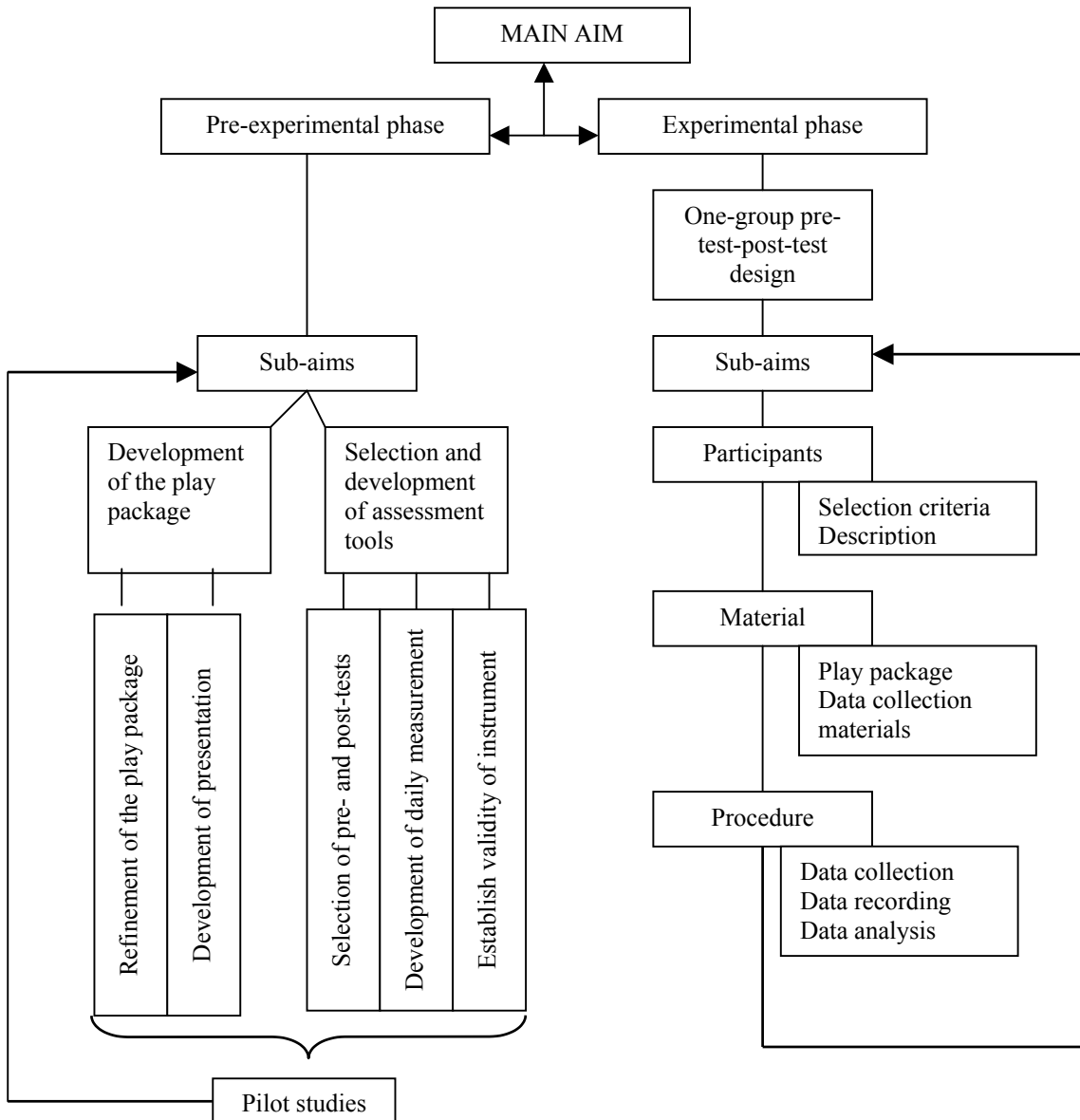
Figure 4.2   Procedural sequence

Each of the stages in the pre-experimental phase will be discussed in the following section.

4.5.1   Initial validation of the play package

Validation allows the evaluation of four specific aspects of the intervention process namely

a)   the goals selected for the intervention

b)   the procedures to be used in intervention

c)   the effects of the intervention

d)   the outcomes of the intervention (Kazdin, 1977; Kazdin, 1982; Wolf, 1978)

The goals for the intervention should be stated precisely. This is done through investigation of theoretical underpinnings of aspects included in the intervention programme. A process of operationalisation (Brink, 1999; Groenewald, 1988) identifies the intervention programme's constructs to be facilitated and measured. Procedural validation includes the type (strategies) and form (presentation methods) of intervention (Kazdin, 1980). The effects of intervention could only be established after data were collected and analysed. Ultimate effects refer to the link between the intervention and the specific changes observed in individuals after measuring the goals of the intervention programme (Schlosser & Braun, 1994). Outcome validation refers to the assessment of perceived changes as a result of intervention (Wolf, 1978). It is divided into proximal-, intermediate-, and distal outcomes.

The first two matters, namely the goals of intervention and the process of intervention were dealt with in the pre-experimental phase, and the effects and outcomes of the intervention will be discussed in the experimental phase of this research, as it is only after the application of scientific research methodology that the theoretical explanations and descriptions can be validated (Groenewald, 1988).

A previously developed play package (Uys, 1997) was used to facilitate communication-related behaviours in children with severe disabilities. However, before the play package could be implemented it had to be refined, graded, and appropriate presentation methods developed.

4.5.1.1 Refinement of the play package

It was decided to refine the play package in order to eliminate duplication of activities with the same characteristics but ensuring that all the relevant communication-related behaviours would still be elicited.

In the previous research (Uys, 1997) ten activities were used. The activities of the play package were validated by a group of experts who rated the activities in terms of the skills they facilitated. Activities were rated according to their strength in terms of three different constructs, namely sensorimotor, cognitive and communication. The determination of elements represented in the activities by experts in the field, is a form of validity by definition

(Lynn, 1986). In Figure 4.3 the results of these ratings are portrayed. Some activities scored higher in certain domains than others, but all the activities stimulated not less than 50% on each domain area.
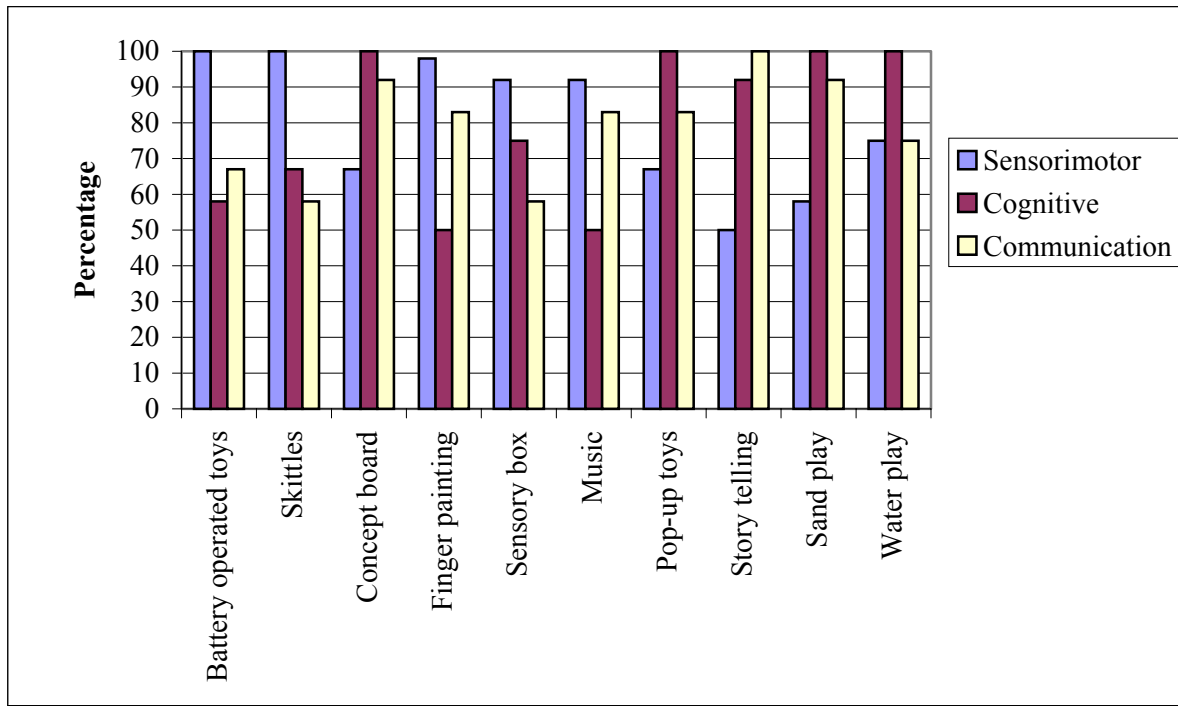
Figure 4.3   **Rating of activities in terms of sensorimotor, cognitive and communication behaviours**

The activities were ranked after an intervention period according to the gains in performance of children with severe disabilities on three different constructs, or developmental domain areas. The ranking was as follows: 1 being the most gain and 10 being the least gain. The results are shown in Table 4.1.

Table 4.1   Ranking of activities

Activities	Rank
Sensorimotor	
Battery-operated toys	3
Skittles	4
Finger-painting	2
Sensory box	6
Music	10
Cognitive	
Concept board	9
Pop-up toys	1
Sand play	7
Water play	8
Communication	
Storytelling	5

These findings indicated that some activities yielded better results than others in different domains or constructs. By selecting the two highest-ranking activities from each construct, except for communication, inclusion into the refined play package was determined.

In Table 4.2 the refined play package is described and a photograph of each activity is provided.

Table 4.2   Activities in the play package

Activity	Description	
Battery-operated toys	Brightly coloured battery-operated toys were selected. This included an elephant, a car, a worm, and jumping fishes. These toys were operated by different switches, namely joystick, large and small pressure switches, and a pinch switch. Corresponding 2D pictures were used.	
Finger-painting	A mirror was set up and shaving foam was used as a paint medium. Three different paint colours were introduced in the session. Imitation of certain movements was required.	
Pop-up toys	Three different pop-up toys were used, namely jumping jack, a "piano" keyboard with hidden animals popping up after pressing a key, and a multimethod pop-up toy where different actions released different hidden animals. Corresponding 2D pictures were used.	
Sand play	The following objects were used: a sandpit of approximately $1m^2$, different tools (spade, cups, bucket, rake, sieves, wooden blocks), a sand wheel, and plastic animals. Hide and seek games were played as well as operating the sand wheel. Construction play was introduced with the wooden blocks.	
Storytelling	A story was told, using the different objects functionally in the process. The objects included a train, a car, an aeroplane, a doll, and a telephone. Corresponding 2D pictures were used at the end of the story.	

4.5.1.2 Grading of activities

To establish a valid procedure of presentation, the activities had to be graded according to a learning phase and an intervention phase. The learning phase (5 minutes) gave the child the opportunity to explore with the materials and objects (Reilly, 1974) and to learn the basic rules of the activity before the intervention phase commenced.

Pre-set criteria for the learning phase were written and a child was expected to reach a 100% performance before continuing with the intervention phase. The criteria were

- understanding basic instructions (verbal and non-verbal) how to execute the activity;
- doing basic movements to indicate active participation; and
- visually focusing on the activity for at least 50% of the time.

4.5.1.3 Establishment of presentation methods

Before principles for facilitation of communication-related behaviours could be selected and applied the different developmental domains had to be identified. This was done in Chapter 2. Strategies and principles to achieve specific goals and objectives as set for the treatment of communication-related behaviours were identified and are presented in Chapter 3.

These principles and strategies were incorporated in planning the presentation of the activities. This presentation was developed in three phases namely:

- How the response from the child would be probed
- Anticipated response from the child
- Principles and strategies implemented to facilitate the response.

The activities had to be presented in a fixed way to each subject to eliminate the influence of external variables. The presentation of each activity was therefore written in a step-by-step manner. The external raters used these forms as a control mechanism for a reliability check. For a complete set of the presentation methods of all the activities see Appendix A.

A time limit of 20 – 30 minutes was set for each activity presented, as short sessions ensure maximum participation (Steenkamp & Steenkamp, 1992).

4.5.2 Selection and development of assessment tools

All experimental research consists of treatment and measurement phases. The independent variable (presentation of the activities) was dealt with in the previous section. The focus was to select and develop specific assessment tools for pre- and post-intervention and to develop a measurement instrument to be used daily during the treatment phase. In this process, the central point was to determine that these evaluation procedures should identify specific behavioural indicators reflecting communication-related behaviours. The thoroughness and appropriateness of assessment tools are critical to the results of the study as it influences the meaningfulness of results. (See 3.3) The following steps were pursued.

4.5.2.1 Selection of pre- and post-intervention assessment tools

Children with disabilities are frequently termed "untestable". Traditional assessment processes may not give a true or realistic picture of the performance of a child with disabilities, therefore a functional assessment approach was deemed necessary (Linder, 1993). The benefit of a functional assessment is that it ensures that each test item is potentially an appropriate intervention target. The child can be observed in a familiar, non-threatening situation which assists in maximal performance. A multitrait-multimethod approach was used in this research to establish validity. Therefore it was necessary to select more than one assessment tool. Various assessment tools were considered for this study and the following were selected.

(a) Transdisciplinary Play-Based Assessment (TPBA)

Functional assessment tools, sensitive to the needs of children with disabilities were identified and after video-conferences with Purdue University and Pennsylvania State University, the TPBA (Linder, 1993) was selected. The TPBA model is developmental, transdisciplinary, holistic, and dynamic. As this research uses play as medium for intervention, it seemed appropriate to assess the children with disabilities through the use of play activities. The TPBA is a natural, functional approach to

---

assessment which could be used for pre-and post-intervention evaluations (Linder, 1993) and it assesses the four domains of communication-related behaviours, namely sensorimotor, cognitive, language and communication, and social-emotional. The TPBA involves a child in a play situation with a facilitating adult, the parents (if possible), and another child. It was designed for children between 6 months and 6 years of age. The TPBA was evaluated against the six operational LINK standards and obtained a high score, proving it to be an authentic assessment tool (Bagnoto, Neisworth & Munson, 1997).

The TPBA is divided into five phases:

Phase I — The play facilitator begins with unstructured facilitation during which the child leads the play and the play facilitator imitates, models and expands the child's play.

Phase II — This phase incorporates aspects of play that the child did not initiate.

Phase III — A peer is introduced to observe child-to-child interaction.

Phase IV — Structured and unstructured motor play is included.

Phase V — To observe oral motor difficulties snack-time is included.

The author suggested incorporating a phase for parent interaction, but this was not done as the intervention occurred during the school day and the parents were not available. The sequence of these phases could be changed to incorporate the child's preference of play situations. After the play session, videotape analysis was done and the summary sheets completed.

Scoring

A summary sheet for data recording which resembles a checklist format is provided. Observers had to complete the form for each item with a [+] if the child exhibited skills within a standard range of development, or with a [−] if the child demonstrated a delay in development or deviation from the normal behaviour patterns (Linder, 1993). These decisions are based on age charts provided.

(b) Symbolic Play Scale (SPS)

The SPS (Westby, 1980) was also selected as it incorporates cognitive and language development, two of the constructs of communication-related behaviours. The SPS

evolved from a Piagetian-based language programme for severely retarded children. Children without developmental disabilities were observed and the normal developmental ages were added to normal play development. The SPS provides a means of assessing a child's representational thought, a major cognitive development during the pre-operational period (Piaget, 1951; Westby, 1980). As the sequence of play stages is the same for children with and without developmental disabilities (Westby, 1980), the SPS could be used for the population targeted for the purpose of this research.

The assessment consists of five groups of toys, arranged in areas within a large room. These include:

-       the infant stimulation area with pull toys, windup toys, talking toys, busy boxes, musical toys, rattles and soft stuffed animals.
-       the household area with dolls, cooking utensils, dress-up clothes and telephone.
-       the store area with play money, miniature food, telephone/ intercom.
-       the creative play area with a sandbox, rubber animals, blocks, village with garage and cars.
-       the gross motor area with a slide, steps, walking board, riding toys, bowling set, beanbags and balls.

Scoring

Children were brought into the playroom individually and one adult interacted with the child while video-recording the session. The researcher and external raters completed the observation form (See Appendix B) after watching the video recordings. To score the child's performance, the observer had to select a specific developmental stage most suitable in description. There are ten stages categorised from 9 months to 60 months.

(c)   Developmental Test of Visual-Motor Integration Test (VMI)

The VMI was selected as these test results correlated with children's mental age. The VMI measures an integrative ability important to adequate functioning beyond visual-motor behaviour. The VMI is a reliable and valid test (Beery, 1989).

The VMI incorporates a developmental sequence of 24 geometric forms to be copied with paper and pencil. The child sits at a table with the test booklet and an pencil without an eraser in front of him. He is instructed to redraw the same forms as given in the booklet. Specific instructions are presented in Beery (1989).

Scoring

Scoring is done as prescribed by the author. A raw score is obtained and then an age equivalent is worked out according to age equivalent tables provided.

4.5.2.2 Development and validation of Daily Multiple Measurement Instrument (DMMI)

Many tests are biased against children who are unable to demonstrate what they know due to physical, emotional, sensory, or other impairments (Bricker, 1989; Brooks-Gunn & Lewis, 1981; Garwood, 1982; Zelazo, 1982). The researcher decided to develop the DMMI to assess children's performance on the different developmental domains during the experimental phase. This could indicate trends in developmental sequences. According to Bricker and Cripe (1995) there should be an active link between the assessment, intervention, and (daily) evaluation components. This linked system allows for "…an efficient and focused approach that maximizes the probability that children… will acquire target skills… that will move them toward independent functioning" (Bricker & Cripe, 1995:89).

Benson and Clark (1982) suggested a process of planning, construction, quantitative evaluation and validation when developing and validating multiple measurement assessment tools to be used on a daily basis, linked to intervention. These suggestions were used in the development of the DMMI.

(a)   Planning the measurement instrument

The purpose of this tool was to measure performance reflecting communication-related behaviours of children with intellectual impairments on a daily basis, using functional activities.

In the development of a valid measurement tool, it was necessary to investigate the theory pertaining to communication- related behaviours. Operationalisation of the theoretical constructs was conducted (Brink, 1999; Groenewald, 1988; Williams *et al.*, 1995). Through a process of naming, defining, categorising, and sorting, the researcher converted the constructs identified (sensorimotor, cognitive, communication and social-emotional) into measurable behavioural indicators. A list of at least twice as many measurable indicators that will be needed for a final measurement tool was the end-result of this process.

Two different teams of experts were asked to observe patterns of behaviour on video-recordings of a child with intellectual impairments engaging in activities of the play package. The experts included six speech-language and hearing therapists, six occupational therapists and one physiotherapist. The aim was for the experts to identify observable, measurable indicators and to document the results. Each expert did this process individually. The researcher then analysed and categorised their observations into the four constructs. The sum of all the experts' identified behavioural indicators was calculated and is represented as percentages. The results of this process are shown in Table 4.3. The values indicate the percentage of all the experts that recorded the measurable indicator.

Table 4.3   Identification of observable, measurable indicators

% of Sensorimotor indicators		% of Cognitive indicators		% of Communication indicators		% of Socio-emotional indicators	
Bilateral hand use, manipulation	92	Cause & effect	58	Expressing likes & dislikes,	33	Enjoyment, expressing emotions	100
Gross co-ordination	58	Attention and concentration	83	Requesting	50	Interaction	100
Fine co-ordination	42	Object permanence	33	Choice-making	50	Initiating	8
Imitation	100	Matching objects & pictures	67	Vocalisation, communica-tion attempts	75	Participation, motivation	67
Exploration	17	Object recognition	67	Verbal imitation	42		
Visual focus, tracing and scanning, eye contact	83	Functional object use	50	Expressive, labelling, verbalisation, commenting	50		
Pointing, reaching	58	Identifying body parts	33	Understanding or following instructions	67		
Averages	**64**		**54**		**52**		**68**

The results indicated that the experts could identify relevant behaviours and that these behaviours proportionally reflected all the constructs to be investigated. These averages indicated that indicators on sensorimotor, cognitive and communication constructs were easier to identify than on social-emotional although all agreed on prominent indicators. These experts were not trained to identify specific behaviours because they were all professionals working in the field of disability. It was therefore interesting to note that the social-emotional construct was not proportionally represented in the identified indicators. Although social-emotional indicators were fewer than the rest, this tendency is also observed in assessment instruments such as the TPBA. There was more agreement between the speech-language and hearing therapists and occupational therapists and physiotherapists on the sensorimotor construct, but less agreement in the cognitive and communication constructs. These results were compared to literature on child development to ensure the reliability and validity of the information (Groenewald, 1988), and were used in drafting the DMMI. As some of these behaviours depend on the same underlying performance

components, one component was selected, representing more than one behavioural indicator.

(b) **Construction of the instrument**

To increase the authenticity of the DMMI, the six operational LINK standards as presented by Bagnato *et al.* (1997) were incorporated in the design and construction. These standards were discussed in Chapter 3. A six-point rating scale was selected, as an even number of responses forced the evaluators not to select a neutral or "middle" response (DePoy & Gitlin, 1994). Operational definitions were drafted for all the measurable indicators on the six-point scale, as an operational definition assigns meaning to a variable and describes the activities required to measure it and is directly linked to literature findings (Brink, 1999). This enabled the evaluators to score the child's performance precisely after observations and increased inter-rater reliability, as it is specific.

The refined DMMI was again scrutinised by a group of experts. The aim was to assess the usability of the form in clinical practice and to determine whether the behaviours observed could be scored successfully. In this instance the experts were presented with definitions of grading the observable behavioural indicators. They were required to comment on the clarity of these definitions.

Results and changes to the DMMI

- In general the objective was to change the definitions to emphasise the behaviour observed. Words like "attempt to…" were replaced with "Can… or cannot…".
- *Requesting:* There were two areas where requesting could be scored (non-verbal and verbal). These two were combined.
- *Choice-making:* The first score's wording was misleading and could be interpreted in different ways. This could influence inter-rater reliability. The scores were changed to make it more user-friendlier.
- *Understanding instructions:* The wording was changed to "Following instructions". The grading was also modified to correlate with the presentation methods of the activities.

- *Attention:* A better sequence of time intervals was provided.

Not applicable column: This column was removed as it did not pertain to all the items. Some activities could not be scored on all the items, e.g. sand play does not have the opportunity to match objects with 2D pictures. Therefore only those items, not applicable for each activity were shaded.

(c)   Quantitative evaluation of the instrument

The quantitative evaluation included a pilot study, testing the DMMI. The pilot study provided the data from which the initial reliability was estimated. Reliability is the consistency with which a measurement instrument performs (Leedy, 1985). Table 4.4 provides the different reliability measures and explains how and during which phase it was implemented. See Chapter 3 for the motivation of the use of these methods.

Table 4.4   Reliability measures

Reliability measures	Motivation	When implemented
1. Stability (Intra-rater reliability)	Stability refers to the extent to which the same results are obtained when the instrument is administered to the same sample twice by the same rater.	During the pilot study.
2. Equivalence (Inter-rater reliability)	Two raters observe 20% of the recorded sessions executed by the researcher and then score the children's performance on the video-recordings independently on the same measurment instrument as the researcher.	During the experimental phase.
3. Internal consistency	This refers to the extent to which all of the items on an instrument measure the same characteristic.	An item analysis could not be done due to the small sample size.

(d)   Validation of the instrument

The validation of a newly developed instrument should be an ongoing process, which is seldom accomplished in one study, or by one researcher (Benson & Clark, 1982). Despite this, validation is essential because validity guarantees that the measurement instrument is accurate in what it measures, given the context in which it is applied. In

---

Table 4.5 a delineation of the different types of validity is provided, and motivated how and when it was applied in this study.

Table 4.5   Measures of estimating validity

Type of validity	Motivation	Implementation
1. Face validity	Experts in the field essentially base face validity on an intuitive judgement. This procedure is useful in the instrument development process in relation to determining readability and clarity of content. It means that the instrument measures what it is supposed to measure (Brink, 1999).	External raters will evaluate the DMMI to indicate whether the communication-related behaviours were adequately represented. This was done during the planning and construction phases of the instrument.
2. Content validity	The steps to obtain content validity include (1) specification of the full domain of a concept through a thorough literature review and (2) adequate representation of domains through the construction of specific items (DePoy & Gitlin, 1994). Each item on the instrument should be evaluated by experts in the field with regard to the degree to which the variable to be tested is represented. The overall appropriateness for it's use should also be evaluated (Brink, 1999).	During the planning phase, external raters were familiarised with the content of the DMMI. Theoretical training was then presented with regards to the content of the DMMI so that they could determine the appropriateness of the measurement instrument. During the judgement-qualification stage the experts judged the content validity of the instrument by determining the proportion of total items (Lynn, 1986). A content validity index will be established through a process of rating the relevancy of each item.
3. Construct validity	Construct validity addresses which constructs the measurement instru-ment actually measures (Brink, 1999; DePoy & Gitlin, 1994). It establishes the relationship between the results provided by the instrument to the underlying theoretical concepts of the instrument. The multitrait-multi-method approach is regarded as the preferred method to establish construct validity. This approach is based on the premise that different measures of the same constructs should produce similar results (Polit & Hungler, 1983; Waltz *et al.*, 1991).	In this research, the TPBA, SPS, VMI and DMMI assessed the four identified constructs, namely sen-sorimotor, cognitive, social-emo-tional and communication. Both the TPBA and DMMI assess all four constructs. The DMMI's results will be compared with the SPS, as it respectively measures cognitive and language develop-ment and the integrative ability of the child. The VMI will be used as a countermeasure as it measures only visual-motor integration. Thus correlation between the measurement instruments could only be conducted in the experimental phase.
4. Convergence	The results of the new instrument (DMMI) were compared to the data of a criterion measure (a known instrument), at the same time (Brink, 1999).	The TPBA's, SPS's and VMI's results will be compared with the newly developed instrument (DMMI). This could only be done after the experimental phase.

Face validity, content validity, construct validity, and convergent validity was addressed in this research. Face validity and content validity were done during the pre-experimental phase and construct validity and convergent validity during the experimental phase.

The pre-experimental phase aimed at developing material and procedures for the experimental phase. Pre- and post-intervention assessments were identified and the DMMI were developed and validated as far as theoretically possible. The intervention material, namely the play package was refined and graded. These materials now had to be subjected to a pilot study.

4.5.3   Pilot Study 1

As a final stage of the pre-experimental phase, the developed and selected play package and measuring instruments were applied in a pilot study.

4.5.3.1 Aims

(a)   The application of the play package
  i) To apply the five play activities to children with moderate to severe intellectual impairments, to identify whether any problems exist with the presentation of the activities.

  ii) To test the usefulness of the DMMI in terms of presentation and scoring of the activities.

(b)   The application of the measurement instruments

  i) To identify any problems with the clarity of definitions given for each score of the multiple measurement instrument to enhance inter-rater reliability during the experimental phase.

  ii) To use the DMMI to assess the child's performance and to ascertain problem areas.

4.5.3.2 Participants

(a)   **Criteria for selection**

    i)   **Home language**

As the researcher is proficient in English and Afrikaans, the subjects had to be either English- or Afrikaans-speaking. This was decided to ensure that language problems did not influence the effectiveness of the intervention.

    ii)   **Sensory abilities**

Children with visual and/or hearing impairments were not included. The application of the play package was oriented towards children with intellectual impairments.

    iii)   **Communication**

All children included in this study had to be communicatively impaired because the use of the play package was developed to facilitate skills related to communication. This included children with little or no functional speech.

    iv)   **Diagnoses of intellectual impairment**

All the children had to be classified as children with moderate to severe intellectual impairment according to the definition (American Association on Mental Retardation, 2001). That included children with an IQ classification of between 25 and 50 who were accommodated in special schools in South Africa (Steenkamp & Steenkamp, 1992). Children with severe motor impairments were excluded. The school where the study was conducted has specific inclusion criteria (Appendix C) of which the primary criterion is mental retardation. During their initial assessment before inclusion, a multi-disciplinary team evaluated the children and a functional age level was established.

    v)   **Chronological and mental age**

The children's chronological age had to be between 4:6 and 7:6 years, with a mental age of between 1:6 and 3:6 years (as obtained from school records).

vi)   **Therapy**

The children should not have received extra occupational therapy, speech therapy or physiotherapy except those which form part of the school programme.

(c)   **Description of the subjects for the pilot study**

Three children who met the inclusion criteria were selected for the purpose of the pilot study. The children's profiles are presented in Table 4.6:

Table 4.6   **Profiles of participants included in the pilot study**

Subject	Diagnosis	Chronological age	Mental age	Gender	Home language
1	Down Syndrome	4:6	2:0	Male	Afrikaans
2	Mental retardation and attention deficit disorder	5:7	1:6	Female	Afrikaans
3	Down Syndrome	6:6	2:0	Female	Afrikaans

4.5.3.3 Material for data collection

-     Transdisciplinary Play-Based Assessment (Linder, 1993)
-     Symbolic Play Scale (Westby, 1980) (See Appendix B)
-     Draft Daily Multiple Measurement Instrument (DMMI) (See Appendix D)
-     Visual Motor Integration Test (VMI) (Beery, 1989)

4.5.3.4 Equipment

a)   Equipment for intervention (SeeTable 4.2)
       Play package (See Appendix A)
b)   Equipment for data recording
       -     Video camera and video cassettes
       -     Tripod
       -     Remote control for the video camera

-             Television monitor
-             Video cassette player

4.5.3.5 Procedures

An overview of the aims, methods, procedure, findings and recommendation are shown in Table 4.7.

Table 4.7   Procedures followed, findings and recommendations of Pilot Study 1

AIM	METHOD	PROCEDURE	FINDINGS	RECOMMENDATIONS
1. To evaluate the clarity of definitions provided for each score.	- Video-recordings of children with intellectual impairments were made while the researcher presented the activities to them according to the examples given in Appendix A. - The researcher as well as two external raters scrutinised the recordings for any problem areas that could still exist.	- The room was set up with the video camera in the corner so that it did not distract the child. - The child was brought into the room - The first activity was put in front of him. - The video camera was switched on with the remote control - The activity was presented for 15 minutes. - The activity was taken away and replaced with the second one. - Only two activities were done each day so that the child could concentrate maximally on each.	- Terminology of some of the scoring was not clear enough and the wording had to be changed to ensure better inter-rater reliability. - The definition for scoring each item included a complexity as well as a frequency level. This influenced the accuracy of the scoring of the children's performance.	- The wording of all the scoring was evaluated again to ensure clarity in the definitions to make it more understandable. - A new form should be developed where each item's scoring was divided into a complexity and a frequency level.
2. To use the DMMI to assess the children's performance and to ascertain problem areas.	Same as above.	The researcher as well as the two external raters watched the video-recordings of the five activities.	The children scored high on the certain items (cause-effect, following instructions). This meant that it would not be able to indicate improvement over time sufficiently. The scoring of those items was not sensitive enough.	The rating was changed of the item indicated not to be sensitive enough.
3. To apply the five activities with children with intellectual impairments, to identify whether any problems exist with the presentation of the activities.	Same as above.	The same procedures were followed as with DMMI.	- Some of the activities did not allow the child to follow more than one-step instructions, making it difficult to score according to the grading on the DMMI. - Only one activity gave the child the opportunity to recognise 2-dimensional (2D) pictures.	- Presentations of the activities were adjusted to allow the child to follow from one step instructions to three step instructions. - The presentation of some of the activities was changed to incorporate so that the children could get more exposure to 2D grading.
4. To test the usefulness of the DMMI.	Same as above.	Same as above.	- Finger-painting and sand play did not include matching of pictures.	- Due to the nature of the activity, it was decided not to change the presentation.

4.5.3.6 Results of Pilot Study 1

The aim of Pilot Study 1 was to implement the materials and procedures to be used during the experimental phase of the research.

-       Two external observers felt that some of the wording of the DMMI had to be changed for the sake of clarity and reliability.

-       A new daily multiple measurement instrument was developed where each item's scoring was divided into a *complexity* and a *frequency* level. This was in accordance with literature stating that the development of skills occurs in a sequential fashion and that the child then uses the opportunity to practise the new skill to gain mastery (Kielhofner, 1992; King, 1978; Schade & Shultz, 1992). This made the form more user-friendly and increased reliability. In Chapter 2 detailed discussions on this topic are provided. (See Appendix E for the final DMMI).

-       The new DMMI was again subjected to the group of experts to establish face and content validity. Table 4.8 shows the results.

Table 4.8   Results of face and content validity testing of the DMMI

1.   Face validity   There was a 100% agreement between the experts that all the items on the DMMI adequately represent communication-related behaviours.
2.   Content validity   The judgement qualification stage of the DMMI entailed establishing a content validity index (Lynn, 1986). The group of experts gave a 100% rating for each of the 21 items on the DMMI as well as for the DMMI as a whole. This meant that the DMMI was evaluated as content valid.

4.5.4   Pilot Study 2

A second pilot study was undertaken to scrutinise the changes made to the measurement instruments and to assess the usability of the TPBA, SPS and VMI.

4.5.4.1 Aims

a) To train the external raters (consisting of an occupational therapist and a speech-language and hearing therapist) in the use of the TPBA so that consensus could be obtained on the scoring.

b) To use the TPBA and SPS on a child with intellectual impairments so that the external raters could be exposed to the assessment procedure and gain consensus on the scoring.

The results of the second pilot study are presented in Table 4.9.

Table 4.9   Procedures followed, findings and recommendations of Pilot Study 2

AIM	METHOD	PROCEDURE	FINDINGS	RECOMMENDATIONS
To use the TPBA, SPS with children with intellectual impairments so that the external raters could exposed to the assessment procedure and gain consensus on the scoring.	Two external raters were trained to use the TPBA and SPS as an assessment tool as this tool is not commonly known in the RSA.	Video-recordings of the children were observed.	- The TPBA only has a [+] or [−] score meaning that the children scored appropriate for their age level, or below their age level. - It was not possible for the parents or caregivers of the children to be present. The execution of the test was done without them.	- The method of scoring on the TPBA should change as performance could not be accurately assessesed. - The use of the SPS and VMI as other forms of criteria to measure against in the multitrait-mutimethod approach will overcome the limitations of the TPBA. - The SPS was used successfully.

4.5.4.2 Results of Pilot Study 2

- One of the children could not do the VMI, the formal pen and paper test. This highlighted the problems that exist when using traditional tests with children with disabilities. Although a raw score could not be obtained, Beery (1989) provided more qualitative norms for the younger child, according to which the child could be allocated an age norm.

- The execution of the TPBA has limitations as the parents or caregivers were not able to attend the assessments due to logistical problems. The test was therefore done without their presence, but the researcher felt that this fact did not diminish the comprehensiveness of the data collected. Similar situations will occur during the experimental phase as the children are brought to school by bus and not by the parents.

- Other limitations were experienced in the scoring method, as the TPBA only has a positive or negative score. Due to the population selected for this study this checklist format was inadequate, as all the children would score [−] on most of the items as a result of their developmental delays. It was decided to rather use the age charts provided for each item to be scored. For consistency these age norms were grouped into 6 developmental levels ranging from 0 months to 60 months.

- Certain items were discarded from the assessment form. The reasons are given in the Table 4.10.

Table 4.10   Excluded items on the TPBA

Items discarded	Reason
Development and mobility in prone, supine, sitting and standing, mobility in hands and knees	Due to the population selected for this study the children were capable of these skills.
Attachment and separation from parents	The children were seen at their school which was a familiar environment, but an objective evaluation could not be obtained.
Development of intentionality, oral-motor skills	The age norms provided for this item only ranged to 12 months and the assessment should allow for up to 60 months.

For the final assessment form for the TPBA see Appendix F.

4.6   EXPERIMENTAL PHASE OF THE RESEARCH

In the pre-experimental phase the play package was refined and specific measurement instruments were developed and selected for validation through experimentation.

4.6.1 Participants

4.6.1.1 Selection criteria for participants

The same selection criteria were used as in the pilot study.

i) **Home language**

As the researcher is proficient in English and Afrikaans the subjects had to be either English- or Afrikaans-speaking. This was decided in order to make sure that language problems did not influence the effectiveness of the intervention.

ii) **Sensory abilities**

Children with visual and/or hearing impairments were not included. The application of the play package was oriented towards children with intellectual impairments.

iii) **Communication**

All children included in this study had to be communicatively impaired as the use of the play package was developed to facilitate skills related to communication. This included children with little or no functional speech.

iv) **Diagnoses of intellectual impairment**

All the children had to be classified as children with moderate to severe intellectual impairment according to the definition (American Association on Mental Retardation, 2001). That includes children with an IQ classification of between 25 and 50 who are accommodated in special schools in South Africa (Steenkamp & Steenkamp, 1992). Children with severe motor impairments were excluded. The school where the study was conducted has specific inclusion criteria (See Appendix C) of which the primary criterion is mental retardation. During their initial assessment, before inclusion, a multi-disciplinary team evaluated the children and a functional age level was established.

v) Chronological and Mental age

The children's chronological age had to be between 4:6 and 7:6 years, with a mental age between 1:6 and 3:6 years (as obtained from school records).

vi) Therapy

The children should not have received extra occupational therapy, speech therapy or physiotherapy, except those which formed part of the school programme.

4.6.1.2 Selection procedures

a)   Several schools for children with disabilities were visited.

b)   A school was selected that accommodated children who complied with the selection criteria.

c)   The researcher met with the headmaster and occupational therapists of the school to obtain permission to treat their pupils and to use the facilities.

d)   Informed consent forms (See Appendix G) were sent to parents of nine potential participants.

e)   Six participants who best met the selection criteria were selected with the help of the school's occupational therapist.

4.6.1.3 Sample size

Five children with moderate to severe intellectual impairments were selected.

4.6.1.4 Description of the participants

Table 4.11 describes the subjects included in the study. The children all attended the same school.

Table 4.11   Profiles of participants included in the main study

Subject	Date of birth	Gender	Home language	Chronological age at time of study	Mental age
1	1996-02-17	Female	Afrikaans	5:0	1:6
2	1995-12-07	Female	Afrikaans	5:2	3:0
3	1994-06-29	Female	English	6:8	2:6
4	1993-11-24	Male	Afrikaans	7:3	2:0
5	1993-02-27	Female	Afrikaans	7:0	2:6

4.6.2   Materials and equipment

4.6.2.1 Materials

- Transdisciplinary Play-Based Assessment (Linder, 1993) (See Appendix F)

- Symbolic Play Scale (Westby, 1980) (See Appendix B)

- Daily Multiple Measurement Instrument (DMMI) (See Appendix E)

- Visual Motor Integration Test (VMI) (Beery, 1989)

4.6.2.2 Equipment

a) Equipment for intervention (See Table 4.2)

   Play package (See Appendix A)

b) Equipment for data recording

   - Video camera and video cassettes

   - Tripod

   - Remote control for the video camera

   - Television monitor

   - Video cassette player

4.6.3 Procedures

4.6.3.1 Data collection and recording procedures

The data collection was done through structured observations. Objective, valid, reliable and accurate description of behaviours observed in a natural setting is invaluable when measuring performance before, during and after intervention. The research was carried out over an eight week period. Table 4.12 presents the different phases of data collection, which assessment tools were used during which phase. See Figure 4.1 for a schematic presentation of the research design.

Table 4.12   Data collection phases

Weeks	Phases		Assessment procedure
1	Pre-intervention assessment	A_1	TPBA SPS VMI DMMI
2	Intervention	$*A_2$	DMMI
3	Intervention	$*A_3$	DMMI
4	Intervention	$*A_4$	DMMI
5	Post-intervention assessment	A_5	TPBA SPS VMI DMMI
6-7	Withdrawal		
8	Post-withdrawal assessment	A_6	TPBA SPS DMMI

*Daily measurements

A description of the procedures of each phase will be presented for clarity.

(a)   Pre-intervention assessment phase procedures

- A quiet large open room was used. Different toys and play materials were prepared for five areas, as described for the TPBA and SPS (See 5.2.1).
- The video camera was set up in a corner of the room and switched on as the children entered the room.

- The children were brought into the room individually and asked with which toys they wanted to play.

- Procedures as set out by the TPBA and SPS were followed (See 5.2.1).

- Snack-times were observed in the classroom.

(b)   Intervention phase procedures

The five activities were presented in the same set every week, namely battery-operated toys, finger-paint, pop-up toys, sand play and storytelling. Only one activity was presented each day. The sequence of the set was determined by the ranking of the activities (See Table 4.1). It was decided to keep the set stable over the 3-week intervention period, so as to decrease the external variables that could influence the results.

The six children were randomly sequenced for each day's presentation.

Daily data collection procedures

– The same room was used for the intervention as described during the pre-intervention phase.

– The researcher started at 08:00 each morning. The area was prepared by organising the toys selected for each particular day. The video camera was positioned in the corner of the room in order not to draw attention to it. A remote control was used to switch it on when the child was positioned in front of the toys.

– The selected child was fetched from the classroom and brought to the room.

– The child was put at ease by telling him that he was going to play with some toys. The child was then positioned in front of the toys.

– The activity was presented in the two phases as described in 5.1.2 (See Appendix A for detailed presentation methods).

– During the first phase the child was taught how to do the activity until he had complied with specific set criteria.

– The second phase of the activity was introduced, namely the intervention phase. Each activity was presented for 20 - 30 minutes.

- The child was then told that the activity was completed for the day and that another activity would be done the following day.

- The video camera was switched off.

- The child was then taken back to the classroom and the next one brought in.

- The researcher scored the video-recordings afterwards. Each day a new sheet of the DMMI for each participant was used for scoring. This procedure ensured that the researcher was not influenced by previous scorings of the subjects.

- The same method was used for the external raters. The raters watched the video-recordings separately so that they could not influence each other.

(c)   Post-intervention assessment procedures

- The same procedures were followed as during the pre-intervention phase.

- The external raters followed the same procedures as described in the intervention phase.

4.6.3.2 Reliability control measures

Two outside observers or raters, professionally trained as a speech-hearing therapist and an occupational therapist, were included for the specific purpose of recording data together with the researcher. External raters are more objective as they have no vested interest in the events they are observing (Williams *et al.*, 1995). It is difficult to observe, participate and record data simultaneously, therefore every session of each participant was video-recorded. All the scores were made after observing the video-recordings.

The raters had to fill in the assessment forms of the TPBA, SPS and DMMI (See Appendices F, B and E) during all the A-phases of the research. (See Figure 4.1). The researcher scored all the sessions and the two external raters observed 20% of all the sessions. These sessions were randomly selected.

The reason for using external raters was twofold. Firstly, to determine whether the activities were presented consistently across all the participants and secondly, to ascertain inter-rater

reliability of the daily evaluations process (See Table 4.4). For the first aim, the presentation of the activities (See Appendix A) was used as a criterion.

The raters watched the videos after the completion of the eight-week period. They watched the videos together, but no interaction between them was permitted.

4.6.3.3 Data analysis procedures

Small group experimentation requires the application of descriptive and non-parametric statistics for the analysis of data (Barlow & Herson, 1984). In this research an association between the ordinal variables was investigated. In a bivariate situation the data are arranged in a cross-classification table to see how scores on one variable are paired with scores on another (Lutz, 1983). This analysis determines whether there is a pattern to the pairing of the scores and describes the character of any pattern found.

This method of data analysis was applied to establish intra- and inter-rater reliability (sub-aim 4.3.2.1), to establish outcome validity (sub-aim 4.3.2.2), and to establish construct validity (sub-aim 4.3.2.3). The Friedman two-way analysis of variance test was selected to analyse the data. This test is designed for comparing three or more related samples on an ordinal variable. It is approximated by the chi-square distribution. The appropriate degrees of freedom are found by the formula of $(k - 1)$, where k is the number of conditions under which the measurements are taken. This test also converts raw scores to ranks. The ranks are assigned across the rows (for each participant), not across all scores. This test uses a statistic labelled χ_r^2 and is known as "chi-r-square". The "r" stands for the use of ranks in the test. The test statistic is found by the following formula:

$$\chi_r^2 = \frac{12}{Nk(k+1)} \; [\Sigma(\Sigma R_i)^2] - [3N(k+1)] \qquad \text{The Statistics for Friedman Two-Way ANOVA}$$

where: $\quad r$ = the use of ranks in the test
$\quad\quad \Sigma R_i$ = sum of ranks in each measurement condition
$\quad\quad k$ = number of conditions
$\quad\quad N$ = total number of cases

T... ... ...
(r_s) was used. This is a measure of association specifically designed for rank-ordered data.

The Spearman's r_s is based on a comparison of the squared differences in ranks between differences. The test statistic is found by the following formula:

$$r_s = 1 - \left[\frac{6\,\Sigma\,D^2}{N\,(N^2 - 1)} \right]$$ The Spearman's r_s Formula

where D = difference between X and Y ranks for each case
$\Sigma\,D^2$ = sum of the individual squared differences
N = number of pairs of ranks

4.7 SUMMARY

In this chapter a delineation was given of the process on validating the play package when presented to children with intellectual impairments (Kazdin, 1977; Kazdin, 1982; Wolf, 1978). A quasi-experimental design for a small group was selected as a control group was omitted. During the pre-experimental phase the selection of authentic assessment tools were described. The development of a valid daily multiple measurement instrument was then addressed (Benson & Clark, 1982). This instrument allowed the researcher to collect and document scientific evidence on which decisions about the effectiveness of the intervention could be based. The experimental phase was carried out over a period of seven weeks and included, as all experimental research should, multiple assessment and intervention phases. A multitrait-multimethod approach (Brink, 1999; Robson, 1994) was used to collect data and reliability procedures were included to enhance the validity of the results.