City Research Online

City, University of London Institutional Repository

Dynamic Assessment and Informed Intervention for Children with Language Impairment

Natalie Karen Hasson

Thesis submitted in fulfilment of the requirements for
the degree of
Doctor of Philosophy

City University
Department of Language and Communication Science

February 2011

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

DECLARATION

I grant powers of discretion to the University Librarian to allow this thesis to be copied in whole or in part without any further reference to me. This permission covers only single copies made for study purposes, subject to the normal conditions of acknowledgement.

ABSTRACT

Although speech and language therapy practitioners commonly place great weight on standardised, static assessment, the procedures may not be fully representative, and reveal little about the child's learning potential or the direction that intervention should take. Vygotsky's theories, particularly his notion of the Zone of Proximal Development, underpin a range of approaches within the complementary assessment paradigm known as Dynamic Assessment (DA) (Vygotsky 1986). The term is used for assessments consisting of '*active intervention by the examiners and assessment of examinees' response to intervention"* (Haywood and Lidz 2007 P1)

The current project investigated the application of Dynamic Assessment to a population of children with previously identified Language Impairments. As in parallel studies of intelligence, both manifest skills of language, and underlying processes used in manipulating and constructing language as a tool, were elucidated. The contribution that such an assessment can make to extending the understanding of language impairment, and in devising intervention programmes was investigated.

This thesis describes the development of a Dynamic Assessment task requiring implicit knowledge of syntactic structure. The construction of the procedure was a novel adaptation and combination of established DA methodologies that are described and evaluated in Chapter 1. The task, which is essentially a sentence anagram, comprised 12 items specifically selected to assess particular grammatical structures reported in Chapter 2 to be problematic in children with Specific Language Impairments (SLI). The details of the task construction are reported in Chapter 3. The measure was employed on 24 children aged 8-10, with identified language impairment, and the results are reported in Chapter 4. Inter-rater reliability of the test measure was 88%, and the sensitivity of the test to change over time was demonstrated. Information about participants' ability to transfer learning between items, their ability to use less directive prompts, their strategy use, and their metalinguistic and metacognitive awareness was extracted, and reported to the speech and language therapists working with the children. Evaluation of the test is discussed in Chapter 5.

The thesis also reports on an investigation of the role of the information derived from the DA in informing intervention programmes (Chapter 6). The same cohort of 24 children with SLI was randomly allocated to two groups. Reports from the DA were used to inform the ongoing language intervention of one of the groups of children. In Chapter 7 the outcomes of therapy from that group were compared to the outcomes of the group receiving regular intervention. Differences between groups were nonsignificant although the gains achieved by subgroups of children were predicted, and in particular children making little progress in their ongoing therapy were shown to derive most benefit from the modified intervention. The information was rated as useful by participating SLTs who altered the nature of their intervention strategies. Discussion of the results

and identification of factors such as emotional and behavioural issues that affect progress in intervention are discussed in Chapter 8. Implications for further development of the DA paradigm are discussed, and conclusions are summarized in Chapter 9.

CHAPTER 1

INTRODUCTION TO DYNAMIC ASSESSMENT

CHAPTER 1   INTRODUCTION TO DYNAMIC ASSESSMENT

1.1   Introduction

The ability to learn and the process of learning were considered essential components of intelligence early in the 20th century. Those who were creating and devising static intelligence tests, and who were credited with the psychometric IQ test, such as Binet, and Spearman (reviewed by Brody 2000) did not however, create the tools and specify the methods for this assessment of learning potential. As a result, the static measurement of specific cognitive skills remained the standard method of measuring intelligence. Underlying the debate about methods is the theory of intelligence itself, the need to define the construct that needs to be measured, and the reason for measurement at the outset. What is it that psychologists were hoping to achieve by measuring an individual's cognitive ability? Educational and clinical aims of assessment are diagnostic, predictive, and remedial. Theorists and practitioners of the 20th century have devoted themselves to the assessment of the abilities of individuals in order to categorize them, predict their future outcomes, and devise interventions to benefit them.

Alternatives to the psychometric approach were developed in response to social needs for more culture-fair tests, and educational needs for better predictive validity. These methods include the measurement of learning potential, sometimes termed 'Dynamic Assessment', an approach that aims to evaluate the process of learning itself, rather than the products of learning that have been transformed into the abilities frequently tested by psychometric tests of intelligence. This approach will be considered in greater detail in the current dissertation. The approach and the insights gained from this model of assessment will be applied to the specific field of

expressive use of language, and the challenges of that application will be a focus of the discussion.

In order to further this application, the relationship between cognition and language, and the parallels between the constructs of thought and language will be considered. Issues of language as a domain specific skill versus an underlying domain general skill will influence the way in which language skills are assessed and the implications for remedial intervention. The specific skills and abilities of those individuals identified as having a 'Language Impairment' will be carefully examined, and linked to evidence based assessments and programmes of intervention.

To this end, the current project investigated the use of a more probing and individualized assessment tool, namely Dynamic Assessment, with a population of children identified as having 'Language Impairments'. Like parallel studies of intelligence, the usefulness of elucidating both manifest skills of language, and underlying processes used in manipulating and constructing language as a tool, will be considered. The contribution that such an assessment can make to furthering the understanding of language impairment and its role in the devising of intervention programmes will also be investigated.

1.1.1 What is Dynamic Assessment?

While the term 'Dynamic Assessment' (DA) is used interchangeably with other terms such as 'interactive assessment' and 'learning potential assessment', *"the constant aspect of the definition is active intervention by the examiners and assessment of examinees' response to intervention"* (Haywood and Lidz 2007 p.1). This broad definition encompasses all types of interventions within an assessment, and any occasion on which the tester "does more than give instructions, pose questions and record responses" (Haywood and Lidz 2007 p.1). Some writers, including Carl Haywood, propose subsets of interactive assessment in which the term 'Dynamic Assessment' might be limited in application to those methods that specifically include planned, mediational teaching, rather than other types of prompting or cuing, but in the current study, the terms will be used broadly and interchangeably.

Dynamic Assessments are usually contrasted to 'static' 'standardised' or 'normative' assessments. This implies that DA procedures are none of these things, but recently developed European assessments of learning potential have striven to be standardised and psychometrically validated (Hessels, Berger and Bosson 2008). Thus the more useful contrast to a Dynamic Assessment is a static test that looks at an individual's independent performance on a given task at a given point in time. Few if any, proponents of DA advocate substituting static tests with DA, rather the two procedures are seen to be complementary and useful in combination to provide classifications and normative evaluations based on the individual's independent level of ability, as well as his potential to learn from instruction, and maximal level of performance with support.

Assessment of the potential to learn from instruction implies a focus on the processes involved in that learning, i.e. *how* the individual

learns in a situation, how he uses his prior knowledge and experience, how he integrates new information, and what strategies he uses to solve problems. The difference between assessment of ability and process is clearly described by Das, Naglieri and Kirby (1994). They see ability as a construct that reflects how much, or how well an individual can manage a task, for example language ability is defined by the success in doing language tasks. Das, Naglieri and Kirby go on to say that the cognitive approach to intelligence emphasizes that *abilities* or capacities interact with *processes* that are dynamic, and the nature of processing can alter performance or functional level of the ability. Processes and strategies are changeable, and may be learnt and replaced by more efficient strategies and processes, thus determining performance on a task. It may be more useful to assess a client's processing than his/her ability, and this has become an important notion in DA. Multiple goals are served by assessment of processing, though not necessarily simultaneously, and different approaches lend themselves to the different goals of diagnosis, predictive validity and informing instruction.

The theoretical bases of DA have been attributed in large part to the writings of Vygotsky, and to the clinical work of Feuerstein. These two frameworks will be examined in greater detail in the following section.

1.2   Theoretical Bases of Dynamic Assessment

1.2.1   Vygotsky's sociocultural theory

The work of Vygotsky continues to be influential in current views of education, especially remedial education, assessment, and multiculturalism, reflecting the ongoing or even increasing relevance of his theories today. Indeed Kozulin (2003) points out that educators have only recently begun asking the questions that make

Vygotsky's theoretical answers relevant. The theoretical concepts most applicable in the current context are the overarching notions of sociocultural mediation, and internalization (Das and Conway 1992).

Vygotsky is credited with the notion that learning in children is not entirely an innate process, but one that is mediated through cultural transmission from others in their environment. The process of learning begins with external stimuli from others, who present, instruct and mediate meanings to the developing child. These external instructions have to be actively internalized and assimilated by the child, who cannot remain a passive recipient of information, but rather has to be involved in making the meanings their own. Vygotsky characterized cognitive development as constantly undergoing change. From a historical perspective, it can be seen that even primitive humans used rudimentary means to control their mental processes, for example in developing primitive memory aids. This leads to the assumption that the '*essence of higher forms of thinking is the individual's intervention in the processes of his or her thinking..*' (Gredler and Shields 2007 p. 29).

From the developmental perspective, Vygotsky can be seen describing the development of a child's thinking as an ever changing system in which new higher mental functions emerge and change already existing lower mental functions (Bodrova and Leong 2003). The sources of this change and development are experiential learning, formal learning, and mediation through other human beings. Experiential learning, or the generalization of everyday personal experiences, results in the formation of 'spontaneous concepts' in young children (Karpov 2003). Such concepts are unsystematic and often inaccurate, described by Vygotsky as 'unscientific', yet form the basis for future acquisition of scientific concepts. Scientific concepts, in contrast, are systematically taught,

and are the basis for further thinking and problem solving, enabling thought that is independent of personal experience and thus theoretical rather than practical.

Higher cognitive processes are developed by the use of *signs* that enable an individual to redirect their thinking. They do not change the cognitive operation, but change the structure of the cognitive activity, for example, rehearsal does not alter the memory for a list, it alters the strategy used to remember the list. Processes and strategies are *mediated* structures - mediated by *signs*. Mediated structures are subject to developmental change or learning, and the signs or stimuli that influence mind and behaviour are seen as psychological *tools* for learning. Language is one of the key tools for learning, and its relationship to the development of cognition is a complex one.

1.2.1.1   Vygotsky on Language

Vygotsky's writings about language are extensive and this is therefore only a brief summary. Nevertheless some key themes have been extracted from the author's reading.

Kozulin, a regular interpreter of Vygotsky's original writing from Russian into English, notes that Vygotsky's use of the term translated as 'speech' more accurately relates to 'language' or 'discourse' (Kozulin 1990 p.151). The term 'speech' will be used here, as used by Vygotsky, with the broader meaning assumed. Vygotsky was concerned with the relationship between language and thought, or how concepts were embodied in words. Along with the occurrence of pre-verbal intelligence (described by Piaget as sensori-motor intelligence), Vygotsky identified pre-intellectual speech, wherein speech was used for social contact. According to Vygotsky, this stage would be followed by the appearance of grammatical forms that

were not used with the full adult understanding, with linguistic explanation of words combined with their non-linguistic referent (Hickmann 1985). Later development enables the separation of language as a linguistic entity, and a tool for reflection, from the concept to which the word refers.

Language use, like other higher functions, develops from 'primitive' or 'pre-intellectual' speech (Vygotsky 1986 p.87) and progresses to 'naïve' speech, in which there is mastery of the surface forms of language, and then to external sign use, or 'egocentric speech' in which a child thinks aloud, using language to solve problems. During this phase, speech does not simply accompany actions, but begins to transform them, as the child begins to use his speech to organize and plan actions (Hickmann 1985). Finally internal sign use develops, and over a period of time during which inner and outer speech interact and overlap, inner silent speech becomes a tool for thought. Thus external language develops first, as a mechanism for, and as a result of, social interaction. It later becomes internalized as a means of self regulation of behaviour and a tool for problem solving. At the first stage, Vygotsky argued that the primary function of speech for both the developing child and the participating adult is for communication, social contact, and influencing the environment. It is a primary channel for the transmission and mediation of social and cultural knowledge and values.

The use of language to facilitate development of more complex notions or concepts was also described by Vygotsky who experimented with various methods of sorting and categorization used by children. The theory is described in more detail by Kozulin (1990 Chapter 5), but briefly links the social and cultural transmission of words by adults in the environment to the development of scientific concepts in place of the child's

spontaneously developed everyday concepts. Language is one of the psychological tools for learning, and like other symbolic systems, its use as a tool needs to be mediated to the child (Kozulin 2003). The development of the tool of language itself is viewed as a sociocultural product (Das and Conway 1992) and acquired through experience and the transmission by others in the environment, but Kozulin (2003) also points out that having the tool of language available, does not guarantee that it is well used to facilitate further learning. This process is enabled by deliberately and specifically guided or mediated activities that emphasize systematic use and generalizability of tools. Ultimately, linguistic signs take on the function of organizing linguistic activity itself, as the individual is able to plan and frame discourse for the transmission of information.

Das and Conway (1992) identify the ingredient contributing to the quality of internalization (p.97) as *reflection,* and note that without adequate reflection, material learnt may not be adequately internalized and become useful for transfer to further learning. External mediation and individual reflection both enable the internalization of the system of signs that can be used by the individual to further his own thought and development, Vygotsky's higher mental functions, cultural development of behaviour, or 'mastery of behaviour by internal processes' (Gredler and Shields 2007).

Vygotsky further noted (1986 p.88) that while thought and speech coincide to produce 'verbal thought' not all thought is verbal, and not all speech derives from thought. Motivation engenders thought, which does not have linguistic form. Verbal meanings are encoded in 'inner speech' the nature of which Vygotsky attempted to elaborate (see Kozulin 1990), but it is only at the final stage of oral or written production that concrete words to convey meaning are selected.

1.2.1.2 The Zone of Proximal Development (ZPD)

One of Vygotsky's best known constructs, the ZPD, is firmly rooted in his theory of child development, the development of scientific concepts, and the role that imitation plays in learning. Chaiklin (2003) noted that one of the earliest appearances of the term in Vygotsky's writings was in an early translation of Thought and Language (Vygotsky 1962, cited by Chaiklin 2003). A later translation of the same text (1986, translated by Kozulin) finds the first use of the term 'zone of proximal development' embedded in an anecdote about the measurement of mental age in two children of the same chronological age, rather than with a definition per se. This places the concept back in the context of devising appropriate instruction for children, rather than as assessment for any other purpose.

The standard definition of ZPD, taken from Vygotsky's later publication (Mind in Society 1978, cited by Chaiklin 2003 p.40) is as follows;

"*the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers*"

The application of the ZPD is specific to devising interventions to further the development of already maturing psychological functions, that are nevertheless still too immature to enable independent problem solving (Chaiklin 2003), but at the same time are present and sufficiently developed to benefit from the assistance of another. The common everyday experience of teaching or instructing a child, allows any insightful adult to see that a child cannot be taught to carry out any skill for which he is not 'ready', and Vygotsky employs

the concept of *imitation* to explicate just such a concept. In order for a child to imitate, there must be some approximation towards the skill being demonstrated, a foundation upon which the behaviour to be imitated can build.

In contrast to the maturational view that made instruction dependent on development, Vygotsky believed that instruction could facilitate development, although not suggesting that there is direct 1:1 correspondence between external instruction and the internal development that instruction brings to life (Gredler and Shields 2007). Vygotsky proposed that one must identify the ZPD, and teach within that to further the development, especially of higher cognitive functions. The concept of *internalization* is important as a developing child gradually internalizes an activity and becomes able to perform the activity independently (Brown and Ferrara 1985). The role of the teacher should be to elicit the collaboration, or co-operation of the child in the learning process, and gradually withdraw his guidance as the child takes control of his learning. The teacher's explanations and questions become the basis for the child's self questioning. Furthermore, the teacher is responsible for developing the child's awareness of his own thinking, enabling him to work out patterns and relationships in problem solving.

Vygotsky uses the term 'mediation' to mean mastering a higher level of behaviour through taking control of signs and tools for learning. The teacher can guide the experiences encountered by the individual, and assist generalisation. According to Das, Naglieri and Kirby (1994 Ch. 9) Vygotskian tradition argues for the importance of the student's ownership of strategy or principle, needing to derive principles themselves from experience. This is *inductive* reasoning, believed to facilitate the insight to transfer learning.

DA is a concept firmly based in Vygotsky's sociocultural theory, and arising out of his theoretical construct of the ZPD, identified and summarized in the foregoing discussion, and by Lidz and Gindis (2003). Vygotsky never advocated abandoning standardised tests completely, yet allowed for assessment of both developed and developing abilities. Vygotsky stated that there was little to be gained from assessment of a child's level of competence, and spontaneous everyday concepts, when it is the acquisition of scientific concepts that is crucial for development. Instruction should be aimed at the upper limit of the ZPD, to stimulate individuals and enable 'retarded children' in particular to achieve greater competence (Brown and Ferrara 1985). Thus assessment should be of scientific concepts, but taking into account the role of socialization, it would be appropriate to focus on the collaborative forms of thinking that facilitate development of scientific concepts. Thus the notion of assessment of the ZPD arose, capturing the difference between a child's spontaneous performance and his assisted or collaborative performance.

Although he did not provide detail of a DA format, Vygotsky described asking a child to imitate the solution to a problem; to complete an unfinished solution; to work with another more accomplished peer on a problem; and he also explained to the child the problem and the solution for him to carry it out. This exemplifies the approach known as the 'experimental-genetic method', defined later by Leontjew (1931, cited by Guthke 1993) as the study of *transition* from one form of behaviour to another, rather than a description of the new form. Vygotsky also described scenarios in which he established the differing ZPDs in children who scored the same 'intelligence age' on Binet's test, and described the assessment of ZPD as more prescriptive of level of achievement than their current levels of development. He did not, however, proceed to test

this assumption empirically, but laid the groundwork for development of DA procedures by post-Vygotskian researchers. The fundamental notion of assessment of the ZPD has been adopted and modified, and features as a well known idea in psychology and educational research. DA has been operationalized in a number of different ways, although Das and Conway (1992) cite Minick's (1987) view that the ZPD does not lend itself to quantitative measurement, and that the essential characteristics may have been lost in recent (US) research. A great deal of care needs to be taken in the application of the original theoretical concept. Section 1.3 will explore some of the diverse methodologies of DA, but the common feature in all is the active role of the assessor in assisting a child to demonstrate his potential to learn. The more recent move towards feedback and cues delivered by computer, suggests, however, that the fundamental priority of socio-cultural transmission has been lost.

1.2.2   Feuerstein's theory of Structural Cognitive Modifiability

Feuerstein's approach to assessment and intervention derive from his theory of Mediated Learning Experience (MLE), which is in turn linked to his theory of Structural Cognitive Modifiability (SCM) (Feuerstein 1980).

"*SCM describes the unique propensity of human beings to change or modify the structure of their cognitive functioning to adapt to changing demands of life situations*". (Feuerstein et al 2003 p.1.1).

 In his more recent view of structural change Feuerstein (2002) took account of advances in neuropsychology, and postulated that certain cognitive activities can effect ongoing changes in the organism, not only at the level of behaviour, but also on underlying neurological structures. He proposed that the *quality of the*

interaction, rather than the nature of stimuli, is responsible for facilitating significant change. The underlying assumption views behaviours as '*states* rather than *traits'*, and as such they are constantly being changed, or modified. SCM '*considers human beings in terms of their biological and sociocultural development and potential for modifiability'* (Feuerstein et al 2002 p.3).

In order to bring about a cognitive change that is not maturational or temporary, but permanent, pervasive and generalizable (Feuerstein et al 2003) a specific quality of interaction between adult and child in the child's environment, is required. This interaction has been characterised as Mediated Learning Experience, and differs from both direct exposure learning and direct instruction, neither of which ensure that lasting cognitive change results from the learning experience. In the development of this theory, Feuerstein moved away from his Piagetian roots, towards the socio-cultural framework associated with Vygotsky, in which the parents or care-givers of a child assume a central role. Although all children have access to some MLE, the quantity and quality of MLE varies, and according to Haywood (1993) the '*immediate source of inadequate cognitive development is lack of adequate MLE'* (p.28).

In a mediated learning situation, the mediator shapes the experience of the learner, by interposing himself between the stimulus, or the experience, and the individual. He is thus able to help the mediatee attend selectively to relevant stimuli, focus on important aspects, process appropriately using comparisons and links to past experiences, and generalize the experience to new situations (Haywood 1993). In order for an interaction to be characterised as mediational it must contain three essential components, namely intentionality-reciprocity, transcendence and mediation of meaning (Feuerstein and Feuerstein 1991).

Intentionality captures the motivation of the mediator to intentionally transmit information or train a skill in order to modify the individual's behavioural repertoire. In order to be maximally effective, the intention should be met by reciprocal readiness to respond, on the part of the individual. During mediation of transcendence, the '*mediator bridges and connects the current, tangible and perceivable experience, to events in the past and future*' (Lidz 1991 p.14), in which the learning may be applied. Mediation of meaning is, according to Feuerstein, the affective component, by which the mediator conveys the importance and reason for the learning experience, and attaches value to the content. These and other aspects of the MLE are thoroughly described in the writings of Feuerstein and Feuerstein (1991), Lidz (1991) and Haywood (1993).

Feuerstein's assessment tool, the Learning Propensity Assessment Device (LPAD), is based on variables in three key areas, the learner, the assessor and the task. The role of the assessor differs most substantially from the role of a tester in static or standardised tests, and is described in terms of the mediations that he employs, as described above. The variables of the learner relate to his cognitive functioning and modifiability, evaluated by Feuerstein through checking the learner's functionality or deficiency in a series of cognitive processes. Feuerstein has labelled and grouped these cognitive deficiencies according to whether they pertain to the gathering of data, (the 'Input' Phase), problem solving (the Elaboration' phase), or 'Output' of the solutions and responses. A list of these cognitive functions may be seen in Appendix I. The objective of this analysis is to profile the learner's strengths and weaknesses in order to determine the needs for intervention and the

nature and intensity of intervention that may facilitate improved learning.

The final component, the task, is analysed in detail by Feuerstein in a framework known as the 'Cognitive Map' (Feuerstein 1979, see Appendix II).   The cognitive map consists of seven dimensions according to which the task can be categorized, namely content, modality, phase, operation, level of complexity, level of abstraction and level of efficiency. Each of these dimensions is manipulable and can serve as an outline for planning further tasks for assessment or intervention. One area of content, for example, can be assessed via a verbal, symbolic or pictorial modality, and relative skills in each modality compared. The phase of the mental act refers to input, elaboration or output, as described above, and the operation refers to the cognitive activity required. Examples of cognitive activities may be recognition, comparison, classification, seriation, analogical reasoning, inferential thinking etc. The level of complexity may be understood as the number of units requiring manipulation and their familiarity to the learner, whilst the level of abstraction refers to the concreteness of conceptual distance between the mental operation and the object to which it is applied. Thus a hypothetical construct is more abstract than an imagined event, which is in turn more abstract than a real object or event. Finally, the level of efficiency, as a characteristic of a task, relates to the speed at which it is performed, the precision required, and the energy that is expended to achieve the particular act.

In summary, Feuerstein's methodological approach to Dynamic Assessment, the LPAD which will be considered in greater detail in the next section, is rooted in his belief that the function of assessment is to identify what may be done to overcome cognitive deficiencies, or to redress the effect of inadequate MLE in the

individual's past. As every individual has the opportunity to experience MLE, so there is also the possibility that every individual may have experienced inadequate mediation of an aspect of learning, and has the potential to benefit from appropriately targeted MLE. Thus the potential for modifiability is present to a greater or lesser extent in every individual, and the task of assessment is to determine the parameters of learning in the individual, as he relates to a specific task and in the context of MLE from a skilled and motivated assessor.

The LPAD is one of a number of methodologies of assessment that have been devised under the umbrella of 'Dynamic Assessment'. The more prominent of these will be considered in more depth in the next section.

1.3 **Methodologies of Dynamic Assessment**

1.3.1 Classic Methodologies

A small number of methodologies have emerged as differentiated prototypes of DA. Numerous studies have adapted and combined these techniques in multiple applications of the principles to different populations and content domains. The four main approaches to DA are described in this section.

1.3.1.1 The approach of Feuerstein

As introduced above, one of the most comprehensive approaches to DA, and one with a more clinical or prescriptive orientation is associated with the work of Reuven Feuerstein and his colleagues. Feuerstein argued for a test procedure that is flexible and individualized, adapted to each individual being assessed. In summary, his battery of tasks, collectively known as the Learning Propensity Assessment Device (LPAD), generates largely qualitative information about the individual, including information about the

peaks of the individual's performance, i.e. the *best* responses that are elicited, rather than his *usual* performance, and attempts to locate the cognitive processes that are the sources of his success or failure, as well as the amount and type of intervention needed to facilitate learning (Feuerstein 1979, Feuerstein et al 2002).

The emphasis on flexibility and unique adaptation to the individual, while theoretically consistent, has given rise to criticism that there is a lack of empirical data enabling studies of reliability and validity (Grigorenko and Sternberg 1998). Followers of Feuerstein, and those basing research on his methods have been striving to introduce demonstrable reliability and validity into their test procedures, whilst retaining the individualized quality which is an essential component of the procedure from the clinical point of view. Construction of the instruments is supported with empirical evidence (Feuerstein et al 2002), and there is evidence in Feuerstein's writing of provision for very explicit rating and record keeping regarding the input of mediation to individuals, and the responses of individuals to the mediation, (see for example Feuerstein et al 2002 Chapter 12). Training in administration of the LPAD, in the experience of the author, is similarly highly controlled and rigorous with regard to instruction in record keeping. For clinical purposes this ought to be sufficient support for the conclusions reached, and to enable practitioners to fully understand the multiple dimensions of the assessment. For research purposes, the instructions for group administration of some of the LPAD instruments, as described later in this section, may be more useful.

The LPAD battery comprises 15 instruments, any combination of which may be selected to build up a composite picture of the skills of an individual. The instruments vary along three parameters;

    i)   the degree of novelty and complexity of the task

ii)   the language or modality of presentation of the task
        materials

iii)   the mental operations required to solve the given problem

These criteria are extracted from parameters of the cognitive map,
described above. Manipulation of the LPAD instruments enables the
examiner to probe the learner's skills, and produce a profile that
demonstrates the propensity of the individual to grasp the principle
underlying a problem, and solve it; the extent to which that problem
solving can be applied to novel tasks that become progressively
more different from the original task; the differential preferences of
an individual for presentation of a task in different modalities; and
the effects of different types of training strategies provided, in
helping the individual solve the problem (Feuerstein et al 2002
p.165). In contrast to the static test's focus on products of learning
which counts numbers of items achieved, the DA carefully grades
items within a test to enable the testee to learn from items as he
progresses. The presentation of graded tasks, along with feedback
about task performance, enables the learner to learn during the test
procedure itself, and thus his performance reflects ongoing change
rather than a snapshot of ability, and reveals the ceiling of the
individual's ability when given the opportunity to achieve his
maximum potential.

The basic framework of the test administration is a test-teach-retest
model, in which the input during the teach phase is mediational in
nature. There is considerable variation however, in the structure of
the presentation of instruments, and Feuerstein's procedures allow
for mediation at different phases in the test procedure. There is, for
example, preparatory or pre-test mediation that enables the
assessor to orientate the learner to the task and materials, as well
as teaching any prerequisite content. Mediation during the test

consists of regulating behaviour, inputting knowledge, providing feedback and facilitating appropriate responses. Furthermore there is mediation after the performance, regulating behaviour, giving feedback and encouraging reflective insights.

The instruments themselves are differently structured, some consisting of a single mediational intervention phase between two standardised administrations of the test (e.g. Organization of Dots), others consisting of several trials enabling the individual to learn successively from repetition and practice with or without explicit feedback (e.g. 16-word memory test), and some providing additional learning opportunities (e.g. the Complex Figure that comprises copy phase, memory phase, mediated learning phase and then further copy and memory trials). Thus the battery allows for probing of responsiveness to different models of learning and cuing, enabling the assessor to gain a composite profile of the learner's abilities. Although some reviewers (e.g. Grigorenko and Sternberg 1998) find the outcomes of studies of improvements in dynamically assessed and mediated children to be inconclusive, Burns et al (1985, cited by Bransford et al 1987) did demonstrate the superiority of mediation over other forms of intervention in a comparative study of the performances of dynamically assessed children on transfer tasks. Similarly, Swanson and Lussier (2001) showed that the effect sizes of pre-test to post-test improvements were greater following strategy training models of DA (including mediated intervention) than following other DA methods.

Scoring of LPAD tasks varies from one instrument to another, according to the administration process. The instruments have quantifiable scores, and in most cases, pre-test and post-test scores can be compared. The numerical scores may also be used as an indicator of memory, efficiency or particular cognitive functions. The

greater value, however, lies in the qualitative observations of changes in the cognitive functioning of the individual, and his use of processes and strategies, both with and without the mediation from the assessor. Feuerstein has devised a format for recording and summarizing an individual's results and performance on the LPAD, a procedure referred to as 'Profiling Modifiability' (Feuerstein et al LPAD Manual 2003). It is also considered vital to keep a record of the mediational interventions used during the assessment, and Feuerstein has devised a 10 point scale of Required Mediational Intervention (RMI) for this purpose (Feuerstein et al 2002 p.530, see Appendix XX).

Finally, provision has also been made for group administration of the LPAD that allows for simultaneous assessment of a group of individuals. It is a preferred alternative when there is concern about an individual's functioning in a group context, but at all times, within child changes are noted, and group assessment does not permit comparisons between members of a group. Modifications of the instruments for group administration are fully described by Feuerstein (Feuerstein et al 2002 p.221-255). One of the advantages cited is economy of time, as the length of time taken for LPAD assessment is frequently raised as a critical issue. Grigorenko and Sternberg (1998) explore the contention that the investment of time in assessing a child fully is justified by the results and quality of the information obtained from the assessment.

1.3.1.2. Budoff's Learning Potential Assessment

Budoff, who published his work during the 1960s and 1970s, was heavily influenced by findings that children from minority and 'poor' homes were frequently misclassified as having some kind of learning disability, and placed in special education classes. He noted that the assumption that all children progress at roughly the same rate was

violated in the case of children with fewer and less frequent 'access to school-preparatory experiences' (Budoff 1987a p.52).

The response of Budoff and his colleagues was to devise a test procedure that offered children the opportunity to demonstrate that they are able to perform at a level average for their peers, on a non-verbal reasoning task, when offered a short period of training on a similar task. Thus he is credited with devising a test-train-retest procedure. The pre-training familiarised children with the task content, and the requirements of the test, as well as reducing anxiety associated with the context of testing and expectations of failure. They aimed therefore to equalize the pre-test experience of the children, to enable them to perform more competently. The tasks were limited to non-verbal abilities in order to minimise the effects of linguistically weak backgrounds, and like Feuerstein, Budoff chose to use tasks minimally related to academic learning (Lidz 1991).

Budoff demonstrated the ability of children previously diagnosed as 'low IQ' to work effectively with cognitively challenging reasoning problems, showing that training leads to improved performance on post-test, and establishing construct validity (Grigorenko and Sternberg 1998). He was also able to identify learning needs in this population, such as impoverished language, and difficulties applying learning strategies spontaneously without prompting. The procedure used '*relatively standardised training procedures'* (Budoff 1987a p.57) and focused on the contribution of the tester and of the training to devise interventions suitable for individual learners.

In his early studies, Budoff identified three patterns of response to his training procedures, groups that he termed 'High scorers', 'Gainers' and 'Non-gainers'. Subsequently, Budoff decided that the

scores of children demonstrating considerable gain could not be directly compared to the scores of 'high scorers' and consideration of pre-training scores was needed. In later work, Budoff used pre-training scores, post-training scores and 'post-training adjusted for pre-test level' as outcome variables. Subsequent studies (e.g. Carlson and Wiedl 1978, Embretson 1987b) have moved away from Budoff's preference for gain scores, and used post test scores only, finding these to be more representative than gain scores for the reasons given above, that gain scores cannot be compared without accounting for level of the pre-test performance. The current study has adapted Budoff's notion of adjusting the gains post-intervention to account for the variability in baseline levels of progress.

Cumulative findings from Budoff's research over several years (see review, Budoff 1987a and 1987b) demonstrated links between the findings of training based assessments and criterion measures related to school achievement, and personality data. As a result of these findings, Budoff coined the term 'educationally handicapped' for children who have not progressed well at school, but have demonstrated potential, and this group is contrasted with those who do not profit from training. High correlations were achieved between Budoff's learning potential tests and subsequent performance on measures of educational achievement, as well as teachers' ratings of ability. Measures of learning potential were more predictive of achievement than IQ scores. Furthermore, correlations were found between learning potential and personality variables, and also with low SES, but not with other demographic data (Grigorenko and Sternberg 1998).

It is possible, however, that altering training strategies could alter the performance of low-performing children, but there is no clear methodology linking the results of Budoff's training based

assessments to intervention programming. This would seem to be a necessary extension to Budoff's work, enabling outcomes from individualised interventions to be evaluated.

According to Grigorenko and Sternberg (1998), what makes Budoff's methodology distinct from other test-teach-retest procedures, is that it is explicitly intended as an alternative to static standardised intelligence testing, and only makes use of adaptations of standardised, reliable and extensively validated tests (p.37). To this end, Budoff and his colleagues developed dynamic versions of several well known tests, for example the Kohs Learning Potential Task, and the Raven Progressive Matrices Learning Potential Test, a battery which assesses problem solving in different modalities. The pre-test and post-test components are administered as static, standardised tests, while the training stage is specific about the role of the examiner. The training directs the student's attention, explains the crucial attributes of the task, and guides the student in mastering the actions needed for finding the solution to the problem. Whilst the content of the training is 'approximate and not absolute' (Grigorenko and Sternberg 1998), it is also not entirely contingent on the performance of the student. However, precisely how the training is standardised is not clear.

Strengths of Budoff's contribution to the field of DA, include the useful differentiation of low-IQ students into those with high learning potential, who may have been disadvantaged, from those with significant learning difficulties. In contrast to diagnostic uses of DA, this enables differentiation within the population of low achieving students, and as such lends itself to the prescription of interventions to facilitate improvements in this population. Specific recommendations for intervention, however, were not elaborated by Budoff, who implied only that deficits identified in assessment could

be addressed in training. Links between the assessor's interventions and training needs were not explored (Lidz 1991). Useful aspects of Budoff's work are the links to personality and demographic factors that place assessments of learning into context. Because tools are adapted from standardised tests, administration of the procedures is easy for examiners to grasp, and instruments have been shown to be robust. Criticisms include the limitation of the test paradigm to a fairly specific population of children with learning difficulties.

1.3.1.3. Graduated Prompting method

The Graduated Prompting method proposed by Campione and Brown was based on the work of Soviet psychologists, including Vygotsky, and on the notion of ZPD (Campione and Brown 1987). Vygotsky described the process by which children learn through social mediation, in which the adult guiding the child's learning and modelling the task solution, gradually reduces their input as the child gains independence. Children who display readiness to learn a task benefit from intervention in that task, while those with less readiness, a 'narrower' ZPD, require more direction to succeed in the task. In the traditional paradigm, the 'width' of the ZPD is represented by the difference between an individual's independent performance, and that which he is able to achieve when assisted by an adult, teacher or more competent peer. The potential for improvement, or the ZPD, reflects the 'immature' concepts that have been acquired by the child that can readily be upgraded or advanced by input or mediation from another. Children with a narrower ZPD may be those for whom a great deal of assistance or intervention is required in order for them to progress, or those for whom only a limited range of skills may be achieved even with assistance from a peer.   This framework influenced the work of Campione and Brown, who were involved in evaluation of the academic progress of weak students.

In summary, the procedure of Graduated Prompting consisted of giving a child one item of a problem, (comparable to an item from a standard IQ test), to solve independently. If the child failed, the assessor systematically added cues, consisting of successively more directive and specific hints, until the child achieved a solution to the task. A count of how much help the child needed to solve the problem was a measure of the child's ZPD. The procedure did not look at how many problems or items a child managed to complete. Once solved, another version of the task was presented, and transfer calculated by seeing how many more hints the child needed to solve this item. Subsequent tasks built on learning acquired in earlier ones. Testing the ZPD like this involved detailed task analysis of a suitable set of tasks, and analysis of possible transfer probes (Campione, Brown and Ferrara 1982).

Initially, the series of training studies was devised as a theoretical tool, rather than for remediation purposes, as Campione and Brown tried to elucidate the nature of intelligence by analysing the nature of the difficulties experienced by 'retarded children' (Campione, Brown and Ferrara 1982 p.393). The logic of their theory begins with the selection of a task, and the question of why some children do poorly on that task. If, in theory, components *A, B* and *C* are required to perform a task successfully, then the hypothesis would be that a 'retarded' child has a weakness in one of these components, for example, *A.* To test this hypothesis, one must train the child on *A* and see if his performance on the task improves (Campione, Brown and Ferrara 1982 p. 412). If there is no improvement, it may be that *A* was not the essential component of the task, or that the child was not weak in *A.* Similarly, if improvement is elicited, then an individual may need additional intervention to improve his skills in *A,* or he may be able to use strengths in a different component skill to compensate for his weakness. The advantage of this approach is

that like Feuerstein's model, it evaluates what is needed in a task, and identifies the processes that are needed to teach to the children.

In addition, Campione and Brown were concerned with facilitating *transfer* of learning, and to this end they determined that problems needed to be situated within rule-governed systems, in which students could learn how to determine which responses to a problem would be appropriate. The system of administering successively more direct 'hints', enabled them to determine the minimal amount of help a child needed in order to solve a problem, and again, how many hints the child needed in order to transfer learning to different problems. In order to maintain a reliable procedure with 'good psychometric properties' (Campione and Brown 1987 p.90) the problem solving hints given to children were kept consistent, in a fixed sequence, and were therefore related to the task, rather than to the child. Later the prompts would be delivered by computer, regulating the procedure and saving time (Campione 1989).

Campione and Brown (1987) reported on a number of studies using their procedure, and noted that they were able to demonstrate both concurrent and predictive validity of the procedure. In particular, the students' performance on transfer tasks were most strongly related to ability measures (p.100) and provided individual information leading to intervention strategies. Grigorenko and Sternberg (1998) affirm that the Graduated Prompting method demonstrated unique quantification and standardization of learning and transfer. Recommendations for intervention were derived from the principles that facilitate transfer, such as training in multiple settings, attention to metacognitive instruction, and explicit teaching of applicability of skills (p.100). Finally, Campione and Brown also noted that the procedure enabled observation of 'personality factors' and variables in approaches to the task and reactions to the prompt procedure,

that were individually informative and potentially predictive of progress. This however is an issue that Grigorenko and Sternberg find problematic, and it emerges that the hints or prompts given are qualitative in nature, and not necessarily of equal value or impact to different individuals. Thus quantification of the amount of help given is not absolute, and may be best interpreted in the light of other cognitive skills such as attention, memory, and learning style.

Having carried out their earlier studies using domain general tasks of reasoning that were associated with scholastic success, Campione and Brown reported that the method was being extended to specific domains, notably maths abilities, to enable tracking of progress and development over a longer period of development. An advantage of the DA procedure, identified by Campione in 1989, was that the procedure of testing also contributes to the instruction process, in other words students learn valuable lessons whilst undergoing assessment. This process of 'learning through testing' is significant in mediational assessments and crucial when considering the psychometric approach to DA and will be discussed again later in this context. Further advantage could be gained by interspersing assessment with regular teaching sessions, evaluating learning progress as it took place. Placing the assessment in a specific subject domain further enabled instruction to be carried out in context.

The procedure of Graduated Prompts differs from other Dynamic Assessments in its focus on counting and measuring aspects of the task, rather than achievements of the individual, but applications and variations of the procedure have been widespread and the method remains central to many current European studies in DA (e.g. Resing 1997, Resing et al 2009).

1.3.1.4   Testing the Limits

Carlson and Wiedl (1992) described the two categories of Dynamic Assessments as those that measure training gains, and those that gauge the effects of approaches that improve performance. Under the rubric of 'Testing the Limits' they addressed the second of these types, and demonstrated the improved performance of children with learning disabilities when tests were administered under different conditions. These different conditions consisted of levels of feedback given to a child during the test procedure.

In their classic study (Wiedl and Carlson 1981 cited by Carlson and Wiedl 1992), the experimenters administered the Ravens CPM to a group of children with learning disabilities, under three conditions. The first of these (C1) was standard instructions for the test. The second (C2) incorporated elaborated feedback to the child after his/her response, explaining the principles for solving the item. In the third condition (C3), the child was required to describe the pattern of the item, describe the answer alternatives and give his/her reasons for the choice made. The examiner then provided elaborated feedback as in condition 2 (p.159).

The child's responses were scored prior to the feedback, and improvements were seen in the transfer of learning from one item to a subsequent example. This scoring method is seen by Haywood and Lidz (2007) as advantageous as the procedures can be applied and scored in such a way as to preserve the normative scoring of the test, prior to implementing modifications to the procedure that alter test performance. The authors demonstrated improved performance by the children, particularly in condition C3 as compared to C1. The test modifications applied by Carlson and Wiedl were also used in conjunction with other cognitive tasks, with similar results, and applied to individuals with a variety of behavioural and personality

traits. The differences in test performance that were seen in impulsive versus reflective children, for example, were reduced under conditions of elaborated feedback and verbalization as described. They concluded that in order to elicit optimal test performance from a child, the administration procedure should be applied differently, as defined by the specific individual or group, being assessed.

Jeltova et al (2007) noted that Carlson and Wiedl set out to find a match between the individual and the test situation, to evoke the best performance from the individual, and that success of the DA is dependent on finding the best fit between the characteristics of the learner and demands of the test. They used this notion of 'fit' derived from Testing the Limits procedures, to inform their Individual Curriculum Based DA (ICBDA).

In an alternative Testing the Limits method, Peña (2001) and Ginsburg (1997) (cited by Gutierrez-Clellen and Peña 2001) incorporated a 'clinical interview' into the test procedure. They formulated questions to help children understand their thinking and explain their approach to the test items, and found that children were better able to demonstrate their knowledge, than during a static test. Peña further moved away from Carlson and Wiedl's method of giving feedback, towards more metacognitive probing questions, such as 'How did you know that?' (Gutierrez-Clellen and Peña 2001 p.214) and found that with this type of questioning she was able to elicit more information about the children's abilities and understanding, and enable them to achieve a better result from testing. Although Peña recommended rigorous record keeping, reliability might be compromised by variation in administration by different examiners. Gutierrez-Clellen and Peña note, however, that there is an apparent increase in the face validity of the test, and

gaining knowledge of children's true competence would seem to the present author to be useful in the formulation of clinical intervention for children with difficulties.

Grigorenko and Sternberg (1998) reviewed the work of Carlson and Wiedl, and others on verbalization effects, and found the Testing the Limits approach to be most appropriate for assessing higher level cognitive functions, especially in those whose performance is low (p.116). Carlson and Wiedl (1992) like Vygotsky and Das (Das and Conway 1992), proposed that one of the key factors underlying improved test performance is metacognitive reflection. Test takers who monitor their own learning and performance, are likely to be more successful in the task, and this metacognitive awareness is in turn facilitated by overt concurrent verbalization. This explanation of one's own thought processes enables new links to be made, restructures thought processes, enables reflection about one's own comprehension, feedback about performance, and perception of one's own abilities that contributes to confidence and self esteem. The difficulty encountered however, as alluded to by Gutierrez-Clellen and Peña (2001) is the requirement that children have sufficient verbal as well as metalinguistic and metacognitive skills to reflect on their performance.

Strengths of Carlson and Wiedl's Testing the Limits paradigm are its inclusion of personality as a variable in DA, and the demonstration that personality and other extraneous individual variables, such as anxiety and impulsiveness, firstly affect test performance, and secondly can be compensated by specific DA methods. The main disadvantage identified by Grigorenko and Sternberg is that all trials were conducted on groups, and no individual data are presented. In spite of this, the procedure does aim to link findings to goals for intervention and make recommendations for the conditions under

which groups of individuals might perform optimally. Studies by Peña (2001) have successfully used clinical interviewing with individual children. Their caution with respect to the Testing the Limits approach was that it had not been shown to empirically differentiate language disordered from disadvantaged children and for that reason it should be used with additional DA procedures. It would appear to the current author to be a procedure that combines well with other procedures, such as Graduated Prompting; the probing of responses would add information about the individual's understanding and use of the prompt information, and the Graduated Prompt scoring would serve to increase the retest and inter-rater reliability of the whole test, while permitting individualized and flexible probing of responses. Furthermore, the Testing the Limits feedback and questioning could be usefully carried out in the mediational style advocated by Feuerstein and Feuerstein (1991), to further promote transfer of learning.

1.3.2 **Other methods of Dynamic Assessment**

Fundamental models of DA described above have been variously modified and combined by major contributors to the field of DA, resulting in a number of unique formulations of assessments. Three significant bodies of research representing shifts in particular directions that are relevant to the current study have been selected for further elaboration in this subsection. These three approaches have been selected because they demonstrate individualised and adaptive testing combined with standardised administration, and provide links to recommendations for intervention.

1.3.2.1   Guthke's 'Lerntest'

Guthke and colleagues (Guthke and Wiedl 1996, Guthke and Wingenfeld 1992, both cited by Guthke, Beckmann and Dobat 1997; Guthke 1993) attempted to combine the individualised testing

priority advocated by Feuerstein with a requirement for psychometric rigour and standardization. Guthke based his research firmly on the thinking of Vygotsky, and the tradition of East German psychologists to follow the post-Vygotskian Soviet thinking and adopt the assessment of learning as the dominant paradigm, embedded in ecologically valid contexts. The earliest result was Guthke's long-term test, using a test-teach-retest design.

The strength of the long-term test was its thoroughness in assessing verbal, numerical and figural modalities, the consistency with established test theory, and the standardization of normative samples. Equivalent parallel tests were constructed for pre- and post training, and instruction was programmed for standardization. Instruction took the form of training thinking skills and metacognitive abilities, with opportunity for a large number of practice exercises, and feedback regarding errors made. Guthke based norms on the performance of students on the post-test, rather than the pre-post test gain, which he found to have low reliability and validity.

In order to overcome the drawback of the lengthy time of administration, Guthke devised the short-term learning potential tests in which feedback and assistance were included within the test procedure which is reduced to one testing session. One of the models made use of Raven's Coloured Progressive Matrices, with feedback and assistance built into the administration. These methods were empirically studied over a long period of time. Findings demonstrated that performance increased after training, and that a period of 90 minutes of training was sufficient to elicit the change. In addition, predictive validity, construct validation and longitudinal effects were empirically studied (summarized by Guthke 1993 p.53) and confirmed. One of the more striking findings was that using the

Raven learning potential test, the learning potential tests were shown to be significantly better than static tests at predicting outcomes in children with below average learning. The authors concurred with Feuerstein (1979), Lidz (1987) and others, that learning potential tests were particularly suited for use with children with learning difficulties and 'irregular learning histories' (p.54). Differential diagnoses of children were made possible by different patterns of response to prompts, and this procedure became the foundation for a series of further tests of learning potential developed in Germany, (summarized by Guthke 1993) and in European schools, e.g. Hessels and Hamers (1993, cited by Grigorenko and Sternberg 1998).

Subsequently, further attempts to increase individualization resulted in 'diagnostic programmes' that combined the learning tests with an adaptive method, in which prompts or additional items were added to the test as required by the individual, whilst core target items that must be attempted by all, were retained. The complex branching structure of administration was managed by computer-assisted application, and became known as the Adaptive Computer Assisted Intelligence Learning Test Battery (ACIL) (Guthke, Beckmann, and Dobat 1997).

The authors demonstrated experimentally that learning tests were more significantly correlated with measures of knowledge acquisition and were more predictive of gains in curriculum related learning than were static test scores. Grigorenko and Sternberg (1998) however, noted that these findings should be interpreted with caution due to the small sample sizes involved. Furthermore, doubts remain as to the generalizations that can be made with regard to curriculum related tasks on the basis of performance on tests using abstract tests of 'intelligence', and in this respect, the criticisms faced by the

authors are the same as those directed at Feuerstein. Nevertheless, the attempt to combine individualized, adaptive testing with standardised aspects of administration that can be scored, is to be applauded.

1.3.2.2 Lidz, Curriculum-based DA

Due to a lack of conviction that approaches to DA provided adequate links to meaningful intervention, (Lidz 2003 p.116) Lidz devised a curriculum-based Dynamic Assessment procedure. Tasks were taken from classroom activities, but methods were based on Feuerstein's MLE in the format of test-mediate-retest. Lidz set out to link the processing demands of the task with the mental processing capacities of the child, and formulated both quantitative and qualitative (but not normative) frameworks for the information elicited. The aspects of processing that are identified in the task and addressed in the learner are the domain general ones of attention, perception, memory and metacognition. Thus while altering the emphasis away from the 'cognitive deficiencies' of Feuerstein (1979), Lidz retains the focus on domain general processing skills, and the mediation of strategies to manage learning.

The Application of Cognitive Functions Scale (ACFS, Lidz 2000; Haywood and Lidz 2007) was devised as a curriculum-based assessment package for preschool children. At the preschool level however, the curriculum is less subject content based than focused on the development of prerequisite knowledge and skills, such as those assessed by the ACFS, classification, auditory memory, visual memory, and pattern sequencing, and uniquely to the ACFS, planning and perspective taking. The package is intended to facilitate development of skills in children who are not yet independent in these areas, and thus may be identified as causing concern in preschool learning environments. Questions pertaining to the DA of

preschool children centre on their as yet undeveloped metacognitive skills, but Lidz and indeed Tzuriel (1997, 2000) address the early developing executive skills that are developmentally typical of children of this age. Although the programme was devised by Lidz with reference to preschool curricula in the US, the concepts are equally applicable to the content of the Foundation stage curriculum in the UK (Department for Education 2006).

The ACFS consists of six subtests, four of which are designated as 'core' and two as supplementary subscales, however the assessor may select which of the subtests to administer, with the recommendation that at least two or three are used to allow comparison of skills. Each subtest is scored using raw scores and percentages correct, in pre- and post testing, although scoring includes points awarded for metacognitive and reflective responses justifying or explaining answers. In addition, after every pre-test and every mediation, Lidz recommends that the assessor complete the Behaviour Rating Scale, a 3-point rating scale based on seven parameters, namely self regulation, persistence, frustration tolerance, motivation, flexibility, interactivity and responsivity. These are accompanied by behavioural descriptors to guide scoring.

In summary, what is unique to the ACFS, other than its application to a preschool population, is the combination of a behavioural rating with pre-test and post-test scoring, and the inclusion of considerable qualitative criteria in the scoring. Detailed task analysis links to teaching objectives, and interventions are linked not only to performance in the six subscales, but also to ratings on the parameters of the behaviour rating scale, so that a child may have targets for auditory or visual memory as well as for persistence or flexibility in a task. The ACFS does not include a subtest specifically on language, but many of the tasks are verbally loaded and although

well developed language skills are not required, the test cannot be used with nonverbal children. Haywood and Lidz (2007) also note that within the test procedure there are multiple opportunities for the child to demonstrate his language and in particular the use of language in problem solving.

In addition, of considerable value is Lidz's generic format for curriculum-based DA, which encourages and facilitates practitioners in a range of fields and contexts to devise and apply DA in their own work. Lidz outlines the steps required for this DA and provides helpful planning and score sheets. She has also introduced the Response to Mediation rating scale that enables practitioners to score behavioural responses quantitatively, instead of incorporating complex statistical computations of valid gain scores from the procedure, which then need not be standardised on the given population.

1.3.2.3   Work of the 'Vanderbilt group', the 'Continuum of Assessment'

Models of DA based on combining traditional formats are advocated by several authors, (e.g. Campione 1989, Guthke Beckmann and Dobat 1997) and the research team of Burns, Delclos, Bransford and Vye, known as the 'Vanderbilt group', devised just such a procedure. Termed the 'Continuum of Assessment Services' model, (Vye et al 1987) the procedure started with administration of a static intelligence test. Children scoring below 1 standard deviation below the norm on this test would be retested using a Graduated Prompting form of DA. Those who scored above criterion on this measure would be deemed 'responsive to instruction' while those who did not would continue to receive mediated intervention in the next stage of assessment. The mediated intervention aimed to assess in more detail the individual's ability to learn from more

intensive individualised instruction, rather than being linked to achievement of any given criterion.

The model therefore drew on both Feuerstein's principles of mediated assessment, and the Graduated Prompt method of Campione and Brown. The mediated assessment was necessarily shortened to be useful within one 30 minute assessment session, and as a result was limited to one domain of testing, such as perceptual performance. If required, a further series of mediation sessions addressing other domains could be carried out. Both the mediations used and the sequence of graded prompts were specifically scripted and uniform, but the mediational prompts were delivered contingent on the child's performance whereas the graduated prompts were based on analysis of the task.

The Continuum of Assessment model followed extensive review of DA methods by the Vanderbilt group who systematically compared the outcomes of different methods of DA and of static tests. They emphasized that the mediated DA procedure of Feuerstein, produced more learning in a child than 'Testing the Limits' instruction, and greater transfer than the Graduated Prompting method of Campione and Brown. They further challenged the assertion that DA is predictive of future performance by breaking down the predictions, and found that DA is predictive of within-domain transfer, but not of across-domain transfer. The differences on transfer performance, however, differed with variables in the children, such that mediated instruction facilitated greater transfer in those who scored less well on criterion measures. The group concluded that the Graduated Prompting method may be more useful as a comparative measure of *relative* learning ability, while mediated methods are better suited for diagnostic-prescriptive purposes.

Burns and colleagues initially planned to investigate the utility of DA for identification of difficulties in preschool children, and to this end focused on identifying the cognitive behaviours of young children. Their observations formed the basis for formulating general teaching strategies for young children that could be communicated to teachers. In another demonstration of the thoroughness of their approach, the Vanderbilt group studied the impact of observations of Dynamic Assessment sessions on teachers, in comparison to their response to static assessments. In general, observation of Dynamic Assessments resulted in more positive ratings of the children than from static assessment sessions, and the authors concluded that expectations of teachers could be altered by their opportunity to observe children in sessions of Dynamic Assessment.

The findings of the Vanderbilt group were indeed comprehensive, as claimed in the title of the 1987 chapter 'A Comprehensive Approach to Assessing Intellectually Handicapped Children' (Vye et al 1987). Their aims encompassed a range of those addressed more specifically by other researchers, such as identification, prediction and the informing of intervention, and they systematically evaluated methods best suited to each primary purpose. They willingly examined the advantages and disadvantages of the different tasks employed, and the utility of testing general cognitive skills versus working within subject specific areas in order to plan for transfer of intervention within that area. They also considered the transmission of outcomes to teachers, and have provided empirical evidence to support the decisions they have taken. Their research is therefore useful and applicable to researchers and clinicians from many fields who can similarly make informed choices between the range of materials and methods available under the rubric of 'Dynamic Assessment'.

1.4   Contemporary Issues in Dynamic Assessment

A review of current papers on the extensive www.dynamicassessment.com website reveals the expanding field of applications of DA and its justification as an approach to assessment in a range of contexts. Those relevant to the field of language will be reviewed in Chapter 2.

A small number of papers reviewed a range of DA methods, in order to evaluate common strengths and difficulties in the field of DA (e.g. Grigorenko and Sternberg 1998; Swanson and Lussier 2001). Certain methodological alternatives, such as domain general versus domain specific tasks, continue to arouse debate. A considerable number of papers are devoted to the issues of reliability, validity and psychometric properties of DA. A review of all of these issues is relevant to the formulation of the methodology of the current study.

1.4.1 Comparative studies

A number of papers review effectiveness of different approaches to DA, according to specific criteria, but as identified in the definition of DA, the aims of process assessment are varied and methods have been developed to meet specific aims, thus may not be directly comparable. For example, in an early review of studies, Vye et al (1987), reported that the mediated DA procedure of Feuerstein, produced more learning in a child than 'Testing-the-Limits' instruction. Further, although there may have been comparable effects on *learning* of a task to criterion level, mediational methods produced greater *transfer* than the Graduated Prompting method of Campione and Brown. This may be the result, however, of the intention of the assessment addressing a specific purpose, and as Graduated Prompting is a method developed for issues related to classification rather than planning instruction, it was not intended to identify transfer to the same extent as mediated learning strategies. Missiuna and Samuels (1988) arrived at similar conclusions, and

commented that Feuerstein's DA approach '*may be the only assessment method which can be directly prescriptive of effective intervention practices for a specific child*' (p.14), when in reality, Feuerstein's approach is one of the few that actually sets out to identify individualised interventions. Campione (1989) also noted that all of the interventions used succeed in improving performance and increasing predictive validity when compared to static tests, but that those aiming to inform intervention and therefore have to demonstrate increased academic success, face a greater challenge and more stringent criteria for evaluating their own success.

Jitendra and Kameenui (1993) again reviewed different models of DA, evaluated their contributions to the field, and arrived at similar conclusions that the value contributed by differing models relates to the purpose for which they were devised. They continued, however, to consider the contribution made by different models to prevailing problems in the field of DA. Thus, for example, the range of models and methods contribute to 'Construct Fuzziness' that is that the features of a DA are not consistent and clear, as well as 'Procedural Spuriousness' which is due to inadequate definition of the nature and complexity of prompts and interventions (p.14). Jitendra and Kameenui further noted difficulties evaluating the generalization of findings based on tasks that are 'Instructionally aloof' (p.14) or otherwise described as 'domain general'. Studies of Graduated Prompting that made use of tasks related to an academic curriculum are those that can be, and have been empirically tested. The difficulty with these, however, is that there would be a need for separate assessments in each academic domain, resulting in a vast number of tests, and many assessment tools in all models are already subject to inadequate operational specificity.

Finally, Jitendra and Kameenui identified a recurrent criticism of all models of DA, namely the time taken and labour intensiveness of DA

procedures. The issue of whether the information extracted from a DA justifies the investment of time and effort in the assessment process is moot, but it is clear that researchers need to aim to devise assessment procedures with shorter administration times if they are to be widely adopted by practitioners.

In a slightly later paper, Grigorenko and Sternberg (1998) carried out a review of dynamic testing studies, summarized the literature, discussed achievements and limitations of various approaches, and focussed on the underlying psychological models and hard empirical data supporting the studies (p.80). In order to facilitate this review, approaches to DA were divided into four clusters, as follows:

i)   metacognitive intervention (Feuerstein)

ii)   learning within the test (Graduated Prompts, Guthke, Campione and Brown)

iii)   restructuring the test situation (Budoff, Carlson and Wiedl)

iv)   training a single cognitive function (Swanson's WM, Spector, Peña)

The first three of the four clusters are the same as those used by Haywood (1997, cited by Haywood and Lidz 2007), and broadly parallel to three out of four categories in Campione's taxonomy (1989). They are, however, labelled according to different descriptors, Campione distinguishing categories according to their use of standardised versus individualized clinical intervention, and their use of general versus domain specific targets of assessment. Thus, Feuerstein's metacognitive intervention is described as Clinical Intervention/General skills. Both Grigorenko and Sternberg, and Campione include the work of Budoff and the testing-the-limits of Carlson and Wiedl in one category, whilst they are frequently

separated, for example by Jitendra and Kameenui, into 'test-train-retest' and 'testing the limits', and by Swanson and Lussier into 'Coaching' and 'Scaffolding'. This difference sometimes makes outcomes of reviews difficult to compare.

The first criterion upon which studies or clusters were evaluated, was termed by Grigorenko and Sternberg 'Comparative informativeness' i.e. '*whether the given method contributes any new information over and above that obtained with conventional measures*' (p.93). The authors found that while different approaches to DA set different aims for their studies, primarily whether they targeted the measurement of change to contribute to research or focussed on the enhancement of change to facilitate instruction; each was able to demonstrate unique information. Thus they concluded that the general claim that DA can elicit additional data from a test situation does seem to have been justified.

However, Grigorenko and Sternberg also identified a drawback common to all approaches to DA, namely the problem of measuring change. Almost all DA methods are subject to practice effects as they employ pre- and post- testing, and complex statistical treatments have as yet not yielded a single widely accepted method of measurement. Furthermore, establishment of predictive validity of various DA tasks has been problematic, but demonstrable to a greater extent when there is a closer match between domain specificity of the learning task used in the DA, and domain specificity of the criterion task.

In 2001 Swanson and Lussier published 'A Selective Review of Experimental Literature on Dynamic Assessment' that included, amongst numerous other important criteria, a comparison of the effectiveness of different types of assessment. Swanson and Lussier

divided studies into 3 categories according to 'type' of assessment namely 'scaffolding', 'coaching' and 'strategy training'.

Swanson and Lussier concluded that studies utilizing 'strategy training' methods elicit greater effect sizes than those using scaffolding, which in turn are greater than those using coaching. This extremely useful finding is compromised by closer examination of the data which reveals that due to studies not publishing the required data, they could not be included in the review, and thus the results represent several contributions by a small number of research groups, who have published the required level of detailed data. The wealth of research by Feuerstein and his associates, for example, is not included at all, due presumably to the lack of empirical information in his publications, as noted by Grigorenko and Sternberg (1998). Research utilizing modifications of methods developed by Feuerstein, for example by Tzuriel and Caspi (1992), Tzuriel and Klein (1985), Keane (1987), Samuels, Tzuriel and Malloy-Miller (1987), (all cited by Missiuna and Samuels 1989), and Peña, and Iglesias (1992), are represented in the strategy training category, which was shown to be maximally effective, and thus inclusion of the achievements of Feuerstein may have been added to this finding. Publications by Campione and Brown referring to their work with Graduated Prompts are also not included.

Of further considerable interest is Swanson and Lussier's report, that some categories and some levels produce such low effect sizes that the results of the DA are not meaningful. Included in these observations were low effect sizes for studies that include verbal measures, studies that include groups of participants with leaning disabilities, studies providing coaching or mediated instruction, and studies using within-subject designs. These criteria would represent those methods most used in clinically oriented studies, such as those in the Feuerstein tradition, and those with emphasis on the

prescription and evaluation of intervention. DAs applied to language impaired children would similarly incorporate all of these aspects, and may be predicted therefore to produce low effect sizes. As several studies have shown maximum usefulness of DA for groups of children with learning difficulties and in special education, it may be that the studies have fallen foul of the specific effect size corrections employed by Swanson and Lussier for the purposes of their study.

Review papers therefore appear to have identified strengths and weaknesses of DA that are common to all the procedures, or indeed to the concept as a whole, with comparative evaluation that would demonstrate superiority of one approach over another being inconclusive. This is in part due to difficulties comparing methods that have aim to serve different purposes, and in part also due to the differing ways in which the authors have grouped methodologies. One of the most enduring debates is that of the use of domain general cognitive skills versus testing in a specific content area, and only Campione's taxonomy explicitly uses this as a criterion to distinguish and describe the classic methodologies. Nevertheless, the dichotomy is worthy of further examination.

1.4.2 Domain general vs Domain specific testing

Feuerstein, in his emphasis on accessing the true learning potential of children from a variety of social and cultural backgrounds, stressed the importance of testing *domain general* learning skills, free from the bias of the structured learning experience of Western schooling. Nevertheless, he acknowledged the need to learn specific content and in particular, language, by including these in his list of cognitive functions (Appendix I). Furthermore, he tapped into learning in different modalities, such as figural, symbolic, numeric and verbal, by employing a battery of tasks. Budoff likewise employed a smaller battery of tasks based on standardised tests of

reasoning, specifying in addition that the tasks should be non-verbal, so as to minimize bias against those with linguistically disadvantaged backgrounds.

Similarly, in a later study, Sternberg and Grigorenko (2001) presented a procedure for the assessment of learning potential in rural Tanzanian children. Being unable to use curriculum-based tasks, they chose to utilize conventional domain general cognitive assessments, with an altered mode of administration. This enabled the children to be tested in their own language, and familiarised with the test procedure until the examiner was satisfied that the children understood the task, and the examiners were therefore reassured that the testing was valid.

In an alternative approach, Swanson (1995) made use of the domain general information processing model, specifically addressing ability in one aspect, namely working memory (WM) in the S-CPT, the Cognitive Processing Test (Swanson 1996). He pointed out that all major information processing models include the component of WM which is related to academic and language related skills, vocabulary, reading comprehension, problem solving and mathematics. Using the assessment of WM in eleven different subtests of verbal or visuo-spatial processing, Swanson was able to construct a comprehensive assessment not only of the components of WM, but also of learning strategies and ability to benefit from intervention. Detailed psychometric analysis of his findings confirmed the construct validity of the S-CPT in comparison to other measures of WM and correlation with measures of general learning ability such as the Peabody Individual Achievement Test and the Detroit Test of Learning Aptitude. Validity studies also demonstrated that WM was modifiable by a prompted DA procedure, and that the scores from the DA improved predictions in reading comprehension.

The overarching challenge facing domain general assessments is the lack of evidence that training cognitive skills impacts positively on curriculum or subject specific learning. Like other domain general DA procedures, the S-CPT was shown to distinguish between learners on the basis of their ability to benefit from greater or lesser amounts of intervention, but Swanson effectively argued the notion that there is a link between assessment and instruction due to the commonalities in the methods used in his test and in the classroom. The supportive teaching role adopted by the dynamic assessor parallels that of the teachers in its efforts to engage with a child and attempt to induce learning, or change.

Swanson (1995 p.681) noted that it would be unlikely that a child would display markedly different responsiveness to teaching cues in a DA test situation and in the classroom. Children who are not responsive to probes and cues would most likely benefit from procedures which place lesser demands on constructing strategies independently, but are more didactic and use drill and practice to teach concepts. Swanson summarized the information derived from the S-CPT as relating to effectiveness of simple feedback on examinee performance, examinee's general knowledge of strategies, degree to which performance is maintained after hints are removed; and finally, examinee's preference for verbal or visuo-spatial information, all of which may inform teachers about the cues needed to support learning in a child.

In spite of this assertion, a focus on domain specific and curriculum-based assessments has been maintained by other researchers, in order to draw conclusions about how teachers might support students' specific difficulties, and in order to evaluate validity by comparison with school based assessments. Curriculum-based assessments have been modified to strengthen the links to required

intervention by construction of dynamic versions of the assessments, for example the CBDA, Curriculum Based Dynamic Assessment (Haywood and Lidz 2007 p.176). Lidz (1991) also demonstrated the presence of general process skills across domains of the curriculum, and thus the need to assess features of attention, memory, reasoning and metacognition throughout a CBDA. According to Lidz, Jepsen and Miller (1997), the objective of the DA materials is '*to reveal underlying correlates of low achievement that are potentially responsive to remediation*' (p.57).

In a model that is typical of the thinking behind curriculum-based assessments, Jeltova et al (2007) have reported use of the Individual Curriculum Based Dynamic Assessment (ICBDA), which addresses competence in the particular curriculum area of mathematics which can be extracted directly from the curriculum of a particular grade level. The pre-test assesses both content knowledge and the learner's relative strengths in cognitive modalities, while the teaching component makes use of specific subject content. As a result, the assessment differentiates difficulties that results from poor reading of the questions and test materials, poor mathematics skills or poor problem solving, and students can be taught in different cognitive modalities, and helped to transfer their skills into weaker modalities that may be needed for success in the given task.

The approach of Hessels (Hessels, Berger and Bosson 2008; Tiekstra, Hessels and Minnaert 2009; Hessels 2009) was slightly different. The test instrument, the Hessels Analogical Reasoning Test (HART) addresses a higher mental process, that of analogical reasoning which has, like WM, been shown to be amenable to training. The HART comprises pre-test, training in analogical reasoning strategies, and post-test. The predictive validity of the HART was demonstrated

first by correlation between post test scores on the HART, and school related criteria in Maths, although correlation was non-significant with French. Subsequently, Hessels (2009) argued that learning potential for school subjects would be more ecologically valid if the test were curriculum related, but used materials that were novel to the child. He therefore devised dynamic tests of Chemistry (the Chemistry Learning test (CLT) and Geography (the Geography Learning test GLT), each of which incorporated a teaching component that was independent of other learning or experience in the subjects, and therefore reflected learning in the specific domain per se. The CLT was shown to be superior to static IQ testing in predicting future learning outcomes (Tiekstra, Hessels and Minnaert 2009).

Further combining learning potential in the domain general skill of analogical reasoning with the curriculum specific learning of geography, the predictive value of the HART was evaluated by correlation with the GLT as well as a static test of Maths. Results were inconclusive for the whole group, but significant in prediction for the special groups of children, namely those in the lowest third of a mainstream class, and those in special education, those for whom others have also shown the advantages of DA over static assessment. Hessels was able to conclude that for these special groups, dynamic measures have greater validity. The studies also demonstrate the complex relationship between assessments of cognitive skill and classroom content based assessments and interventions.

Hessels has contrived to devise tasks that are independent of previous experience, and has trialled his materials on different population groups, as well as including age related variations and accounting for developmental change. In addition, he has preserved standardised forms of administration and psychometric rigour. In

doing so, he has circumvented the problems of culture and experience that are typical of specific and static test items. It was these issues that Feuerstein, Budoff, Carlson and Wiedl, and Grigorenko and Sternberg, have tried to avoid, by modifying both materials and methods of testing, and producing tests of domain general cognitive skills, and in some cases, non-standardised interventions. Separate assessments like the CLT and GLT for each curriculum subject that incorporate training materials that meet these rigorous requirements however, may not be practical to devise.

Guthke and Wingenfeld identified this issue in 1992, when they pointed out that in effect, even a single school subject comprises a range of skills and competences, for example that the solving of algebraic problems requires different skills from solving geometric problems. Thus a large number of very specific curricular tests need to be devised. The resulting difficulties in task design, administration and interpretation, as well as the links to external criteria, led the authors to return to the concept of general intelligence testing as well as some domain specific procedures.

Similarly, Bransford et al (1987) identified many of the issues involved in choosing domain specific vs domain general tasks. In the context of school programmes of teaching and assessment, it seems important to use tasks closely related to specific content areas. The Vanderbilt group further explained that DA of specific areas enabled them to explore, for example, the reasons for students' problems in Maths, and strategies that could help them to learn maths. They noted, however, that information about thinking and cognitive skills is vital, and understanding the relationship between thinking skills and domain specific competencies is not simply additive. Rather 'competencies in a domain, and the ability to think about that domain seem to develop hand in hand' (p.492). Bransford et al

68

concluded that principles of mediation should be used in content based lessons, to facilitate independent learning and problem solving skills in every content area.

The debate continues up to the present. Kaniel (2009) linked the domain specific versus domain general testing debate to the need for evidence based DA, and concluded that adopting domain general assessments as a rule, leads to assessors tending to overgeneralize the findings of DA. Domain specific skills and domain general ones have a complex relationship, and skills might usefully be arranged on a continuum. In an application reminiscent of Bransford et al in 1987, and in agreement with the recommendations published by Haywood and Lidz (2007), Kaniel, (Personal communication 2009) commented that any task can be used as a DA, as long as the assessor mediates. Furthermore, because the predictive validity of a DA is difficult to establish, potential should be assessed in each domain.

Kaniel recommended that selection and construction of assessment tasks should be principled, theory based, and should use standardised instruments. Haywood and Lidz (2007) concur, advocating the application of DA to a wide range of educational and clinical contexts. It would seem also, that the assessment of specific content areas bypasses the problems of the definition of intelligence itself. Arguments that have been tautological about the process of learning to learn, are diverted to the process of deconstructing the learning of specific content, and the products of learning are more distinct from the processes. Some of the foregoing research has shown that linking the learning of cognitive skills to outcomes in curricular content is not straightforward, but the process of learning per se, can usefully be correlated with subject specific outcomes or products. This application makes it possible to use the principles of

DA to evaluate the learning of language, with the performance in various levels of language as a tangible product of that learning.

1.4.3 Validity and reliability issues

Vygotsky proposed that the modifiability of an individual's performance is not only a more interesting construct than their achieved learning, but that the validity of predictions of further learning actually increases. According to Embretson (1987b) issues of test validity are related to the test design, and dynamic tests have a particular effect on construct validity. Primarily, the testing procedure of a DA changes the construct representation of the task, and as a result, affects the 'nomothetic span', or relationship of the test score to individual differences.

The range and scope of different methodologies of DA leads to questions about the face validity of some procedures. Karpov (2008) questioned whether some graduated procedures are tests of learning potential at all, describing them instead of measures of a particular problem solving strategy. The debate goes to whether the task taps into the extent of pre-existing knowledge, or whether learning of a *brand new* (p.416) problem solving procedure is needed to access learning potential. Embretson (1987a and 1987b) presented this issue as one of the advantages of DA, stating that one is able to make a better estimate of ability if testees are trained in the use of a particular strategy, rather than being allowed to solve a problem via any number of their own processing strategies. When all examinees are trained to use a single strategy, learning rate is a better measure of general ability. Similarly, DA procedures enable the tester to eliminate irrelevant variables of testees' prerequisite experience, and tasks containing multiple components and processes.

Whilst simplifying the task requirement to a single strategy learning in this way contributes to greater test validity, and enables more reliable comparison between individuals, it may also be seen to be a disadvantage with regard to comprehensive assessment of the individual. Earlier discussion has raised the issue of differentiation within a group of individuals with learning difficulties, and consideration of the pre-requisite experience and idiosyncratic strategy use must surely contribute to understanding of the individual and formulation of the most appropriate intervention strategies. A degree of flexibility would be useful in the assessment to evaluate learning of problem solving strategies individually suited to a learner, rather than the use of a pre-determined strategy. The approach suggested by Karpov might usefully contribute to rigorous research into the relative learning abilities of a group of participants, but clinical or educationally motivated investigation of an individual's ability would not be served well. Furthermore, Embretson (1987a) pointed out that if the intention of the DA is to improve the estimate of ability, then validity is improved by a DA that is more focussed on the processes and consistency between the test items and the model of the targeted ability.

Construct validity may be demonstrated by only partial correlation with criterion tests, as the educational criteria are frequently not the targets of the DA. This was the case in the comparison of the findings of the HART with the Ravens SPM (Hessels, Berger and Bosson 2008). The study of reliability and validity of the group administration of the HART demonstrated low-moderate correlation with a static test, with which only some of the same dimensions are assessed.

Traditional intelligence tests are able to achieve some predictions of school success, with moderate correlations achieved between static

intelligence tests and school results, which suggest that up to half of a child's performance at school may be predicted by their intelligence test performance. Thus there is a challenge to demonstrate the advantage of DA, which may not be easily supported. (Embretson 1987a p.164). However in Embretson's own study of the modifiability of spatial ability (Embretson 1987b), post-test scores were shown to be a better predictor of performance than pre-test measures, demonstrating the advantage of DA over static baseline measures. European researchers such as Guthke, Hessels, and Resing, place a strong emphasis on the psychometric properties of their DA methods and strive to meet requirements of reliability and validity (Hessels, Berger and Bosson 2008). Their procedures have been found to account for an additional 20% of variance in school attainment measures, over that achieved by static measures (p.44). The advantage of the DA test is greater in children who perform poorly on static tests, namely those with learning difficulties or any impediment to their reliable performance on a test.

Hessels, Berger and Bosson were also able to demonstrate concurrent validity, by low correlations with non-cognitive behavioural measures. Developmental validity was demonstrated by linear increase in performance with age. In a further paper published by Hessels in 2009, significant improvement in all age groups was shown when the children were familiarised with how to solve the task, rather than given a short introduction, and Hessels concluded that employing a DA with only limited instructions for the task limits construct validity. Aiming to establish psychometric properties of DA is justified by the outcomes of this series.

In spite of his assertion that it is unjustified to even seek 'validity of DA' when the concept of validity cannot be applied to a particular test or paradigm, and is dependent on the interpretation being made

of test scores (p.35) Beckmann (2006) agrees with Embretson with regard to predictive validity of DA. He asserts that the *qualitative* differences in the prediction of variance constitute the superiority of DA over static tests.

Several authors agree with Beckmann (2006) that the validity of intelligence tests is threatened by definition of the construct of intelligence itself, and the concept of intelligence versus learning ability needs to be clarified, if assessments of the two constructs are to be constructively compared. However, the position is not the same in language, where static tests are measuring language ability (the product) in contrast to Dynamic Assessments which measure the ability to learn language. Expressive Language as a product is more easily defined than 'intelligence' and domain specific tasks may be devised to assess the extent to which knowledge and skills in multiple levels of language have been mastered. The use of static standardised language tests to validate DAs of language is, however, limited in usefulness.

Embretson (1987a) went on to outline the problems in reliability and validity that are inherent in DA procedures. Retest reliability established by repeated administration of a procedure is complicated by the intention of the test procedure itself to induce temporary, if not true change in the individual over time. Many studies rely on the format of test and retest, using the gain in scores as the measure of learning. Swanson and Lussier (2001), however, point out that '*changes in post test scores may be attributed to practice effects as well as to the notion that any reasonable treatment improves post-test scores*.' (p.323). Gains in performance at post test result from an upward bias in effect size, due to practice effects as well as the benefits of intervention, which a meta-analysis cited by Swanson and Lussier (2001) has shown to be in the order of 0.76. This trend,

which results in a greater standard deviation in the sample, and as a result a smaller effect size may therefore be an artefact of the design of a DA. Swanson and Lussier recommend that the magnitude of the effect size in pre-test-post test studies be considered to differentiate true gains from design effects. Practice effects rather than true effects of intervention, result in greater variation and larger standard deviation for the group on post-test, which leads to a smaller effect size, while treatment effects increase post-test scores with similar standard deviation at each time of testing and a larger effect size for the change between pre- and post-testing. Swanson and Lussier used comparative effects sizes across tasks to determine the relative ability of different DA measures to detect 'true learning' as opposed to reflecting the statistical outcomes of retesting.

Inter-rater reliability may be a useful procedure for qualitative ratings of behavioural criteria or responses if the test procedure is sufficiently standardised in administration. It may be less useful in the individually mediated interventions advocated by Feuerstein, but Tzuriel and Samuels (2000) showed moderate reliability for ratings of deficient cognitive functions and mediational strategies according to Feuerstein's frameworks. Lidz verified the inter-rater reliability of the Behaviour Rating scale component of the ACFS in three independent studies, and found that raters achieved levels of agreement on the pretest ranging from 67% to 74% (Lidz and Haywood 2007 p.124).

1.5 Methodology of the present study

The form and content of any assessment should depend on the diagnostic questions one wants to answer (Resing 2001), and tests of learning potential are no different. For diagnostic purposes, one would set out to establish a procedure that discriminates in categorical terms between individuals, while for educational

purposes, methodologies aim to find out what mediation or teaching is required to produce the most effective and beneficial change in individuals. In contrast to educationalists, therapists focus consistently on the individual, and his performance towards functional competence, be it mastery of handwriting, balance, emotional control, speech or language. Assessment is always linked to intervention, and standardised measures are both criterion referenced and norm referenced. The aim of a Dynamic Assessment would be to determine intervention needs, for any individual whose performance is less than typical. This emphasis on the imperative to help individuals is described by Kaniel (2001), who believes that to clarify this as an aim would be to create clearer guidelines for diagnostic and therapeutic directions. What does become clear is that the amount of pre-post test change, or the actual gain score of an individual is not the most important piece of information obtained (Haywood and Lidz 2007 Ch.1). Rather, the goal of the DA is to extract information that will be maximally useful in devising intervention for individuals that will itself be maximally effective. Children with previously identified language impairments require detailed assessment to enhance understanding of their strengths and limitations, and to inform language therapy interventions.

Language impaired children, have been defined as those exhibiting a deficiency in language that is not predicted by their non-verbal cognitive skills, although component processes such as attention and memory and processing skills may play a role. Kaniel (2009) believes that in the absence of clear evidence linking domain general difficulties to the manifesting impairment, assessment should focus on the domain specific skills, in this case, language, and this philosophy is pursued in the present study.

Although the application may be new, one can learn from the experience and expertise of those who have gone before. Tried and tested approaches to DA have provided a wealth of valuable information and the particular strengths of each methodology have been evaluated. A way forward now would be to draw from a variety of methods in order to construct the most appropriate methodology for accessing specific information, i.e. for addressing a new application of the paradigm.

In devising a methodology for Dynamic Assessment of language, and to inform intervention programmes in language impaired individuals, several strategies may be selected and adopted, in combination, to serve specific purposes. From the outset, group trends will be less useful than individual traits, when the intervention is likely to be individualised, and because the profiles of children with language impairment are so varied. Thus the DA procedure would need to differentiate the individual strengths and needs of particular children within the category of 'children with LI'. Increased inter and intra-individual variation was elicited by giving children feedback during testing, in a study by Berger (2004), and this procedure was therefore adopted in the current procedure with the intention of eliciting a similar outcome.

Techniques based on counting prompts or hints, as in the Campione and Brown model of Graduated Prompts, or on applying the identical standardised intervention in the training phase of the DA would be informative if uniquely quantified for each individual, and comparative information would be enabled by a numerical or quantitative scale. Its primary value would be to improve the precision with which testers can establish *relative* learning ability of children (Vye et al 1987). This is important in the current study to demonstrate the variability within the population of children labelled

as 'LI' who are heterogeneous in their ability, but score uniformly low on standardised tests. Furthermore, it highlights the relative amounts of intervention required to achieve gains or criterion scores, by different children that may influence their prioritisation for SLT services. Furthermore, use of a Graduated Prompt procedure facilitated identification of strategy use by children from different backgrounds (Resing et al 2009) and thus may be incorporated into a procedure intended to identify the range of strategies used by individual participants.

However, Graduated Prompts are less likely to extract unique information and maximum learning and transfer than techniques based on individualised mediational interventions. Mediated learning, as advocated by Feuerstein, is recommended for inclusion into the test procedure to facilitate formulation of individualised interventions that elicit most improved performances. In addition, restructuring the test situation to facilitate improved performance as demonstrated by Carlson and Wiedl (1992) would certainly elucidate the benefits of verbalization and feedback for the individual, and although this may not benefit all language impaired individuals, investigation of those who might benefit is advantageous.

1.6 Summary

The field of DA is broad in its methodologies and applications, and this very breadth has resulted in difficulties of definition and evaluation that are necessary to validate the procedure. Nevertheless, there is an inherent logic and face validity to many of the assessments, and it is this that makes application of the principles to a new context, appealing. Theoretical principles and methodological frameworks presented in this chapter are variously applicable to aspects of the assessment of language, and because this application is in its infancy, a vast number of choices have had

to be made.   Although research will be subject to methodological constraints, and issues of reliability and validity will persist, the domain of language is one that is well defined in linguistic terms, and some definitional pitfalls have been avoided. The nature of language impairment is complex, and development of different and probing techniques of assessment may be able to add to the understanding of the functioning of children with language impairments, that can only enhance the planning of grounded interventions. The nature of language impairment, and its assessment and management, will be presented in the next chapter.

CHAPTER 2

LANGUAGE DISORDER

CHAPTER 2   LANGUAGE DISORDER

2.1 The Nature of Language Impairment

Specific Language Impairment (SLI) has been broadly defined as the failure of language '*to follow a normal developmental course for no apparent reason'* (Bishop 1997 p.21). This broad and inclusive definition encompasses a large and heterogeneous population of children with communication and educational difficulties. One of the ways in which the heterogeneity was addressed in traditional models of language impairment was to specify sub-types of SLI. Those with differentiated clinical profiles of presentation, were classically described by Aram and Nation (1975) and Rapin and Allen in 1987, (cited by Adams, Byers Brown and Edwards 1997). Rapin and Allen used clinical data to identify six sub-types of impairment, based on the most salient features of expressive language, interactive behaviour and apparent comprehension of the children. The categories, which the authors do not claim to be exhaustive, were phonologic-syntactic deficit; lexical-syntactic deficit; semantic-pragmatic deficit; verbal dyspraxia; phonologic programming deficit, and verbal auditory agnosia, which represented children with severe receptive language disorders. In a more recent study, Conti-Ramsden, Crutchley and Botting (1997) used a battery of psychometric tests and teacher interviews to determine six robust subgroups of children with SLI that approximately paralleled those of Rapin and Allen. The evidence from a battery of psychometric tests lent weight to the finding of earlier authors that children with SLI fall into distinct subtypes. The groupings were thought to be useful for the determination of appropriate treatment plans and prognosis for recovery.

The categories utilised all the levels of language to some extent, but were not always sharply defined, and were not intended to be

inclusive of all presenting profiles of impairment. Furthermore they were not age related and changing profiles were identified by Bishop (1997) as a problem for classification. Longitudinal follow up of the children in the classification sample (Conti-Ramsden and Botting 1999) showed that while the features of language difficulty were stable, and the data fell into the same clusters as previously, 45% of the children were found to be in a different cluster a year after their initial assessment. The shifts between clusters reflected changes in the clinical profiles of the children that occurred in both directions. Relative strengths and weaknesses among the levels of language vary, such that children with impairments of language can manifest different profiles, and these profiles may change over time. Preschool children, aged between 3 and 5 are thought to be the most variable and many have transient developmental difficulties that resolve with maturation (Bishop 1997). While it is an important finding that distinct subgroups are stable in different age groups, the finding that children's strengths and weaknesses change in multiple different ways suggests to the current author that the categories are not necessarily useful for prediction of treatment needs and outcomes, or for prognoses about improvement.

Further threats to traditional classification were posed by Leonard (2009) who challenged the receptive/expressive dichotomy, which has persisted as a distinction in characterizations of SLI, and formal assessments of the relative strengths and weaknesses of individuals. Furthermore, Tomblin and Zhang (2006) in a dimensionality study, assessed children on a battery of standardised tests, with each task identified as receptive or expressive and also vocabulary or sentence level. Statistical analysis revealed the presence of these four dimensions as latent variables underlying test scores. Tomblin and Zhang concluded that although standardised test items identify a measurable language impairment as a whole, they are less useful in

characterizing profiles of strengths and weaknesses, and sub-grouping language impairments. The measures used in their study were not likely to reflect reliable differences between receptive and expressive modalities, but were useful to identify differentiated abilities between vocabulary and sentence use, particularly in older children. The authors noted that the unreliability of measures of discrepancy between receptive and expressive modalities may underlie findings of Conti-Ramsden and Botting (1999) that membership of groups was unstable. However, the lexical/sentence level distinction was shown to be a more useful dimension for description of language skills.

While the heterogeneity of children with SLI continues to be an important issue in characterization of the disorder, recent summary papers about the nature of language impairment (e.g. Rutter 2008, Law 2008, Hulme and Snowling 2009, Rice 2007) continue to present explanatory theories of SLI either as a manifestation of a neurodevelopmental disorder, intellectual disability, processing impairment, or as a discrete entity, and in addition, as a disorder that repeatedly manifests comorbidities with other developmental disorders such as impaired motor skills, reading abilities, social skills, memory, attention and executive functions. Furthermore, the existence of discrete modules of language that are selectively impaired in SLI, are controversial (Leonard 2009). Compelling evidence has been presented on either side of the domain specificity debate. In the interim, clinicians proceed with assessments that are designed to identify the features of the presenting condition, and interventions that are similarly based on symptomatic management, or make use of the evidence based interventions that are available. A summary of selected, linked research into the nature of language disorder, assessment and management, that relate specifically to

methodological considerations, case studies and findings of the present study will be presented in this chapter.

Karmiloff-Smith (2009) set out the case for a neuroconstructivist approach to developmental disorders including SLI. She argued that the brain starts with a greater degree of plasticity in infancy, and responding to an interaction of genes, behaviour and the environment, becomes specialized over time into modules. The disproportionately poor performance in language should not be taken to mean that cognition or other domains are intact, but rather that they may be impaired to a lesser degree or at a different time. Furthermore, Karmiloff-Smith noted that performance within the normal range in some skills does not mean that they should be assumed to be normal, but should still be investigated as performance may have been achieved by means of different strategies or atypical trajectories of development. The evidence of plasticity described by Karmiloff-Smith is consistent with Feuerstein's theory of Structural Cognitive Modifiability (1980) in which he maintains that the functioning of individuals continues to be modified by ongoing cognitive experiences.

Neuroconstructivist explanations contrast with mentalist approaches that view language as innate, and language impairments as deficits in the linguistic system itself. The debate relates to whether language impairment can be identified as 'specific' and distinct from other developmental disorders such as autistic spectrum disorders, Williams syndrome, (Karmiloff-Smith 2009) developmental speech disorders, dyslexia, and auditory processing disorder (Pennington and Bishop 2009; Dawes and Bishop 2010), as well as from general cognitive impairments and deficits in processing. In addition, the two viewpoints address the specificity of domains within language and whether modules of phonology, vocabulary, grammar and

pragmatics, and indeed parts of the grammatical system such as certain morphemes, can be differentially and independently affected in different types of language disorders.

There are multiple sources of information to inform these theories, such as evidence from genetics, imaging techniques, and the identification of clinical markers. Genetic studies, based on twins and on evidence of a familial condition, have been used as evidence for the dissociation of levels of language and support for the domain specificity of some impairments (e.g. van der Lely 2010). Furthermore, studies have linked chromosomal sites, associated with children with SLI with difficulties of non-word repetition, and other known markers of SLI (Vernes et al 2008). The development of magnetic resonance imaging (MRI) has enabled the study of early brain development providing support for theories of modularity. Regions of the brain thought to have circuitry appropriate for the development of specialized processing functions, such as language have been identified. The findings of studies using structural MRI have also contributed to differential diagnosis of SLI by enabling the detection of anatomical patterns common to children with SLI. Furthermore, Functional MRI (fMRI) scans reflect changes in brain activity over a period of time, and have been employed to explore regional activation of the brain during particular processing activities. In this way, localization of areas specialised for particular language processing tasks, and measurements of response latency have been made possible, and age related changes have been recorded.

2.1.1 Clinical Markers

A comprehensive review of the data on language performance by children with SLI was carried out by Leonard (1998). He demonstrates that language impaired children manifest problems in every level of language, and across many years of their

developmental history. Many of the error patterns described are those used by typically developing children at a younger age, with significantly retarded development of specific language structures. Some particular language structures, such as certain grammatical morphemes and argument structures have been noted to cause consistent difficulty for a large proportion of children with disorders of language, and have become identified as markers for SLI that can be used to differentiate between those with and without language impairment.

Conti-Ramsden, Botting and Faragher (2001) investigated the relative usefulness alone and in combination, of features thought to be clinical markers of SLI. The key finding of the study was that the tasks involving short term memory i.e. sentence and non-word repetition, were better markers for SLI than the tasks involving tense marking, and were robust even in children in whom earlier language difficulties had largely resolved by the time of testing for this study. These findings supported the findings of Gathercole and Baddeley (1990) that memory deficits were prominent in children with language disorders, a notion which will be explored further in the next section.

The consistency of grammatical morpheme errors and in particular verb tense morphemes, and noun-verb agreements are still considered to be typical markers of SLI. They feature prominently in explanatory theories of language impairment as a specific linguistic deficit in the individual, which contrast, as mentioned earlier, with those that attribute the difficulties to more domain general limitations in processing capacity. These two schools of thought will be considered in more detail.

2.1.2 Linguistic accounts of SLI

Explanations of language disorders as related to particular weaknesses in the learning of the language system mostly make use of the pervasive grammatical difficulties of children with SLI to demonstrate the apparent failure to learn the features, rules and constraints of language. The current study has chosen to address the syntactic abilities of children with language impairments using the Dynamic Assessment framework, and as such specific grammatical features associated with SLI will be assessed. Although the study does not aim to contribute to theoretical accounts of SLI, support for some of the features may be revealed, and a review of the dominant theoretical positions is appropriate here. Such theories assume the modular nature of language, which enables discrete grammatical difficulties to be described independently of other abilities. Leonard (1998) has expanded on six such explanatory models, key observations of some will be covered briefly here.

The *extended optional infinitive* model, attributed to Rice and Wexler (1995 cited by Leonard 1998) was based on the observation that young children go through a stage in which they omit the obligatory marking of infinitives on verbs as if it were optional, but at the same time do not add inflections inappropriately. Whilst typically developing children have mastered the adult system and use the structure correctly by the age of five, children with SLI have persistent difficulties marking tense and agreement, and formulating sentences and questions with auxiliary and copula verbs. The problem is thought to be one of knowledge, i.e. that children with SLI do not know the rules for use of the structures. Other aspects of syntax, such as word order and pronoun marking may be relatively unimpaired in children with SLI. Rice (2007) further summarised recent ongoing research which showed the protracted period of delayed acquisition of finiteness marking in children with SLI, relative to typically developing children of the same and younger

ages. She concluded that this particular part of grammar was disproportionately affected in language growth, and that further insights might be gained from re-examination of maturational models. In this framework, it may be seen that *delayed onset* of a particularly disrupted component such as finiteness marking may be characteristic of the clinical group.

In a similar vein, the inability to learn rules was thought to underlie the grammatical problems of individuals with SLI in the *implicit rule deficit* theory of Gopnik and Crago (1991). A study of the database of grammatical skills of sixteen members of a single family, led to the proposal of a single explanatory theory of 'developmental dysphasia'. In the absence of rules, individuals learn alternative structures, such as inflected forms of the verb, which is a slow and protracted method of acquisition, and accounts for the apparent delays in learning such structures. Alternatively, individuals might learn explicit rules that must be applied on a case by case basis, leading to slower and less consistent performance. Further support for the theory came from the studies of grammatical judgement, which showed individuals with SLI unable to recognize sentences containing errors as incorrect. This suggests their lack of *representation* of grammatical inflections, rather than an inability to encode them in expressive language. However, the theory implies a complete inability to learn rules and does not easily account for the inconsistent performance of children with SLI and the structures (e.g. plurals) that they do not in fact have difficulty with.

Ingram and colleagues (cited by Leonard 1998) focussed rather on the productive application of rules, and proposed a theory of limited or *narrow rule learning* that allowed for the acquisition and learning of some rules, but a failure to apply them in the range of possible contexts. Strengths of this theory are its ability to explain limitations

in rule use in phonology, semantic combinations and sentence construction, as well as morphosyntax. However cross linguistic data are not fully supportive, and evidence of inconsistent structure use at different ages is not conclusively explained.

Amongst theories that are based on structural relationship problems (Leonard 1998), Cromer hypothesized that children with SLI should have selective impairments in understanding sentences that involve *hierarchical* structure, or those in which the meaning is not superimposed on the linear sentence structure. Bishop (1982, and Bishop 1979 cited by Bishop 1997) provided some support for this notion by showing that children with severe receptive SLI struggled with semantically reversible sentences, passives and post-modified phrases. Subsequently, van der Lely proposed the *representational deficit for dependent relationships.* This notion identified the source of the linguistic deficit in the computational syntactic system, that is in the syntax proper (van der Lely 1994, 1998, cited by van der Lely and Battell 2003) and contends that the core deficit responsible for children's grammatical deficits is in the *optional and inconsistent* use of the rules of movement of syntactic elements in a sentence. For example, question formulation requires the movement of the auxiliary verb to the front of the sentence, but children with SLI do not use this consistently correctly (van der Lely and Battell 2003).

This theory was subsequently modified to the *Computational Grammar Complexity* (CGC) model, (van der Lely 2005), which stated that children with grammatical forms of specific language impairments (G-SLI) are '*impaired in the computations underlying hierarchical, structurally-complex forms in one or more components of grammar'* (p.55). Specifically, complexity that arises from 'non-local dependencies' causes difficulty, and as a result errors are seen in questions, passives and reflexive pronouns. Non-canonical

structures requiring syntactic analysis, which relate to the thematic roles were also shown to be difficult (e.g. 'on the paper is the book' as opposed to 'the book is on the paper', and 'give the girl the boy' rather than give the boy to the girl) (examples cited by Bishop 1997 p142).

Structural complexity according to the CGC model extends to phonology where studies of impairment reveal difficulties with structurally more complex morphemes, for example those with atypical phonotactic structures. These phonological difficulties may be independent of, or may co-exist with, grammatical difficulties. Children manifesting persistent primary impairments in the grammatical system beyond the age of nine, along with normal non-verbal cognitive skills have been identified as G-SLI, a relatively pure form of language impairment. Studies of the performance of children with G-SLI in comparison to other clinically defined groups or to typically developing children have enabled specific patterns of performance to be identified, and differential patterns of response have been confirmed in fMRI studies (van der Lely, Jones and Marshall 2011). This is taken by van der Lely (2005) as evidence for the existence of discrete, domain specific deficits of language.

Issues of working memory, more commonly associated with domain-general models to be considered in the next section, do however, play a part in structural complexity hypotheses. van der Lely, Jones and Marshall discuss the similar *derivational complexity hypothesis* (DCH) that attributes a greater role in language deficits to limitations in working memory or processing resources. As a consequence, the DCH might predict relatively unimpaired stored representations of structures, in comparison to the output of those representations, whereas CGC would predict deficits in performance on both input and output tasks. A grammaticality judgement task revealed that

children with G-SLI did not recognize errors of question formation, while maintaining accurate judgements of semantic violations and of syntactically correct questions. This suggests that the lesser processing load required for sentence recognition does not benefit children with G-SLI, who make the same errors in judgement tasks as in elicitation tasks, and lends support to the CGC hypothesis that children with G-SLI have impairments in the underlying computational syntactic system itself.

2.1.3 Cognitive theories of SLI

Despite having earned the label of 'SLI' which according to the discrepancy definition, implies impairments in the language system whilst other skills are relatively spared, children with SLI are reported to have some additional difficulties in non-linguistic areas. Co-occurring difficulties suggest that more general learning skills may be implicated in SLI, possibly alongside specific weakness in language. Cognitive theories seek the inherent weakness underlying language impairments in basic mechanisms that may be either general processing resources, viewed within the general framework of information processing (Montgomery 2002) or specific processing skills that are essential for the learning of language. Candidates for deficits of general processing include speed of processing limitations, capacity limitations, and memory, while specific processing requisite skills include auditory processing and phonological memory. 'General processing' skills are correlates of general or global intelligence ('g-IQ', Sternberg 2000), and are addressed in Dynamic Assessments of cognition, from which many of the concepts included in the methodology of the present study are drawn.

Montgomery (2002) believes that '*a wide range of language problems exhibited by children with language impairments can be profitably viewed and understood within an information processing*

framework' (p.64), and furthermore that some comprehension problems may be seen to arise because of weak representation or inefficient processing. The *generalized slowing hypothesis* (Kail 1994, cited by Leonard 1998) assumes that general limitations on speed of processing should have effects on tasks other than language, including non-verbal processing. Various studies have reported just these results. Miller et al (2001, cited by Hulme and Snowling 2009) found response times of language impaired children to be slower than IQ matched controls, for both linguistic and non-linguistic tasks, and Hayiou-Thomas et al (2004 cited by Hulme and Snowling 2009) simulated the difficulties of children with SLI in typically developing children by increasing processing demands.

2.1.4 Auditory Processing deficits

This theory is associated with the work of Tallal and colleagues, who classically demonstrated the difficulty experienced by children with SLI in processing brief, rapidly changing auditory stimuli. Using non-verbal responses to non-linguistic stimuli, Tallal and Piercy (1973) demonstrated that children with SLI had difficulties in discrimination of rapidly produced stimuli, and of stimuli of brief duration, in comparison to age controls. Subsequent use of verbal stimuli confirmed that children with SLI experienced particular difficulty with sounds of brief duration. Tallal (1976) further noted that the performance of children with SLI was not the same as that of typically developing children of any age, in relation to their particular sensitivity to stimuli with the shortest interstimulus intervals, but was in fact poorer than the performance of even the youngest TD children studied. The assumption of auditory processing deficit in children with SLI was the basis for intervention delivering training in processing of auditory stimuli, a programme which will be discussed in the 'Intervention' section of this chapter.

Whilst evidence for processing deficits in children with SLI is consistent and robust, there are evidently aspects of impaired language, and indeed impaired non-linguistic skills, that are not dependent on rapid auditory processing, and the causal effect of processing on language impairment is not proven. Montgomery (2002) has subsequently argued for the failure of children with SLI on processing tasks not to be seen as temporal processing alone, but a combination of needing to process perceptual information and perform linguistic analysis in a timely fashion. Experimental studies of response times to processing of sentences containing inflected and stem sentences, containing both high and low salience morphemes (e.g. present progressive –ing vs third person –s) confirmed hypotheses that children with SLI fail in processing tasks that require the processing of the grammatical function of short duration morphemes and not the presence of the short duration morpheme alone.

Alongside deficits in specific processing of auditory stimuli, children with SLI are thought to have specific deficits in phonological short term memory. Gathercole and Baddeley (1990) found children with SLI to be particularly impaired in the repetition of non-words, in the presence of normal skills of speech discrimination and word memory, and argued that the poor performance in non-word repetition in SLI is attributable to an underlying impairment in short-term memory. They hypothesized a deficit in the ability to represent phonological material in working memory. Deficits in non-word repetition have been consistently reported in children with SLI, and it has been regarded as a highly robust predictor of SLI, differentiating children with SLI from control groups.

Word learning deficits have been further described as prevalent in children with SLI, with slow acquisition of lexical items especially in

contrast to the rapid learning by typically developing children. The phonological working memory deficit has been hypothesized by Gathercole and colleagues to have an impact on the ability to learn words, but the evidence is not conclusive. Oetting, Rice and Swank (1995, cited by Hulme and Snowling 2009) found a particular difficulty in the learning of verbs as opposed to nouns, in children with SLI, and concluded that there is an effect of grammatical constraints on word learning as well as the phonological limitations.

Gathercole and Baddeley (1993) elaborated the effects of working memory on the learning of language and provided examples of sentences that would require more linguistic processing than could be carried out in working memory, and were therefore dependent on the retention of an accurate representation in short term memory. Examples of such sentences were semantically reversible sentences, and passive constructions. Similarly Montgomery showed children with SLI to have more difficulty comprehending longer than shorter sentences, and interpreted the correlation with phonological working memory as suggesting that the inability to store as much information compromised the comprehension of sentences. Montgomery however, found the relationship between working memory and comprehension difficulties to be complex and slower linguistic processing itself may be implicated in comprehension difficulties (Montgomery 2002, p.74).

Subsequently, Archibald and Gathercole (2006) summarized previous work and followed up on the following key assumptions. Firstly that short term memory is required to support the phonological representations of speech events and enable their storage as lexical items in long term memory, and secondly that working memory deficits in SLI represent general inefficiency in information processing that limits language development and

function. Archibald and Gathercole found greatest effects on memory in tests of working memory and verbal short term memory, and proposed that although the two short term memory deficits are not necessarily part of a single underlying disorder, children with SLI may be affected by difficulties in both. Thus their ability to retain phonological forms for processing as well as their longer term storage may be compromised.

Leonard (1998) presented *the surface hypothesis* to account for the interaction of general processing deficits with features of the language itself. In relation to English in particular, Leonard explained how the auditory processing limitations described above combine with the brief and unstressed nature of many grammatical morphemes in English to reduce the perception of these morphemes by children with SLI, who then have difficulties formulating the rules for morpheme use. Leonard notes that the errors made by children with SLI do not differ from those made by younger typically developing children, whose hypotheses about grammar are immature, and there are no actual fundamental difficulties with grammar of children with SLI per se.

The surface hypothesis is also consistent with accounts of limitations in working memory as processing of bound morphemes may be incomplete while additional material from the sentence string is incoming, and incomplete representations of stems plus morphemes may be stored, leading to omission of morphemes in subsequent productions (details of the mechanisms may be found in Leonard 1998).

Nevertheless there are shortcomings of the theory, particularly in its inability to explain frequent findings of auxiliary inversion errors in questions, and the prevalence of past tense morpheme errors.

Furthermore, children with SLI experience few problems with the plural morpheme –s, in comparison to the third person singular –s, yet the surface properties of these would be the same, suggesting that grammatical accounts of deficit may yet be relevant. Also relevant would be the psycholinguistic model of Stackhouse and Wells (1997) that would make use of both comprehension and production data to locate the difficulty in a component of the processing model. Speed of processing limitations superimposed on the psycholinguistic model might go some way towards accounting for performance limitations in children with SLI. Conversely, probing of the linguistic difficulty in the absence of pressures of online processing, by removing speed and memory constraints, as in a mediated learning context, may shed further light on the specifically grammatical or linguistic deficits.

2.1.5 Executive Functioning

Executive functions have been identified as a domain general component in Baddeley's model of working memory (Baddeley 1996, 2000). As such, a crucial aspect of executive function is the control of selective attention to a stimulus, and the ability to inhibit responses to distracting stimuli. Children with SLI perform poorly on various tasks of working memory, in both verbal and non-verbal modalities, but the role of executive function and attention control in a range of tasks reported by Marton, has not been investigated. Gathercole (2010) noted that children with SLI perform poorly in tasks of verbal working memory and verbal short term memory, but not in visuo-spatial memory tasks, suggesting that the deficit is not in central executive function, which is not, therefore, the primary source of SLI.

Marton (2008) reported on two studies designed to examine executive function in children with SLI in comparison to typically

developing children, the first in relation to visuo-spatial abilities, and the second in relation to performance on selected neuropsychological tests. The findings of the second study in relation to the Wisconsin Card Sorting Test (WCST-64, Kongs et al 2000) showed children with SLI to be significantly impaired, making more errors, and showing greater perseveration and difficulty forming clear concepts than TD children. In the Tower of London test (TOL, Culbertson and Zillmer 2001), children with SLI showed more impulsive behaviour, and violated more rules than the TD peers, although they completed the same total number of moves and in the same amount of time.

Marton concluded that children with SLI do have poorer performance in criteria related to executive functions, specifically attention control, inhibition, switching from one idea to another, and planning their problem solving. These weaknesses affect their verbal and visuo-spatial processing, as well as working memory, although the author points out that further research is needed to clarify the relationship between these aspects. Henry (2010) explored specific executive functions in verbal and non-verbal tasks, to investigate whether executive function limitations in children with SLI were general or language specific. Five areas were assessed, namely working memory, inhibition of inappropriate responses, planning, fluency, such as generating names, and flexibility to switch between tasks. Results showed performance poorer than chronological age matched controls on four out of the five areas, in both verbal and non-verbal tasks, supporting the hypothesis of general executive function deficit.

Investigation of individual executive and higher order skills have previously shown children with SLI to be impaired in hypothesis testing abilities (Ellis Weismer 1991). Language age was also shown to be the most significant predictor of performance in analogical reasoning tasks (Masterson, Evans and Aloia 1993), although

variations were evident in relation to the type of analogy, with children with SLI scoring better on synonyms and antonyms, and poorer on analogies requiring category membership, functional relationships or linear order.

2.1.6 Metalinguistic Awareness

Although not frequently included in accounts of the abilities of children with SLI, a few studies focus on metalinguistic awareness, primarily in older children and in relation to their acquisition and development of literacy skills. A review of metalinguistic awareness in children with SLI is relevant to the current study in the light of the established links between Dynamic Assessment and metacognitive, thinking skills intervention, primarily in the work of Feuerstein, and the essential role of metacognitive awareness in mediational interventions. However, investigations of metalinguistics are complicated by differing definitions of the scope of the term 'metalinguistics' and the behaviours that it encompasses. The issue is explored in more detail by Gombert (1992). For the purposes of the current paper, the term will be used only to refer to the higher levels, or later stages of metalinguistic development, in which processes are accessible to conscious access, and it is possible for the knowledge to be encoded verbally (termed tertiary explicit knowledge by Karmiloff-Smith, cited by Gombert 1992, and 'automation of metaprocess' by Gombert).

Ravid and Hora (2009) summarized the nature and development of metalinguistic awareness in three general statements drawn from key theoretical positions (p.11)
'*First, metalinguistic awareness constitutes an inherent part of natural language acquisition from early on. Second, metalinguistic awareness is not a uniform ability: it gradually develops, becoming*

more verbal, explicit and flexible with age and schooling. Finally, the acquisition of written language is a crucial landmark in learning to think about language, to focus on its components, and to conceptualize form-meaning relations.' Ravid and Hora went on to note that metalinguistic awareness in children could be a useful tool to promote language development and an appropriate target of intervention.

A summary of research papers by Gombert, describes a range of results, and a conclusion that the conscious identification of aspects of syntax is apparent around the age of 6-7 in typically developing children. Intuitive knowledge and functional use of language is evident prior to the development of conscious reflection and intentional control. Efforts to stimulate such functioning through training are seldom successful before the age of five, although some believe that the onset of literacy teaching in school promotes the development of metalinguistic awareness. Metalinguistic ability has been shown to correlate with general cognitive development and specifically with metacognition, literacy, and oral language skills (Benelli et al 2006), and in view of this last, may be poorer in children with LI. There is little research, however, into the metalinguistic abilities of children with LI.

 Magnusson and Nauclér (1993) attempted to clarify the roles of cognitive ability and language ability in metalinguistic awareness, by testing children identified with disordered language matched on age and cognitive ability with children with normal language. In general, the LI group demonstrated poorer phonological awareness than the TD group, but inconsistencies in the data led the authors to analyse further, and find that non-verbal cognitive skill played a greater role in phonological awareness for children with LI than for TD children. Furthermore, the type of language impairment was important, with

significant correlations emerging between phonological awareness and receptive and syntactic skills. Thus inconclusive results from the study suggest that children with grammatical difficulties may be at most risk for metalinguistic difficulties, which may in turn be mediated by strong nonverbal abilities.

In a similar vein, Smith-Lock (1995) used children with SLI and language and age matched controls in order to investigate the relationship between cognitive ability, language ability and metalinguistic awareness. Focusing on metalinguistic awareness of morphological forms, Smith-Lock tested children on sentence completion tasks, requiring both real and nonsense words, comprehension of inflected non-words, and response to morphological errors by means of grammatical judgement and repair of errors. Despite some variation between tasks, results showed the children with SLI to perform at a similar level to language-matched controls, and significantly below the age and cognition matched group. The author interpreted this as evidence that linguistic awareness is associated with expressive language level, and that children with SLI do not have specific deficits in morphological awareness in comparison with their peers matched on language performance.

Marinellie and Johnson (2002) investigated the skills of children with SLI on word definition, a task considered to be metalinguistic in nature (Nippold 1998). The children with SLI performed significantly less well than their TD counterparts, in relation to both the type of information they chose to use, and the syntax used to encode their responses. The children with SLI used more functional definitions, and fewer formal ones, although both groups used many 'transitional terms (e.g. 'something you wear') rather than making use of superordinate category names. The children with SLI were noted to

make use of immature syntactic forms. The authors supplemented their procedure, however, by asking children with SLI for further information after they had given their answer, by prompting them with the question 'what else do you know about X' (the item to be defined). Adding the scores obtained from these responses to the original responses, resulted in the scores for the children with SLI approaching those of the TD group. The authors interpreted this as an indication that the children with LI benefited from more time to retrieve words from memory, as well as from a less formal and broader remit for the task. It was suggested that the children with SLI lacked the metalinguistic knowledge of how to define words when asked. The additional prompt procedure, consistent with the practice of Dynamic Assessment, enabled the tester to analyse more specifically the relative contributions of the child's linguistic, metalinguistic and lexical access abilities. It also highlighted the potential for improvement of the children's definitional skills through training, which was in turn thought to have potential to promote metalinguistic skills (p.256).

2.1.7 Summary

In summary, it would seem that no single account of impairment or avenue of investigation has produced a conclusive explanatory or predictive theory to account for specific language impairments. Instead recent research seems to be uncovering more overlapping features between SLI and autistic spectrum disorder, dyslexia and auditory processing disorders, and advances in genetic and neuroimaging techniques are similarly uncovering commonalities in brain function. Tomblin (2009) concluded that '*there is no unique causation for SLI, but rather it is a constellation (conspiracy) of common factors, and poor language arises out of the number of risk factors and the nature of their interactions*'.

The interaction of multiple factors, changes over time, and evidence from studies supporting diverse approaches to explanation, challenge the traditional assumptions of the nature of (specific?) language impairment. Accepted and clinically useful classifications of subtypes, are not necessarily stable or durable, but vary with time. Persistence of SLI, historically a notable and diagnostic feature, has emerged as a potentially shifting paradigm over the lifespan, which may manifest as late onset of skills, in which some difficulties (most notably phonological processing) may resolve, and which may re-emerge as literacy difficulties in school age children, and social difficulties in adults (Rutter 2008). Apparent independence of language deficits from non-verbal intelligence similarly shows change over time with non-verbal IQ falling in children with SLI from school age (Botting 2005), which appears to threaten the very definition of SLI, and the criterion by which individuals are classified as having SLI for the purposes of research.

This review of the literature suggests that the direction for further research should be to consider individual presentations and constellations of factors, as well as links between cross-sectional and longitudinal data (see also Joffe, Cruice and Chiat 2008), taking into account trajectories of both normal and disordered development in multiple skill areas. This approach has implications for both assessment and intervention, which will be considered in the following sections.

2.2 The Assessment of Language Impairment

2.2.1 Traditional assessment of Language Impairment

The primary means of accessing children's language abilities for the purposes of diagnostic identification of language impairment or for elucidating the nature and severity of the impairment is through the use of standardised, norm-referenced tests. Those most widely used,

according to Dockrell (2001), are 'global' tests, comprising several subtests, such as the Clinical Evaluation of Language Fundamentals (CELF Semel, Wiig and Secord 1987), and tests of vocabulary such as the British Picture Vocabulary Scales (BPVS II, Dunn et al 1997), although specific tests for different levels or domains of language have been developed.

One alternative to formal tests that is traditionally implemented alongside or in place of formal tests are questionnaires and language samples. Questionnaires addressing specific difficulties, such as developmental language, stammering, or pragmatics are regarded by practitioners as useful tools. While reliability and validity of some scales has been verified, Dockrell noted that their use as predictive measures has not been established. Language sampling, advocated primarily by David Crystal in the 1970s and 80s (Crystal 1982, Crystal, Fletcher and Garman 1976), provided a more ecologically valid sample of an individual's spontaneous use of language in context. Samples however were not always found to be fully representative of the range of language structures in the individual's repertoire, and the skill of the examiner in the elicitation, transcription and analysis of the sample affected reliability. Furthermore, clinicians were not convinced that the investment of time in the procedure was justified by the quality of information obtained. Nevertheless, Crystal, Fletcher and Garman (1976) did formulate recommendations for management of individuals found to have typical profiles on the LARSP procedure. A summary of surveys of clinicians in the US by Caesar and Kohler (2009) reported that although the majority of SLTs relied on formal tests, language sampling was the most frequently used informal assessment procedure.

The selection of assessment tools depends on the purpose of the testing. The primary aim of assessment is frequently identification of a language impairment, to determine the need for management, or eligibility for services. Assessment for clinical management and the planning of intervention is considered to be a slightly different process. In making identification decisions, distinct categorical criteria with clear definitions and numerical boundaries often need to be established. Research and epidemiological studies similarly require predetermined and fixed definitional criteria. This last was clearly explicated by Tomblin, Records and Zhang (1996) in their study to establish inclusion criteria for SLI in kindergarten children to inform their study of prevalence of the condition in this population (Tomblin et al 1997). The system they employed consisted of five composite scores derived from norm-referenced tests in vocabulary, grammar, and narrative, and in both comprehension and expression. Children who scored below −1.25 standard deviations on two or more composite scores were considered as children with language disorder. It was shown that the children identified by this diagnostic system were consistent with those identified by clinician rating and previous results.

Tomblin et al made use of cutoff criteria that yielded an acceptable level of sensitivity and specificity. These notions have been further recommended by Spaulding, Plante and Farinella (2006), who tested the assumption that children with language impairments will score significantly low on tests designed to identify language impaired children, by investigating the statistical properties of 43 tests. They found that the cutoff scores for language impairment, which are arbitrarily set for each test, did not satisfactorily identify children who had been identified as language impaired by clinical criteria. In total 56% of language impaired children scored within 1.5 standard deviations below the mean score for the whole population. Spaulding,

Plante and Farinella recommended that clinicians scrutinize the normative data presented in each test manual to ensure that sensitivity and specificity data are presented before deciding to use a particular test.

The difficulties with cutoff scores were earlier identified by McFadden (1996) who highlighted the difficulties resulting from tests being based on normative samples that are 'truncated' and do not include children with language impairments, who then score very low or below the first percentile on standardised tests. These data led to ASHA's recommendation (cited by McFadden 1996) that clinicians should use tests without rigid cutoff scores, and in combination with data gathered from other sources. Dockrell (2001) also noted that standardised test results are inadequate even for making '*only screening and placement decisions*' (p.79) and information from other sources should be sought. Law et al (1998) have reported screening tests to be inaccurate, but Gardner et al (2006) published results of a new screening test, the Grammar and Phonology Screening (GAPS) which the authors have devised to specifically target core grammatical and phonological skills which are known to be impaired in children with SLI. The GAPS was shown to be reliable in identification of children in need of further assessment for language impairment or for literacy deficits. SLT practitioners continue to rely almost exclusively on the results of tests for their clinical decisions, and quantitative norm referenced data are required by those making decisions regarding educational placement and provision.

In a more positive vein, Friberg (2010) like Mikucki and Larrivee, (2006) found that in general, psychometric reliability and validity of standardised assessment tools has improved since the benchmark criteria were set by McCauley and Swisher in 1984. These criteria

related to the necessity for the authors of tests to provide detail of their standardisation sample, statistical data and administration instructions, as well as evidence for item validity, concurrent validity, predictive validity, re-test reliability, and inter-examiner reliability. However, Friberg's study examined only those tests already determined to be discriminating for language disorder by Spaulding, Plante and Farinella. Friberg concluded, like Spaulding and McFadden, that it is the responsibility of practicing SLTs to be aware of the psychometric characteristics of assessments that they select if they are to make well informed clinical decisions.

Accurate diagnostic information is also derived from the comparison with different skills, and none is more pertinent than the discrepancy between verbal and non-verbal skills which is frequently thought to be fundamental to the definition of specific language impairment (SLI) itself (although note comments earlier in this chapter). The *discrepancy definition* is based on distinguishing SLI from global difficulties that include language, by requiring a difference (WHO, 2007 ICD-10) between non-verbal IQ, and language measures. This of course requires accurate assessment of both intelligence and language, and Bishop (1997), Lahey (1990), Botting (2005), and Aram, Morris and Hall (1992), outlined a number of difficulties arising from this. Primarily the extent of the difference between non-verbal IQ and language may be arbitrarily set, may vary considerably depending on which measures are used, and may change over time.

Lahey (1990) pointed out several problems with calculating 'mental age', for example that the amount of development in one year is not constant in every year, and comparison of a child with another who is younger but has an equivalent MA is not informative with regard to performance of age equivalent peers. A different combination of

tests of both non-verbal IQ and language could result in a range of estimates of ability for both, with a difference that may or may not meet definitional criteria. Rigid cutoff scores defining criteria might result in borderline abilities, for example a child may score an IQ of 86, meeting the criterion for 'normal' intelligence, by being within 1SD of the mean, and score 80 on a language test, meeting the criterion for impairment, when there is actually minimal difference between these verbal and non-verbal scores. Conversely, a child may score 86 on a language scale and achieve a non-verbal IQ of 105, thus showing as within normal limits on both scales, yet with a considerable discrepancy between the two. In both these scenarios, the implementation of a fixed cut-off score obscures the detail of the comparison of abilities. Lahey (1990) noted that although comparisons by standard deviation are better than comparison by MA, the accuracy of scores are still subject to measurement error. Furthermore, Bishop cited several studies in which these score profiles did not match clinical observations or identification of disorder by clinical criteria, and Lahey believed that ability shown on test performance may not be representative of a child's actual performance particularly under conditions of stress. Aram, Morris and Hall (1992) examined various formulae for calculating discrepancy between non-verbal IQ and language, but concluded that whichever formula was used, approximately half of children identified with developmental language disorders by clinical descriptors would not be identified by discrepancy measures.

Standardised formal test scores have been used to try and predict future change in children with language impairments, and discrepancies between skills have again been employed for this purpose. A review by Olswang and Bain (1996) reported that discrepant scores between cognitive and language abilities, and between receptive and expressive language scores have been taken

as indications of potential for change. Children with larger differences between nonverbal and language skills have been recommended for services on the basis of greater potential to benefit from intervention. The results of the Tomblin and Zhang (2006) dimensionality study however, suggested that the individual traits measured in different subtests are not truly independent, and there is not as much information lost in the presentation of a single composite language measure as one might imagine. In particular, they found that the receptive/expressive distinction was not reliably measured by standardised tests.

Some research attempts to define and investigate as pure a group of individuals with SLI as possible, in order to ensure that features and performance limitations identified are attributable to deficits of language rather than intelligence. Similarly, intervention studies would prefer to be able to ascertain the effects of therapy on language impairments rather than on global difficulties. From the perspective of clinical management, there are also concerns that resources for individuals with language impairments are limited and inclusion of children with language impairments secondary to or associated with other types of difficulties would overstretch the resources and dilute the services available (Botting 2005). Botting further pointed out that the relationship between cognition and language impairment is complex and even the direction of influence is not clear. Indeed, it is not fully understood whether cognitive limitations affect language learning, or whether language mediates the learning of more complex intellectual concepts.

Botting (2005) presented the findings of a number of studies that suggest that the relationship between IQ and language is not a stable one, and her own results showed a fall in non-verbal IQ over time in children with SLI. Aside from the contribution of these

findings to the understanding of language impairment as a domain general as opposed to a specific impairment, there are implications for management which suggest that cognitive difficulties may need to be targeted in assessment and intervention alongside language skills. Similarly, Montgomery noted the potential need to address processing skills alongside language, and also the confounding effect that processing limitations may have on verbal language tests. He cited, for example, the extreme demands placed on working memory by the Semantic relationships subtest of the CELF-3 (Wiig, Semel and Secord 1987) that may result in poor scores not attributable to the language material itself (Montgomery 2002 p.77). Campbell et al (1997) demonstrated that children from minority backgrounds who performed poorly on traditional language tests, did not differ from the majority population on processing dependent measures, and thought that this may be a way to distinguish children with language impairments from those with different experiential backgrounds. Although the battery of measures used by Campbell et al was not tried on children diagnosed with language impairments, the authors assume that children with LI would perform poorly on both the vocabulary weighted knowledge dependent standardised test, and the psycholinguistic based measures, which included non-word repetition. They tentatively '*suggest that the diagnosis of language impairment should rest on processing dependent measures designed to minimize the impact of background*' (p.523).

A further difficulty identified by Camarata and Nelson (2002) related to the dependence of many nonverbal or performance tests, as well as verbal scales, on comprehension of language. Failure to consider functional language comprehension could lead to children's cognitive and performance abilities being underestimated. Camarata and Nelson hypothesized that if language competence could be increased though intervention, a noticeable improvement in non verbal test

scores would result, but the authors conceded that the required improvement in language would be unlikely to be achieved in children with language impairments, and teaching the skills would not automatically improve the abilities of children with language impairments. Furthermore, the recommendation of the authors that goals of intervention ought to be to ensure that a child has the requisite language skills to enable reliable assessment of their cognitive abilities does not seem to be a clinically valid or functional approach to intervention. In the opinion of the current author, it is the test methodology that should be altered to enable more accurate assessments of children with difficulties, rather than persisting with 'well-normed and valid' instruments, and taking care to interpret the performance of children with caution.

With regard to the use of language tests to devise intervention programmes, Rutter (2008) described the need to use *dimensional* assessments, in contrast to the *categorical* answers required to determine eligibility for therapy. This includes the assessment of multiple risk factors relevant to the individual, and detailed consideration of the severity and duration of the risk factors as well as their nature. Dockrell (2001) made the point that a single unitary assessment such as a specific test to assess a single aspect of language on either receptive or expressive levels, for example receptive grammar, narrative, word finding etc., would be an inadequate assessment of a child's language ability. Clinicians need to rely on general tests with multiple subtests, or a battery of individual tests in order to make decisions about the management of the individual. Montgomery (2002) recommended the inclusion of tests of information processing in order to gauge how these might relate to comprehension abilities. These assessments might include measures of short term memory, digit span, word span and non-word repetition, as well as varying the rate of presentation of stimuli

to assess speed of input processing. However, even these would be subject to some of the limitations mentioned earlier.

None of the language tests or measures are sufficient on their own to fully diagnose and describe language impairment in a child, and most do not attempt to formulate recommendations for intervention. Building a composite profile of abilities from standardised tests, case history data, observations and questionnaires would provide the best baseline from which to make management decisions, and despite the apparent reliance on standardised tests, clinicians in practice do incorporate case history data and their own observations into their assessments (Caesar and Kohler 2009). These tend to address developmental progress up to the present, gauge risk factors, and evaluate functional and informal skills such as play, attention, non-verbal and interpersonal communication. They also contribute information about contextual understanding of language, for example the following of directions, and the use of language for various functions in conversation, but do not specifically address skills of receptive and expressive vocabulary, semantics and syntax, and for these formal aspects, tests are the main source of information.

Lloyd and Blandford (1991) highlighted the inadequacy of formal tests for planning instruction in special education. Dockrell (2001) further noted that the information gained from standardised tests is inadequate for the planning of intervention in language therapy, because of the lack of detail. However Friberg (2010) found that despite the recommendations that assessment of language impairment should be based on multiple sources and types of data, many clinicians do seem to give weight to the quantitative results of standardised tests to determine eligibility for management and plan intervention. In the recent study by Law et al (2008) test results

were used by almost one-third of SLTs as rationales for their intervention, yet in summarizing, Law noted that 'despite a level of homogeneity in the profiles of the children – *based on standardised test performance at least – there is no parallel homogeneity in the interventions provided for this group of children'* (p256). This suggests that the tests used do not differentiate well between the children, and do not lead to specific recommendations for intervention that are taken up by SLTs. Instead practitioners are using poor scores to justify the need for therapy, but are devising that intervention based on other sources as described by Law et al.

2.2.2 **Alternative Approaches to Assessment of Language Impairment**

Mislevy and Yin (2009) explored the multiple levels of spoken language that take account of characteristics of the speaker, the conversation partner, the context, both current and historical, as well as perceptual, social and cognitive constraints. Their interactionist approach to assessment based on this model, is one that is defined by the context, targets and purpose of the assessment. Thus formal language tests are construed as having lean context, predetermined targets and a purpose to support learning. Task based curriculum assessments may be rich in context, make use of opportunistic targets and also focus on supporting learning. Assessments that focus on the individual's abilities tend to generalize observed traits to behaviours thought to be typical of the individual in many contexts, while task oriented assessments focus on the requisite skills required in a particular context without regard to whether the individual has the capability to act in that situation. These two angles are reconciled by the assessor, whose prior knowledge of the task and of the examinee should play a role in assessment. Thus there is a recommendation that the assessor

bears a responsibility for inferential interpretation of the performance of a particular learner on a given task.

Mislevy and Lin conclude that assessment of language is a gathering of information, to ground inferences about an individual's ability to use language in a range of real-world contexts. It would seem that this broadly conceived approach would lead to both functional assessments and functional interventions for individuals with impaired language, but the theoretical model has yet to be operationalized in clinical and educational terms. In a similar, though less theoretically complex manner, Lloyd and Blandford (1991) propose that the assessment of children for the purposes of instruction, or intervention, in the context of special education, should include four specific aspects; identification of the areas of need for instruction, the skills to be taught that are the starting point of instruction, the delivery of instruction and finally monitoring of whether instruction is succeeding (p.46). The first of these, identifying the need, is the area currently served by formal tests, which determine whether the performance of the individual is age appropriate in a given area. However, the second requires more detailed assessment of the skills that need to be addressed in order to achieve improved performance, and in this area, Lloyd and Blandford recommend that trial tasks are used to sample skills and subskills to determine the precise level of breakdown at which intervention should be introduced.

With regard to the delivery of instruction, the authors recommend that in addition to evidence based practice and examples of best practice, intervention should be tailored to meet the needs of the individual, and information about these needs should be obtained by manipulating variables of pace, materials, and reinforcements. Finally, Lloyd and Blandford stress the need for ongoing monitoring

of progress that is traditionally incorporated into the cycle of SLT intervention. A significant part of the proposed model is assessment not only of the individual, but of the environment in which instruction should take place, and Lloyd and Blandford stress that location and timing of intervention influence a child's responsiveness to intervention, as well as the type and clarity of instructions, use of models and examples, and amount of practice given. Tailoring intervention according to these variables maximises the success of the learning programme in children with learning disabilities, and comprehensive assessment at the outset facilitates that planning.

A model known as 'holistic assessment' was advocated by Gillam and McFadden in 1994. This model also aimed to provide therapists and educators with a broader more contextualised assessment of students with special needs. Three methods of holistic assessment were described by Gillam and McFadden, each of which altered and extended traditional assessments to increase the amount of information obtained. The authors believe that in holistic assessment '*results from standardised testing should not comprise more than a quarter of the total information that is used to build an understanding of a learner*' (p.39). The methods suggested by Gillam and McFadden include an expanded reading miscue analysis that appears to be extremely useful in identifying the use of semantic information in reading, an authorship sequence that although lengthy allows qualitative assessments of written narratives as a whole, as well as through their constituent parts, and Dynamic Assessment of language. The argument for such holistic assessment to underpin understanding of a child and inform intervention that will have the greatest impact on his functioning, is convincing.

The theme of these papers, and of the constructivist approach, described by Meltzer and Reid (1994) is towards a broader, more

representative assessment of individuals, which highlights more specifically the direction that future intervention might take. Their purposes are varied, addressing the needs of children with learning disabilities and second language learners, but the principles might equally be applied to populations with language impairments, and SLTs might do well to learn from the experience of psychologists and educators in other fields. An area which has prompted the development of more representative assessments of language is the need to differentiate those performing poorly on assessments for reasons of cultural difference, from those with learning or language impairments (Laing and Kamhi 2003).

Alternatives to traditional assessments that are thought to be more culturally fair to those from different cultural backgrounds include assessments of processing, and criterion referenced assessments, as well as the Dynamic Assessments to be considered in more detail in the next section. Tests of processing ability place less emphasis on prior language knowledge and experience (Laing and Kamhi 2003), and include memory and perceptual tasks, which are also thought to be impaired in children with language impairments as outlined previously in this chapter. Many of the alternative approaches and theoretical positions on assessment have recommended Dynamic Assessment procedures to elicit the broader based, culturally fair assessments that give rise to recommendations for intervention. These papers specifically related to the Dynamic Assessment of language skills will be reviewed in the next section.

2.2.3 Dynamic Assessment of Language Impairment

Although developed by psychologists, DA has been adopted by educators in both mainstream and special education, and to a lesser degree by Speech and Language Therapists working with children with language impairments (Hasson and Joffe 2007). Much of the

research in DA of language aims to differentiate children with LI from other populations, and increase the accuracy and predictive validity of diagnostic classifications. A few papers have addressed modifiability of language skills in clinical populations and many advocate a combination of this clinical approach with the use of some formal tests for the purposes of normative assessment and educational placement. Merritt and Culatta (1998) explain some of the features that make DA and the combined approach particularly applicable to the study of children with language disorders within the context of mainstream school education.

DA methods have added to the body of knowledge about language development and performance in typically developing children and other clinical populations, and it is these studies that will be considered first. Larson and Nippold (2007) used DA methods to probe the understanding of derivational morphology in typically developing 12 year old children. A series of graded prompts was used in order to ascertain whether students explicitly analysed the individual morphemes in morphologically complex words, and whether prompting to look at separate morphemes would facilitate their understanding of the words. The procedure elicited a wide range of performance scores that correlated with aspects of literacy, and demonstrated the potential to identify low performing students in need of intervention in this area. There is a need for a similar procedure to be applied to children with language impairments to identify their needs for intervention to support morphological awareness, language comprehension and expression and contribute to improvements in literacy.

The use of DA methods to investigate the pragmatic function of requesting information (RI) in children with ASD was explored by Donaldson and Olswang (2007). Whilst the static method was useful

in enabling comparison of the performance of ASD children and TD children in their spontaneous use of requests, and in establishing a baseline of performance in all populations, the dynamic procedure enabled a more representative reflection of the abilities of the lower functioning children with ASD when more facilitative prompts were made available to them. Information about which linguistic prompts best facilitated production of RI would enable these to be incorporated into an intervention programme. Schwabe et al (1986, cited by Donaldson and Olswang 2007) found that children with LI demonstrated limited RI, which was in part attributed to the linguistic demands of the RI task. The static/dynamic procedure employed in the Donaldson and Olswang study may similarly have potential to be useful for investigating RI in children with LI.

Likewise, the DA procedure used to assess the receptive language skills of children with Down syndrome (Alony and Kozulin 2007) may be applied to children with language impairments. Alony and Kozulin found that the performance on a test of receptive vocabulary, the PPVT-R could be improved with the minimal mediation of focussing the child's attention on the task. Furthermore, verbal mediation for those children who were deemed to be in need of it, facilitated developmental trajectories comparable to those of the normative sample, despite the subjective decision making regarding the need for intervention.

Strong parallels may be drawn between the work of Swanson and Howard (2005) on reading disabilities, and the field of language impairment. The problem posed by Swanson and Howard is the distinction between those with reading disabilities (RD) who have genuine information processing difficulties, and those who are poor readers for reasons of poor experience and instruction. Proponents of domain general processing limitations as a key component of

language impairments may be confronted with similar questions. Dynamic Assessments of working memory were used to determine whether the procedure could increase differentiation of RD from poor readers and whether the responsiveness to mediation of children with RD was poorer than that of poor readers. The tasks used were of phonological working memory, i.e. rhyming, and of semantic memory i.e. digit and sentence recall, and prompting sequences to facilitate performance were constructed for each. Results of the study were complex and inconclusive due to the small sample sizes. Nevertheless, there were indications that although the children with RD were not clearly differentiated from the poor readers, strategy learning was maintained in poor as well as skilled readers, but not in children with reading or reading and maths disorders. Thus the study highlighted the poor response to treatment in clinical groups, which may be useful as a classification feature. The authors concluded that DA procedures do have some use and that although working memory is an important component of achievement, DA of content material should also be considered.

As discussed in Chapter 1, key studies into the use of DA to assess culturally and linguistically different children (hereafter CLD), and differentiate typically developing children from those with LI have been published by Elizabeth Peña and colleagues. Gutierrez-Clellen and Peña (2001) described the tendency of children from diverse cultural and linguistic backgrounds to under-perform on standardised tests, resulting in over-diagnosis of language impairment. DA was found to be a more culturally fair means of assessing the responses of such children to learning experiences, providing opportunities to familiarize them with test expectations and probe their responses. Peña and Iglesias (1992), used a DA including Mediated Learning Experience (MLE), and found that post-test scores and modifiability ratings were useful to differentiate between typically developing

children and those with language disorder. Peña (2000) further probed the modifiability ratings used in her earlier research and found the examiner ratings of children on the Modifiability scale, comprising such criteria as examiner effort, responsiveness and transfer, differentiated TD from low language ability children. Furthermore, the expanded Learning Strategies Checklist (Peña 1993 cited by Peña 2000) found all criteria with the exception of motivation, to be significantly different in the two groups. Peña concluded that clinician modifiability ratings were a useful, non-biased means of determining diagnosis of LI in CLD children.

Similarly, Peña, Iglesias and Lidz (2001) examined the performance of preschool CLD children, using a word learning task, with a pretest-teach-posttest method. The teach phase consisted of mediated strategies for naming, and the children's performance during these sessions was also rated for modifiability. Posttest scores and ratings differentiated the typically developing children from those with low language ability, who were less able to benefit from the short-term MLE intervention. Typically developing CLD children markedly improved their performance on posttest, and were also able to transfer learning to other areas of language, showing improved scores on other tests of language that did not specifically tap naming abilities. Dynamic Assessment methods were more predictive in this differentiation than static pretest scores, which have been shown to overdiagnose children with CLD as language impaired.

Further insights into the nature of training in the teach phase of DA that best facilitates gains from pre- to post-test, were investigated by Kester, Peña and Gillam (2002). Fifty-two low SES, 3-4 year old children, of whom fifty were 'CLD' were randomly assigned to Direct Instruction (DI), MLE, Hybrid, or no-treatment (Control) groups. The

children were considered to have normal language development according to observations and reports. The Expressive One Word Picture Vocabulary Test-Revised (EOWPVT-R, Gardner 1990) was used to assess vocabulary before and after an intervention phase of three sessions spread over six weeks. Children in the DI group learnt to name items through play, imitation, repetition and practice, while the MLE and Hybrid programmes focussed on metalinguistic strategies for naming, and incorporating labels into communication. Although children in the group receiving Direct Instruction improved significantly from pre to post-test, the gains made by children receiving MLE and Hybrid interventions were greater. The gains made by the population of low SES and CLD children suggested that this group were underachieving in the static pre-test alone, and at risk of over-referral to special needs services. However the paper was most useful in demonstrating the advantages of individually determined, mediational intervention focussed on cognitive and metalinguistic strategies in facilitating improvement in language in CLD children.

Further studies using DA to identify children with language impairments include a study of receptive vocabulary by Camilleri and Law (2007). In the first study, a DA of receptive vocabulary was developed in order to compare the performance of monolingual English speakers with children with English as an additional language (EAL), and of typically developing children with those referred to SLT services. The static administration of the BPVS (Dunn et al 1997) was followed by a DA procedure aiming to facilitate learning of vocabulary by strategic use of relevance, discrepancy and mutual exclusivity criteria, rather than actual teaching of a new word. A hierarchy of mediational prompts was employed to lead the child to the strategy use. The DA procedure was found to differentiate between children with normally developing language and those

referred to SLT services, but to equate monolingual children and children with EAL, whose static scores on the BPVS differed. This suggests that the static test may not be suitable for children with EAL and risks overdiagnosing them as language impaired.

In a follow up study Camilleri and Law (in preparation) reassessed children from the original study after six months, and found the DA to be predictive of the change in receptive vocabulary, in the group of children identified as Low Scorers on the static BPVS. Most significant was the predictive validity of one of the expressive tasks in the DA procedure, which suggested that the reinforced lexical representation of the item required for the expressive task was important in word learning. Furthermore, a modified procedure (Camilleri and Botting, in press) demonstrated the reliability and predictive validity of the DA as well as its value in accessing clinically useful diagnostic, predictive and intervention information.

Moving away from vocabulary studies, Peña et al (2006) examined the classification ability of a DA of narrative ability in first and second grade school children. Two wordless story books, found to be parallel were used as pre- and post tests, and two sessions of intervention targeting story components, were carried out in the 'teach' phase. Intervention was mediational in nature, and slightly individualised for each child. Ratings of modifiability according to the earlier 3-criterion Modifiability Scale were also carried out after the second intervention session. In general, all children performed better on the post-test after the two sessions of MLE, but the TD children showed greater gains than those with LI. Pre-test measures of narrative did not accurately classify TD and LI children. The authors identified which measures of narrative on the post test were the most discriminating, and found that measures of 'Story Components' and 'Episode Structure' yielded the best classification accuracy, but

that the best single predictor was the clinician's modifiability rating, which was seen as consistent with the aims of DA which are to assess responsiveness to instruction. Thus it can be seen that the results of the study using narrative parallel the findings of earlier studies using naming tasks, and confirm the advantage of DA over static tests for classification purposes, as well as the significant usefulness of modifiability ratings.

Subsequently, Peña, Resendiz and Gillam (2007) formalised the role of modifiability ratings in a study in which examiners blind to the classification of children as TD or LI, rated aspects of their functioning in MLE sessions. The study aimed to determine the extent to which measures of modifiability could predict language ability and gain scores. The nature of modifiability was also scrutinized by comparing the interactions between clinicians and TD or LI children. The rating scale used, the *Mediated Learning Observation,* consisted of 5-point ratings on twelve criteria, divided into the four domains of Internal Social-Emotional (or Affect), Cognitive Arousal, Cognitive Elaboration and External Social-Emotional (or Behaviour). The study identified the two criteria of metacognition and flexibility as being the most reliable differentiators of TD and LI children, as children's awareness of their errors and ability to modify their output in response to instruction was variable. Cognitive factors were shown to be related to improvements in aspects of narrative, and may be important in both identifying language impairments and facilitating improvements in therapy.

Case study data further elaborated the nature of the responses rated during the MLE sessions, and demonstrated that a child with LI was asked many more questions than the TD child, in order to enable him to grasp and generalize concepts, and was given many

more cues to enable him to apply his learning to the material. Thus a great deal more examiner effort was recorded in facilitating the gain of the child with LI. Carrying out MLE sessions and using ratings of cognitive skills were not only reliable indicators of classification, but were extremely informative with regard to obtaining detailed assessment of the skills available to children with language impairments and identifying the prognosis for improvement of those skills through intervention.

Studies without a focus on diagnostic classification but which use DA to probe the language abilities of children with LI, inform further intervention, and make prognoses for improvement, include investigations of phonology, phonemic awareness, receptive vocabulary, and syntax. Glaspey and Stoel-Gammon (2007) used the Graduated Prompt paradigm to construct the Scaffolding Scale of Stimulability (SSS) an assessment of phonological skills that not only records the accuracy of a child's productions, but evaluates the cues and manipulations that support his performance. Clinicians have informally used assessments of stimulability for sound production to gauge prognosis as well as to select targets for remediation, and the current procedure incorporated that notion into a DA paradigm. The SSS rates stimulability on a scale of 1 to 21, obtained by manipulating the linguistic environment in which the sound is produced, and the cues used by the clinician to support the child's production. Whilst still untested in a large sample of children to ascertain whether the cue hierarchy is valid in all children, indications are that the SSS enables detailed and clinically relevant assessment of phonological skills. The authors' intention to provide a means of mapping incremental progress while a sound is still not stable enough to register correct on a static test is clinically valid and advantageous.

In an earlier study, Spector (1992) devised a DA of phonemic awareness, specifically segmentation skills, that was found by positive correlation with subsequent tests, to predict progress in reading.   It appeared that children who showed the most improvement in word recognition across the study period were those who benefited most from the prompts and cues supplied during the DA. Although the explanatory value of the study is limited as it is not clear whether the DA was a more sensitive measure of phonemic awareness than static phoneme segmentation or whether it addressed a different skill altogether, that of responsiveness to intervention, there is potential for the assessment to be developed into a prognostic indicator of the need for intervention to facilitate progress in reading, or for its use to probe in detail a child's prerequisite skills of segmentation to underpin his reading abilities.

Olswang, Bain and Johnson (1992) applied Vygotskian theory and Feuerstein's DA methods to gauge the learning potential of young children in the language acquisition process. The authors constructed a hierarchy of prompts and transfer tasks to assess the potential for children at the single word stage of development to progress to combining two words in various semantic relationships. Two children, aged 32 and 35 months, both using single word utterances only and thus exhibiting language delays in comparison to their chronological age norms, were investigated using the DA protocol. Although the children performed similarly on the static assessment, their response to prompting during the DA differed markedly. One child was seen to produce several two-word constructions of the agent-action, action-object, action-location, and entity+attribute types, in response mostly to direct modelling and elicitation techniques, but also on occasion in response to shaping, direct and indirect models. The other child produced only one 2-word structure, in response to a direct model plus elicitation prompt. Thus the procedure

demonstrated the differing potential for immediate improvement in the two children.

The children subsequently received three weeks of intensive (one hour sessions, three times a week) direct treatment designed to teach the two-word utterances. As predicted from the DA, the rate of change of the two participants differed, although both showed some gains in production of targeted semantic structures. Given the very small experimental study, conclusions are tentative, but do signal the value of the DA procedure to determine a child's potential to benefit from intervention. Indeed, one subject showed the propensity for imminent change that may have occurred spontaneously without intervention, while the other showed little ability to benefit from instruction at the time it was given, and the authors postulated the existence of a third profile between those two, that of a child showing gains, but only with substantial instruction. Olswang, Bain and Johnson appropriately identified a need for further investigation of precursors to linguistic development and follow up studies determining the longer term development of both stimulable and non-stimulable behaviours.

Olswang and Bain (1996) later compared the predictive strength of static and dynamic assessments, by correlating each with measures of immediate change in children with LI. Using similar procedures to those in the previous study for the intervention phase, extended static and dynamic assessments were carried out on a sample of 21 children of 31-36 months, all at the single word stage and identified as having expressive language impairments. Language was measured in terms of MLU, and change over the intervention period was calculated using the Proportional Change Index (PCI, Bain and Dollaghan 1991, cited by Olswang and Bain 1996). PCI was correlated with static and dynamic measures taken in the baseline

phase of the study. Results showed that individual static tests were not good predictors of language change, but discrepancy measures did show improved correlations. Dynamic Assessment scores however, had the highest correlations with PCI, and appeared to accurately predict children whose language would or would not change significantly during the treatment study.

Olswang and Bain discussed in more detail the children who did not improve during the intervention phase and hypothesized that the targets, treatment techniques or duration of the treatment phase may have been inappropriate for some of the children. The need for further research into the nature of language learning in children with LI was identified, as well as the need to investigate precursors that enable progress in language to take place. However, the authors did not explore the role of DA in identifying suitable interventions, rather for the purposes of the research, children were placed in a standard intervention programme. The wider potential of DA as a procedure to inform intervention practice was therefore not maximised.

Few research papers apart from the Bain and Olswang studies addressing two-word structures, have utilised DA to assess grammar or expressive syntax. Gummersall and Strong (1999) investigated the use of complex sentences within narratives, showing significantly improved performance by TD children following modelling and requests for imitation in comparison to standard story retelling procedure. The second experiment reported, showed children with language impairments who were assessed in the assisted condition, producing fewer complex sentences than the TD children in the facilitated story telling condition, but more complex sentences than the control group TD children in the simple story retelling condition. The authors concluded that typical story-retelling narrative assessment methods may not be sensitive to the full potential of

children with LI to use complex sentences, and may underestimate their abilities as even a simple manipulation such as that employed in the study elicited performances superior to that of the control TD children. However, as no group of LI children in the control condition was used, there is no means of estimating the ability of LI children to benefit from the intervention, and the results of the study are inconclusive.

In 2001, Peña adopted the elaborated feedback technique to explore the knowledge of words of a child with CLD, following assessment on the EOWPVT-R (Gardner 1990), in comparison to a child who did not receive feedback. The feedback provided greater insight into the semantic organization, retrieval and concept formation of the child, while the child in the no-feedback condition reached a plateau sooner. The face validity of the procedure was demonstrated, although Peña cautioned that reliability had not been proven. Similarly, Peña illustrated the usefulness of information gained through clinical interview of a young child with language impairment, in which the opportunity to probe responses added insight into the child's abilities.

Finally, Peña and Gillam (2000) published a series of case studies illustrating the clinical information to be gained from DA and MLE in vocabulary, narrative and explanatory discourse. Procedures for assessment, mediation and scoring of responses were described, based on previous research by the authors into the areas studied, and findings were used to demonstrate how to determine needs and useful strategies for intervention programmes.

In summary, despite improvements in the psychometric rigour of standardised tests, and calls for clinicians to attend to statistical properties that would make the use of testing more valid, there is an

ongoing dissatisfaction with standardised tests and repeated calls for more holistic and broad based functional assessments of language. In addition, the issue of culture fair tests has become more prominent, and there is an urgent need for development of these. Dynamic Assessment of language has begun to address some of the issues, and in combination with static test results has produced promising outcomes with regard to differential assessment of children, and recommendations for intervention. Whether the promises for improved intervention can be delivered, remains to be researched.

2.3   **The Management of Language Impairment**

2.3.1 **Traditional Interventions**

Few evidence based intervention studies have addressed children of school age, and even fewer in the last ten years, since expectations of methodological rigour have increased. The meta-analysis published by Law, Garrett and Nye in 2004 included only one paper published in 2000. Earlier published studies, from the 1970s, 80s and into the 1990s have as the focus of their research, methods of intervention that may be shown to be effective with children with LI. Indeed several studies compare one method with another, for example Friedman and Friedman (1980) compared a 'Programmed approach' that used a pre-planned structured hierarchy of stimuli to address specific syntactic targets with an 'Interactive approach' that embedded training in an approximation of normal conversational interactions. Weismer and Murray-Branch (1989) compared simple Modelling, with the enhanced 'Modelling with evoked production' technique, that required children to repeat the structures that were modelled to them, and Yoder, Kaiser and Alpert (1991) compared a Milieu Language teaching programme which emphasized modelling and reinforcing new forms as they occurred in functional language with the 'Communication Training Programme' (CTP) that

predetermined the number of trials and the targets of intervention. Finally Nelson et al (1996) compared Imitation with Conversational Recast methods which correct a child's errors by repeating back to him an expanded or corrected model of his own attempt. Each of these studies selected particular aspects of language for development, basing their investigations on structures absent or poorly developed in the language of the children, or in fact on measured aspects such as MLU or poor test results. A review of these studies is contained in Ebbels (2008).

A comprehensive review of intervention methods by McCauley and Fey (2006) enables practitioners to compare approaches on the same criteria, and select which may be applicable to their practice. Traditional methods of Focused Stimulation, Enhanced Milieu Training and Conversational Recasts, such as those contained in the preceding paragraph, are presented with their empirical basis, but are all to be identified as 'Targeting Prelinguistic Behaviour' and applicable only to older children with language impairments whose language is at a level of early grammatical development. Approaches for older children tend towards the development of literacy, and aside from intervention for phonological awareness, there are few documented strategies for targeting grammar in older school age children.

Some early studies also attempted to relate the methods of intervention to the characteristics of the population to be addressed. For example, Friedman and Friedman (1980) found that children achieving higher baseline IQ, Developmental Sentence Score and visuomotor integration scores improved more on the Interactive approach than with the Programmed approach, while the converse was true for children scoring lower on the baseline IQ, DSS and visuo-motor integration scores. As a consequence it was thought

that more didactic approaches may be suitable for lower functioning children. Yoder, Kaiser and Alpert (1991) however, found that the lower functioning children in their study benefited more from the Milieu teaching than the more structured and targeted Communication Training Programme, which they thought might be related to the locus of attention of the children, and also the language structures being targeted, namely vocabulary for lower functioning children, and syntax or morphology for more advanced children.

Thus the authors debunked the 'one size fits all' notion that any one treatment method was superior for all children to any other method. Nevertheless, taken together, studies of methodology indicated that grammatical structures can be improved through intervention for young children with disorders of expressive language. In addition, Fey et al (1993) found that the gains in grammatical development that could be shown as a result of interventions were robust enough for the intervention to be delivered by the parents of children with language impairments, when they were instructed by SLTs.

Early studies contributed significantly to the body of knowledge about treatment for children with language disorders, and traditional methods are incorporated into the more holistic methods represented in more recent intervention studies. More recent accounts of good practice e.g. Fey, Long and Finestack (2003), Kamhi (2006) and Balthazar and Scott (2007) note that clinical decisions relating to selection of therapy targets, sequencing of goals, methods used to implement therapy, length and frequency of intervention sessions, and role of the SLT are based on research evidence as well as clinical expertise, client values, theoretical perspective of the clinician, and service delivery constraints. Fey, Long and Finestack (2003) place the emphasis first on the selection

of the goals for intervention, and then the selection of methods to facilitate development of those targets, as well as the functional application of learnt behaviours. Their ten principles for grammatical intervention for children with SLI primarily address morphology and syntax, whilst simultaneously accounting for the concomitant social, behavioural and academic difficulties of this population. Their recommendations thus pertain to the selection of goals and sub goals, as well as to the methods used to address targets, and the contexts in which intervention is conducted. These principles may be revisited as they relate to intervention studies detailed in this section.

2.3.2 Issues in Intervention research

A selection of papers directly relevant to the current study is reviewed here. For the purposes of the current study, interventions addressing sentence level structures only will be reviewed. Interventions targeting phonology are extensive and beyond the scope of this review, however, it is important to consider cross-domain interventions, such as the study by Tyler et al (2002) in which intervention that targeted morphosyntax was found to have beneficial effects on the phonology of the participants, while interventions that targeted phonology significantly improved phonology, but had no effect on improving morphosyntax. Although the findings are not in agreement with outcomes of all other studies cited by the authors, important implications for the sequencing of targets and interventions for children with both morphosyntactic and phonological deficits were identified.

Camarata et al (2009) similarly demonstrated cross domain effects in a study somewhat uniquely considering receptive language. Growth in receptive language skills were detected as a result of intervention targeting expressive language. Therapy addressing a range of morphosyntactic structures through a combination of direct

and indirect methods of modelling, imitation, recasting and milieu teaching, over a period of 12 weeks, resulted in gains in receptive language, as measured by the Auditory comprehension subtest of the PLS-3. Camarata et al emphasized that the improvements resulted from traditional grammatically targeted intervention, in the absence of any auditory processing training. Again, important implications for intervention programming were suggested.

The issue of auditory processing relates to the publication of a computer delivered training programme known as FastForWord (FFW), that was one of the first to address receptive language difficulties (Tallal et al 1996, Tallal et al 1998, Tallal 2000). The intervention includes auditory training of children to improve auditory discrimination, consistent with theories that implicate auditory processing deficits in language disorders. Tallal et al (1996) reported that training children with modified speech improved their gains on language tasks. The subsequently developed FFW enabled delivery of an extremely intensive daily auditory discrimination and comprehension training programme of activities embedded in computer games. Subsequent studies have called into question the assumptions of the programme and the findings of the FFW. Bishop, Adams and Rosen (2006) for example, pointed out that the programme would not necessarily be applicable to all children with SLI, nor to other clinical groups that it had been used with. Gillam et al (2008) carried out a randomized control trial on 216 children assigned to FFW and three other intervention conditions. They found that while all the children improved as a result of their intervention, there was no advantage of the FFW programme over the other interventions on measures of language skill and of temporal processing. The authors noted that their results call into question the temporal processing hypothesis as an explanatory theory of SLI. Similarly, in the UK, Cohen et al (2005) carried out a randomized

control trial with 77 children and again found no advantage of FFW over commercially available computer activities and regular SLT without additional intervention.

Bishop, Adams and Rosen (2006) studied the efficacy of computerized training of grammatical comprehension, with and without auditory processing training components. Intervention aimed to improve sentence comprehension accuracy through training of specific constructions namely the prepositions over/under and above/below, passive constructions, and sentences containing comparatives. These were selected as ones in which word order in sentences was crucial, as this is a known area of difficulty for children with LI. In brief, Bishop et al did not find any benefit to grammatical skills of children with LI as a result of the computerized intervention, irrespective of whether modified speech was used. Some modest improvements were noted particularly in the understanding of passives, but improvements in speed of responding were difficult to interpret.

Implications arising from Bishop's study contribute to understanding of language impairment. Firstly, the authors noted that children with LI do not have a complete absence of understanding of particular structures, but a fragile grasp that is subject to disruption when they are asked to process structures online. Intervention therefore needs to target structures in context, rather than training morphemes such as prepositions, in isolation. Secondly, even syntactically simple sentences were incorrectly interpreted when reversible structures such as 'above' and 'below' were used. These structures are not easily remediated through use of repeated practice with corrective feedback, and research into effective interventions for this type of structure is still required. Finally, Bishop noted that several aspects of the intervention study were not consistent with recommendations

for good practice by Fey et al (2003) in isolating grammatical structures from functional contexts.

An issue identified in the FFW programme and one commanding attention in current intervention research, is that of 'dosage' (Cirrin et al 2010, McCauley and Fey 2006). This refers to the length and intensity of therapy sessions, as well as the duration of the period of intervention. A recent series of studies, by Leonard et al (2008) explored in detail the acquisition of three particular morphemes, namely third person singular (3S), auxiliary is/are/was, (AUX) and past tense –ed, as a result of intervention specifically targeting the first two of these. The first study in the series, (Leonard et al 2004, cited by Leonard et al 2008) demonstrated the gains made by young children in the 3S and AUX structures, representing an awareness of *tense+agreement* features that may not have generalised to the *tense only* aspects of the past tense. The second and third studies expanded the intervention programmes first by increasing the number of treatment sessions, and finally by including a condition in which intervention was less focused, and also investigating maintenance of learning after one month post termination of treatment. Of concern to the current author is the reporting of only modest gains in learning of 3S and AUX, up to 45 or 50% accuracy, after a series of 96 sessions of intervention. The final study confirmed the advantage of focussed intervention over a programme of general language stimulation, and reassuringly also provided empirical evidence that the facilitative effects of therapy continue after the intervention period. However, the modest gains persisted, and the authors proposed the influence of age and maturational readiness as an explanation. Nevertheless, it does not seem an efficient use of resources when gains are modest after a number of intervention sessions that would not be viable in most packages of

care offered in UK services, and positive outcomes of the intervention programme should be interpreted with care.

Cirrin et al (2010) attempted an evidence based systematic review of different service delivery models addressing speech and language therapy in school children. Unfortunately, only five studies met the criteria for acceptance for the review, and results are necessarily inconclusive. There was some evidence that classroom based services were equally effective as 'pullout' intervention, and that trained SLT assistants working under guidance of an SLT can achieve good outcomes from their intervention. However, generalization of any findings is limited owing to the specific criteria employed in each study, and it is apparent, for example, that not all clinical environments and schools would be employing assistants as highly qualified and experienced as those shown in the current study to produce good outcomes of intervention. The current recommendations therefore are that *'lacking adequate research-based evidence, clinicians must rely on reason-based practice and their own data until more data become available'* (p.250).

The systematic review by Law, Garrett and Nye (2004), was more successful in establishing both the effectiveness of SLT interventions, and aspects of service delivery, that found parents to be equally effective at delivering therapy as clinicians for some types of language difficulty (mainly receptive), and that a therapy programme of more than eight weeks duration was significant in achieving good outcomes. Unfortunately none of the interventions identified as metalinguistic or metacognitive met the criteria for inclusion in the meta-analysis, and a summary of these is contained in the next section.

2.3.3 Metacognitive interventions

Metacognitive interventions have been recommended for a range of curricular domains, including second language learning (Williams and Burden 1997), and applications to language therapy are apparent. This section contains a brief review of metacognitive approaches that may be seen to have relevance to language processing programmes.

Metacognitive or thinking skills programmes introduced into the National Curriculum in England, were reviewed by Burden (1998), who highlighted systemic difficulties in integrating holistic thinking skills programmes into the educational system. A subsequent meta-analysis was carried out by Higgins et al (2005) to evaluate the impact of thinking skills interventions on teaching and learning, and indeed to find quantitative evidence for effects on pupils' attainment in schools. The working definition of thinking skills programmes adopted by the reviewers was similar to that given above for metacognitive skills, i.e. the programmes were ones which *identify for learners translatable mental processes and/or which require learners to plan, describe and evaluate their thinking and learning'* (Higgins et al 2005 p7). The Instrumental Enrichment programme devised by Feuerstein (Feuerstein, Rand and Hoffman 1980) was named as one of the most established and extensively researched thinking skills programmes.

The meta-analysis revealed that thinking skills programmes were effective not only in improving performance on cognitive measures, such as the Ravens Progressive Matrices, but also on curricular outcomes, in maths, science and reading. The authors concluded that thinking skills programmes are effective at improving pupils' attainments, with an effect greater than that demonstrable from other educational interventions, and their use in schools should be supported. Further research is required to pinpoint the effects of

content versus teaching methodology, and to relate specific outcomes to variables such as age and curriculum area. Nevertheless, the present author can see parallel benefits for including thinking skills or metacognitive approaches in language therapy.

Metacognitive approaches for clinical populations include the use of Instrumental Enrichment (IE, Feuerstein, Rand and Hoffman 1980), as well as the IE Basic programme (Feuerstein and Feuerstein 2003, cited by Kozulin et al 2010), the adaptation of IE type interventions for younger children. Kozulin et al, in a multicentre study, showed that it was possible to improve the fluid intelligence of children with a range of developmental cognitive impairments, using instruments from IE Basic, along with the mediated learning experience techniques associated with Feuerstein's theory. Similarly, the Bright Start Cognitive Curriculum (Haywood, Brooks and Burns 1992) also based on principles of Feuerstein's cognitive modifiability and MLE, has shown positive results from intervention in children from low SES backgrounds (Tzuriel et al 1999) and in children with Down syndrome (Garcia and Conte 2004).

Process Based Instruction (Ashman 1992) was developed to meet a demand for a learning skills intervention that could be applied in mainstream and special needs classrooms, and was linked to ongoing monitoring assessments. The programme addressed problem solving strategies, the understanding of problem solving principles and planning for problem solving. It incorporated aspects of reciprocal teaching methods, in which responsibility for learning is gradually transferred from teacher to student, and also principles of MLE in which the teacher is responsible for framing learning in a way that is accessible to students. The programme consisted of a sequenced programme of steps towards task orientation and

performance, instruction, impeded learning, and generalization. Ashman reported the results of a study of the effects of the PBI for training simultaneous processing strategies or sequential strategies in 34 students with mild intellectual disabilities. He found that the strategy training did improve their processing performance in relation to the tasks on which they were trained. There was no transfer of training from simultaneous to sequential tasks, but interestingly, training on sequential processing had positive effects on language processing tasks. Positive outcomes from the programme included improved attitudes towards learning in the students, and demonstrable value from the integration of assessment with instruction.

2.3.4 Metalinguistic interventions

A prominent metalinguistic approach to intervention for children with speech and language impairments is the Metaphon programme (Dean et al 1990) which addresses both assessment and intervention for phonological awareness and sound production. Dean et al (1995) provided a rationale for the approach by citing relevant findings about developmental progression in awareness of linguistic segments, and the role of metalinguistic awareness in facilitating second language learning and literacy. Their central assumption is that awareness of the way in which phonemes contrast would facilitate changes in the (disordered or delayed) processing of speech sounds. Thus children are taught to identify and label features of phonemes, in order that they can reflect upon and alter their own phonological output. Dean et al presented a series of case studies to illustrate the effective use of Metaphon intervention.

Strengths of the Metaphon approach are its theoretical grounding and attempt to involve the child as an active participant in learning. In addition it attempts to meet the needs of a heterogeneous

population by a multilayered design that encompasses a number of different dimensions, for example contrast training, matching and categorizing sounds, or experimentation with articulatory configurations. It proposes as its target the improvement of phonological awareness as well as sound production. Critics of Metaphon however, point out the difficulties caused by identifying phonetic features that are not representative of sounds in connected speech (Grundy 1995), and the resulting productions of atypical sound sequences that are associated with children receiving articulation therapy. Bleile and Hand (1995) have difficulty with the metalinguistic approach of teaching component parts of the behaviour, and express concern that this analytical process does not facilitate correct use in connected speech. Miccio (1995) focused on the missed opportunity to facilitate generalization of training by predicting relationships between sounds and selecting targets according to these. The clinical concerns of Klimacka (1995) relate to the minor role attached to concomitant vocabulary and syntactic impairments in the programme.

Although the current review does not intend to focus on phonological interventions, there is no parallel metalinguistic intervention programme to Metaphon that addresses disorders of syntax. It is thought that such a parallel would be useful and avoid some of the pitfalls of Metaphon, while preserving the strengths. The use in Metaphon of vocabulary to name the attributes of sounds raises the level of awareness from an automatic level at which children may spontaneously self correct their utterances, or in fact play 'word games' to the highest level at which verbal reflection is enabled. This would be a level at which metalinguistic interventions targeting syntax might similarly be useful. While sentence construction tasks may not be fully facilitated by 'bottom –up' word or morpheme selections, sentence construction is a more explicitly rule governed

system that may be explicated through reflection on rules, and lend itself to generalizations across structures that are subject to the same rules. Nettelbladt (1995) suggested that Metaphon could be interpreted within a Vygotskian socially constructed framework, and indeed many of the principles of guided learning and the zone of proximal development are consistent both with the Metaphon approach and that of Dynamic Assessment and Mediated Learning Experience principles.

Other areas of language addressed by metalinguistic approaches include the explicit teaching of phonological and morphological awareness to typically developing kindergarten children in a study by Casalis and Cole (2009). The metalinguistic interventions showed positive effects of training in phonological and morphological domains on the trained domain, but few and very specific cross domain transfer effects. Guterman (2002) similarly showed positive effects of metacognitive strategies in enhancing the reading performance of fourth grade children, and Justice and Ezell (2004) presented an evidence based strategy known as 'print referencing' that enhanced emergent literacy in young children. Ravid and Geiger (2009) addressed morphology intervention via metalinguistic means in TD children, using aspects of humour, as did Yuill (2009), who showed improvements in receptive language of TD children after metalinguistic discussion of ambiguity. Graham and Perin (2007) targeted sentence combining in the writing skills of TD children, and Kohnert and Danahy (2007) addressed rule learning for morphemes in TD children learning a second language. As none of these papers specifically target children with language impairment, no further review is contained here.

Metalinguistic therapy approaches to address the grammar of children with LI are few, but the body of research is growing. Bryan

(1997) devised the Colourful Semantics system which used colour to help children identify components of sentences, and compose argument structures. Case studies providing anecdotal evidence for the success of this approach were carried out, but Ebbels (2008) noted that no experimental studies using control groups were available. The use of colour for different questions was also included by Spooner (2002) in an intervention designed help children organize sentence components. One of the two children in her study increased her sentences from predominantly verb plus one argument structure, to sentences containing two or three arguments. Both children improved in the accuracy of their lexical verb selection and inflection, and gains were detected by formal tests after therapy. However, despite the colour coding that made question forms salient for the children, there was little use of explicit metalinguistic training in the intervention.

Ebbels (2007) however, used a system of colours, shapes and arrows and made the rules of sentence construction completely explicit for children. Colours were used for the basic parts of speech, (noun, verb, adjective) and shapes were used to encode phrase structures according to their role or position in a sentence, in order that these shapes could be moved around, or even embedded inside each other. This allowed for encoding of subject versus object forms and active versus passive sentences. In addition, verb morphology is coded by a system of arrows to prompt children to add morphemes including those for agreement. At each stage in the therapy (details of which are given in Ebbels 2007) rules are explained to the children, who become aware that words answering a particular question are coded within a particular shape. Complexity is built up slowly and sequentially in response to the levels of the child.

Ebbels and van der Lely (2001) reported gains in the use of passive sentences and wh- questions in response to targeted shape coding therapy, that was significant in three out of four children. Ebbels (2007) studied the effect of therapy using shape coding on the understanding of prepositional and dative forms in three of these children. Previous therapy had established the prepositional form (give X to Y), but not the dative (give Y X) in two of the children, who were shown to benefit from shape coding and grasp the dative form after intervention. The third child differed in his pre-therapy performance which showed unstable understanding of the prepositional form, and did not benefit from the intervention, a finding which may have been due to auditory memory difficulties. Further study of two of the same children showed shape coding therapy to be effective in teaching comprehension of comparative questions.

Finally, shape coding therapy addressing past tense morphology was shown to be effective in six out of nine children in a group, and helpful to a further two children when delivered in paired therapy. Ebbels concluded that shape coding is a useful method, flexible enough to be used to address a range of grammatical targets, but like most other approaches, cannot be assumed to be effective with every child for every structure, and thus individual differences need to be pursued in further research. What distinguishes shape coding from the intervention approach of Bishop, Adams and Rosen (2006) is the delivery by clinician rather than by computer, and the explicit teaching with feedback that enables children to recognize their own errors, and either one of both of these would appear to be the crucial factor in eliciting positive results from intervention.

Finestack and Fey (2009) captured what the current author believes to be the essence of the rationale for metalinguistic intervention,

phrasing their intervention as a deductive procedure. Children with SLI fail to progress sufficiently in language even when exposed to adequate environmental models, and though they might benefit from increased models and recasts, efficiency of intervention seems most likely to be improved by provision of explicit feedback about how to correct their language, or teaching of rules and principles, such as through shape coding. Finestack and Fey used two groups, one of whom was provided with examples from which they were to learn, i.e. the *inductive* method, and the other provided with awareness of the target patterns underlying grammatical constructions from which they were to deduce rules, the *deductive* method. Swisher et al (1995) used a similar paradigm in their experiment. They found, however, that children aged 4-6 with SLI had greater difficulty than TD children when required to abstract a rule and apply a bound morpheme to a nonsense noun, in response to explicit instruction as compared to implicit learning. While all the children had difficulty generalizing their learning of the morpheme, children with SLI had more difficulty in the explicit instruction condition. The task involved having to learn novel lexical items with novel affixes, requiring stable comprehension of the big/little concept, as well as generalization of the rule governing its application. One possible explanation for the findings might be the overload of the task requirements. Finestack and Fey, in contrast used slightly older 6-8 year old children, and real vocabulary, although with novel morphological markers on the verb for gender of the subject (a structure that does occur in many other languages). Gains in production of the novel morpheme after training by the deductive method, were compared to gains elicited by the inductive method, and were shown to be significantly better for production of the morpheme, generalization, and maintenance. As ever the results of this early efficacy study need to be extended to other morphemes, linguistic contexts and probes, but the results were suggestive of a positive finding.

Levy and Friedmann (2009) carried out a comprehensive and detailed case study of metalinguistic intervention with a 12 year old boy with SLI. The language of the subject was thoroughly analysed and interpreted in the light of linguistic theories, culminating in a diagnosis of specific difficulty with structures involving movement, specifically those resulting in non-canonical word orders. Intervention made use of explanations of linguistic concepts such as argument structure and movement, and all structures were trained in written language first, in order that a concrete representation may be preserved, and then in oral language. After an intervention comprising 16 sessions over a six-month period, improvements were shown in all structures, and generalization to some non-treated structures took place. Results were compared to the positive results obtained in aphasic subjects, and suggest that extension of this approach to therapy for children with SLI has merit.

Metalinguistic therapies such as those described have made use of detailed and explicit assessment and analysis of the structural language abilities of the participants involved. In many, the intervention targeted a single linguistic structure, such as a particular prepositional structure or bound morpheme, or a rule governing word order in non-canonical structures. Detailed syntactic analysis of the client's use of the structure or rule was carried out in order to accurately ascertain the precise weakness, and identify needs for intervention. In general standardised tests of developmental language do not contain sufficient examples of structures, or possible permutations of case or word order for this level of detail to be extracted, or for rules such as those governing movement to be elucidated. Further probing of the possible structural variations, or indeed questioning to find out the extent of the individual's understanding of the use of structures or rules,

would increase the specificity of targeted intervention and potentially improve outcomes.

In summary, metalinguistic approaches to intervention for children with language impairments figure more prominently in more recent studies, and the findings, though limited, are generally positive. Law et al (2008) noted the age trend for metacognitive therapies to be used for older children while younger children are taught more specific skills. Traditional skill based methods of intervention continue to have a place in the toolkit of practicing clinicians, but issues of individual variation along with the range of other practical considerations in planning intervention, result in eclectic selections of intervention methods. Furthermore, Law noted that rationales given by SLTs for their interventions are based more on deficits and tacit knowledge of therapies than theoretical models of language impairment per se. Theoretically grounded interventions remain the domain of the researcher, and there are many more directions for continuing research.

2.4 <u>The Present Study</u>

The present thesis reports two main research studies, the first regarding the development of a novel Dynamic Assessment procedure for the assessment of the language of children with SLI; the second, an effectiveness study to explore whether the input of information derived from the Dynamic Assessment enhances intervention in this group. Chapters 3, 4 and 5 consider the former study, whilst chapters 6, 7 and 8 report the second investigation.

CHAPTER 3

DYNAMIC ASSESSMENT: DEVELOPMENT OF THE TEST

CHAPTER 3   DYNAMIC ASSESSMENT: DEVELOPMENT OF THE TEST

3.1   Aims of the project

The aim of the research reported in this dissertation is:

To formulate a valid and reliable procedure for Dynamic Assessment of language that yields useful information for planning intervention for children with SLI.

Specifically:

i)   To explore the concurrent validity of the DA test compared with a static criterion test

ii)   To examine the retest reliability of DA scores over time, and sensitivity to change over time compared with a static test.

iii)   To develop a DA with good inter-rater reliability

iv)   To develop a DA test with good predictive validity

v)   To explore the content structure of the DA test

vi)   To create parallel forms of the DA test

The stated aim is consistent with a research goal stated by Budoff (1987a) which is for DA to assess those who have been correctly diagnosed, but whose potential for improvement has not been gauged.

In addition, the procedure needs to be replicable and teachable, in order that any demonstrated utility could then be adopted by practising SLTs in the field. Similarly, the demand in terms of administration time and scoring simplicity need to be controlled.

3.2   Design of the DA

3.2.1 Structure of the procedure

The dynamic procedure designed for the current study was a hybrid of procedures, with the training prompts and mediation directly interposed into the procedure, and without a post-test measure. Pilot testing of the procedure used a test- train- retest format (see Hasson and Botting 2010), and revealed that post testing was not useful for some language impaired children who gave a poor performance in the standardised test situation, regardless of their achievements in the intervening training.

The structure of the test was fundamentally a Graduated Prompt procedure, (adapted from Resing 1997), that enabled quantification of the number of prompts required by the individual in order to solve the given language task. While the essential number and nature of the prompts was standardised, administration was flexible within the prompt hierarchy, and cues were delivered in an individualized mediational style, enabling feedback to be given and responses to be probed. This combination of feedback adapted from Carlson and Wiedl's 'Testing the Limits' procedure (1992) and mediation (as recommended by Feuerstein) was intended to facilitate maximum transfer between items and times of testing, and also to highlight metalinguistic knowledge and strategy use by the participants.

Standardised tests would be retained alongside the DA for their normative value, and as an opportunity to get information about the child's independent performance on a selected task, and his ceiling level of achievement (Lidz 2003). Individual needs with regard to knowledge, strategy use and metacognitive or executive control were elucidated by the DA, as well as personal variables of attention, motivation and responsiveness. The whole procedure enabled the examiner to observe demonstrated abilities of the child, and allowed

planning of mediation, matching the demands of the task to mediations that may be facilitative of improved performance in the child.

Rating of the child's response to mediation using Lidz's Response to Mediation Scale (2003 see Appendix III) was carried out at the conclusion of each session, in order to rate behavioural responses in a systematic way. Clinicians' ratings of modifiability have been shown by Peña, Resendiz and Gillam (2007) to be a valid means of forming hypotheses about what might help a child benefit from intervention.

3.2.2 General content of the test

Children with Specific Language Impairment (SLI) manifest a specific deficit in the domain of language. This suggests that the target of diagnostic assessment ought to be within the domain of language, however Guthke's notion of multiple competences within a single subject area certainly applies to language. Multiple skills and processes are involved in the understanding, retention, processing and production of phonology, vocabulary, semantics, grammar and pragmatic areas of language, or combinations of these.

3.2.3 Rationale for the task

An area that is notoriously difficult to assess, and one for which there are few standardised tests, and even fewer standardised for older school age children, is that of expressive syntax. Those that do exist tend to target morphology, (e.g. CELF Word Structure, Semel, Wiig and Secord 1987; TEGI, Rice and Wexler 2001), rely on memory for structures by providing a model for imitation, (e.g. CELF Recalling Sentences, Semel, Wiig and Secord 1987) or employ a prompt to elicit a targeted structure in response (e.g. RAPT, Renfrew 1988, TEGI, Rice and Wexler 2001, ACE Syntactic Formulation,

Adams et al 2001). Expressive sentence construction is accessed
either through asking the subject to make a sentence with a given
word (e.g. CELF Formulated Sentences, Semel, Wiig and Secord
1987) or through analysis of spontaneous language samples (e.g.,
STASS, Armstrong and Ainley 2007, or profiles such as LARSP,
Crystal, Fletcher and Garman 1976).   Several approaches have
targeted narrative production, with the syntactic analysis utilising
length of T-units and number of subordinated clauses as a measure
of complexity (e.g. ERRNI, Bishop 2004, ACE Narrative Syntax,
Adams et al 2001) and some tests have used specific tasks such as
arranging words into sentences (e.g. CELF Sentence Assembly,
Semel, Wiig and Secord 1987).

Task analysis would reveal each test requiring a slightly different
skill, and a combination of subtests would provide a range of
information about an individual's abilities. It is essential, however,
that detailed task analysis is carried out, in order to tease out
exactly which component of any task causes problems for an
individual. Did a child failing the RAPT, for example, have difficulty
recognizing the events in the picture, making the appropriate
inferences, understanding the vocabulary or the grammar of the
prompt question, or formulating the response? Restricting a subtest
to isolate a very specific skill would result in a minute area of the
child's overall language ability being elucidated, and a risk of losing
the greater picture, especially with regard to functional use of
language. Thus, when selecting and employing a single test, the
examiner should be aware of the nature of the demands being made
on the respondent.

3.2.4 Task Design in the present study

The task was based on the Sentence Assembly subtest of the CELF-3
(UK) (Semel, Wiig and Secord 1987). Earlier versions of the CELF

(Semel and Wiig 1981) were criticized for a lack of theoretical or developmental basis for the subtests, and the omission of presented data regarding construct validity of the subtests (Spekman and Roth 1984, Muma 1984). Subsequent versions have improved standardisation and replaced subtests, but the present author was unable to find the authors' rationale for development and inclusion of the Sentence Assembly subtest in the test manual or elsewhere. This task was chosen for the present study as it enabled sampling of a number of underlying componential skills and processes, found in a pilot test (Hasson and Botting 2010) to be accessible through probing of responses as permitted by a dynamic style of assessment. Kahn and King (1997) similarly used the CELF Sentence Assembly task, giving no reason for its selection, but demonstrating its utility for accessing and assessing cognitive functions.

The materials utilized the same format as that in the CELF-3, with words presented visually, printed on a single card, (i.e. not separately in order that they could not be moved about manually) in random order, requiring the child to formulate two possible sentences from the given words. In addition, the grammatical structure of the possible sentences was controlled, requiring different linguistic constructions to be extracted, and presenting items in order of increasing difficulty, and/or increasing length/number of items in the sentence, for each grammatical structure (see Appendix IV).

The structure of the task removed the variable of comprehension of grammar, requiring the child only to understand the meaning of the single words presented, (which in the case of verbs did entail some knowledge of the argument structure) and the dynamic procedure permitted checking that the child was familiar with all the vocabulary items, or these could be explained if necessary, as was shown to

occur on a small number of occasions in the pilot test. Reading difficulties were similarly compensated by checking and helping the child to read each word, which would not affect the procedure, but conversely would provide additional information about the individual's needs for support.

Four parallel versions of the materials for the DA were developed to enable repeated administrations for the purposes of evaluating reliability of the procedure and for the investigation into the role of the DA in intervention (see Chapter 6). Each of these contained the identical sentence constructions, with alteration only of the vocabulary inserted into the sentences. The syntactic and morphological structure of sentences remained the same, and the number of words was exactly equal. All the vocabulary used was everyday vocabulary, of nouns and verbs thought to be well within the experience of children of primary age. The DA procedure, however, also allowed for word knowledge to be checked during the procedure, and any unknown vocabulary to be explained to participants. The versions were therefore considered to be equal (see Appendix V).

The task also reduced the demand on short term memory, by having the words written and in view throughout the task. There was still, however, a demand for working memory as the subject had to hold the sequence of words in mind as he formulated the sentence. Presenting the words on separate cards that could be moved around would in fact have eliminated this factor as well, but it was decided that working memory is a key skill in linguistic formulation, and one that has been shown to be poor in children with SLI (Gathercole and Baddeley 1993), and therefore qualitative observation of the child's performance in this area would be informative. The response to mediation could in fact elucidate strategies that the child used to

manage the memory demand of the task, and grading of the tasks in terms of length, or number of elements in the sentence would give further information about the limits of the individual's ability.

The knowledge requirements of the task pertained to the rules of combination of elements into a grammatical sentence, and the constraints of particular vocabulary items with regard to the argument structure that they command. In this respect, the task was both syntagmatic, i.e. the grammatical items stand in a sequential relationship to each other, and also paradigmatic in that alternative constructions can be extracted that encode a different meaning through different arrangement of the items (Huddleston 1984). For example, the sentence '*The boy is washing my car'* and the question '*Is the boy washing my car?*' can be formulated from the same set of words. In fact by requiring two different sentences to be made, the task taps directly into the individual's ability to manipulate linguistic elements to encode different relationships, within the rules of grammar which must be implicitly known. The *knowledge* of linguistic rules must be applied via effective *thinking skills* (Sternberg and Grigorenko 2001). Furthermore, Gopnik (1990) elaborated on the use of grammatical judgement tasks, which are implicit in the current task, as a reliable indicator of the individual's ability to process linguistic features (p.147).

The task in some ways parallels that described by Gredler and Shields (2008) as an example of the paradigm preferred by Vygotsky. They describe Vygotsky's preference for the "experimental-genetic" method, which uses '*concrete tasks that externalize the key principle under investigation*' (p.44). The example cited is that of a test of number concept, in which a child is given 10 tokens, and asked to make 4 in a number of different ways. The use of several different

strategies such as taking 6 tokens away from 10, suggests a more fully developed concept than simply being able to count out 4 tokens.

Finally, the specific linguistic structures that were tested by the developed materials were selected in part to represent structures known to be impaired in the language of children with SLI, although it was not possible to accommodate some structures within the format of the test task. Structures were sequenced in order of presentation taking account of grammatical complexity and sequence of age of acquisition, within the constraints of the test format. This degree of manipulation differs from the original CELF-3 subtest, where the sentence types are randomly mixed. This is in order to facilitate training through the presentation and practice of the test items themselves. In addition, Feuerstein (cited by Tzuriel 1991) ensured that tasks included in the LPAD battery began with simpler examples in order to give children confidence and ease the fear of failure and reluctance to participate that affects their performance on assessment tasks.

3.2.5 The Test Items

The items utilized in the DA were devised in order to assess the knowledge of a number of grammatical constructions as well as the individual's ability to manage variables such as semantic constraints, argument structures and length of sentence.

Processes and strategies employed to construct the first sentence from the given words, in some examples facilitated construction of the second sentence, by simple reversal of semantically equivalent elements (e.g. the Mum is../ the Dad is..). Semantically reversible sentences have been shown to be of particular difficulty for children with SLI (van der Lely and Harris 1990, Bishop 1997). Alternatively, the only possible second sentence may have been a question,

requiring auxiliary or copula verb inversion, also an area known to be difficult for children with LI (Rice and Wexler 1995, van der Lely and Battell 2003, van der Lely 2005) or may have required manipulation of an entire clause around a subordinating conjunction, (e.g. *'Cassie screamed because the door banged' / 'Because the door banged, Cassie screamed'*) with subordinated clauses, as well as the resulting non-canonical structure, a source of difficulty (van der Lely and Harris 1990).

The items were arranged in pairs with a common structure in each pair of items, but with the second increasing the level of difficulty in some way. This was either by increasing grammatical complexity by modifying the phrase structure, for example modifying the present continuous tense verb (*is painting*) to a future tense verb group (*is going to score)*, or by increasing the overall length and number of elements to be manipulated by the child, by adding arguments (e.g. *mum is eating*, to *mum is picking the flowers*). Strategies used to formulate two sentences from the words in one item should facilitate the solving of the subsequent item in the pair, i.e. if a question form was required in Item 1, Item 2 can also be solved by forming a question as one of the sentences.

The internal pair wise arrangement of items comprised a transfer task. This gauged the child's ability to transfer learning to items that are at times more abstract, and at other times more complex, but not more abstract, criteria defined by Feuerstein's 'cognitive map', and used in the construction and analysis of tasks in the Learning Propensity Assessment Device (LPAD).

The items used in the A version of the test are presented in Table 1

Table 1. Items used in the Test.

Item No	Structure of 1st and 2nd possible sentence	No of words	Item content
1	SVO Declarative with auxiliary plus main verb + PossPron NP Interrogative with aux reversal	6	The man is painting my wall Is the man painting my wall?
2	SVO Declarative with auxiliary plus main verb + future tense VP Interrogative with aux reversal	7	Billy is going to score a goal Is Billy going to score a goal?
3	Declarative with coordination SVcSV Reversible content	7	Mum is eating and Dad is drinking Mum is drinking and Dad is eating (etc)
4	Declarative with coordination SVOcSVO Reversible content with some semantic constraints	11	Mum is picking the flowers and Dad is cutting the grass Mum is cutting the grass and Dad is picking the flowers (etc) NB that 'picking the grass'* is unacceptable
5	Dative SVOdOi N and N in Subject position Reversible content	10	The girl and the boy showed the monkey a banana The boy and the girl showed the monkey a banana
6	Dative SVOiOd N and N in Obj position	10	The girl gave the boy a drink and a biscuit

			The boy gave the girl a biscuit and a drink (etc)
7	SVA Declarative with copula verb and Adj NP Interrogative with verb reversal	5	The big boy is hungry Is the big boy hungry?
8	SVA Declarative with copula verb +, contracted neg Interrogative with verb reversal	5	The bird's cage isn't broken Isn't the bird's cage broken?
9	SVA Declarative with modal auxiliary plus main verb and prepositional adverbial phrase Interrogative with aux reversal	6	Susie will hide under the table Will Susie hide under the table?
10	SVOA Declarative with modal auxiliary plus main verb, Object, and prep adverbial phrase Interrogative with aux reversal	8	You can hang your coat on the hook Can you hang your coat on the hook?
11	SVOsSVO Declarative with (temporal) subordinate clause (Use of anaphoric pronoun)	9	Joe brushes his teeth before he goes to bed Before he goes to bed, Joe brushes his teeth NB He goes to bed before Joe brushes his teeth*
12	Declarative with (causative) subordinate	7	Debby cried because the window broke

clause SVsSV (Semantic constraints)		Because the window broke, Debby cried NB the window broke because Debby cried*

The scoresheet completed during the session with the child (see Appendix VI) was annotated during and after the session. This enabled analysis of the problem solving or trial and error strategies employed by the child, and could give insights into his linguistic knowledge. For example, if the child attempted to formulate 'the man painting my wall'*, this suggests that his knowledge of auxiliary verbs may be weak. He may offer this as a complete sentence to the assessor, or self -correct on the basis of checking that he has used all the words. Similarly the child may substitute the present tense for past, regular or irregular 'the girl show the monkey'* 'the girl give the boy'*.

Some of the items were selected to assess morphosyntactic structures commonly demonstrated to be underused or inaccurately used in children with SLI at different ages. These include those in Table 2 below:

Table 2. Linguistic Structures assessed by test items

Structure	Example	Item	Reference
Auxiliary verbs	Present progressive 'is painting' future tense, 'is going to'	1, 2, 3, 4	Crystal et al (1976) Loeb and Leonard (1991) **-ing shown to be not impaired in SLI children

			Rice and Wexler (1996)
Modal auxiliary	will hide, can hang	9,10	Leonard (1995)
Auxiliary inversion in Questions	Is X painting? Modals 'Can-?' and 'Will--?'	1, 2, 9, 10	Rice, Wexler and Cleave (1995)
Copula verb	Boy is hungry + negative isn't broken	7, 8	Rice and Wexler (1996), Leonard (1995), Loeb and Leonard (1991)
Copula inversion questions	Is X hungry?	7.8	Rice, Wexler and Cleave (1995)
Dative	Give the monkey the banana	5,6	Bishop (1997) Ebbels (2007) van der Lely and Harris (1990)
Genitive / Possessive pronouns	Bird's, my	8	Leonard (1995)
Determiners	the/my - 'my man*	1	Gopnik 1990, Rice and Wexler (1996) Leonard (1995)
Pronominal-ization	Joe / he you/your	11, 10	Loeb and Leonard (1991) Leonard (1995)

Attention was also given to the child's treatment of a small number of lexical items, for example prepositions, which were identified as an area of difficulty by Bishop, Adams and Rosen (2006), Ebbels (2007), and van der Lely and Harris (1990), but were conversely

not found to be impaired by Watkins and Rice (1991). Responses of the children to sentences containing subordinating and co-ordinating conjunctions were also to be explored.

Some structures reported to be impaired in children with SLI could not be included as no possible example could be formulated in which the same words could be arranged into two possible sentences, for example Wh- questions, and relative clauses. Other structures known to cause difficulty for children with LI were not assessed but were included in the test items, for example the third person singular –s in 'brushes' and 'goes' as well as the regular past tense morphemes on 'showed' and 'cried' were used, as were the irregular past tenses 'gave' and 'broke'. The structure of the task did not lend itself to assessing whether the children were able to generate these structures independently.

A considerable amount of difficulty particularly affecting the grammatical elements identified in the table would suggest a specifically linguistic basis to the child's impairment, which is consistent with domain specific linguistic accounts of SLI. The data elicited by the current study procedure would not, however, be sufficient to provide evidence in support of any particular linguistic theory of language impairment. Nevertheless, predominance of inaccuracy, substitution or struggle behaviour associated with these constructions would signal that the child might belong to the subgroup of grammatical-SLI (van der Lely 2005).

Furthermore, implicit, and sometimes explicit in the task was the need for children to make judgements of grammaticality, which was identified as an area of deficit by Gopnik and Crago (1991) and Wulfeck and Bates (1995 cited by Leonard 1998). In the present study, after the participants had arranged the words into sentences,

they were asked to judge whether the sentences were correct, and this judgement task was frequently facilitated by the examiner repeating the sentence back to the child in order that they could make judgements based on a fluent and appropriately inflected auditory model. van der Lely, Jones and Marshall (2011) recommend grammaticality judgements as a useful methodology for distinguishing impairments in syntactic knowledge from those induced by the load of processing expressive sentence constructions.

Alternatively, the child more obviously affected by the number of items, semantic constraints on sentences, or who fails to transfer learning from one item to the next, may be identified as having more domain general processing problems, or limitations on working memory capacity (Marton and Schwartz 2003). While the intention is not to seek support for theoretical positions, nor to label or categorize children, the information derived from comparing the performance of individuals to documented patterns of performance should contribute to the process of devising appropriate intervention that is directed towards the specific areas of weakness of the individual.

3.2.6 Administration

The test/training is carried out systematically, in the design of Guthke's (1993) Short-Term Learning Potential Test in one, or possibly two sessions if the child has not completed the items due to lengthy training time being required, or due to fatigue. All 12 training items are presented, as they cover a range of grammatical structures that are in themselves a source of information about the child's linguistic knowledge. The tasks are presented in a given sequence, due to the increasing level of grammatical complexity of the items. In addition, Tzuriel (1997) points out that bridging from

concrete to abstract levels of functioning is facilitated by the grading of items with regard to complexity, and 'the mediation given on one item should prepare the child to solve the following item' (p.85).

Items are presented to the child for him to solve independently, and cues are provided only when required to help the child solve the problem item. There are five levels of help available, based on the method employed by Resing (1993) with cues being prescribed in as far as the general type of information and nature of assistance is described, but which are mediated in a flexible and individualised way, dependent on the responses given by the child. Testing is adaptive, in that cues are applied only if and when required (Guthke, Beckmann and Dobat 1997). The cues are graded from general metacognitive direction, or no specific prompt (level 1) to more specific strategy based suggestions (levels 2-3), breaking down the task into components and using specific feedback (level 4) and finally to item specific feedback and instruction (level 5; See Table 3).

In addition, repeated use of the procedure, using parallel forms with identical linguistic structures, enabled the experimenter to elicit small amounts of improvement within a child over time. For example, a child may require less directive prompting to solve an item after intervention, than before, but still not be able to solve the problem independently. Thus small changes in task performance may be measurable via the DA procedure. Similar use of a DA to record progress by means of the amount of scaffolding required was reported by Glaspey and Stoel-Gammon (2007).

Table 3. Prompt Sheet for administration of DA

Cue Level	Description of Cue	Example of Phrasing
1	Metacognitive direction Spontaneous response	Do you know what you have to do?
2	Drawing on previous knowledge	How did you do this before? Do you know all the words? Is that right? Can you fix it?
3	Finding strategies	Which one can you start with? You need to start with something different this time Can you make little groups of words? Can you make a question? Can you swop the words around?
	Problem solving	Have you used all the words? What have you left out? Reminder – 'Last time you said..'
4	Breaking down the task.	Which one shall we start with? Which one can you start with to make a question? Start with..X. What comes next?
	Using specific feedback	You have left this one out – where does it go? Repeat part of answer already used Giving part of answer
5	Learning from feedback and instruction	Scaffolding sentence bit-by-bit Presenting cloze task. Explaining. Identifying errors

		Modelling for imitation
	Reflection - when the answer is correct	Is that the right answer? Why was it not ok? Can you tell me how you did that? How did you know how to do that? Was it easy or hard? Why?

In Resing's training procedure, a 'correct' response also requires the child to appropriately justify his answer, or verbalise his strategy in solving the problem. Similarly, in the current study, when a correct solution is presented, the examiner prompts the child for a reflective response, primarily a judgement of grammaticality and sense, as well as an evaluation by the child of the strategies he used, and the level of difficulty he experienced. This procedure was also described by Gutierrez-Clellen and Peña (2001) as a variation on the 'Testing the Limits' procedure, and found to enable children to better demonstrate their knowledge, and the examiner to better understand the child's thinking and approach to problem solving.

The procedure of prompting the child for a reflective response enabled the investigator to informally assess the child's metacognitive awareness, and use this information to make recommendations for the planning of intervention programmes. The procedure of 'reflect when prompted' (Bannert and Mengelkamp 2008) has been identified in metacognitive research as a valid tool for assessment of metacognitive skills, and has also been shown to be an 'adequate intervention' in metacognitive training (p.46), as it enables participants to activate their metacognitive knowledge and skill. It is intended to facilitate improved generalization as recommended by Peña, Resendiz and Gillam (2007, p.335) and to provide opportunity for further metacognitive mediation, which promotes transfer of gains (Keane 1987). Thus the use of mediated prompting after level 5 of the current procedure, while not scored or

included in the Graduated Prompt format, was considered part of the mediational intervention of the DA.

3.2.7 Mediation

The style of interaction adopted by the examiner should be mediational, as recommended by Feuerstein, and described by Lidz (1991). Gains and transfer performance following mediational intervention have been shown to be greater than those following graduated prompting and other methods (Vye et al 1987, Keane 1987, Swanson and Lussier 2001, Grigorenko and Sternberg 1998). Vye et al (1987) structured their Dynamic Assessment sequentially, offering Graduated Prompting to those children who perform poorly on standardised testing, and mediation to those who perform below criterion following the Graduated Prompting. The current study, however, assumed that children already identified as having language impairments are unlikely to benefit from prompting alone, and thus the individualised mediation was included from the outset.

The session should incorporate essential mediational components, in that there should be clear transmission of intentionality and meaning to the child, mediation of competence and task-regulation throughout, and transcendence or bridging of the skills to other applications. Mediation of any component during the training procedure is allowed, but should not detract from the essential sequence of the cues, or distract the child or examiner from the task. As stated by Lidz (2003 p.121) other MLE components are assumed to automatically be included by the examiner. Mediation of metacognitive strategies such as planning, self-regulation, and checking, that are domain general and not limited to the task at hand, have been reported to facilitate generalization. Nevertheless, there is a need to be able to administer and score the procedure objectively, therefore a need to operationalize mediation.

3.2.8 Scoring of the procedure

As there were two responses required for each item, scoring of the number of cues required to facilitate a response, was based on a total of 24 responses.

The procedure was scored in several ways, each of which will be elaborated below:

Quantitative measures were used to measure change within an individual over time.

i)   Total number of cues (max 24 items x 5 cues = 120).

ii)   Number of each type of cue required.

Qualitative measures were used to inform intervention in the experimental intervention phases of the procedure:

i)   Identification of grammatical structures which caused the child the greatest difficulty

ii)   The effect of amount of content (sentence length) and nature of semantic content on the child's construction of linguistic structures

iii)   The child's ability to transfer, or generalise learning or strategies ie. item - to item transfer, as well as which items benefited from transfer effects

iv)   The child's *metalinguistic* knowledge, ability to label, explain and manipulate linguistic concepts

v)   The child's *metacognitive* ability i.e. awareness of the processes and strategies that are used to solve the given task

Recording:

The number of cues presented was recorded on a scoresheet (see
Appendix VI) as the procedure was carried out. The number of each
type of cue could be totalled immediately after completion of the
assessment session. The Response to Mediation Scale (Lidz 2003)
was completed within a few minutes of the end of the session. In
addition, all sessions were videotaped in order to verify the count of
cues. The video was used to check ratings on the Response to
Mediation scale, and for verification by an independent rater.

Quantitative measures:

i) The number of cues, that was required to facilitate each response
to an item, out of a maximum of 5 as described previously, was
recorded during the procedure. As the examiner presented
successive prompt questions, she was able to check the next box on
the scoring checklist. The number of prompts reflects the amount of
help required, and indicates the degree of difficulty the child has with
the task. This type of score has been considered to be the inverse of
the learning potential (Resing 1993 p221). The number of prompts
required differentiates between individuals, and across test times
within an individual.

ii) The number of each type of cue required across the 24 response
procedure is added up at the end of the test. It gives composite
information about the learning needs of the individual throughout
the whole procedure, elucidating whether he requires mostly
strategy training or item specific application of knowledge.

Qualitative measures:

i)   Identification of grammatical structures which caused the child
the greatest difficulty, was carried out by considering the scores and
extracting those items that required the greatest amount of
prompting, i.e. level 5 cues. This was supported by the examiner's

perception of the child's difficulty in grasping particular linguistic structures, which was noted down at the time of testing by circling the structure on the scoresheet.

ii)   The effect of amount of content (sentence length) and nature of semantic content (reversibility) on the child's construction of linguistic structures was extracted from the scoresheet at the time of writing of the child's report, and noted in simple terms only as whether length and reversibility affected the child or not.

iii) Instances of the child's ability to transfer, or generalise learning or strategies i.e. item - to item transfer, were noted on the scoresheet by the examiner at the time of testing. Both apparent use of a structure or strategy that had been supported earlier in the session, and those elicited by questioning the child were included. These gave rise to qualitative assessment of transfer ability that was commented upon in each child's DA report.

iv) The child's *metalinguistic* knowledge, ability to label, explain and manipulate linguistic concepts was described with reference to specific linguistic structures that were probed. For example children were asked to identify or formulate questions, find the verb (or 'doing word' or 'action'), person, 'describing word' etc. in a sentence, and responses to these were recorded as metalinguistic knowledge. The items probed varied from one child to the next, according to the responses made by the child to the task items, and as a consequence, the content reported about each child varied.

v)   The content of the report of each child's *metacognitive* ability i.e. awareness of the processes and strategies that are used to solve the given task was again individually determined as a result of individualized probing and questioning carried out during the administration of the DA. Reflective responses elicited after correct responses were given and at the end of the test, also contributed information to each child's report.

Rating of child's response to mediation using Lidz's Response to Mediation Scale (2003) was employed in order to give information about personality, individual responsiveness, and qualitative behavioural and interpersonal factors that will enable a clinician to accommodate a child effectively in therapy.

3.3 <u>Reliability and Validity of the Test</u>

Content and face validity were established by the adoption of a task structure from a standardised language test (Sentence Assembly subtest of CELF-3 UK), and interposing intervention that is apparently directed at improving task performance, and has measurable features of mediated intervention as described by Lidz. Criterion validity, however, cannot be entirely established by comparison with other tests because the DA procedure has a different aim from standardised tests, setting out, as it does, to measure learning potential, and only in part, achieved learning. Nevertheless, correlation between DA scores and those obtained from a standardised test such as the CELF would suggest that the DA does tap into the basic language abilities of the child. Similarly, comparison to other dynamic measures, such as cognitive tasks or scales from the LPAD, would not be useful as specifically language based tasks are required to capture the specific weaknesses of children identified as having language impairments. No appropriate measures of language learning potential were identified, that would be useful as measures of concurrent validity.

The majority of standardised language tests have a diagnostic function, as indeed do the DA of language procedures published by Peña (Peña and Iglesias 1992; Peña 2000; Peña Iglesias and Lidz 2001, Peña et al 2006 and Peña, Resendiz and Gillam 2007). In order to establish diagnostic value, the procedure would have to be applied to a population of typically developing children, or those with

difficulties of a different type. The use made of the DA in this instance, however, was to differentiate within the group of children all labelled as 'Language Impaired' and highlight differences in learning potential within this category, known to be, and frequently described as 'heterogeneous'. As a result, the use of an additional population for diagnostic purposes was unjustified.

Predictive validity is a criterion highly applicable to tests of learning potential. The responsiveness of a child to prompting and mediation should be predictive of the rate at which they can progress in the given skill area. According to Kaniel (2009) DA instruments do not have sufficient reliability and validity demonstrated, and the effectiveness of Dynamic Assessments is best established by their predictive validity.

Many children with identified LI, have the benefit of individualised or small group intervention programmes delivered by speech and language therapists, and thus this is a format through which progress can be monitored. The difficulty however, lies in the variability of the content and means of delivery of that intervention, as well as the means of measuring the changes. In other words, progress of individual children is not comparable as they do not have the same or necessarily equivalent opportunities to develop particular skills. Furthermore, their progress is traditionally measured via standardised language tests that may or may not capture the changes achieved by the child.

'When no obvious criterion is available for validation' (as in the current test) 'we may resort to construct validity' (Pring 2005 p.178). The current procedure is based on theoretical constructs shown to be useful in other fields, and adapted to the current task. Practical use by SLTs over time will establish the usefulness of the procedure to

inform intervention that is consistent with the theoretical basis of the test design.

Establishing reliability of the procedure is problematic, primarily because the procedure is not a static or standardised test in which results should be replicable over time or over repeated administration. The intention of the procedure is to mediate and improve the performance of the individual during the test. Thus split half reliability is inappropriate as learning is intended to take place from item to item, and transfer of learning to a subsequent item is a desirable feature. Similarly, retesting using a parallel form would also be confounded by the intention to induce change as a result of the procedure. Inter-rater reliability, however, is essential to establish unequivocal evaluation of the level of prompting required to elicit correct responses in the child, especially as some variation in the content of prompting is permitted.

3.4 Summary

In summary, the methodology for Dynamic Assessment that has been devised is a hybrid of a number of established methods, drawing on the strengths of each, to serve the purpose of diagnostic assessment of expressive language in children with specific language impairments. The procedure has been condensed into a single session that is easily scored as the assessment proceeds, but yields additional qualitative data from systematic analysis of a transcription of the session. The materials are easily accessible and themselves give important information about the child's linguistic ability, as well as their potential to learn and benefit from intervention.

CHAPTER 4

DYNAMIC ASSESSMENT:

EVALUATION OF THE TEST

CHAPTER 4   DYNAMIC ASSESSMENT: EVALUATION OF THE TEST

4.1 <u>Study Design</u>

The Dynamic Assessment of language was administered to a group of children with previously identified language impairments. For clinical purposes, only one administration of the test is required to inform intervention, although the same or alternative versions of the test may be used later to measure change in performance, so one test trial was used to evaluate sensitivity of the test to individual differences, and a further trial to detect sensitivity to change over time. Two further trials were linked to the investigation of the role of the DA in informing intervention (see chapter 6), and the data were used to investigate predictive validity. The order of administration of the different versions of the test was randomised across time of administration and different participants.

The nature of the DA test procedure required that all the DA testing was carried out by the investigator. In order to control for experimenter bias, the assessments were videotaped, and scored by an independent assessor (see section 4.4).

4.2 <u>Participants</u>

4.2.1 Procedure for recruitment

Before identification and recruitment of participants could be undertaken, the study was submitted for ethical approval by the National Research Ethics Service. Approval was gained from Redbridge and Waltham Forest Local Research Ethics Committee in May 2008, and identification of potential participants began immediately thereafter.

Speech and Language Therapy managers from several PCTs across London were approached by letter. Speech and Language Therapists known to have previously expressed an interest in Dynamic Assessment were personally approached, and managers and SLTs from three trusts and an independent school agreed to take part in the research and refer children from their caseloads. At the outset, the requirement that SLTs be willing to continue ongoing intervention programmes that would be monitored, and to collaborate in redesigning intervention based on the outcomes of DA, was specified, and therapists agreed to these conditions themselves, before identifying potential child participants.

SLTs were asked to identify children who met the following criteria:
- Aged 8-10 years old, and in Years 4 and 5 of school.
- Known to Speech and Language therapy services or Language units/resource bases, or special school.
- Language disorders identified as the primary disorder, likely to score <1SD on a standardised language test.
- English as a first language.

In addition, SLTs were asked to exclude children with:
- Moderate or severe learning difficulties, (IQ<70).
- Hearing impairments.
- Attention deficit disorders.

35 children were initially identified by SLTs who also sought permission from the parents for their names to be passed to the researcher, or who undertook to obtain informed consent themselves. Parents of all the participants were supplied with an information sheet (see Appendix VII) which had been approved by the Research Ethics Committee, and a similarly approved consent form. All parents were offered the opportunity to contact the researcher

directly to ask any questions, but none of them took up this opportunity, and 31 consent forms were returned to the researcher via the schools, or the self addressed envelopes provided with forms.

Permission to conduct the research at schools was obtained from the headteachers of each of the seven schools involved in the study.

Testing of potential participants to confirm eligibility for inclusion in the study began at the start of the school term in September 2008.

4.2.2 Measures to confirm eligibility

Although criteria for referral were fairly specific, children were screened to confirm eligibility for inclusion in the study. The following measures were employed:

1. The Coloured Progressive Matrices (Raven, Court and Raven 1990) was used as a measure of non-verbal intellectual ability. Children were considered eligible for inclusion in the current study if they scored above the 25th percentile, signifying 'intellectually average' (Raven, Court and Raven 1990 p.CPM30). This ensured that the language difficulties of participants in the study were not secondary to global learning difficulties. Although those with global learning difficulties can be effectively assessed by DA methods, the current study aimed to assess the language abilities rather than the cognitive abilities of children with specific impairments of language.

2. The Strengths and Difficulties Questionnaire (SDQ) (Goodman 1997) was used to detect features of ADHD. Questionnaires were distributed to class teachers or SLTs who were familiar with the children. Children scoring more than 7 on the items pertaining to hyperactivity were excluded on the basis that attention deficit altered the mediational needs of the children. Pilot testing

showed children with ADHD requiring a higher proportion of mediations addressing behavioural control (Hasson and Botting 2010).

3. CELF-3 (UK) (Semel, Wiig and Secord 1987) to confirm the presence of LI, by a score of 1.25 SD below the norm for CA.

Four children were excluded from the cohort on the basis of the screening tests. Two of these scored below 25th percentile on the Ravens, one scored within normal limits on the CELF-3, and one was functioning at too low a level to be able to complete any of the formal tests. A further child was excluded as injury resulted in absence from school during the testing period, and another after the first phase of testing as prolonged illness resulted in intermittent school attendance.

One further child was thought to be eligible on the basis of the screening tests, but was subsequently excluded on account of behavioural difficulties that prevented experimental tasks being completed.

4.2.3 Characteristics of the sample

The final sample consisted of 24 children, aged between 8;2 and 10;9 at the time of the first test. The mean age of the whole group was 9;3. There were 21 boys, with an age range of 8;2 to 10;9, and 3 girls with ages ranging from 8;2 to 10;7.

Six children were drawn from a special school for children with speech, language or communication needs, in South London. The remainder were referred from six language units or resource bases, two of which were from an Inner London borough, accounting for six

children, and the remaining thirteen were from across East and West Hertfordshire.

A summary of the characteristics of the sample is shown in Table 4.

Table 4. Characteristics of Participants

Child No	Age at T1	Gender	School	CELF Std Score	Ravens CPM %ile	SDQ-H	SDQ-Total
1	8,2	Female	BS	64	90	2	4
2	8,5	Male	BS	64	25-50	3	16
3	8;1	Male	BI	64	25	7	8
4	8;4	Male	BS	64	75	2	3
5	8;6	Male	BI	64	95	5	15
6	8;9	Male	BS	64	75-90	3	6
7	8,10	Male	BS	64	95	0	6
8	9;1	Male	BS	65	25-50	2	3
9	9;1	Male	SS	65	50	4	9
10	9;1	Male	BS	65	95	1	10
11	9;2	Male	BS	65	75-90	4	18
12	9;3	Male	SS	76	75	7	19
13	9;3	Male	BI	65	75	3	17
14	9;3	Male	BS	65	75-90	5	9
15	9;7	Female	SS	65	25-50	2	9
16	9;7	Male	SS	74	50	5	8
17	9;8	Male	SS	65	50-75	4	14
18	9;8	Male	BS	80	90-95	4	6
19	9;10	Male	BI	71	50	6	22
20	10;0	Male	BI	64	25-50	5	5
21	10;0	Male	SS	64	90	6	12
22	10,3	Male	BS	64	25-50	3	6

| 23 | 10,7 | Female | BS | 64 | 50-75 | 2 | 4 |
| 24 | 10;9 | Male | BS | 64 | 75-90 | 0 | 6 |

Key to schools: SS=Special School;

BS=Language Base, Suburban; BI=Language Base, Inner London

Standard scores on the CELF-3(UK) showed 20 out of the 24 children to be at least 2.5 SD below the mean for age, scoring below the 1st percentile. Within that group which is indistinguishable by their standard scores, raw scores ranged from 46-93 in children aged 9-10, and from 28-87 in the children aged 8 at Time 1, suggesting a considerable difference in ability.

Evaluation on the Ravens CPM, revealed one child on the 25th percentile, five between 25 and 50th percentiles, as well as six children over the 90th percentile, suggesting a considerable range in non-verbal reasoning skills and differing profiles of verbal:non-verbal skills within the cohort of participants.

4.3 Procedure

4.3.1 Administration

Appointments were made for all children to be assessed over two sessions, usually one week apart. In the first, the Raven's Coloured Progressive Matrices was carried out, followed by subtests of the CELF- 3, for the purposes of confirming eligibility as well as baseline performance. In most of the children, all six subtests of the CELF were not completed, and the remaining subtests were carried out in the second session. In the intervening week, participating Speech and Language Therapists or the teachers of the children were asked to complete the SDQ, and teachers also completed a behavioural questionnaire. The DA procedure was carried out in the second testing session, after the CELF-3 was completed, and this component

was videotaped. The DA took between 25 and 40 minutes, and in all children was completed in one session.

The DA was carried out as determined in section 3.2.7. One version of the materials was selected at random. All 12 items were presented, in sequence, with two responses required for each. Prompting was given as required, according to the predetermined hierarchy of cues.

4.3.2 Mediation

At the start, children were asked 'Do you know what you have to do?' and 'Can you tell me what you have to do'. The responses were used to evaluate metacognitive awareness of task requirements.

The participants were then allowed time to respond to each item without prompting from the assessor. Incomplete attempts were supported with 'Yes', 'Go on' or repeats of parts of sentences that the children produced. When participants indicated difficulty verbally or non-verbally, the prompts were introduced sequentially. Wording of the prompts was variable, and mediational in nature, allowing participants to find their own solutions as far as possible, and to make judgements or justify their attempts. Intentions of the prompting were identifiable as use of previous knowledge or experience (prompt level 2), searches for strategies (prompt level 3), simplification of the stimulus (level 4), or feedback and instruction (level 5).

When the items were completed, participants were asked reflective questions, such as 'How did you know how to do that?' 'Was it easy or hard?' 'What was hard?' or 'What made that one easy?' These responses were not scored, but used to evaluate metacognitive awareness. In addition, some children were asked questions relating

to linguistic structures, in order to identify metalinguistic knowledge that could inform intervention, or for the assessor to gain insights into the problem solving and grammatical awareness of the children. For example, some participants were asked 'Which is the person/doing word in this sentence?' and several were asked what was meant by the possessive apostrophe ('s), in one item, to which several replied that it signified a plural.

4.3.3 Scoring

Cue levels were recorded on a prepared scoresheet (see Appendix VI) as the assessment progressed. Although the scoresheet was kept out of the line of vision of the children, it was not hidden and if they asked about their scores, it was explained to the children that although they had already got the answer 'right', the numbers represented whether they were given 'clues' by the assessor, or solved the problem 'all on their own'. Thus children were encouraged by scores of 2 and 3, being shown that they only needed a little clue and then they were able to arrange the sentence. They were also shown evidence of transfer, the assessor saying, 'Look I helped you a bit here, but you've remembered what to do and done it alone here'.

Total scores and number of cues at each level were totalled after the session was complete and the children had left the room. Any uncertainties were checked by subsequently reviewing the video.

4.4 **Inter-rater reliability**

An independent rater was given approximately one hour's explanation of the cue rating system used in the project, and practice using examples taken from the recording of one child. Criteria used for cue levels were explained and discussed. She was then asked to rate a sample of the tapes for practice and to measure

inter-rater correspondence with the tester. A sample of 25% of the tapes at Time 1 were used to establish consistency of definition and to identify any ambiguities in the scoring criteria. Following Time 3, the independent assessor scored the DAs of all of the participants (100%). The assessor was blind to the group allocation of the participants, which is relevant to the investigation of DA in Intervention study (see Chapter 6).

4.5   **Response to Mediation Scale**

The Response to Mediation Scale (Lidz 2003) was completed immediately after the Dynamic Assessment session. Eleven criteria were each rated on a scale of 1-5, according to qualitative descriptors (see Appendix III). Ratings captured general behavioural factors, responses to the material and to the assessor during the dynamic testing, and problem solving skills. This information supplemented that gained through the DA, and was used in conjunction with information from the DA in the reports supplied to SLTs to inform intervention.

The full sample of RtM rating scales for all participants at Time 3 were rated by an independent rater, blind to group allocation, from the videos of the DA sessions.

4.6 **Data Analysis**

The total number of cues required for the 24 sentences (two sentences required from each set of words on a cue card) was totalled and used as the participant's DA score in statistical analysis. The range of scores from the full cohort was described statistically and compared to the scores obtained from a standardised measure of language, the CELF-3(UK) to estimate concurrent validity with regard to the DA as a measure of language ability. Similarly, scores were correlated with scores on the Response to Mediation scale (Lidz

2003) to establish criterion validity of the DA as a test of responsiveness to mediation, although the RtM is not a standardised measure.

Subsequent retest scores, after a 4-month period of language therapy (Time 2) were used to evaluate reliability as well as sensitivity to change over time, as change in a positive direction would be expected in the children after periods of intervention. The scores from the initial test (T1) were also correlated with the changes in scores on the standardised CELF-3 over time, after two or three periods of intervention, to determine the predictive validity of the DA.

Inter-rater reliability of the DA instrument was evaluated, and internal consistency of the test items was examined, along with qualitative examination of the response to particular items. Equivalence of the four versions of the test used at different times was also established.

4.7 Results of the experiment to evaluate the DA procedure

4.7.1 DA Scores at Time 1

The scores of the whole cohort at T1 were widely distributed, with a mean score of 61.83 and standard deviation of 20.72. The distribution is shown in Figure 1.

Figure 1. Distribution of DA scores at Time 1.
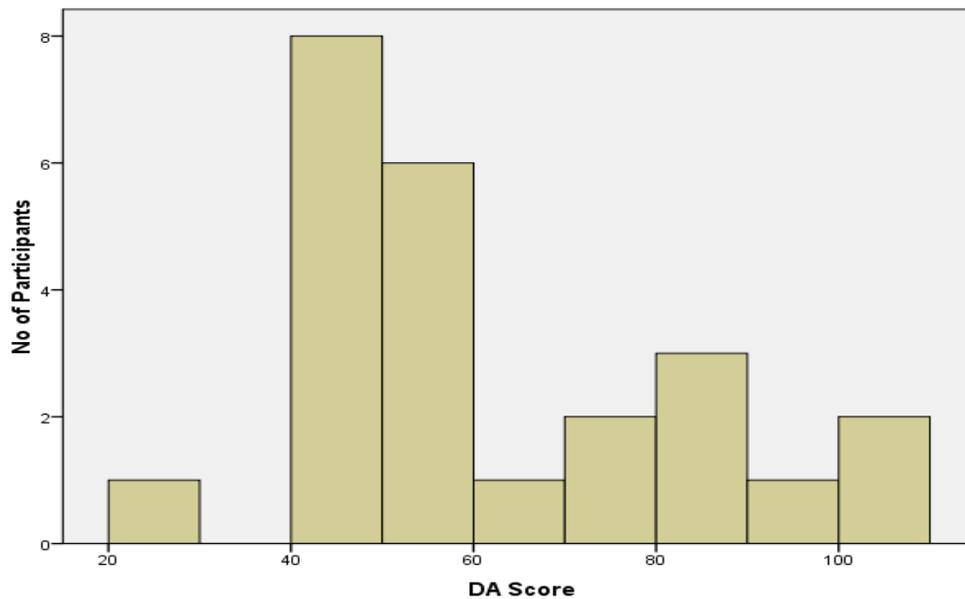
DA Score. Range Min 24-Max 120

4.7.2 DA at T1 in comparison to CELF

A significant correlation was found between participants' scores on the DA and total raw scores on the CELF-3 (UK) at Time 1 (r_s = -.481, p = .017) suggesting a relationship between performance on a standardised test of language and the test of learning potential. The correlation is in a negative direction as favourable performance on the CELF obtains a greater score, while stronger performance on the DA is shown by need for fewer prompt cues and hence a lower score. The moderate correlation observed would be expected as the CELF-3 links only with the language knowledge component of the DA.

Figure 2. Correlation between DA scores and CELF-3 Total raw scores at Time 1
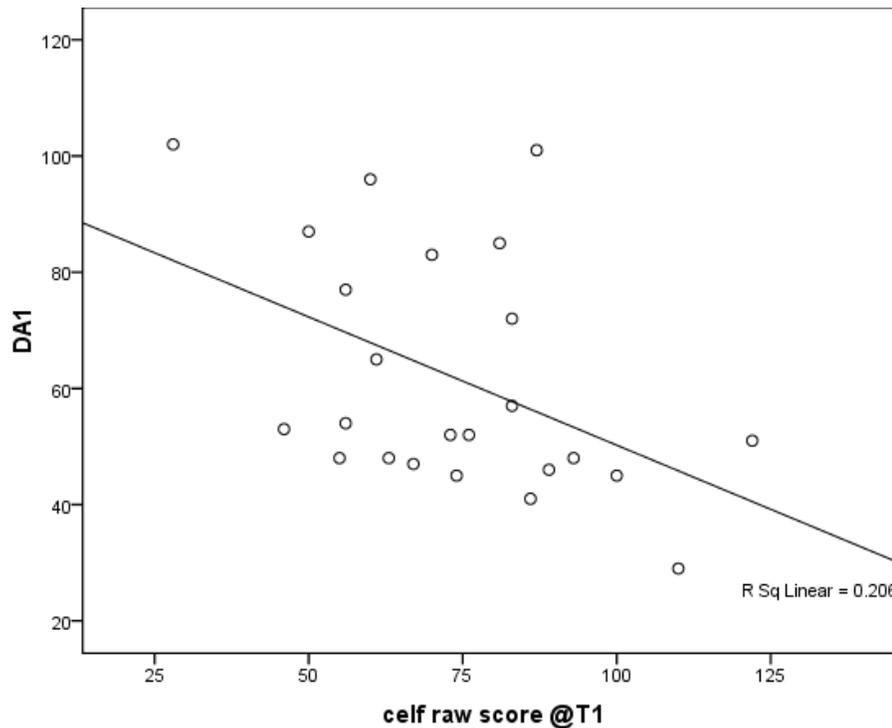
4.7.3 DA at T1 in comparison to Response to Mediation (RtM)

Correlation between ratings on the RtM scale (Lidz 2003) and the DA at Time 1 were just short of significance (r_s = .401, p = .052) and although this result is borderline, it shows a moderate association. The Response to Mediation links only to the strategic learning and responsiveness aspects of the DA and again only a moderate correlation would therefore be expected.

4.7.4   DA scores at Time 2

Table 5 shows the shift in scores from Time 1 to Time 2 in the DA task. The sample mean has decreased from 61 to 52, reflecting the overall improvement in performance of the group, with similar variance.

Table 5. DA at Time 1 and Time 2

	No of participants	Mean	Std deviation
DA scores at Time 1	24	61.8	20.7
DA scores at Time 2	24	52.2	19.0

The Wilcoxon test shows that DA scores at T1 and T2 are significantly different, (Z = -2.786, p = .005), but a highly significant correlation was found between the scores at T1 and T2 (r_s = .706, p < .001), reflecting reliability of the procedure and sensitivity to change over time.

Figure 3. Correlation between DA scores at Time 1 and Time 2
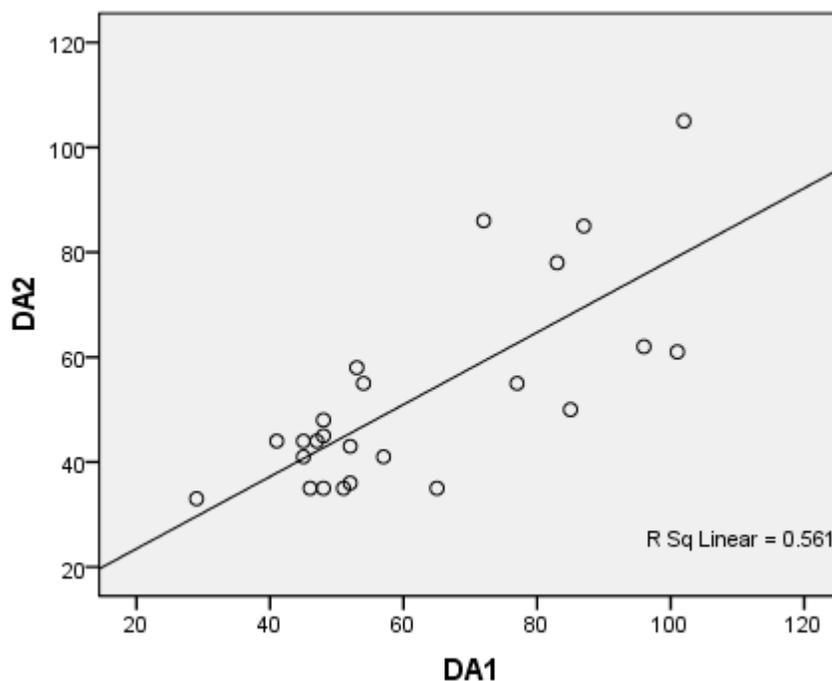
4.7.5   DA at T2 in comparison to CELF

A significant correlation was found between the scores obtained by participants on the DA at T2 and the raw scores on the CELF-3 (UK)

at Time 2, (r_s = -.705, p = < .001) suggesting a mapping of gains on the standardised test of language with those on the test of learning potential.

4.7.6 Inter-rater Reliability Measurement

An independent rater was supplied with 6 randomly selected videos of the DAs from the cohort at Time 1. Each item in the test was rated on a scale of 1-5 pertaining to the level of assistance given (as described in section 3.2.7). A significant correlation was found between the total scores for each participant (n=6) for each rater (r_s = .886, p = .019).

Scores for each item of the DA test procedure (n=24) were compared for the 6 participants at T1. Highly significant correlation was found (r_s = .805, p < .001).

Scores for all participants at Time 3 were rated independently by the same rater, blind to the group allocation of participants for intervention (see Chapter 6). Correlation between ratings by the experimenter and blind rater are shown in Figure 4.

Excluding the 2 outlying scores, shown in red in the upper right quadrant of Figure 4 significant correlation was found between ratings by the experimenter and independent rater (r_s = .874, p< .001). The two outliers represent two participants who continued to require large numbers of prompt cues (89 and 99) on the DA at Time 3, all other participants having improved to DA scores between 29 and 65. Both the experimenter and independent rater identified these exceptionally high scores, and inclusion in the correlation improved the correlation to r_s = .902. However as the

outlying figures may have affected the correlation it was thought to be safer to exclude them in the calculation of the correlation.

Figure 4. Inter-rater agreement for all participants at Time 3.
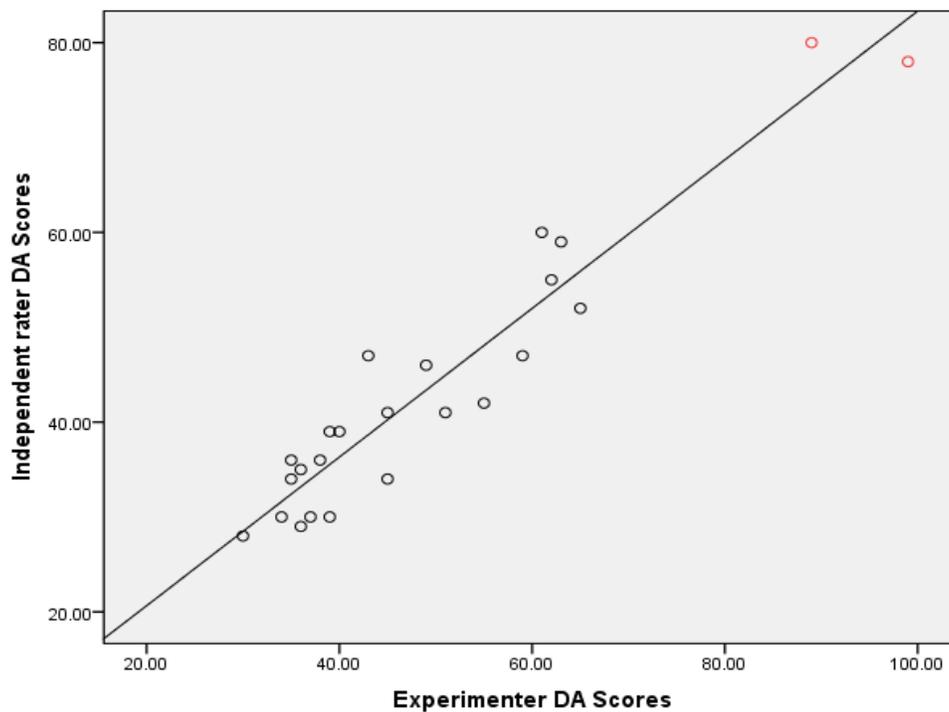
Inter-rater correlation was also calculated for the ratings on the Response to Mediation Scale (Lidz 2003). No prior discussion or training was carried out, the independent rater filled out the RtM rating form for all participants at Time 3, based on videos of the DA session. Correlation between the experimenter and the independent rater was more moderate (r_s = .570, p = .004) and subsequent findings will be considered in light of this. The correlation is represented in Figure 5.

Figure 5. Inter-rater correlation for Ratings on RtM
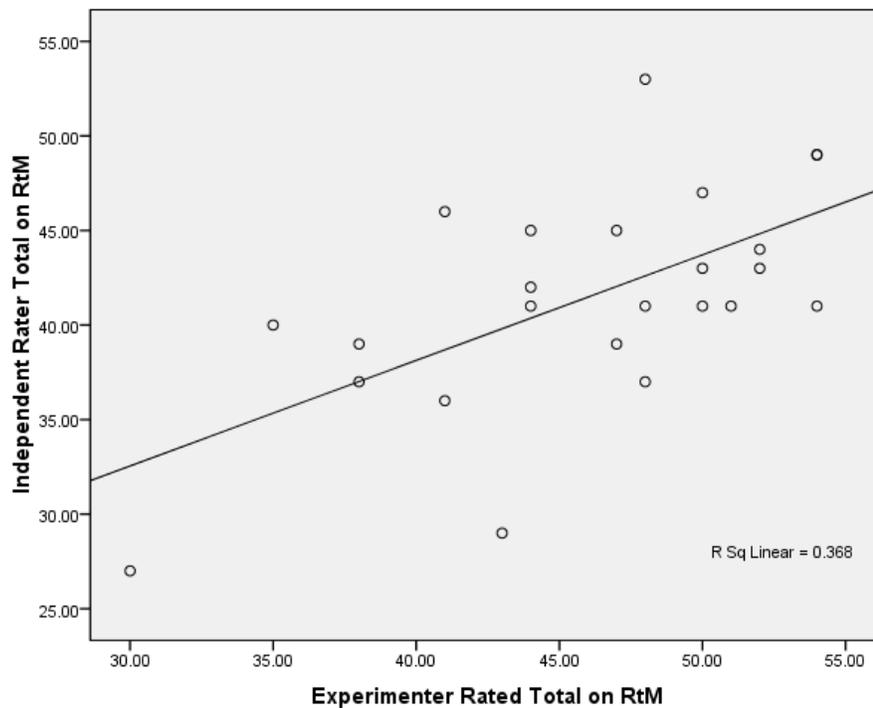
4.7.7 DA as a predictor of gains in therapy over time

Dynamic Assessments have been used to increase the predictive validity of assessments by identifying the potential of an individual to benefit from intervention. In the current paradigm, the number of prompt cues required in the DA procedure is regarded as a measure of the individual's ZPD, or their learning potential. Given that there was opportunity for all participants to continue in prescribed and individualised intervention for the duration of the study, the DA could be related to gains made during that intervention.

In comparing the score obtained in the DA at Time 1 with the progress made over the study period, it is useful to make use of the percentile ranks of the CELF-3 scores, which like the standard score totals are normative and therefore adjusted for age over the study period, but which represent a measurable shift in achievement

relative to the normal population. The changes in standard scores were not found to be informative as almost all of the participants at every time of testing achieved the two lowest standard score bands, and a change in standards cores from 65 to 71 was not thought to be meaningful. The scores of children in whom there was any change in percentile rank, are recorded in Table 6.

Table 6. Change in Score on CELF-3 over Time

			T 1		T 2		T 3		T 4	
Child Name	Age at T1	DA T1	Raw score	%ile	Raw score	%ile	Raw score	%ile	Raw score	%ile
D2	10.3	41	86	**1**	104	**2**	113	**4**		
CH1	9;1	45	74	**1**	91	**1**	100	**4**	122	**12**
D5	10,7	48	93	**1**	100	**2**	120	**2**		
R4	9;3	48	63	**1**	96	**5**	117	**19**	111	**2**
CP1	9,2	52	76	**1**	98	**4**	109	**5**	132	**10**
BH7	9;8	54	56	**1**	73	**1**	98	**1**	108	**2**
TA2	10;0	57	83	**1**	95	**2**	103	**2**	138	**13**
TF5	9;11	46	89	**2**	103	**2**	111	**2**	135	**10**
BH4	9;7	45	100	**4**	126	**25**	121	**4**		
BH3	9;3	51	122	**5**	144	**25**	141	**23**		
R2	9;8	29	110	**9**	114	**4**	121	**5**	127	**14**

The data show that eleven children improved sufficiently to alter their percentile ranking on the CELF, by Time 3 or Time 4 of the study period. All of these children had scores in the lower range, i.e. below 60 on the DA at Time 1. Examination of the raw data for all participants shows that all except two of the participants with low DA scores (below 60), which signalled good learning potential improved by an increase in raw score of more than 20 points. These two children were R2, shown in Table 6 as having scored in the 9th

percentile at the start of the study, and child TF4 whose scores fell after he turned 9 and completed different subtests. Despite the gains, four of the children remained in the 1st percentile, having had very low scores on the CELF at T1.

No children who had higher range DA scores, i.e. poorer learning potential, improved from their lowest percentile ranking. Of these children (n=9), two showed raw score gains over 20 raw score points, the remainder improved to a lesser degree, according to their raw scores. This included four further children who were aged 8 at the start of the study, whose scores improved and then subsequently fell.

The analysis by change in percentile rank was therefore thought to be an informative reflection of the correlation between low DA scores and gains on the CELF relative to the normative population. The analysis was less reliable for the children who were aged 8 at the start of the study whose raw scores were low relative to the expectation for age, and for whom the change in subtests selection affected their overall gain scores. All except one of these children also scored poorly on the DA at Time 1, although their were substantial improvements in their performance on the DA on later retests.

Children with greater learning potential as shown by low DA scores were able to benefit substantially in relation to the normal population, from the learning inherent in the DA procedure and from the subsequent intervention. At T4, 16 participants remained in the study, and all those with low DA scores, showed gains in the CELF-3, suggesting that all those obtaining low DA scores at T1, eventually confirm good learning potential.

Eight children were lost to the study at Time 4, for various reasons. Three were in Year 6 at the start of the study, and therefore changed schools, two were from one of the language resource bases that was no longer able to accommodate the study, and three children were from the experimental group in the special school for children with communication needs, who no longer wanted to participate in the study, but who allowed the experimenter to retest the children in the control group in order to increase the data. The T1-T3 gains made by the children who were lost to the study were compared to the remaining 16 participants and the difference was found to be non-significant (t(22) = .444, p = .661) suggesting that the findings of the remaining participants in the study at T4 were representative of the performance of the whole cohort.

The DA score, however is at least in part attributable to language ability, so children scoring high on the DA tended to be those with more severe language difficulties. Children with more severe linguistic difficulties might also be assumed to be less likely to show measurable improvement on retest of the CELF. However, scores of 7 of the 11 children who did improve, were in the 1st percentile on the CELF at T1, indicating severe language difficulties, and the CELF would not on inspection, have predicted the improvement of these children. The low DA scores, however, suggested good potential for improvement.

A significant correlation was found between the DA at T1 scores and the change in total raw score on the CELF-3 from T1-T4 (r_s = -.534, p = .033), but not between the CELF at T1 scores with the T1-T4 change scores (r_s = .105, p = .7). Non-significant correlations were found between DA1 scores and CELF change scores at T2 (r_s = -.267, p = .207), and T3 (r_s = -.273, p = .198) but the longer term outcome was predicted. Correlations between CELF

scores at T1 and CELF change scores at T2 and T3 were also non-significant, T2 change ($r_s = -.034$, $p = .874$), T3 change ($r_s = -.215$, $p = .312$),

In order to tease out the predictive components of the DA, the scores of all participants were adjusted by subtracting the number of ratings of 1 in the DA. In other words, items in which the children spontaneously achieved the target sentence were removed, and remaining scores reflected only the amount of cueing received. Comparing this measure to the change scores on the CELF improved the correlation slightly to ($r_s = -.588$, $p = .017$), although correlations with changes at T2 and T3 were still non-significant. Although the difference is small, this manipulation confirmed the finding that it is the need for prompting and cueing that predicts future attainment in language learning, rather than the score for achieved language, such as that measured by the CELF-3.

In the children scoring poorly on the DA, with scores 60 and above, the gains in CELF scores at T2 showed improvement in 4 out of 9 of the participants. Two further children moved from the 8 year old tests to the 9 year old tests, and their raw scores as a result went down. The scores of these 2 children subsequently recovered and gains were shown at Time 3 or Time 4. A further participant began to show improvement at Time 3, but two participants did not show more than a few points of improvement on the CELF in any time period, and it would seem that the task remained very difficult for them, and they were not able to benefit from the amount of mediation given (see case studies in section 7.5).

This analysis is complicated, however by the factor of age. At the start of the study, seven children were 8 years of age, and six of these scored poorly (above 60) on the DA. The trajectories of

improvement as measured by the CELF-3(UK) were complicated by
the switch from subtests designed for 6-8 year olds, to those
designated for 9+ years, which affected the raw scores obtained by
the subjects, and obscured the measurement of gain in raw scores.
The use of standard scores would be useful, but not sensitive enough
at the floor levels of the test to detect any changes as even raw
score gains of 10 -20 points were insufficient to alter the standard
scores or percentile ranks. The progress of children aged 8-9 years,
is explored in more detail in the Discussion in section 8.1.2.2.

4.7.8 Internal consistency of the test

Internal consistency was recommended as a measure of reliability by
Swanson (1995) for procedures in which test-retest reliability is
variable due to the expected change in retest score, such as in DA.
Internal reliability of the DA was tested using responses of all
participants to the first item in each pair, i.e. the number of cues
required to arrange one sentence from the given words. This was in
order to assess the consistency of difficulty of the grammatical
structure of items. The second sentence produced was more
dependent on metalinguistic knowledge and strategic problem
solving. The responses of 24 participants at T1 to 12 sentence items
produced an internal reliability Cronbach's α=.833

4.7.8.1 Analysis by Item

The number of cues required by the whole group of children to
achieve the correct response to each item, was totalled, as shown in
Table 7. This illustrates the degree of difficulty experienced by the
cohort as a whole, in response to each sentence structure. In
general, the degree of difficulty increased throughout the test, as
intended by the developmental sequence of the structures included.
Certain structures, however, were observed to cause particular
difficulty for the children, and this is reflected in the scores. For

example, dative sentences containing both Direct and Indirect Objects (Items 5 and 6) were difficult, although many children observed that 'there is no "to"', suggesting that they might have succeeded better on the prepositional format.

Table 7. Total number of prompts required to elicit sentences according to grammatical structure, and over Time.

Item No	Content Structure	No of prompts T1	T2
1a	Declarative with auxiliary plus main verb + PossPron NP	31	32
1b	Interrogative with aux reversal	62	48
2a	Declarative with auxiliary plus main verb + future tense VP	67	38
2b	Interrogative with aux reversal	62	41
3a	Declarative with coordination SVcSV	48	41
3b	Reversible content	35	31
4a	Declarative with coordination SVOcSVO	49	49
4b	Reversible content with some semantic constraints	33	29
5a	Dative SVOdOi NP in Subject position	81	82
5b	Reversible content	71	47
6a	Dative SVOiOd NP in Obj position	86	88
6b	Reversible content	65	47
7a	Declarative with copula verb and Adj NP	44	33
7b	Interrogative with verb reversal	56	44
8a	Declarative with copula verb, contracted neg + possessive NP	61	53

8b	Interrogative with verb reversal	67	40
9a	Declarative with modal auxiliary plus main verb and prep phrase SVPrepP	51	43
9b	Interrogative with aux reversal	51	37
10a	Declarative with modal auxiliary plus main verb, Object, and prep phrase SVOPrepP	90	68
10b	Interrogative with aux reversal	67	58
11a	Declarative with (temporal) subordinate clause SVOsSVO (Use of anaphoric pronoun)	72	63
11b	sSVOSVO reversal of subordinating conjunction	78	71
12a	Declarative with (causative) subordinate clause SVsSVO (Semantic constraints)	58	62
12b	sSVSV reversal of subordinating conjunction	99	96

Effects of sentence length, counted in number of words in the sentence (rather than number of morphemes, as that is how the task appears visually on the test materials) did not appear to influence the performance of most of the children. Correlation between number of words in the item and number of cues required by the whole group in order to solve the item, at Time 1 was non-significant (rs = .317, p = .131).

Qualitative comments on item difficulty, relating the greater numbers of cues required on some items to their grammatical structure, are contained in section 5.2.5. At time 2, learning of several of the items was evident by the smaller number of prompts

required overall, but correlation with the number of words in the item was still non-significant.

4.7.9 Equivalence of parallel versions of the test

For the purposes of evaluating reliability of the DA over time, and monitoring improvements in the children over time (see Chapter 6), four parallel versions of the test were created (as described in Section 3.2.5). Although it was considered that all versions were equivalent, mean scores obtained by all children taking each version of the test at Time 1, were compared, as presented in Table 8.

Table 8. Mean scores obtained on each version of the DA test, at Time 1

Version	No of administrations	Mean score (Std Dev)
A	8	58.9 (26.6)
B	7	70 (21.5)
C	5	53 (7)
D	4	66.5 (17)

The results of Univariate ANOVA with one between factor (Version of test, A, B, C or D) indicated no statistically significant effect of versions of the test, (F(3,20) = .81, p = .503), confirming equivalence between the 4 versions administered.

These findings will be explored in greater detail in chapter 5.

CHAPTER 5

DISCUSSION: THE DYNAMIC ASSESSMENT AS AN ASSESSMENT TOOL

CHAPTER 5   DISCUSSION: THE DYNAMIC ASSESSMENT AS AN ASSESSMENT TOOL

In summary, although some minor modifications to the test would have improved the procedure, it would appear that the DA devised in the current experiment was successful in achieving the first part of the stated aim, to formulate a valid and reliable procedure for Dynamic Assessment of language. The DA of language was evaluated as an assessment tool by administration to a group of 24 children with previously identified language impairments. One test trial was used to evaluate sensitivity of the test to individual differences, and a further trial to detect sensitivity to change over time. All the DA testing was carried out by the investigator, and all sessions were videotaped, for scoring by an independent assessor which enabled inter-rater reliability to be established. It was found that the test was sensitive to individual variation and to changes over time, whilst retaining re-test reliability and high levels of inter-rater consistency. Results obtained on the DA were compared to those obtained on other standardised measures, and correlation between the DA and the CELF-3 were significant whilst correlation of the DA score with that obtained on the Response to Mediation Scale was marginally significant. Results obtained at Time 1, at the start of the study, were correlated with changes over time, to establish the reliability and usefulness of the DA as a measure of incremental change over time. Correlation between the DA at T1 and the CELF-3 over four time points, suggests prediction by the DA of longer term outcomes of intervention that are not predicted by the score on the CELF at T1. Furthermore, internal consistency of the test was high. These findings are explored in more depth below.

5.1 Discussion of Results

5.1.1 Comparison of the DA with other measures

DA procedures have a different aim from standardised tests, because they set out to measure learning potential, and not achieved learning, and as a result standardised tests cannot be used to fully establish criterion validity of a DA. Nevertheless, results obtained on the DA were compared to those obtained on other standardised measures, to gauge the criterion validity of some of the aspects of the DA.

5.1.1.1 Comparison with the CELF-3(UK)

The criterion validity of the dynamic test devised for this study was established by significant correlation between the scores obtained on the measure, and those obtained on the standardised CELF-3(UK). The correlation was weak however, as anticipated by the assumption that only part of the achievement on the DA measure is related to achievements in language, as also measured by the CELF-3. This would be consistent with the interpretation of Embretson (1987b), who asserted that validity of a DA may be demonstrated by only partial correlation with criterion tests, as the achievement criteria are frequently not the targets of the DA. The remaining variance is influenced by the learning potential measures, in this instance the number of cues required for the child to produce the targeted response. Similarly, Hessels, Berger and Bosson (2008) demonstrated low-moderate correlation (.45) between the Hessels Analogical Reasoning Test (HART) and a static test, the Ravens SPM with which only some of the same dimensions are assessed. The validity of the HART was further supported with other measures of reliability and validity. Likewise Camilleri and colleagues (Camilleri and Law 2007; Camilleri and Botting, in press) reported statistically significant but moderate correlations between their dynamic measures of word learning and the static BPVS. Brown and Ferrara

(1985) explored this relationship in further detail and found that a standardised IQ test did not predict learning potential factors such as speed of learning and ability to transfer in almost 50% of their sample.

The first administration of the DA was used to evaluate sensitivity of the test to individual differences. Evidence was sought that the test would identify that the participants had a range of abilities which could be differentiated by the task. The range may be used to predict behaviours, in this case predict the improvements in language in response to intervention, which would be one way of validating the test itself (Pring 2005 p.178). The results elicited a wide range of scores from 29 to 102 (possible range 24-120), with a mean score of 61.83. The histogram (see Figure 1) illustrated an asymmetrical, negatively skewed (skewness 0.712, mean 61.83, median 52.5) unimodal distribution, with no ceiling or floor effects, thus representing an appropriate range of difficulty of the procedure for the population described. The scores suggest sensitivity of the test measure to differentiate within participants in the population.

In contrast, within group variation identified by CELF *standard* scores appeared to be limited, showing considerable floor effects in 20 of the 24 children achieving the lowest possible standard scores of 64/65, with a range of standard scores of 64-80, representing a range of percentiles from 1st to 9th. In the light of the floor effects evident in standard scores, the range of total *raw* scores on the CELF-3 was examined, and seen to be wide (28-122) although raw scores are considered less useful than the normative standardised scores and percentiles (Pring 2005).

5.1.1.2 Comparison with The Response to Mediation Scale (Lidz 2003)

The RtM was completed alongside each episode of testing with the DA. As reported, ratings on the Response to Mediation Scale at T1 were at best marginally significantly correlated with the DA scores at T1. The use of the Response to Mediation scale alongside the DA was intended as a supplement to contribute quantitative and qualitative data regarding the children's responsiveness from a behavioural perspective, separated more clearly from their responsiveness to language learning. The majority of criteria therefore tapped into markedly different aspects, and correlations would not have been expected.

One of the particular items of the RtM on which children scored poorly across the cohort, was 'Strategic problem solving', reflecting whether participants actively planned their responses, and this may have evaluated behaviours influencing performance on the DA. Children who were able to respond to level 2 and 3 cues on the DA tasks were those with relatively intact problem solving skills, and would accordingly have been highly rated on this criterion on the RtM scale. Similarly, rating on the criterion 'Response to Challenge' would concur with the amount of mediation a child required to complete the tasks, and the RtM items 'Use of the Adult as a Resource when Child Needs Help' and 'Responsiveness to Initiations of the Mediator' would be consistent with the levels of cuing required in the DA task. Thus a small number of items (4 out of 11) would be expected to be correlated, and indeed correlation between scores on these four items and the DA was (r_s = .414, p = .044) which is slightly more significant.   However this area of the task will be discussed later and needs further investigation to maximise its usefulness.

5.1.2 Sensitivity to change over time

The comparison between DA scores at T1 and on retest at T2 showed an overall improvement of the group, reflecting the effects of intervening intervention and of experience with the test, coupled with the mediation offered during the procedure at Time 1 which was intended to benefit performance at Time 2.

Gains in the DA at Time 2, and indeed on subsequent retests as well, may be attributed to the therapy in the intervening period or to natural development, as well as to learning of problem solving strategies and approaches to the task which were mediated during the test procedure itself. The ongoing programmes of intervention devised by the SLTs were unlikely to have specifically addressed the anagram task of sentence assembly. The task is not one that is specifically functional as a language task, although it elucidated the participants' problem solving skills and facilitated the mediation of useful and generalizable problem solving strategies, and was therefore useful as an assessment, and specifically a Dynamic Assessment task. However, intervention may have addressed other sentence construction tasks, and indeed grammatical judgement tasks involving word order. Furthermore, the mediation used in the DA was intended to improve long term problem solving, and may have been retained at T2 when the same task was presented, whereas the intervening intervention would not necessarily have facilitated improved strategy use that would have benefited the children in subsequent trials of the DA.

As a result, improved scores were anticipated in the repeated administration of the DA, in part due to learning of language structures during therapy and natural development, and in part due to learning from previous trials of the DA. The difference between group mean scores at T1 and T2 was significant, in the predicted

direction, of improvement over time. This confirms that the DA was sensitive to the amount of learning that had taken place. There was also significant correlation between the scores, confirming that performance at T2 was related to original baseline performance, and providing a measure of reliability of the procedure. Retest reliability cannot be fully established due to the intention of the test procedure itself to induce change in the individual over time, but confirmation was sought that retest scores were at least related to the original test scores.

Swanson and Lussier (2001) explained that improvements at post-test that were due only to a practice effect from exposure to the test at the pre-test stage would result in improved scores at post-test, but with greater variation, reflected in a larger standard deviation, and therefore a smaller effect size between pre- and post-testing. If however, post-testing reflected the true effects of treatment, and by this Swanson and Lussier were referring to the treatment contained within the DA procedure, then as well as post-test scores of the cohort as a whole being improved, the standard deviation of the group should be comparable between that obtained on the pre-test and the post-test and the effect size between pre- and post-testing greater. In the current study, the DA design did not include pre- and post-testing, however, as noted above, the retesting at Time 2 should reflect learning as a result of the mediated intervention of the DA (as well as the intervening treatment), and is some ways is therefore comparable to post-testing. The mean scores at T2 were in fact improved, but the standard deviations were similar, in fact slightly smaller at T2 (SD at T1 20.7; SD at T2 19). In relation to Swanson and Lussier's interpretation, this suggests that, as hypothesised, true learning had taken place in the intervening period.

In the CELF-3 scores, the means again improved from 73.7 at T1 to 86.33 at T2, but with a slightly larger SD at T2 (23.83) compared to the SD at T1 (21.29). Although this is in the direction of greater variability, signalling a practice effect, the differences are probably too small for any clinical significance to be drawn.

The demonstrated retest reliability of the DA suggests that although the intended clinical use of the DA would entail only one administration of the test, which is sufficient to extract recommendations for intervention, the test could be repeated after a period of a few months and the outcomes would be a useful indicator of retention of the concepts mediated as well as a measure of incremental progress that is not reflected by standard scores of static tests. This is a clinical use of DA documented in procedures such as the Scaffolding Scale of Stimulability (SSS, Glaspey and Stoel-Gammon 2007).

5.1.3 Inter-rater reliability

Inter-rater reliability was identified as essential to establish unequivocal evaluation of the level of prompting required to elicit correct responses in the child, especially as some variation in the content of prompting was permitted. Results of inter-rater reliability assessment showed high levels of correlation between the ratings of the examiner and the independent assessor. Only a short period, lasting approximately one hour, of explanation and practice was required before the independent rater scored a sample of videos from participants at Time 1. This was important to the experimenter in the design of the DA, as the need for lengthy training in DA would impede the uptake of DA by SLTs, and it was intended that the test methodology used in the current study would be accessible to practitioners without further training.

The agreement between the experimenter and the independent rater at T1 was excellent, but further discussion followed to define scoring criteria more tightly, and resolve discrepancies. Flexibility in the wording used by the experimenter during the sessions caused some discrepancies in rating until it was clarified that the wording was less important than the *intention* of the examiner in using the cue. For example, the examiner might have used a phrase intended to be orientating the child and facilitating a spontaneous response, which the rater interpreted as a level 2 cue. Recognizing the timing and general facilitative nature of the prompt as orientating rather than as a specific search for a strategy increased the understanding between raters.

Subsequent ratings of all participants by the independent assessor blind to group allocation, at Time 3, revealed a high correlation with the ratings of the examiner, slightly better than that obtained at T1, but not improved a great deal by the discussion following the T1 scoring. This may be due to greater variation in the group of children and the wider range of cues implemented by the examiner, which were difficult to rate. Nevertheless the correlation was considered to be at a high level, and verified that the scoring was reliable, while the procedure retained the individualised quality essential to the mediation of problem solving skills to individual children.

Inter-rater reliability was also examined with regard to the scoring of the RtM. Correlation between the ratings of the examiner and the independent rater at Time 3 revealed significant correlation but of a lesser magnitude (.570) than the inter-rater correlation for the DA (.874) There was no prior discussion or trial of the RtM rating, to resolve discrepancies, because both the experimenter and independent rater were using this scale for the first time and both had access to the same instructions from the original source of the

scale (Lidz, 2003) according to which they interpreted the behaviours. However, the experimenter rated children's behaviour and responses on the basis of live interactions with the children, and the independent rater used videoed samples only. The independent rater revealed in discussion that she was aware that online decisions were being taken by the experimenter, to pursue or terminate interchanges with the children, and the experimenter's ratings may have been influenced by subjective feelings evoked in the session. For example, in the item framed as follows there is little difference between ratings 2 or 3, and subtle signals that determined the experimenter's rating may not have been apparent to the independent rater, on the video.

USE OF ADULT AS A RESOURCE WHEN CHILD NEEDS HELP	
1	Does not refer to adult
2	Nonverbally, passively signals need for help
3	Nonverbally actively seeks help
4	Verbally asks for help
5	Actively seeks help and seems to appreciate help provided

However, the same applies to the DA procedure, where the experimenter made online decisions to mediate or cue the child, that may have slightly altered the procedure, so although the ratings of cue level would appear to be reliable, it is still possible that another assessor might have obtained different results. Nevertheless, the relative ratings that enable comparisons between children or within a child over time, should be consistent and should similarly be consistent with the ratings of an independent rater. Under the same conditions of live versus video observation, the DA ratings obtained higher levels of inter-rater agreement than the RtM scale, suggesting that the criteria were sufficiently well defined and the procedure was executed consistently.

This result is critical, as mediational interventions are thought to be too individualised to enable standardization or replicability of the procedure, (Grigorenko and Sternberg 1998) but have been used in the current paradigm to elicit individualized information, whilst retaining some level of reliability in scoring. In general, individualised mediational DA makes an important qualitative contribution to remedial and educational planning for individuals, without the need to demonstrate psychometric rigour, while reliable quantitative data enables planning for services to deliver intervention in a manner most likely to benefit the population.

5.1.4 Predictive Validity

Dynamic Assessments have been used to increase the predictive validity of assessments by identifying the potential of an individual to benefit from immediate intervention. In the current paradigm, the number of prompt cues required in the DA procedure is regarded as a measure of the individual's ZPD, or their learning potential. Cautions directed at the use of DA point out that the prediction is only as effective as the implementation of intervention that follows (Elliott 2003). Using DA to make predictions for the progress of children in educational settings is not useful unless poor prognoses can be redressed by specific interventions to improve a child's performance. Given that the children were all engaged in regular ongoing therapy programmes, and the SLTs were available and willing to implement recommendations of the DA, the study provided an opportunity to identify whether the DA (at T1) was able to predict which participants were able to profit most from SLT intervention.

As well as evaluating the gains from the current intervention programme in which the children were enrolled, the DA might predict further gains from intervention based on the increased information

elicited from the DA. Furthermore, the information about metalinguistic and metacognitive awareness that is not usually available from standardised assessments might potentially enable SLTs to plan explicit linguistic or metacognitive activities, which are recommended by advocates of DA such as followers of Feuerstein. In the light of the heterogeneity of the population of children with LI however, it was recognized that only a proportion of participants in the study were in a position to benefit directly from the specifically linguistic and metalinguistic/metacognitive recommendations arising out of the DA. This line of thinking is pursued in the investigation of the role of the DA in intervention (See chapters 6-8), in which children were monitored by retesting with the DA and the CELF-3 over a longer period of either DA-based, or non-DA based intervention. However, the usefulness of the DA at Time 1 as a predictor of gains made over the longer term is considered in this section.

The results showed significant correlation between the initial DA score, and the measured progress over time, reflected by the gain in CELF-3 score from Time 1 to Time 4. Although this is a weak correlation, it suggests that the DA is able to identify to some extent the prognosis for improvement in grammar of the children with LI, receiving regular intervention. Manipulation of the DA scores to remove the component attributable to known language and leave the score that reflects prompting only, served to emphasize the relationship between the individual's learning needs identified by the DA, and his future gains, and improved the strength of the correlation slightly. The CELF-3 (UK) scores did not achieve significant correlation with change scores, being unable to predict the change in performance over time. This would concur with studies (see review by Grigorenko and Sternberg 1998; Hessels, Berger and Bosson 2008; Vye et al 1987) showing the weaker predictive validity

of static test scores in comparison to DA post-test scores. The DA served to differentiate *within* the group of 'children with LI', separating those more likely to progress in grammar in the immediate future from those for whom gains would be smaller, or for whom gains may manifest over a longer period of time.

All of the 24 children in the current study remained until Time 3, although only 16 could be followed up in the following school year, at Time 4. The remaining cohort was shown to be a statistically representative cross section of the original cohort, despite the oldest children having been lost to the study. In the light of the fact that the group identified (in section 4.7.6) as showing percentile rank changes contained no high scorers on the DA and none of the youngest children in the cohort, the predictive value of the DA would seem to be best for children with the greatest levels of difficulty whose gains were of a smaller magnitude and elicited over a longer period of time, and therefore less apparent on retest of standardised tests.

Nevertheless, the comparison between DA scores and those showing percentile rank shifts was interesting and clearly illustrated that the DA captured the potential of the children to make progress in relation to the normative population. This is an important criterion, as special needs provision is intended to maximise the potential of individuals and to enable them to be *included* in mainstream educational settings. Some of the children appeared to make gains on the CELF at Time 2 or 3, and then decrease again. This may be in part on account of the timing of their birthday, as tests immediately after a birthday meant that children were evaluated against the norms of an older sample, and gains seem to be lost. This is a feature of normative tests with standard scores which need to evaluate children against their age-matched peers, and this could be

an important aspect of assessment. However, for children whose standard scores and percentiles do not appear to improve from one year to another, particularly those who score very low on standardised scores, examination of raw score gains may at least capture an element of growth (Hasson & Botting, 2010).

Those showing shifts in standard scores and percentiles represent a positive outcome of intervention, and it is this that has been more successfully predicted by the DA at T1, than by the CELF-3. Low scores on the CELF-3 need not predict that a child will remain at a low level, rather, an assessment of learning potential might indicate his likelihood of improving from intervention. Predictive validity is a criterion highly applicable to tests of learning potential. The responsiveness of a child to prompting and mediation should be predictive of the rate at which they can progress in the given skill area. This predictive validity is an important confirmation of the effectiveness of the current DA, according to criteria described by Kaniel (2009).

5.1.5 Internal consistency of the test

Internal reliability of the DA was verified statistically, using the first sentence in each item in order to assess the consistency of difficulty of the grammatical structure of items. The high internal reliability measured by the Cronbach's α confirmed that no particular items were inappropriate in their level of difficulty with respect to the test as a whole. Efforts were made in the design of the test, however, to sequence the items according to order of acquisition of the structures, as an indicator of difficulty. The results of the DA, where the number of prompts required by the whole group can be taken as a measure of difficulty of the item, suggest that the items were not arranged exactly in order of difficulty. Some items, appearing earlier in the test, for example the dative structures in items 5 and 6,

caused greater difficulty for the children than later appearing items, such as the copula verb in item 7. In the presentation of the test, however, moving from a more difficult item back to an easier one may have motivated the children to continue, rather than them becoming demotivated as items became more and more difficult, and it was felt that the sequence of items worked well. Further detail of the responses to individual items is presented in the next section.

In summary, statistical evaluation of the DA points to its strength in key aspects of validity and reliability, and aspects of prediction that suggest clinical utility. The hybrid design, incorporating mediation into a graduated series of prompts has not, to this researcher's knowledge been reported, and neither has DA of language been applied to the range of syntactic constructions included. Qualitative analyses of the choices made in design of the methodology are discussed below, in the light of the potential to develop the DA as a useful instrument for SLT clinicians.

5.2   Evaluation of Test Design

5.2.1 Dynamic Test methodology

5.2.1.1 Use of the Graduated Prompt framework

The use of Graduated Prompts as the basic design (modelled on the work of Resing 1997, and Guthke's short-term test 1993) enabled the relative performance of the participants to be quantified, and it emerged that the five levels of cues resulted in a spread of scores ranging from 29 to 102 out of a possible range 24-120. Thus there were no maximum or minimum scores, and the scale was thought to be discriminating. In addition, there was sufficient quantitative data to carry out statistical analysis, and capture both group trends and individual changes over time. Incremental progress could therefore be captured in children still unable to reach criterion scores in sentence formulations. It would have been preferable to reverse the

score allocation however, and have highest scores represent 'best' performance, with lower points awarded for the greater amount of prompting required, for no reason other than consistency with the majority of tests, and a more logical interpretation.

A breakdown of the occurrence of each cue level across the whole cohort at Time 1 is shown in Table 9.

Table 9. Incidence of use of each cue level in Graded prompt framework

Cue Level	Example	Occurrence
1	Metacognitive direction, Spontaneous response	278
2	Drawing on previous knowledge	19
3	Finding strategies, Problem solving	57
4	Breaking down the task. Using specific feedback	113
5	Learning from feedback and instruction	109

The table shows that level one, representing those occasions on which the child spontaneously formulated the required sentence with no input from the examiner, occurred many more times than any of the other prompt levels were required. This score also reflects the knowledge of language component of the DA. It would be possible to remove this part of the scoring altogether, leaving the scores reflecting levels of assistance only, and indeed this manipulation improved predictive value of the test, as represented by the correlation with change in scores on the CELF-3 over time. However, in the opinion of this researcher, it is useful to retain a record of items that were completed spontaneously correctly. The apparent bi-

modal distribution of scores seen in Table 9 is not important, as the cue levels are based on an ordinal rather than an interval scale, and scores are not those used in any statistical analysis, but only in this section to demonstrate the prompting requirements of the children with LI.

Of the remaining cue levels, it can be seen that the greater amounts of prompting, levels 4 and 5, were used substantially more than the less directive cues rated 2 and 3. This suggests that children found it difficult to arrange the sentences assisted by reminders to use their previous experience, or search for a problem solving strategy. Even when they knew how to approach the problem, e.g. to make a question, start with a different word, or swap the words around, the children with LI still struggled to complete the anagram, and the examiner was most often required to break the task down for the child, give him feedback or scaffold the sentence construction in detail. This would suggest that the participants' linguistic skills required specific scaffolding, but the improvement on subsequent retests, after mediation of strategic problem solving, would suggest that their learning skills were at least in part contributing to their performance. Recommendations for intervention would therefore consist of targeting both learning strategies and linguistic structures.

The use of the Graduated Prompting also enabled identification of spontaneous and facilitated strategy use by some of the children. For example, a few of the children used their fingers to cover words on the cards that they had used in their sentence, to support their own working memory. One child repeatedly counted the number of words, and used that as an indication of difficulty of the item although when prompted to reflect, he agreed that longer sentences were not always harder. Many instances of trial and error were observed, and children tried saying sentences aloud, before making their own

judgements or asking the examiner for feedback. Cues given at level 3 and 4 in particular, encouraged children to select a first word for the sentence, or to search for the verb, in order to formulate the arguments, and to rearrange words in order to find a second sentence from the same words. Resing et al (2009) reported the use of a Graduated Prompting procedure, similar to the one used in the current study, to examine the pre- and post-test use of problem solving strategies in her participants. Those in the experimental group, who received training, decreased in their trial and error strategy, and increased in their use of systematic measurements. Resing's study, however enabled examination of only two possible strategies, and it was considered that the more productive task of sentence formation and the structure of the present study maximised the opportunity to observe a wider range of approaches to the task, used by the participants. However, the current study did not systematically examine the change in strategy use as a result of the DA teaching, which would have been difficult given the lack of static pre and post-tests. Nevertheless, the CELF-3 Sentence Assembly subtest could have been adapted for this purpose, and this would have elicited an additional layer of information about each child.

5.2.1.2 Use of the Mediational Intervention strategy

The use of mediational strategies was incorporated in order to maximise the amount of learning that would take place during the DA (Vye et al 1987, Swanson and Lussier 2001). Rigid graduated prompts would not, it was felt, allow the assessor to probe the children's abilities sufficiently, or facilitate their learning to the maximum. Mediations included in the sessions varied, but all children were reassured of the examiner's intention to help them learn how to succeed in the task, and transcendence or bridging of strategy use to other tasks was included at opportune times. For

example, children were helped to realise that checking carefully that they had used all the words was the same as checking that they had all the necessary books for school, or that sequencing words correctly was as important as sequencing numbers correctly in Maths. They were also shown how varied sentence forms would make their class based stories more interesting. Encouraging feedback was given at all times, and the scoring was used to motivate children if they asked, by showing them how they had needed help for one item, whereas they had managed another item on their own. Above all, mediational strategies of questioning rather than modelling were used to help children derive their own strategies or solutions to problems, and to evaluate their own responses. Attempts at sentences were always reflected back to the child in order that they made a judgement of accuracy before their judgement was confirmed or questioned by the assessor. These strategies, adapted from the recommendations of Haywood (1993) and Lidz (1991), have been shown to facilitate the maximum amount of learning within a DA, and also to enable transfer to other items.

Although there were large numbers of level 5 cues used, there were very few instances with any participant, in which the assessor needed to model the complete sentence for the child to imitate. Correct answers were always completed before the next item was presented, and even when maximum facilitation and cuing were required, children were encouraged to fill in the final words in the sentence (see transcription in Appendix VIII as an example). In this way, the mediated learning ensured that children were active participants in the learning process at all times. It was thought that the integration of mediational teaching strategies into the procedure was a valuable component of the assessment design.

5.2.1.3 Use of 'Testing the Limits' strategies

Principles of the 'Testing the Limits' approach of Carlson and Wiedl (1992) were also intended to be incorporated into the procedure, as it was thought that the verbalization and feedback might benefit some of the children. In the event, administration under the second condition described by the authors as follows, was adopted in the majority of DA sessions.

'elaborated feedback involving the principles involved in solution given by the examiner after the child's response (C2)' (p.159).

In fact, this approach was consistent with, and part of, the mediation of learning to the children. The third condition, in which the participant verbalizes the question, the answer alternatives and reasons for their response, was shown to facilitate the greatest transfer in the Wiedl and Carlson study (1981, cited by Carlson and Wiedl 1992). This was not used in the present study, primarily because verbal explanation is an area of weakness for individuals with language impairments, and would have required scaffolding by the examiner, and in part because additional time spent on each item would not have been accommodated in the DA sessions. However, it is recognized that the higher functioning children in the study may have benefited from the learning resulting from the complete 'Testing the Limits' procedure.

5.2.1.4 Use of 'clinical interview' techniques

What was in fact carried out was more consistent with the 'clinical interview' used by Peña (2001) which permitted not only reflection on the problem item, but probing of related knowledge, which was mainly metalinguistic in nature. For example, some children were asked to identify the verb (doing word, or action) in the group of words, or select the 'person' to start the sentence with. Questioning was also directed at finding out how the children solved items, if they were able to identify their own strategies. The strengths and

limitations of the clinical interview identified by Peña (2001) were applicable to the current study. It is recognized that the process was subjective and dependent on the assessor, and that there would be little consistency between examiners. In accordance with the recommendations made by Peña, recordings of the sessions and notes made during the session were retained for inspection by others involved in the management of the child (at least for the duration of the study period, according to ethical guidelines).

The additional probing questions and responses obtained in the 'clinical interview' were not scored as they were carried out outside of the Graduated Prompting structure, after the item had been solved and the cue level recorded. They would not therefore have interfered with the basic procedure and the reliability of scoring. The benefits of the additional information obtained, for increasing understanding of the children's performances, and adding to targets and strategies for intervention were thought to be considerable, and the clinical interview was considered an essential component of the current procedure.

The combination of various methods of DA was intended to maximise transfer of learning, and the items were devised in order to facilitate near transfer between items of similar grammatical structure. While instances of transfer were reported for each child individually, no measure of overall transfer was possible, and the usefulness of the procedure in facilitating transfer cannot be evaluated. However, the priority of the procedure was to elucidate the performance of individual children in detail in order to inform intervention planning, and this aim was met.

5.2.2 The standardised test

It was decided to make use of the CELF-3(UK) (Semel Wiig and Secord 1987) as a measure of both baseline and ongoing progress for various reasons. Firstly the CELF-3(UK) is a battery of subtests that taps into both receptive and expressive abilities in semantics and syntax, which is standardised on a UK sample of children up to the age of 16. It was essential to test children on a normative standardised test to confirm their eligibility for the study. Also, as the study period began at the start of the school year, when it is usual to retest children in schools and plan new targets for the school term, it was timely to combine the school's need for a standardised test with the study's need for a baseline, and share results. The CELF-3(UK) was routinely used in many of the schools participating in the study, and all were in agreement that this instrument could be used at that particular time. As promised, full results of the CELF-3(UK) for all participants were made available to the SLTs and parents of the participants.

Secondly, the DA procedure was based on the Sentence Assembly subtest of the CELF-3(UK) and therefore the original static test served as a pre-test or baseline of the children's abilities, and subsequent retests could be used to measure transfer to untrained items, and to problem solving in the unassisted condition. There were some reservations regarding using a static test to assess outcomes of the intervention (see Chapters 6-8), but no other criterion based tools were available. When results of the DA were presented to the SLTs and parents, findings with regard to difficulties with specific grammatical structures that were *in addition* to those identified by the CELF-3 were stated, and CELF-3 findings were regarded as reliable for all structures that they addressed. The ability of the Sentence Assembly subtest of the CELF-3(UK) to differentiate between groups of children was demonstrated and its

useful structure for adaptation to a DA framework was sufficient rationale for its uptake by the current author.

A considerable issue that occurred in relation to the use of the CELF scores in the current study was the structuring of the subtests that requires a slightly different battery of subtests to be administered to children up to and including 8 years of age, to those administered to children of 9 and above. This meant that during the year long period of the study, the children who were 8 at the start, were reassessed at some point on different tests, which were not entirely comparable. If they were to be tested on the same (age 5-8) subtests, scoring would not have been valid. The normative standard scores and percentiles can be compared over the subsequent retests, but raw score totals were not directly comparable, and in fact would be seen to go down as new subtests were introduced (see Appendix XVI). Furthermore, as the Sentence Assembly subtest is not part of the battery for 5-8year olds, it was carried out on those children as an additional subtest, in order to obtain a baseline raw score, for comparison purposes, but not included in the total CELF-3 raw score. The experimenter acknowledges that there are some inaccuracies in the comparison of total raw scores within children, over time.

It was considered important to retain the use of a standardised test alongside the DA as advocated by many authors, as the DA cannot provide normative data, and is not intended to replace standardised testing, but rather to supplement it (Lidz 1987, Missiuna and Samuels 1988, Deutsch and Reynolds 2000). It is also considered inappropriate clinical practice for clinicians to 'dynamise' standardised tests, and thus invalidate their use for future assessments (Haywood and Lidz 2007). In the field of SLT where Dynamic Assessment instruments are few, practitioners might be encouraged to develop their own training materials based on

standardised tests, in order to assess the potential of children to learn on these materials, but teaching of the actual test items, and use of standardised tests with a teaching component interposed, is not advised. The use of DA by SLTs in the UK is not widespread, whereas standardised tests are widely used in clinical settings, thus it would be inappropriate to invalidate their use for any children on a clinical caseload. Haywood and Lidz (2007) and Lidz (2003) in fact describe a generic procedure for developing dynamic procedures to assess any content area, in order that practitioners can develop DAs rather than using standardised tests in a dynamic way. In the experience of the author, this would be an important research area for SLTs to develop a battery of dynamic language assessments to supplement the existing standardised tests.

5.2.3 **Response to Mediation**

 The Response to Mediation (RtM) rating scale (Lidz 2003) was included on account of recommendations arising from the work of Peña, Resendiz and Gillam (2007) that has repeatedly demonstrated the usefulness of such clinician judgements, particularly in making differential diagnoses, and assessments of CLD children. However, differential diagnosis was not required in this study, and although children came from different cultural and SES backgrounds, all were first language English speakers. Furthermore, the selection criteria for the study specifically excluded children with attention difficulties and other developmental disorders. This limited the use of some of the sections of the rating scale e.g. Self regulation of attention, motor activity and emotion, in which only criteria at the 'milder' end of the spectrum, levels 3-5, were likely to be used. Many of the other sections overlapped with observations that it was possible to make from the DA, but the RtM focussed the examiner's attention on these criteria, and it was referred to when reports of findings for each child were written. In this respect it was very useful. However,

inter-rater comparison showed poor reliability of the scale, and correlation with the DA scores was marginal, and in the light of this, inclusion of the RtM alongside the DA would not be considered essential. It may have been useful to make use of the Mediated Learning Observation Checklist (Peña, Resendiz and Gillam 2007) as an alternative, and this is recommended as an area for further study.

5.2.4 Content of the Test

Results suggested that the content of the DA was appropriate to elicit a wide range of responses from the cohort, and the anagram task was one that was suitably challenging, but not too difficult for the children to achieve any measure of success. The amount of information gleaned about the children's linguistic knowledge was extensive, and will be explored in detail in the next section. In addition, the procedure enabled the examiner to assess reading ability and working memory limitations as well.

The anagram task is one that has been seldom used in assessments or intervention programmes, and tends to be associated with reading. Weaver (1979) identified the link between sentence organizational skill and reading comprehension, and aimed to improve reading comprehension by training of sentence organization skills using an anagram task, and in particular, teaching of a 'word grouping' strategy. Her results showed that it was possible to train sentence organization, and increase the length of sentences that children were able to sequence, by using the metalinguistic word grouping sequence of prompts. Application of Weaver's findings to the present study are limited as she made use of typically developing children, with different reading levels, but language skills in the normal range. She also used simple active sentences ranging in length from 5-15 words, but not controlled for grammatical structure. Finally the training used by Weaver was dependent on the children's ability to

read, not only the words to be arranged into sentences, but the strategy prompt sheet. However, the metalinguistic approach that instructed children to search first for the verb, and then attach argument structures, prompted by asking Wh- questions, as well as making grammatical judgements (Is the order sensible?) and strategic evaluations (Have all the words been used?) was shown to be effective with third grade children, and thus may have had some application to strategies to be mediated in the present study.

The children with language impairments in the current study, were not, however, required to be able to read. Three of the participants had significant difficulty reading even simple single words, and several other children struggled intermittently with unfamiliar words. The test procedure permitted the examiner to assist the children, read for them, or remind them of words as required. However, the non-readers were forced to rely more on short-term memory for the words than the good readers, and as a result the task was more difficult for them. It was not intended to assess working memory, and in fact the task was constructed to minimize the effects of memory by having the written words available throughout the task. Some of the children, however, relied more on external strategies of manually covering the words they had used, rather than relying on their grammatical knowledge to assume that the sentence was complete, and for these children, issues of memory became more prominent. In order to focus the task on linguistic knowledge, it may be more applicable to ensure that children suitable for assessment via the DA, have basic reading skills.

An unexpected bonus, however, was revealed in the attempts of several of the children to deal with unfamiliar names, which functioned for them as non-words. Although the examiner attempted to use simple and common names for 'people' in the sentences,

some of the selections were unfamiliar or culturally inappropriate for the children in the study. One child, for example, substituted 'Jon' for 'Jos', [miri – ri-miririun] for 'Miriam' and [mil] for 'Neil', as well as 'claws' and then 'poorly' for 'paw'. Phonological processing of unfamiliar or non-words was apparently impaired in this child, and this observation was relayed to his SLT, although non-word repetition and phonological processing were not directly addressed in this study. Video and transcription of responses facilitated this observation which was not entirely focussed on right/wrong marking as many static tests are.

5.2.5 The Test Items

Group data pertaining to response to each item of the test at Time 1 and Time 2 will be examined in this section as part of the evaluation of the test itself. Individual variables on performance will be discussed in Chapter 8. Consideration of the data from Time 2, after which all children had received four months of regular intervention from their SLT at school, was used to ascertain if there was any retention or transfer of solutions or strategies gained from Time 1, and to monitor retest reliability of grammatical structures.

For ease of expression, items on which the participants as a whole group required the least amount of prompting and cuing, and thus received the lowest scores, are designated '*easier*' (easiest) and those for which many children required higher levels of prompting are designated '*difficult*'.

5.2.5.1   Grammatical Structures, and Effects of Transfer of Learning

Item No.	Finding
1a Simple active declarative sentence	Easiest item in the test -only 3 children required cues to solve at T1. None of the children had difficulty with the auxiliary verb and present progressive verb ending (cf. Rice and Wexler 1996). Two needed prompting to complete the sentence by adding the object noun phrase (NP) One was confused by the pronouns, producing '*my boy'**.
1b Interrogative	14/24 participants needed some level of prompting to produce. Half of the children who needed help benefited from the prompting to formulate a question as a strategy. Half needed more specific support, even given the declarative sentence, and the intention to formulate a question, consistent with reports that question formation with auxiliary verb movement is difficult for children with SLI (Rice, Wexler and Cleave 1995).
2a Active declarative sentence with future tense verb group	'*is going to*' caused confusion for 13 of the children, all needed level 4 or 5 prompts, to break down the sentence or scaffold parts of it. Marked improvement at T2, total score dropped from 67 at T1 to 38 at T2, with only 4 children experiencing difficulty. Possible practice effects or learning from experience. Difficulty with this structure is not a reliable finding.
2b Interrogative	Comparative level of difficulty as item 1b. 3 children required level 2 cues, e.g.'think of how you did this before', and 3 level 3 cues, but 8 children still

	required linguistic supports.
	Some transfer of question strategy, but the formulation of question form remained difficult for several children.
3a and 4a SV and SVO clauses joined with '*and*'	Few problems even with the increased length and complexity of conjoining SVO structures. 6 children needed prompting in item 3, but improved in item 4, and 4 children struggled with the length of item 4 having managed item 3 without help. Co-ordinating conjunction is developmentally early (Owens 2001) and not the source of difficulty.
3b and 4b Semantically equivalent items	Children achieved more easily than the first sentence, suggests transfer of structural organization from the first to second sentences. Bishop (1997) notes ability to substitute semantically equivalent words in a sentence suggests knowledge of the thematic roles in the sentence.
5a Dative	One of the most difficult items in the test, only 7 participants achieved spontaneously. Dative forms identified as problematic for children with LI, (van der Lely and Harris 1990). Additional noun in Subject NP, added difficulty and possibly obscured the manipulation of the dative construction itself. Several children commented 'there is no '*to*'', signalling recognition of the argument structure of the verb, but only in the prepositional format, in agreement with Ebbels (2007) that the prepositional format was understood more easily than the dative form.
5b	Transfer from previous strategic use of reversal of

Semantically equivalent elements	semantically equivalent items in items 3 and 4, as well as transfer of the dative structure from the first to the second sentence was anticipated. Small amount of transfer, observed but children's grasp of the dative appeared insufficiently stable to permit even near transfer.
6a Dative	High number of prompts, children struggling again with dative plus the object NP. 'Chunking' strategies appeared to be only marginally helpful. Presence of four determiners (2x 'a' and 2x 'the') added difficulty – many children omitted determiners, or used them randomly with no sense of definiteness- consistent with the reported findings of Gopnik, (1990), Rice and Wexler (1996) and Leonard (1995).
6b Semantically equivalent elements	Scores quite considerably lower, and at least 5 of the children who required level 4 or 5 prompts in 6a, were able to reverse the elements and solve 6b without assistance. Suggests emerging grasp of dative that permits transfer of learning.
7a and 7b Sentence containing copula verb	Relatively easy for the children, possibly facilitated by shorter sentence length of 5 words. Rice and Wexler (1996) noted children with LI inclined to omit the copula verb, but when present, it was likely to be in the correct form. Accordingly, given the verb 'is' to be arranged in the sentence, few errors were made. Adjective-Noun structure was a source of error for the children. On probing, many children did not recognize interchangeable adjectives, with same roles in the sentence e.g. '*the old man is tired*' and

	'*the tired man is old*'. Clinical interview format permitted probing of metalinguistic knowledge, but few children identified role of adjectives or 'describing words'.
8a Copula with contracted negative	Verb '*isn't*' not the main source of error Possessive NP most difficult. Many children used genitive spontaneously but when asked what the contracted suffix '*s* meant , many responded with *plural*. Agreement was not managed by children - produced 'cats' paw'* Metalinguistic questioning revealed problems with possessive.
8b Interrogative	Question forms with movement of the copula verb, could benefit from transfer of the question formation strategy from items 1 and 2, but little evidence of transfer noted. Children required reminding to find the appropriate strategy, but had little difficulty with the question formulation. Few children required high levels of support, despite Rice, Wexler and Cleave (1995) that use of copula BE in questions was poor in children with SLI.
9a Sentence with modal auxiliary and Prep Phrase	Verb group with modal auxiliary '*will'* caused difficulty for 7 children at T1 and 6 at T2, although number of prompts was not very high, and the item not considered particularly difficult.
10 a	Verb group with modal auxiliary '*can'* was most difficult item for children to find a first sentence.

Sentence with modal auxiliary and Prep Phrase	Achieved without prompting by only 4 children. Possible sources of difficulty were i) semantic concepts seemed to confuse children who did not spontaneously recognize the sentence meaning, ii) increased sentence length, iii) prepositional phrases though correct in previous sentences, iv) individual vocabulary items which were explained and clarified where necessary, v) presence of both nominative pronouns and genitive pronouns. Children struggled to identify the subject of the sentence with which to start constructing their sentence. Confusion between nominative and accusative case pronouns in children with LI, with nominative case vulnerable to error (Loeb and Leonard 1988, cited by Leonard 1998). Improvement at T2, and greater number of level 3 strategic prompts were used.
10b Interrogative	Question formation managed better than the formulation of the active sentence. Having grasped the basic sentence meaning, the question was not too difficult grammatically. Informally many children confirmed familiarity with simple questions beginning with 'can' and were able to generate some. Transfer from item 10a to 10b, reflecting semantic organizational issues with the first sentence.
11a Temporal subordination	Difficult but managed by the children slightly better than anticipated, especially with canonical structure required in first sentence.
12a	Causal concept easier than the temporal conjunction

Causal Subordination	in item 11
	In addition to grammatical placement of '*because*', children struggled with semantic sequencing of events in sentence.
	Difficulty for children in understanding non-reversibility of 'The window broke because Debbie cried*'
12b Causal Subordination	Formulation of the non-canonical '*because the window broke, Debby cried*' achieved by only one child without support, and was appropriately situated as the last item in the test.

It was predicted that the strategy of question formation would transfer to successive items, but there was little clear evidence of this in the quantitative results. It may be that the awareness of the need to formulate a question transferred, but children were still unable to formulate the movement required for question formation. This would be consistent with the observation that the declarative sentences did not seem to help facilitate formation of questions, the problem would seem to be in the movement of elements more than in the argument structure of the verbs, many of which were very simple.

Further observations pertaining to specific items of the test included the exploration of the difficulty children encountered with the possessive in item 8. Having identified the possessive '*s* as a plural, children were willing to accept the second noun in the phrase as singular, producing 'cats' paw'*. Inaccuracy in test construction however resulted in inconsistency across versions of the item, such that interpretation of 'girls' room' and 'birds' cage' would be acceptable where 'cats' paw'* would not. Children were questioned to establish their interpretation of the morpheme, but the issue of

agreement in plural marking was not further investigated, despite the reports of errors in plural marking by children with SLI, in Gopnik and Crago (1991). The content of the item could have been improved in more than one way, to assess the negative copula independently of the genitive, to assess agreement, and to make different versions of the test consistent with regard to semantic constraints. Nevertheless, the opportunity for metalinguistic questioning revealed clinically interesting data.

Item 10 presented as a good indicator of learning skills. The reason for this conclusion is that children managed the interrogative form more easily than the declarative, having in many cases struggled with the basic sentence formulation. As noted above, the prompting required revealed that the combination of several grammatical and semantic elements combined to confuse children, despite the basic concepts being relatively simple. This is apparent in the transfer to the interrogative which was managed well once the children had grasped the basic sentence. Prompts required were level 3 strategic prompts, such as 'Can you make a question?' which enabled the children to make use of prior mediation and learning, and demonstrate transfer of their learning.

Inconsistency between versions of the test was again evident in item 11, in which children were required to sequence two actions joined by the temporal conjunctions 'before' or 'after'. The items were intended to have a 'correct' sequence, for example in version A, it was clear that 'Joe brushes his teeth before he goes to bed'. However, some examples emerged as reversible to some of the children, and although the intended structure was 'Dan eats his dinner after he washes his hands' the examiner had to accept the reverse as correct as well. This meant that in formulating a second sentence, children were able to reverse elements, when the intended

structure of the test was the non-canonical 'before Joe goes to bed, he brushes his teeth'. In this instance, the scoring of the second sentence was adjusted so that the prompt to 'start the sentence with after', was not counted as a level 4 cue, but this is acknowledged as unreliable. The total cohort score still reflected a great deal of difficulty and need for support in formulating this structure.

Conclusions that can be drawn from this detailed review of data relate to evaluation of the selection and organization of test items. Evidence of difficulty encountered by the group in relation to specific grammatical structures was in the main consistent with structures identified in key papers to be difficult. Thus question formation with auxiliary inversion and dative constructions were difficult for many of the participants, while pronouns and articles were vulnerable to omission or substitution in more complex sentences. Simple argument structures and reversal of semantically equivalent elements were not exceptionally problematic, though non-canonical word orders did cause difficulty. These findings confirm that the test methodology was accurate in eliciting grammatical weaknesses in the children with LI, which are consistent with the findings of other reports.

In addition, the sequencing of items in the test was appropriate in the placing of easier items earlier, and more complex, difficult ones, at the end. The difficult dative constructions were items numbered 5 and 6, however this worked well as there was opportunity for children to regain confidence with the more manageable copula verb sentences in items 7 and 8, before the difficulty increased again later in the test. Furthermore, the planned opportunities for observing transfer of strategies from previous items was effectively realised, along with the analysis of retention and transfer of structures and strategies from Time 1 to Time 2.

5.2.5.2 Sentence Length

Effects of sentence length, counted, as previously noted, in number of words in the sentence, was not statistically shown to influence the performance of most of the children, although many commented that an item was 'hard' because there were lots of words (e.g. item 4, containing 11 words). This assumption by the children was not borne out by their performance and difficulty was more related to grammatical complexity than sentence length. There were possibly some complex interactions, as noted above when shorter or simpler sentences might have facilitated the grasp of particular complex grammatical structures e.g. the datives in items 5 and 6. Similarly, Item 7, which contained a copula verb, was achieved with little difficulty despite the nature of the verb, and this was thought to have been facilitated by the shorter sentence length. The greater effect of syntactic complexity than sentence length on performance parallels the findings of Marton and Schwartz (2003) who found that syntactic complexity was more demanding on working memory capacity than sentence length, in both children with SLI and typical language development.

Although in the current task the materials remained on view for the children, and thus reduced the short term memory requirement, a degree of WM was still required for children to formulate the words into sentences, and rehearsal of sentence fragments was observed in many of the children as they struggled to formulate additional parts of the sentence. However they were more able to check that all the words had been used when fewer were on display, and encountered difficulty in longer sentences, even when they were syntactically simple e.g. '*Mum is driving the car and dad is riding the bike'*. Some children required prompting to check that all words had been used, and in particular that auxiliary verbs and articles were in place. This

highlighted the structures that were vulnerable to processing limitations, i.e. the grammatical morphemes rather than the sequential arrangement of arguments. The opportunity was also afforded for the examiner to observe domain general cognitive functions under executive control, such as planning and checking behaviours. Marton and Schwartz similarly identified the role of executive functions in performance, although the function of attention switching was most implicated in their task.

Montgomery (2002) highlighted the potential interaction of a deficit in linguistic knowledge and deficient information processing in children with SLI, and pointed out that assessment should attempt to determine both of these aspects. Other than standardised language tests, the methods recommended are a range of informal measures, and detailed task analysis. The structure of the task in the current procedure is such that there is no time limitation or speed of processing requirement, and this would be consistent with Montgomery's recommendations for ascertaining linguistic skills separately from speed of processing abilities. In the light of this, it may be assumed that the difficulties experienced by the children were due to linguistic or executive function deficits, rather than to processing limitations.

5.2.5.3 Grammaticality judgements

All of the participants were asked to evaluate their own responses before they were given feedback by the examiner, a task that required them to judge the grammaticality of their own productions. Observations of each child's ability to recognize correct and incorrect sentences were made and recorded on the report supplied to SLTs and parents after the DA. In this way, the information about the individual's skill in this area was linked to recommendations for intervention. No group results are available with regard to the

performance of the whole cohort in grammaticality judgement tasks as it was felt that this would not yield information relevant to the evaluation of the test procedure. The fact that the clinically relevant observations were facilitated is a strength of the DA procedure, and it would be recommended that this be retained as part of the mediated intervention aspect. Indeed, feedback about the child's accuracy in judgement was also fed back to him, and if necessary, identification of the metacognitive skills of checking, and the need for precision and accuracy were mediated, and linked to other contexts in which these skills are important. These strategies are consistent with the recommendations for mediation formulated by Haywood (1993) and Lidz (1991). The general observation that many of the children were uncertain and lacked confidence in their own judgements is consistent with the findings reported by Gopnik and Crago (1991).

5.3 **General Discussion**

The foregoing discussion identified some of the alterations to the test methodology and items that might have improved the utility of the procedure. However, the data presented have demonstrated that the aim of the project which was '*to formulate a valid and reliable procedure for Dynamic Assessment of language for children with SLI'*, has been achieved. The research goal stated by Budoff (1987a) for DA to assess '*those who have been correctly diagnosed, but whose potential for improvement has not been gauged'* would also appear to have been met.

In addition, '*the procedure needed to be replicable and teachable, in order that any demonstrated utility could then be adopted by practising SLTs in the field*' and high reliability of scoring reflected in the correlation between the experimenter and an independent rater with minimal training, suggested that this would be possible. The

demand in terms of administration time and scoring simplicity were controlled as the test was completed within a single test session in every instance, and scoring was completed during the session. An evaluation of the final part of the stated aims, to formulate a procedure '*that yields useful information for planning intervention for children with LI'* will be explored in Chapters 6-8.

The application of DA principles to the assessment of previously diagnosed children with language impairment, for the purposes of finding out more detailed information about them, was a novel one. Further information about the way in which they approach a language task, problem solve and self evaluate, and their potential to learn from input from the examiner was sought, and found. No attempt was made to differentiate the children from any other population of their peers, the intention was purely to gain information that would be useful to inform intervention for the children, and in this respect the aims of the study were different to previous studies of DA of language. In addition, the targeting of explicit syntactic knowledge in these school age children was also original, previous studies having addressed word knowledge, (Peña 2000, 2001, Camilleri and Law 2007) narrative, (Peña et al 2006, Miller, Gillam and Peña 2001), expository discourse (Peña and Gillam 2000) or younger children, (Olswang and Bain 1996).

For this novel purpose, a unique combination of established DA techniques was devised. In some respects, the procedure of Graduated Prompts was altered and the recommended standardisation of the procedure (Campione and Brown 1987) was undermined. Graduated Prompting was devised with the intention of facilitating transfer of learning to different problems, but the burden of this was instead placed with mediational intervention. Nevertheless, the fundamental principle of using the number of hints

required by an individual to solve a problem, as a measure of their ZPD was retained, as were the psychometric properties of the procedure. Furthermore, the application of the Graduated Prompting method to a curriculum area was consistent with the recommendations of the authors, and Campione (1989) also noted that the procedure contributes to the 'instruction process' enabling learning from the prompts delivered as part of the assessment.

Similarly, the mediated learning experience (MLE) recommended by Feuerstein (Feuerstein and Feuerstein 1991) was not intended to be delivered in any kind of pre-determined, task related way, but instead focussed on the individual mediational needs of the child. Although Feuerstein was firmly in favour of mediation through domain general skills, not limited to any particular content area, the verbal modality was considered a fundamental one, and some of the LPAD instruments specifically address verbal functioning (e.g. The 16-word memory test, and the Verbal Similarities test).

The feature common to both methods used is that support for the individual starts with more general, metacognitive and strategic hints, and only becomes more directive if it is needed. Intervention is adaptive and individuals are facilitated in both paradigms to discover their own solutions to problems. This does in fact contrast with some traditional language therapy interventions (e.g. Weismer and Murray-Branch 1989) that model complete and correct structures at the outset, from which a child should learn, or which a child should imitate (Matheny and Panagos 1978). The common principle enabled the hybrid procedure to work, and the combination of methods yielded the intended results with the procedure managing to contain individualised mediation while still being reliably identifiable as consisting of graded cues. Thus, extracting the fundamental principle underlying the teaching inherent in the two

procedures, emphasized their similarity rather than their difference and enabled the two to be combined into a hybrid procedure.

It is the same principle of adaptive support which is advocated for language therapy approaches that are metacognitive in nature. If maximum learning and transfer are facilitated by such methods in the assessment procedure, they should surely be effective *for the same children*, in the intervention process. This does not imply that these methods are effective for all children, and there are those that do not exhibit a great deal of learning in the DA process, that must therefore require more directive or intensive intervention in order to learn, and a useful DA must be able to identify these children as well. Thus a range of scores must be elicited, and the current procedure produced this as an outcome.

5.4 <u>Implications of the Study</u>

The effective combination of DA methods as described above adds to the growing body of literature about the methods of DA implementation, and certainly adds to the fields to which it has been applied. The philosophy is that of Haywood and Lidz, (2007) who advocate the spread of DA principles in just such a way, through adaptation and application, in both educational and clinical contexts. In addition, the advantages of an eclectic approach to DA are elucidated.

From the point of view of language impairment, the study has much to contribute in terms of methodology and scope of structures addressed. DAs of language are few, and the method may be used as a template for other studies to be developed. It answers the call for assessments to be more adaptive, holistic and representative, allowing, as it does, the probing of responses. It sits alongside standardised tests, and increases the information obtained.

Assessment in this manner of expressive syntax was an ambitious goal. Few tests of expressive syntax are available, and all are limited in some way to a range of structures they can target. Even language sampling techniques may not be representative of the language structures that were not elicited (Crystal 1982). The current study is no different, although the procedure can be used to target any linguistic structure, as long as the examiner can formulate an appropriate test/training item. The current procedure did not set out to test a comprehensive range of structures, nor to target all those most consistently identified as clinical markers of LI. Furthermore, it did not set out to search for empirical support for any one theory of language disorder. However, the items that were seen to present relatively greater or lesser difficulty for the children with LI in the sample, were consistent with those described in the literature, and contributed some understanding of the difficulties faced by the children. Again, examination on an individual basis would have the most value in contributing to understanding of a particular child's knowledge and skills, as performance of individuals was very varied. Nevertheless, the selection of syntax as the object of assessment was an important one, and despite limitations described below, the study has opened up an avenue for assessment of this linguistic level.

Intervention for children with LI could be based more on an understanding of the presenting symptoms than on the features themselves. Thus for example, recognizing that knowledge of argument structure is intact, but syntax is difficult would lead to a different approach to intervention than needing to address argument structure per se. Approaches to intervention would be more theoretically grounded, and less targeted on morphosyntax in isolation. Probing responses, examining transfer and asking metalinguistic questions all contribute information towards the

planning of effective intervention. Thus the current study contributes a method of assessment, and a means to access clinically relevant information, to the field of syntactic assessment.

5.5 <u>Limitations of the Study</u>

The greatest limitation of the study evaluating the Dynamic Assessment tool was that all testing was carried out by the one examiner, and there is as yet no evidence for the effectiveness of the procedure when carried out by other clinicians. This is an area for further study.

A further limitation relates to the participants, and primarily to the age of the children, who ranged between 8 and 10 years. At the lower level, some 8 year olds were able to manage the test adequately, while a small number struggled a great deal, and more than any participants aged 9 and above at the start of the study. Furthermore, the three non-readers in the study were all aged 8. In spite of this, no participant scored at the maximum level of cues, the floor of the test, so whilst caution should be applied to children for whom the test would be useful, the age of 8 would probably be an appropriate cutoff. The upper age range may be more flexible, as children aged almost 11 by the final tests still did not score at ceiling, but younger higher functioning children may in fact reach the point at which all items are achieved spontaneously, without prompting. The age range of application of the procedure is therefore restricted to approximately 3 years.

Linked in part to the issue of age, is the use of the standardised CELF-3(UK) as part of the procedure. A static standardised test such as the CELF is thought to be unreliable for the assessment of children with SLI, whose responses to tests are unreliable and not necessarily representative of their maximum ability. The use of such

a measure for participant selection is justified by the normative and discriminant properties of the CELF-3, but for monitoring performance, it is theoretically inconsistent with the rationales for DA development.

A number of limitations related to the items in the task were identified in the close examination of task items in the previous section. Four parallel versions of the test are unlikely to be needed in the clinical context, and the issue of consistency of versions is an artefact of the present study. Furthermore, some important grammatical structures were omitted from the test protocol, and an item addressing passive constructions would have been compatible with other items on the test.

5.6 <u>Future Directions for Research</u>

Further research to address limitations of the current study and to extend and apply the research questions more widely is recommended. In the first instance, in relation to the clinician administering the test, it would be important to ascertain the reliability of the procedure when administered by other SLT clinicians, who have not been specifically trained in mediation. It is recognized that different examiners will pursue different avenues of mediation, and attempt to mediate different strategies, so outcomes will always be qualitatively different, but research into whether the measure of prompt levels is similar, and equally reliable when rated by an independent observer, remains to be evaluated.

In relation to the behaviour of the children with language impairments, closer examination of the strategy use by the children would be a valuable extension of the research. This may describe and possibly quantify the nature and range of problem-solving strategies used by individuals, in a manner similar to that outlined

by Resing et al (2009). Similarly, a measure of efficiency of strategy use, such as that suggested by Swanson (1995) might be incorporated into the procedure.

With regard to procedural variables, experimentation with varying the grammatical structures targeted and the age and severity of language need of the children assessed is also recommended to ascertain range of applicability and potentially extend the range of applicability of the basic DA procedure.

Finally, predictive value of the DA is an avenue for further research. The current procedure aimed to elicit information that would assist in the planning of appropriate intervention for individuals, and investigation of whether the information does in fact, benefit intervention planning and indeed the outcomes of intervention, would be a logical follow up. A first attempt at this is contained in chapters 6-8, but in the opinion of the current author, this can best be evaluated on a case-by-case basis, and a series of case studies would be welcomed. Similarly, case studies would provide the opportunity for clinicians to try the assessment procedure, and evaluate its reliability as well as clinical usefulness. The current study benefits from both the advantages of a cohort study and a case study series (McCartney 2004). Overall effects achieved by the DA on the selected group indicate its potential usefulness, while individual profiles highlight trends in predictive utility, and both of these avenues could be further researched.

Having established the reliability and validity of the DA tool itself, the next task was to explore assumptions often made in the literature that more sensitive testing would lead to improved intervention. The next sections of the study will therefore examine

the usefulness of the DA procedure in informing and enhancing therapy for the same group of children.

CHAPTER 6

IMPLICATIONS OF DYNAMIC ASSESSMENT FOR INTERVENTION: METHOD OF INVESTIGATION

CHAPTER 6   IMPLICATIONS OF DYNAMIC ASSESSMENT FOR INTERVENTION: METHOD OF INVESTIGATION

6.1   Design of the experiment to explore the impact of information derived from a DA on intervention

A study was designed to evaluate the progress made by participants in different phases of SLT intervention, with the critical difference between the phases being that the intervention was based on information derived from static only, or static plus Dynamic Assessments. It was planned to analyse the data set in two ways, first at group level using RCT methodology, but also because the sample is relatively small and heterogeneous with regard to type and severity of language impairment, using a case-series approach.

The current study included a baseline which consisted of the progress measured during a period of regular ongoing intervention that had been, and continued to be, the cornerstone of the management of the child. It was not possible or appropriate to include a true 'no treatment' phase.   Results of the tests for eligibility for the study i.e. the Ravens CPM and the baseline CELF-3(UK) were made available to the SLTs, as were copies of the teachers' questionnaires, and these were available for the SLT to use in planning intervention.

In the second phase, half of the subjects, the control group, continued with their ongoing programme of intervention, devised and implemented by the Speech and Language Therapist with whom they work in their school, and informed by standardised language tests and reports from their teachers.

The experimental group participated in revised intervention in the second treatment phase, which was informed by the outcomes of the

Dynamic Assessments, at Time 2 as well as the static tests and teachers' reports. For this group, DA results as well as recommendations for targets and strategies to be used in the intervention were made available to the participating SLTs, as described below in 6.2.

Finally in the third intervention phase, the results of the DA at Time 3 for the control group children were also made available to their SLTs, and all children participated in 'revised' intervention until they were reassessed at Time 4. The results of DA at Time 3 for the children in the experimental group were not shared with the SLTs and the experimenter did not collaborate in the updating of the targets.

There were 4 months between each assessment time point. The sequence of events in the study can be summarised in Figure 6 as follows:

Figure 6. Summary of stages of experimental design

	T1	Phase 1	T2	Phase 2	T3	Phase 3	T4
E group	Static test results to SLT	Regular ongoing Intervention	Static and DA results to SLT	Intervention informed by DA	Static test results to SLT	Intervention informed by DA	Retest static and DA
C group	Static test results to SLT	Regular ongoing Intervention	Static test results to SLT	Regular ongoing Intervention	Static and DA results to SLT	Intervention informed by DA	Retest static and DA

Change in performance measured over time was thought to be attributable to the effects of practice of both tests, and the effects of the intervening learning from SLT intervention. Thus comparison of changes in score from Time 1 to Time 2, and from Time 2 to Time 3, for each child, reflected the difference in learning rate resulting from the differing interventions in those periods. In the control group, rate of learning was expected to be similar across the two time periods, as the intervention remained based on the same information and planning criteria, and was then subject to change in the third phase when intervention was altered.

It was hypothesised that little change would be measurable by standard scores in the repeated administrations of the CELF-3(UK). Pilot testing (Hasson and Botting 2010) showed children with language impairments to score low standard scores, and even considerable gains in the raw score, did not raise the standard score. Raw scores were therefore used to measure change. It was anticipated that scores on the Sentence Assembly subtest may improve more than the other subtests, as a result of the mediation of strategies for that task that were delivered during the DA.

It was anticipated that there might, however, be positive changes in the scores obtained on the DA over time. These would reflect not only gains in the language knowledge aspect tested by the DA that might improve as a consequence of SLT intervention, but also a decrease in the amount of mediation or assistance required to complete an item, as a result of practice and the mediation received in previous DAs. Changes in performance may be measurable via the DA procedure. Similar use of a DA to record progress by means of amount of scaffolding required for the child to achieve a criterion is reported by Glaspey and Stoel-Gammon (2007). For this reason, it was decided to use the DA at each of the times of testing, even

though the findings of only one DA would be made available to SLTs working with the children (at Time 2 in the experimental group and Time 3 in the control group children).

6.2 Construction of the Experimental Intervention based on the outcomes of DA

The current study aimed to investigate whether intervention based on the results of a Dynamic Assessment together with standardised tests and teacher reports would be any more effective than intervention based on the results of the standardised tests and information gathered from teacher reports alone. It is vital to link intervention to both results of detailed assessment, and to a theoretical model. Both the selection of theoretically grounded targets and methods, and the comprehensive assessment of children are essential components in planning therapy.

The current study based intervention on assumptions from both domain specific models of grammar as a primary source of difficulty in SLI, and domain general models of cognition and learning skills influencing language learning. The Dynamic Assessment elicited information about language structures that were problematic for the child, and determined that intervention should have linguistic targets. However, as the DA also highlighted difficulties in processing certain common features, learning from examples, and metalinguistic understanding, these could be used to identify more process based targets for intervention as well. In addition, as Feuerstein's principles of Mediated Learning Experience (MLE, Feuerstein 1991) were incorporated into assessment procedures, the same principles could be used to implement targets that addressed cognitive awareness of the tasks, reflective thinking and problem solving, and also the nature of the intervention, that was mediational in style.

SLTs working with the participants in their schools were not specifically trained to carry out this intervention programme. The reason for this was the preference for the procedure to be accessible to all professional SLTs without the need for additional training that would impede the uptake of DA-MLE management in clinical settings. Instead, a manual was provided, that contained a small number of basic principles of mediation, with emphasis on the need for explicit, cognitive intervention. The major change for the participating SLTs was the amount of individual information that was made available from the assessment. Thus the SLTs were in a position to devise much of the intervention independently when they received the detailed results of the DA (see Appendix IX - Manual for SLTs).

6.2.1 The sources of information to inform intervention

Linguistic Information was derived from:

> i) standardised tests
>
> - receptive and expressive scores from CELF-3 (UK)
>
> - item analysis from CELF-3 (UK)

> ii) DA
> - effect of semantic content, semantic constraints, amount of
>   content
> - syntactic structures from DA
> - transfer, generalisation of learning i.e. item - to item
>   transfer
> - metalinguistic knowledge, ability to label, explain and
>   manipulate concepts

Behavioural Information was derived from:

> i) teacher reports
> ii) DA
> - attention / activity/ emotion variation according to task

- motivation / attitude to learning / interest / responsiveness
- use of strategies / including reliance on others for help
- learning needs of the individual, i.e. requiring metacognitive monitoring vs strategy training vs item specific application of knowledge
- metacognitive awareness, recognition of learning strategies and processes used to solve the given task.

6.2.2 Targets of the intervention

The differentiation of the experimental intervention stemmed from the additional information made available by the DA that enabled the clinician to construct the kind of principled intervention described by Fey, Long and Finestack in 2003. Their ten principles of grammar facilitation began with target selection that looked beyond the specific goals of individual grammatical structures, and aimed to enhance language learning strategies, and included a recommendation that specific intervention goals should be based on 'functional readiness' (p.7), a concept that may be clarified by the assessment of potential or modifiability embodied in DA.

Standardised tests would have been used by the SLTs prior to the present study to determine goals for intervention, along with knowledge of the children gained from experience, and liaison with class teachers. Outcomes of that intervention, along with results of retests, would be the starting point for targets in the baseline phase of the present study.

Although developmental order is less prominent in children at the upper stages of primary school and in the later stages of developmental language learning, Balthazar and Scott (2007) suggested that target selection should take account of the following two questions which are felt to be relevant here. 1) *Has the child*

learnt the basic set of grammatical structures expected for school-age children? And 2) *Does the child comprehend and produce an adequate array of complex syntactic forms for the pragmatic needs of school?* (p.149). Formal assessment will in most cases have identified the expected grammatical structures, and teachers' reports will draw attention to the weaknesses the children may be experiencing in curriculum access.

Linguistic targets were also related to findings of the DA regarding *structural* difficulties, i.e. focused on sentence construction, rather than on lexical or morphological levels. Common or recurrent difficulties were identified as targets, that is linguistic or structural areas of difficulty were grouped into broader, themed targets, i.e. *a feature or principle common to more than one linguistic structure* e.g. use of auxiliary verbs in declarative sentences to carry tense/aspect marking as well as to form questions; verb arguments as a means to sequence and structure elements of a sentence (rather than addressing individual elements by means of answering questions 'did *what*?'). This method of goal selection is comparable to the principle described by Fey, Long and Finestack (2003), who suggested 'intermediate goals' derived from categories of linguistic targets (p.5).

Identified targets were to be extended in application, i.e. targets for the period of intervention should include use of the identified linguistic structures in a variety of linguistic contexts, such as, for example, the construction of narratives, and all targets were to include functional use of linguistic constructions (see Appendix IX for example contained in Manual for SLTs).

6.2.3 Strategies and methods of intervention

With regard to methods and strategies based on findings of DA, treatment protocols were recommended which should address the particular strengths and weaknesses in learning of the individual. In addition, the evidence of explicit teaching being useful in SLI (Ebbels 2007, Bishop, Adams and Rosen 2006), combined with the demonstrated progress in the 'teach' phase of DA being best facilitated by mediated learning interventions (Bransford et al 1987), suggests that mediated and metacognitive teaching would be an effective strategy to apply. This is a notion compatible with Feuerstein, and also with process-based learning by Ashman (1992), both of whom advocated a type of treatment that is flexible in its administration and accommodates a range of teaching targets, to be made applicable to a wide range of individuals.

All intervention was to be delivered via individually tailored mediation, individualised to address maintenance of attention, regulation of emotional and motor responses, as well as problem solving strategies. Materials were chosen to engage the individual, as well as to challenge. The intervention was recommended to be mediational in nature, incorporating essential components of mediated intervention. The essential components were provided to the participating SLTs with explanations and examples, in the Manual (see Appendix IX). The essential components were described by Lidz (1991) as:

- mediation of intentionality – conveying to the child that you intend to help him improve
- mediation of meaning – sharing the purpose of the activity
- mediation of transcendence – linking the activity to other contexts in which the skill can be used,
- mediation of a feeling of competence – targeting praise so that the child learns what he has done well, learns that the

tester has confidence in him, and gains confidence in his own ability

Mediation of meaning and transcendence imply explicit, metacognitive teaching, making sure at each stage that the child grasps the principle that he is learning, its importance and application to the task and wider, functional use. The method involves a high level of interaction between therapist and student, with reciprocal checking of understanding, and verbalizing of processes (c.f. Ashman's Process based Instruction and Feuerstein's Instrumental Enrichment).

There is less need to modify the context of teaching to ensure naturalistic opportunities for structures to be used, than in mileau teaching. This is because learning is intended to be based on generalizable principles, rather than contextually bound examples. Similarly, there is little emphasis on elicited imitations, with the preference being on self-discovered and explicitly verbalised principles, rather than the need for practice.

For the purposes of the current study, the experimental intervention was randomly sampled for monitoring by the investigator. A sample of 12% of the participants' sessions were videoed and rated for the presence of mediational interventions, and the implementation of cognitive approaches to the process based targets.

6.3 **Participants**

The recruitment and selection of participants for investigation of the DA was described in Chapter 4. The cohort of 24 children with previously diagnosed language impairments that participated in the trial of the DA also participated in the study to explore effects on intervention.

6.3.1 Allocation to groups

After completion of all the eligibility and baseline testing at T1, when 24 participants had been confirmed for the study, participants were randomly allocated to the experimental and control groups. This was done on a school by school basis, in order that all the SLTs of participants were provided with the manual containing information about mediated learning techniques, and had opportunity to implement them. Comparison of the intervention in the two phases could then be based on the whole group of SLTs of the participants who would have equivalent training. In addition, the samples of children from special school vs language resource base, and inner London vs suburban schools were evenly distributed between the experimental and control groups.

Because the participants from each school were randomly allocated to groups, this resulted in an equal number of children from each school in each group, unless there were an odd number of children, in which case they were randomly allocated to a group. At the time of allocation to groups, the experimenter was notified that the SLT in one school, with only one child participating in the study, was to leave her post. It was unclear whether the child concerned would be able to complete the study, or would have to be excluded. After the other 23 children were allocated to a group, when child CH1 was confirmed as remaining in the study, he was randomly allocated to a group, and became the 13th member of the experimental group, leaving 11 in the control group. The allocation of children to groups is summarized in Table 10.

Table 10. Allocation of the sample to groups

Group	Participant	Age (at T1)	Gender	School
C	D1	8,2	F	BS
C	D2	10,3	M	BS
C	D5	10,7	F	BS
C	R3	9;1	M	BS
C	R4	9;3	M	BS
C	CP2	8;9	M	BS
C	BH11	9;1	M	SS
C	BH7	9;8	M	SS
C	BH9	10;0	M	SS
C	TA2	10;0	M	BI
C	TF3	8;8	M	BI
E	D3	8,5	M	BS
E	D4	8,10	M	BS
E	R1	8;4	M	BS
E	R2	9;8	M	BS
E	CP1	9;2	M	BS
E	CP3	10;9	M	BS
E	BH6	9;7	F	SS
E	BH4	9;7	M	SS
E	BH3	9;3	M	SS
E	TA1	9;4	M	BI
E	TF1	8;2	M	BI
E	TF4	8;7	M	BI
E	Ch1	9;2	M	BS

Key to schools:
BS=Language resource base, suburban school
SS= Special school for children with communication difficulties, suburban
BI = Language resource base, inner city school

6.3.2 Characteristics of the samples

The experimental group consisted of 13 children, age range 8;2 to 10;9 with a mean age of 9;2. The control group consisted of 11 children, age range 8;2 to 10;7 with a mean age of 9;5.

6.4 Testing Procedure

6.4.1 Summary

Children participated in testing procedures at 4 stages in the project:

Time 1

In addition to the testing employed to confirm eligibility for inclusion, the static standardised language test, CELF-3 that was used to confirm eligibility, also provided a baseline score, and a basis for the planning of the regular ongoing intervention programme.

Teacher reports were elicited, and both these and the CELF-3 results were presented to the SLTs working with the children. CELF-3 results and results of eligibility tests were made available to the parents of all participants.

The Dynamic Assessment procedure was also carried out and used to measure the degree of support needed by the child to solve the specific task at the outset of the intervention period. These results were not made available to the participating SLTs, schools or parents.

Time 2, after the baseline intervention period.

Static, standardised testing with CELF-3, and the teacher report were repeated to assess progress after intervention. Dynamic Assessment was carried out using a parallel form, to assess progress, and to inform the following period of intervention for the children allocated to the experimental group only. In all instances, DA was carried out after the CELF-3, so that mediation delivered during the DA did not affect the Sentence Assembly subtest of the CELF-3. Testing was carried out over two sessions, with four or five subtests of the CELF-3 completed in the first session, and the remainder,

along with the DA in the second session. For children in the experimental group, results including the DA report and static test were made available to SLTs, school and parents, while for children in the control group, only the CELF-3 results were shared.

Time 3, after the second intervention period.

Static, standardised testing with CELF-3, teacher report and DA were carried out with all participants to assess progress after further intervention. DA and static test results for children in the control group were shared with the SLTs, schools and parents, enabling use of all the information in the planning of further intervention for participants. Static test scores were shared for children in the experimental group.

Time 4, after a further 4 months of ongoing intervention

Assessment of all children by means of static standardised test and DA, to gauge whether there was maintenance of the rate of progress facilitated in the second intervention stage, for children in the experimental group, and to measure progress in the third stage of intervention for children in the control group, who would have had one term of revised intervention.

6.4.2 Measures

6.4.2.1 Static, Standardised language tests

All children were tested on the 6 core subtests of the CELF-3 (Wiig, Semel and Secord 2000), namely Concepts and Directions (CD), Word Classes (WC), Semantic Relationships (SR), Formulated Sentences (FS), Recalling Sentences (RS) and Sentence Assembly (SA).   Children under 9 years of age who completed the age appropriate subtests, namely Sentence Structure (SS), Concepts and Directions (CD), Word classes (WC), Word Structure (WS), Formulated Sentences (FS) and Recalling Sentences (RS) were also

tested on the Sentence Assembly subtest, to elicit a raw score that could be used for measurement of progress in that task in particular. Receptive and expressive language scores as well as Total Language Scores and percentiles were computed. Standard scores for each subtest were calculated, item analysis completed, and specific subtests causing most difficulty were identified. The completed test form was copied for the records of the SLT at the child's school.

6.4.2.2 Teacher reports

A questionnaire was devised for teachers, (see Appendix X) intended to elicit their main concerns about each participant, and their opinion about the participant's performance in a range of language and behavioural situations. The information elicited was copied for and shared with SLTs at T1, to enable them to make use of the information in the planning of their intervention. The scale was then repeated after each phase of intervention to elicit any perceived changes in the child as a result of the intervention.

The reason for the inclusion of teacher reports was that intervention by Speech and Language Therapists is seldom devised on the basis of test results alone, but takes into account functional difficulties of the individual, and the way in which their impairment might impact on their social and educational performance as well. Priorities for therapy targets may be based on these factors, and measurement of outcomes may be based on change in functional abilities.

Furthermore, the DA procedure already described included completion of the Response to Mediation Scale (Lidz 2003) which describes 'non-intellective' or behavioural factors, and change in the individual may be identified by means of this scale. Information derived from the RtM scale was incorporated into the reports of the findings of the DA, and made available to SLTs when they used DA

information to plan intervention. Thus in order to account for the use of these factors in the therapy programmes that were not informed by the results of the DA, information was collected in a systematic way, and presented to the SLTs for their use in intervention planning.

The questionnaire (see Appendix X) consisted of 15 items, each in the format of a rating of the child's behaviour on a scale of 1-5, with a qualitative descriptor for each level. The items requested the teacher's own opinion of the child's performance, based on their experience and knowledge of the child. In order to avoid 'response set' the wording of items is varied so that in some items level 1 may reflect the greatest difficulty, while in other level 5 may be the 'poorest' score.

The items addressed performance on functional tasks, such as following rules, attending to a language versus a practical task, or telling a story, that could be likely targets of language and communication programmes for children of this age group. The items were devised by the experimenter with input from an experienced teacher of children with Special Educational Needs who commented on the likelihood of teachers having access to the required information, and the clarity of wording of the items on the questionnaire.

6.4.2.3 The Dynamic Test

The Dynamic Assessment as described in Chapter 3 was carried out at four time points in the study. Parallel versions of the test were used, allocated in the first instance on a random basis, and thereafter according to which versions had not yet been seen by the participant.

Feedback of the results to the SLTs and parents was in the form of a qualitative report, giving information as described in section 6.2.1 above, 'Sources of information'. Quantitative scores from the DA were not supplied as these were not thought to be informative to the SLTs in planning intervention. Instead a report was written, in a predetermined format that was supplied to the SLTs in the Manual, (Appendix IX) as follows:

1. Detail of language structures that the child has difficulty with, that is additional to that obtained from the standardised tests
2. The effect of amount of content and nature of semantic content on the child's construction of linguistic structures
3. The child's ability to transfer, or generalise learning or strategies i.e. item - to item transfer
4. The child's *metalinguistic* knowledge, ability to label, explain and manipulate linguistic concepts
5. The child's *metacognitive* ability i.e. awareness of the processes and strategies that are used to solve the given task

The DA also contributed information about the child's:
• attention /activity/ emotion while engaged in the presented task
• motivation / attitude to learning / interest / response to input, while engaged in the presented task
• use of strategies, including reliance on others for help

6.4.2.4 The Response to Mediation Scale

The Response to Mediation scale (RtM) (Lidz 2003) was completed after each session of DA, with the responses based on observations of the participant during the DA. Eleven criteria were each rated on a scale of 1-5, according to qualitative descriptors (see Appendix III) giving a total score out of 55. Like the DA, quantitative scores and record sheets were not supplied to SLTs or parents, but the findings

were incorporated into the reports supplied to SLTs to inform intervention.

6.4.3   The intervention phases

In the first intervention phase, all SLTs working with participants were asked to continue ongoing regular programmes of intervention. All SLTs were asked to complete a short questionnaire in order to record the intervention currently in use with participants (see Appendix XI). The method of asking SLTs by questionnaire to report on their practice was used by Law et al (2008), to investigate the practice and theoretical assumptions of SLTs working with children with receptive language impairments. Open ended questions about activities that SLTs were engaged in were used to identify the nature of therapy. In the same way, open ended questions were used to identify both the targets and methods used by the SLTs of the children, identified with LI in the current study.

The questions addressed to the SLTs were as follows:

1. Outline 3 targets that you are currently working towards with the child

2. 'Outline 3 activities that you have recently or currently engaged in with this child' (Law et al 2008 p.249).

For each activity, specify at least one method of facilitation used, and how change was measured.

The responses to both questions 1 and 2 were used to identify:
1. The areas of language being addressed, which were placed in one of the following 7 categories;
-   vocabulary,

- morphology,

- semantic organization,

- sentence comprehension,

- sentence construction,

- narrative,

- pragmatics.

In addition areas not pertaining to specific levels of language were identified, e.g.

- auditory memory,

- attention and listening,

- phonology,

- speech production

2. Whether targets and activities reflected 'skills acquisition', 'metalinguistic activities', or 'meta-cognitive activities' adapted from Law et al (2008).

Information gained with regard to the facilitations was used to assess whether the activities made use of content based prompts, metalinguistic activities or metacognitive cues, to confirm the judgements previously made about the nature of activities.

Definitions used to determine the categories of activities were:

Skills acquisition - the specific training and practice of language targets.

Metalinguistic knowledge – the ability to label, explain, or manipulate *linguistic* concepts, and encouraging the child to reflect on the *language* process itself.

Metacognitive awareness - the recognition of learning strategies and processes more widely generalizable to tasks other than language.

In addition, a sample of 12% of ongoing interventions was videoed to confirm treatment fidelity and these were rated in order to ascertain any difference between these therapy programmes and those implemented in later phases of the project.

Videos were scanned and rated in the same way as the questionnaires i.e. in terms of areas being addressed, and skills acquisition versus metalinguistic or metacognitive emphasis. In addition, the videos were checked for the presence of behaviours specifically identified as 'mediational' in the manual supplied to SLTs for the experimental phase of intervention. This was in order to identify whether therapists spontaneously used mediation in the therapy, prior to its introduction following the DA.

In this way, methods used in the first intervention phase could be directly compared to interventions used in the second, experimental or control phases, when the participating SLTs had been given guidance regarding Mediated Learning Experience, which they were asked to incorporate into their therapy. Presence of some mediational behaviour was anticipated in the sessions during the experimental intervention, and an increase in the use of mediation would confirm that changes had been implemented by SLTs, as intended by the experimenter.

Videos were viewed and the experimenter identified each instance of mediation of the four types identified in the manual i.e.

- mediation of intentionality
- mediation of meaning
- mediation of transcendence
- mediation of a feeling of competence

In addition, six techniques were presented in the manual, provided to SLTs in the experimental phases of intervention under the heading of 'How do I mediate?' These were:

- Ask process questions - usually containing 'How?'
- Bridge to different applications
- Challenge the child to justify his answers
- Teach about Rules
- Emphasize order, system, sequence and strategy
- Create Task -Intrinsic Motivation

Each technique was further elaborated by sample questions and strategies that could be used (see Appendix IX). Each time that one of these questions or prompts appeared in the videoed intervention session was counted by the experimenter.

The Second Intervention Phase

In the second intervention phase, all SLTs working with participants allocated to the control group were asked to continue ongoing programmes of intervention, as prior to the study, and in Phase 1. Results of the CELF-3, were copied to SLTs, although as they reflected change and practice effects that resulted from repetition of the test within a 4-month period, they did not yield the standardised information that was provided by the first administration of the CELF-3.

The SLTs working with participants allocated to the experimental group were contacted personally when the testing was completed, and an appointment made to discuss the findings. Some of the discussions were held by telephone instead of personally. The SLTs were provided with the 6 page 'Manual for the SLT delivering the experimental intervention' (Appendix IX), as described in section 2.2. The experimenter reviewed the manual briefly, and focused the

discussion on the findings from the DA that were to be used in planning the intervention.

The findings of the DA were presented to the SLT in a written report, of approximately 1-2 A4 sheets in length. The structure of the report mirrored the structure of information gained described in the Manual, and information was presented under five headings as described above. The findings were then summarised and recommendations for intervention included (see sample report in Appendix XII).

The experimenter discussed her findings with the SLT, which enabled her to compare perceptions of the child. The SLT was able to confirm or counter the experimenter's observations, and the experimenter was able to present aspects that had emerged from the DA. These were frequently in the areas of metalinguistics and metacognition which were explicitly addressed in the DA, but may not have been previously assessed. Furthermore, the SLT was able to identify findings that emerged from the DA which were not typical of the child.

The experimenter and the SLT then discussed modifications to the existing targets that the SLT had supplied, the addition of syntactic or metalinguistic targets, or the maintenance of existing targets. In some cases, methods or materials were also discussed. Specific targets for intervention were not written collaboratively, but by the SLT alone. SLTs were encouraged to incorporate mediational methods of intervention, as explained in the manual, but this was given less emphasis than the use of further information derived from the DA. The experimenter assured SLTs that she was available to explain, clarify or discuss any of the information at any time during the experimental therapy phase.

The SLTs participating in the study were able to do so as minimal disruption to their usual timetables was required. SLTs were asked only to maintain a regular timetable of intervention across both phases of therapy, in order that progress in the second phase would be comparable with that in the first, baseline phase, and both intervention phases spanned an equivalent length school term. Within this requirement, however, there was some variation in the frequency of therapy sessions carried out with each child. Furthermore, some children had individual therapy sessions while others worked with their SLT in pairs or small groups. In these groups, children from the experimental and control groups sometimes worked together, on similar or common targets. As the SLTs had access to the DA feedback for the experimental group children only, they could have made use of that information to modify targets or strategies for the individual children, but in some instances control group children would have also worked on modified targets and/or with modified strategies of intervention.

 A few weeks after the testing was complete and results had been supplied, SLTs were again asked to complete a short questionnaire in order to record the intervention currently in use with participants in both control and experimental groups. These were analysed in the same way as in the previous stage, and results compared to those obtained in intervention phase 1.

6.5   **Data Analysis**

The total raw scores of the CELF-3 were taken at four time points and used to statistically determine whether the children's language changed over time. The scores were also used for the control and experimental groups separately in order to evaluate whether the changes were different in the groups after intervention programmes had been modified. DA scores were similarly compared to measure

changes in the two groups over time, and the RtM and teachers' questionnaire results were used qualitatively for the same purpose, to detect changes in the children's behaviour or functional skills.

The perceived usefulness of the data derived from the DA in planning intervention for the children was assessed by comparing the data about targets and methods supplied by the SLTs for each intervention period. Each target was rated according to which area of language was addressed, and whether the methods targeted the acquisition of skills, metalinguistic knowledge or metacognitive awareness. The numbers of these were totalled and compared at each time and for each group to determine whether the SLTs had been able to incorporate information supplied from the DA reports, and made changes to their planned intervention. Data from the video recordings was used to determine whether the instructions for mediational intervention supplied in the manual had been sufficient to enable SLTs to adopt more mediational style in their intervention.

Finally, in line with the design of the study as a series of case studies in which each child's progress during the first period of intervention should serve as his own baseline according to which further progress could be measured, individual differences were taken into account in the analysis of results. Changes in scores during the periods of experimental intervention were compared to progress during baseline therapy. Factors affecting the performance or profiles of individual children during the study were considered in the case study series.

Non-parametric statistics have been applied wherever possible because of relatively small sample size and especially when examining the DA scores which use an ordinal rather than an interval scale. However, for more complex analyses in Chapter 7, where no

equivalent non-parametric analyses are available, mixed ANOVA and ANCOVAs have been used.   Statistical advice on this matter has been sought and suggests that the analyses are robust enough without transformation. Repeated measures / mixed ANOVAs have used the linear model option (because a linear trend in intervention and development data is expected) which minimises lack of power through sample size.   Means and SDs have been reported throughout for ease of reading.

CHAPTER 7

IMPLICATIONS OF DYNAMIC

ASSESSMENT FOR

INTERVENTION: RESULTS

CHAPTER 7   IMPLICATIONS OF DYNAMIC ASSESSMENT FOR INTERVENTION: RESULTS

7.1   Was the altered intervention effective?

7.1.1 Standardised measures, the CELF-3(UK)

All 24 children with SLI identified for the study to evaluate the DA, were tested at times 1,2 and 3. Sixteen participants were followed up at T4, six months after testing at T3. Nine of the 11 control group participants were retested on the CELF-3, the remaining 2 having transferred to secondary school where there was no means of ensuring consistent continuation of the intervention they had been receiving from SLT at their previous school. Seven of the original 13 experimental group participants were retested, one having transferred to secondary school, and the remainder having been retested at their own schools who were unable to continue participating in the project.

The data derived from each group's performance on the CELF-3 at each of four time points is summarized in Table 11. A two factor mixed ANOVA, with one related samples factor of Time (T1, T2, T3, T4) and one independent samples factor of Group (Control vs Experimental) was conducted on CELF-3 raw scores. A statistically significant main effect of Time was found, $\underline{F}(3,42) = 20.79$, $p < .001$, effect size $\eta_p^2 = 0.598$, indicating that in general the participants improved in their CELF scores over time. No significant effect of Group was found on CELF-3 scores, $F(1,14) = .301$, $p = .592$. No significant interaction was found between Time*Group, $F(1,14) = .879$, $p = .364$ The data are represented in Figure 7.

Figure 7. Total Raw scores on CELF-3 for each group, over time.
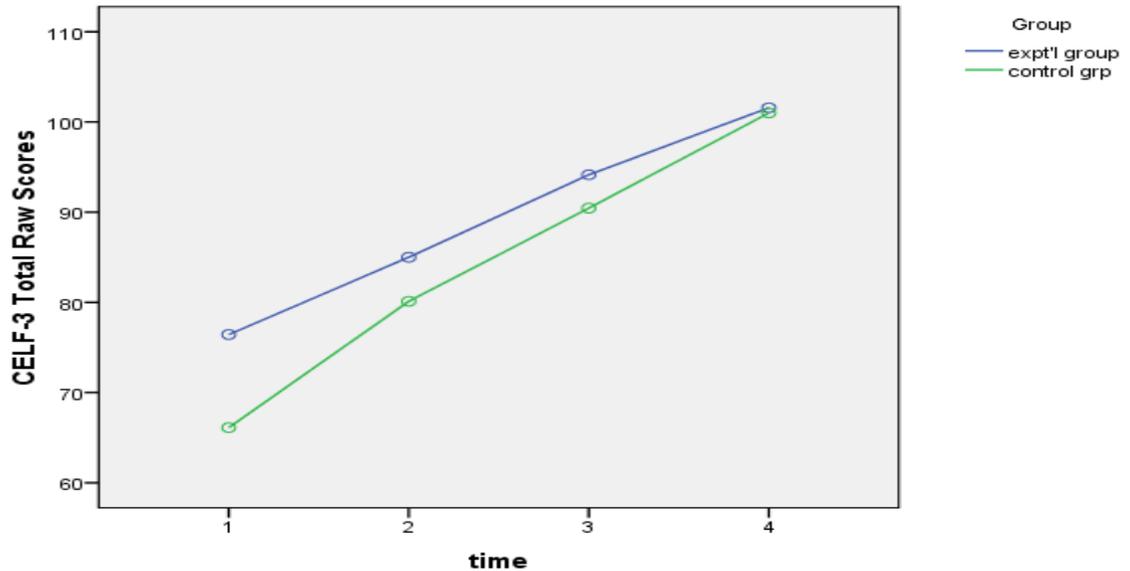
Independent samples t-tests with groups as the independent variable were performed at each time point and the results are contained in Table 11.

Table 11. Mean raw score on the CELF-3 for each group at each time point.

	Time 1		Time 2		Time 3		Time 4	
	Mean	Std dev	Mean	Std dev	Mean	Std dev	Mean	Std dev
E Group	76.43 n=13	20.19	85.00 n=13	23.03	94.14 n=13	18.80	101.57 n=7	25.78
Control Group	66.11 n=11	14.63	80.11 n=11	13.94	90.44 n=11	18.56	101 n=9	24.78
Group Compare	t	p	t	p	t	p	t	p
	.7	.491	.416	.681	-1.88	.852	.045	.965

The results confirm that there is no significant difference between the 2 groups at any time, suggesting that the differing interventions

between T2 and T3 did not result in differentiation between the groups.

Further within-group tests using related samples t-tests comparing scores between each different time point for the whole cohort showed significant differences between the scores at Time 1 and Time 2, $t(23) = 5.4$, $p < .001$, Time 2 and Time 3, $t(23) = 4.02$, $p = .001$, and between Time 3 and Time 4, $t(23) = 2.26$, $p = .039$. These results indicate that the whole cohort improved in their performance on the CELF-3(UK) at each time of testing, suggesting that their regular ongoing intervention (combined with a possible practice effect of testing and mediation from the DA) was improving their performance.

Despite raw score gains in 22 children, in only 8 participants did the percentile rank change, in 14 children the improvement was not sufficient to alter their overall standard score. Thus if using and reporting only the standard scores of the standardised test, these children would not be seen to have improved. Time 1 to Time 2 change can be used as a baseline level of improvement for each child for this length of time, against which further improvements can be measured.

7.1.2 DA measures

The DA was used in this analysis to detect differential changes between groups at each time period. The test would reflect changes in both language learning affecting the current task, and learning or problem solving strategy.

The data derived from each group's performance on the DA at each of three time points is summarized in Table 13. Mann-Whitney U test for independent samples with groups as the independent variable

were performed at each time point and the results are presented in Table 12.

Table 12. Group Performance over time on the Dynamic Assessment

	Time 1		Time 2		Time 3	
	Mean	Std dev	Mean	Std dev	Mean	Std dev
Experimental Group (n=13)	62.62	23.613	56.85	23.731	52.31	21.246
Control Group (n=11)	60.91	17.790	46.82	9.816	45.82	11.007
Group Comparison	U	p	U	p	U	p
	68	.865	62	.608	65.5	.733

The results show that there is no significant difference between the 2 groups at any time.

A two factor mixed ANOVA, with Time (T1, T2, T3) (related samples) and Group (Control vs Experimental) (independent samples) as the two factors was conducted on DA scores. A statistically significant main effect of Time was found $F(2,44) = 16.199$, $p < .001$, indicating that in general the participants improved in their DA scores over time. No significant effect of Group was found $F(1,22) = .681$, $p = .418$, and there was no significant interaction was found between Time*Group $F(1,22) = 1.147$, $p = .296$. The data are shown in Figure 8.

Figure 8.   DA scores for groups over time
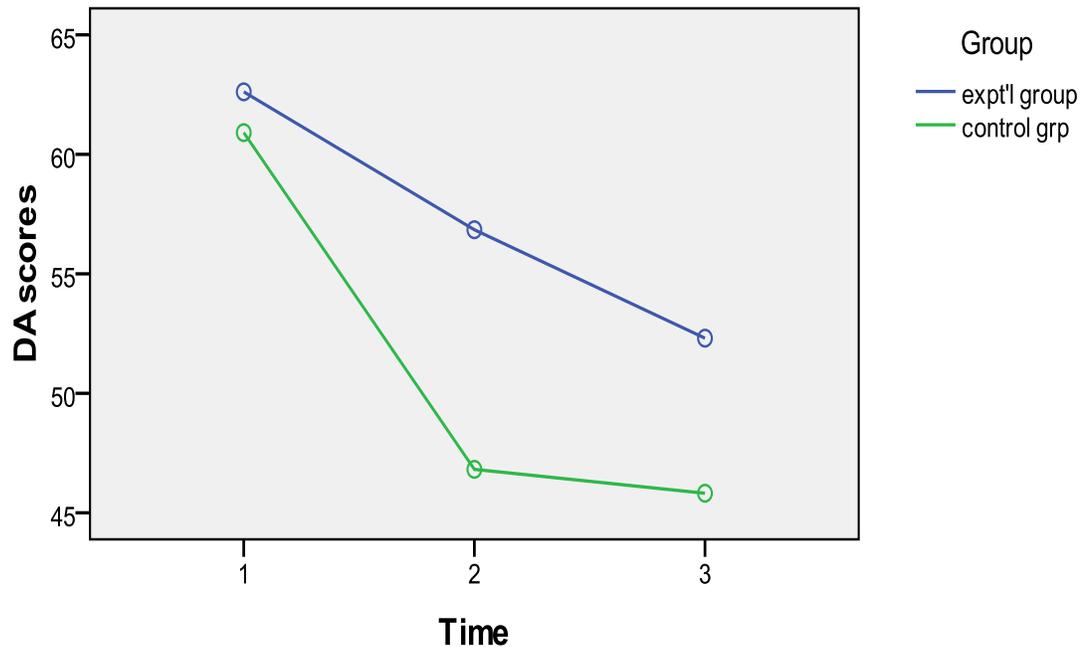
Follow up tests using non-parametric related samples Wilcoxon signed ranks test showed significant differences between the scores at Time 1 and Time 2, Z = 2.786, p < .05. However difference between Time 2 and Time 3 was non-significant, Z = 1.548, p = .122. This suggests that the improvement in language and learning evident after the first phase of baseline testing was not further changed after the second phase. As the T2-T3 change in the CELF score was significant, we might conclude that the change in language ability improved, whilst the strategy learning component plateaud after the second administration of the DA. There was, however, considerable individual variation within the cohort, with some children improving to a greater extent between T2 and T3, while others plateaud and a small number scored less well (a higher

score) on the DA at T2 or T3. The raw data is presented in Table 13 and the variation is explored further in section 7.4.

Table 13. DA scores for all children at Times 1, 2 and 3

Child	Age at T1	DA at T1	DA at T2	DA at T3
BH3	9;3	51	35	35
BH4	9;7	45	41	39
BH6	9;7	83	78	59
BH7	9;8	54	55	38
BH9	10;0	65	35	51
BH11	9;1	53	58	49
CH1	9;1	45	44	34
CP1	9,2	52	43	43
CP2	8,9	85	50	61
CP3	10,9	47	44	45
D1	8;2	96	62	64
D2	10.3	41	44	39
D3	8.5	87	85	89
D4	8.1	72	86	62
D5	10,7	48	35	30
R1	8;4	101	61	63
R2	9;8	29	33	37
R3	9;1	77	55	55
R4	9;3	48	45	39
TA1	9;3	48	48	40
TA2	10;0	57	41	35
TF1	8;2	102	105	99
TF4	8;7	52	36	35
TF5	9;11	46	35	43

In summary, straightforward intervention effects were not evident in this study. Statistical analysis therefore supports the null hypothesis, i.e. that the outcomes of intervention programmes informed by the additional data derived from the DA are not significantly different from the outcomes derived from the ongoing therapy programmes previously devised by the SLTs working with the participants.

Since the literature on Dynamic Assessment often inherently assumes that DA will improve intervention, and because intuitively, more detailed diagnostic information seems likely to enhance intervention, this result was unpredicted. Therefore, rather than leave the result at this, the study progressed to explore possible factors in explaining this result. These were:

- Sensitivity of measures – would different sub-tests or alternative measures reveal change?
- Application of DA information into therapy – did therapists report changing therapy, and was this evident in treatment validity video monitoring?
- Individual variation – were there some children for whom the added DA information was beneficial?

These factors will now be examined in turn.

7.2. **Which measures were most sensitive to change?**

7.2.1 **Sentence Assembly subtest of CELF-3(UK)**

Not all subtests of the CELF-3 would have been equally sensitive to the particular changes facilitated by the DA procedure and the subsequent interventions. The DA test procedure itself addressed the task of Sentence Assembly, and all children received mediation in this task during the test procedure. Improvement in this subtest over time would have been expected in all children as problem solving strategies for this task were mediated to children in both groups equally during the DA.

However, information regarding the specific difficulties experienced by children in the experimental group, and interventions to which they were responsive, were supplied to their SLTs at T2, and differences between the groups resulting from changes to their intervention, may have been seen at T3 and T4. Improvement over time was seen in Table 14.

Table 14. Means of raw scores in the Sentence Assembly Subtest of the CELF-3 over time

	TIME			
	T1	T2	T3	T4
Sentence Assembly Subtest Raw Score C. group mean	4.45	6.36	9.18	10.67
Sentence Assembly Subtest Raw Score E group Mean	4.77	6.92	8.23	10.57
Sentence Assembly Subtest Raw Score Full cohort Mean	4.60	6.67	8.67	10.60

The results of a two-factor mixed ANOVA with 1 within-subjects factor, Time of measurement (T1, T2, T3, T4), and 1 between-subjects factor, Group (Experimental vs Control) indicated a statistically significant main effect of Time, $F(3,42) = 22.857$, $p < .001$, $\eta_p^2 = 0.62$, indicating that the participants scored in general higher at each successive time (M=4.437 at T1, 6.270 at T2, 8.111 at T3 and 10.619 at T4; SEM= 0.928 at T1, 0.954 at T2, 1.277 at T3 and 1.134 at T4). The main effect of Group was not significant ($F(1,14) = 0.263$, $p = .616$). In addition, no significant interaction was found ($F(3,42) = 1.474$, $p = .235$).

Follow up tests using independent samples t-tests comparing scores of the   whole   cohort   at   each   time   point   showed   significant

differences between the scores at T1 and at T2 *(t*(23) = -4.192, *p* < .001) the scores at T2 and T3 (t(23) = -3.483, *p* = .002) and the scores at T3 and T4 (*t*(15) = -2.81, *p* = .013)

These results indicate that the performance of both groups improved significantly with each successive retest, but that no difference between the rate of improvement of the two groups was statistically significant. Non-significant but slightly greater change was detected in the control group at T3 and in the Experimental group at T4, thus mirroring (and having contributed to) the results obtained for the CELF-3 as a whole.

7.2.2 Formulated Sentence subtest of CELF-3(UK)

Mediation of sentence formulation strategies during the DA, and subsequent interventions addressing sentence level grammar, may have been expected to transfer to the 'Formulated Sentences' subtest of the CELF-3. The processing required by this task is close to that accessed by the task used in the DA procedure, and scrutiny of this subtest individually might be able to detect improvements in the study participants. Results of the subtest over time are presented in Table 15.

Table 15. Means of raw scores in the Formulated Sentences Subtest of the CELF-3 over time

	TIME			
	T1	T2	T3	T4
Formulated Sentences Raw Score C. group Mean	14.18	17.45	19.81	20.89
Formulated Sentences Raw Score E group Mean	15.46	17.07	17.84	20.00
Formulated Sentences Raw Score Full cohort Mean	14.87	17.25	18.75	20.50

The results of a two factor mixed ANOVA with 1 within-subjects factor, Time of measurement (T1, T2, T3, T4), and 1 between-subjects factor, Group (Experimental vs Control) indicated a statistically significant main effect of Time ($F(3,42) = 12.585$, $p < .001$, $\eta_p^2 = .473$), indicating that the participants scored in general higher at each successive time (M=13.619 at T1, 16.230 at T2, 18.278 at T3 and 20.444 at T4; SEM= 1.605 at T1, 1.648 at T2, 1.557 at T3 and 1.675 at T4). The main effect of Group was not significant ($F(1,14) = 0.01$, $p = .923$). In addition, no significant interaction was found ($F(3,42) = 0.602$, $p = .617$)

Follow up testing using independent samples t-tests comparing scores of the whole cohort at each time point showed significant differences between the scores at T1 and at T2 ($t(23) =2.737$, $p = .012$) and the scores at T3 and T4 ($t(15) = 2.15$, $p = .048$), but only borderline significance between the scores at T2 and T3 ($t(23) = 1.9$, $p= .07$)

These results indicate that the performance of both groups improved significantly in the initial baseline period of intervention, and also in the final period of intervention when both groups were receiving informed intervention, but not in the intervening period of therapy when the control and experimental groups received different programmes, the SLTs of the experimental group having received additional information and recommendations for intervention, but the control group continuing in their ongoing baseline intervention programmes. There was no significant effect of group, however, and no significant interaction.

The similarity between the results in the SA and FS subtests and the Total CELF scores suggests that the Total CELF scores are

representative of the performance of the participants in expressive sentence construction, and not unaccountably altered by variation in the participants in receptive, semantic or memory subtests of the CELF.

Alternative measurements, examining different criteria for change over time, might reveal progress in areas of language not addressed by the CELF.

7.2.3 Response to Mediation Scale.

The Response to Mediation Scale (Lidz 2003) was completed alongside each episode of testing with the DA (see section 4.5). Each criterion was rated on a scale of 1-5, in which 1 was least responsive, and 5 closest to typical behaviour. The scores for all 11 items (Max score =55) were totalled for each child at each time, and the mean scores for each group are shown in Table 16.

Table 16. Mean scores on the Response to Mediation Scale over time.

	Time 1	Time 2	Time 3
Experimental group mean (n=13)	41.85	43.77	46.50
Control group mean (n=11)	47.36	47.23	45.23

The results of a mixed design ANOVA with one within subjects factor, Time of measurement (T1, T2 and T3), and one between subjects factor, Group (Experimental and Control), indicated non-significant main effects of Time, ($F(2,44)$ = 0.571, p = .569) and Group, ($F(1,22)$ = 1.526, p= .230) but a significant interaction ($F(2,44)$ = 4.128, p = .023, η_p^2 = .158). The interaction was followed up by performing independent samples t-tests comparing experimental and control groups at each time point. The results showed a significant difference between the groups at Time 1, ($t(22)$ = -2.419, p = .024),

but no significant differences at Time 2 (t(22) = -1.329, p = .197) or Time 3, (t(22) = .486, p = .632). This indicates that the groups were unequal in their responsiveness according to the Response to Mediation scale at the start of the study, with the children in the experimental group rated less responsive than the control group children, but that the difference between the groups reduced over the course of the study. T-tests comparing changes over time for each group showed non-significant changes in the control group from T1 to T2, (t(10) = .119, p = .908),   T2 to T3 (t(1) = .944, p = .367),   and overall change from T1 to T3 (t(10) = 1.194, p = .260),   In the experimental group, the results showed a non-significant difference between ratings at T1 and T2 (t(12) = -.965, p = .353),   difference just short of significance at Time 2-Time 3 (t(12) = -1.939, p = .076) and a significant difference overall from T1 to T3, (t(22) = .486, p = .012). This indicates that the experimental group showed a significant benefit from therapy compared to the control group.
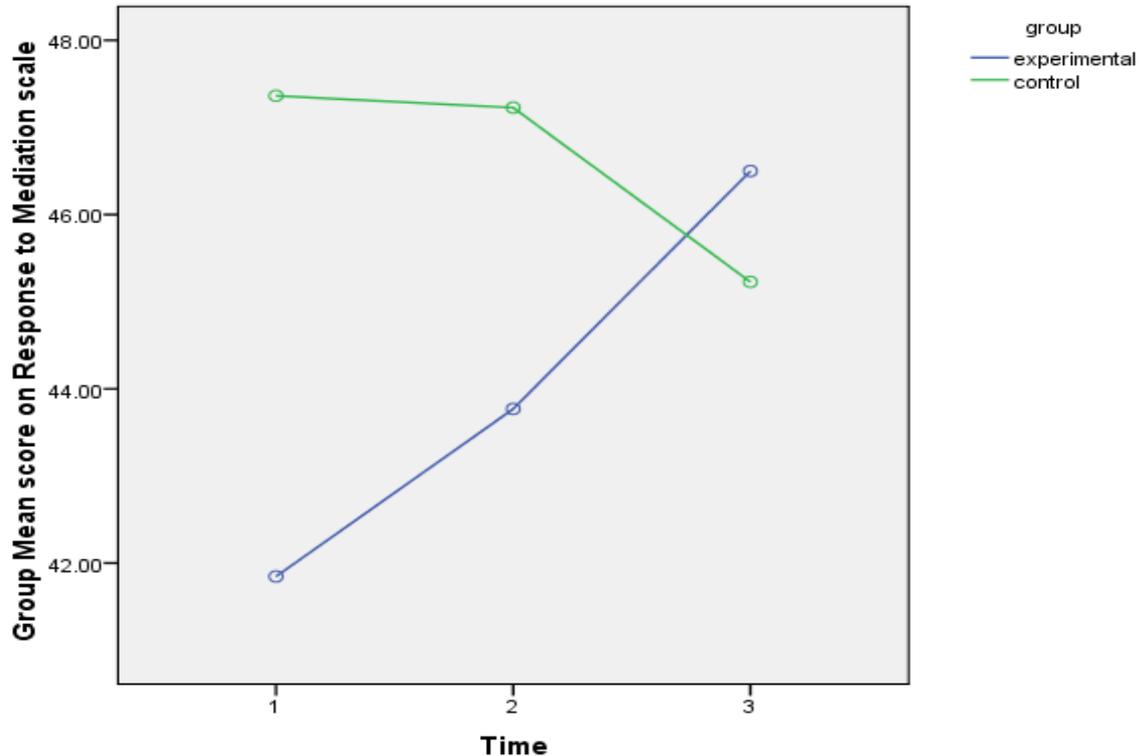
Figure 9. Group Mean scores on the Response to Mediation Scale over time.

The items on the Response to Mediation scale were independently rated, as described in section 4.7.7 by a rater blind to group allocation. A moderate, significant correlation was found between the scores of each rater, based on total scores for each child (r_s = .570, p = .004).

Changes in some behaviours occurred in both directions in many of the children. While the majority (15/24) children scored more positively at T2, several children scored less well at T2, with the rating reflecting sessions in which they were less attentive, more emotional, reflecting anxiety when they found the task difficult, or less communicative with the examiner. Familiarity with the examiner and the task altered the nature of the sessions considerably with

many children being less shy and inhibited, and more inclined to converse with the examiner, but at the same time less anxious and less focused on succeeding at the task.

Correlations were performed between the Response to Mediation at T1 scores and the T1-T4 change scores, using Total Raw scores on the CELF. Correlations were non-significant at the .05 level, indicating that the Response to mediation scale did not independently predict longer term outcomes on the CELF.

In general however, the scores obtained on this scale were high, with many approaching ceiling (25% of participants scored between 50 and 55 at T1, although statistical exploration of the whole cohort showed normal distribution). This may be a result of stringent selection criteria for the study that specifically excluded children with learning, behavioural or socio-emotional difficulties such as ADHD or ASD, although criteria for evaluating attention and emotional control are part of the Response to Mediation scale.

Evaluation by Item.

On two items in particular, scores across the group of children were particularly poor. 'Strategic problem solving', reflecting whether participants actively planned their responses, elicited a low overall score across the whole cohort of participants. Similarly, 'Response to Challenge' reflected that most of the children required some degree of encouragement to persist with the task, rather than being intrinsically motivated by the task, or sufficiently able to persist in the task. Finally, the lowest score overall was the rating for 'Evidence of self-talk when working on a challenging task'. Many of the children did not verbalize their own problem solving, as would be expected in a group of children with LI for whom verbal mediation is a weak channel for learning.

Inspection of Table 17 reveals a slight trend towards improvement over Time for each of the criteria, the trend being more marked in the experimental group than the control, even at T2, when both groups had continued in their ongoing intervention programmes.

Table 17. Mean ratings on 3 items of the Response to Mediation scale, over time.

	Problem solving			Response to challenge			Self talk		
	T1	2	3	T1	2	3	T1	2	3
Mean scores C. group	4	4.2	4.4	3.7	3.9	4	3.6	3.7	3.8
Mean scores E group	3.2	3.8	4.2	3.4	3.7	4	2.8	3.1	3.5

Of the 6 children with the highest (i.e. poorest) DA scores, only 2 were scored poorly on the Response to Mediation scale. These 2 emerged as the weakest participants that were unable to improve at any time on the DA, due to difficulty with the task itself. Children who struggled with the task and needed high levels of prompting at T1, but who made significant gains at T2, scored well on the Response to Mediation scale, i.e .demonstrated good skills of engaging with the task and examiner, and in particular, 'Response to Challenge'.

In summary it would appear that the Response to Mediation scale was sensitive to an interaction that reflected a trend towards greater improvement in the experimental group after the first DA, and after

experimental intervention. The children in the control group performed significantly better at the start of the study, suggesting that they had better strategies for learning, although the CELF did not find that they had significantly better language. This difference was reduced at the end of the baseline intervention period, when the children's responses during the DA showed changes in the measures of responsiveness. It was found, however, that only some items in the Response to Mediation scale were sensitive to the changes in learning potential shown by the children, and this will be further explored in the Discussion.

7.2.4 Teachers' reports

A scale for detecting and rating functional difficulties according to the teachers of the children, was devised (see section 6.4.2.2 and Appendix X). Rating scales were completed by teachers for all 24 of the participants at T1, however at T2, one form was not returned, and at T3, the teacher of 3 children in one school failed to complete the sheets fully, and the teacher of one child refused to complete the sheet, thus full sets of data are available for only 19 of the children.

As the rating scale contained some items for which 1 reflected the 'best' score instead of the 'worst', to decrease the likelihood of response set, the scores for these items have been reversed in the collation of this data. Lower totals (minimum possible=15) reflect the greatest degree of difficulty. Findings for each group at each time of testing are recorded in Table 18.

Table 18. Mean ratings on teacher's questionnaire, over time

	Time 1	Time 2	Time 3
Control group mean score	44.1 (n=11)	47.1 (n=10)	46 (n=10)
Experimental group mean	43.7 (n=13)	49.9(n=13)	45.95 (n=10)

The results of a mixed design ANOVA with one within subjects factor, Time of measurement (T1, T2 and T3), and one between subjects factor, Group (Experimental and Control), indicated significant main effects of Time, ($F(2,34) = 3.093$, $p = .058$, $\eta_p^2 = .154$) but non-significant effect of Group, ($F(1,17) = 0.064$, $p = .803$) and non-significant interaction ($F(2,34) = 0.572$, $p = .569$)
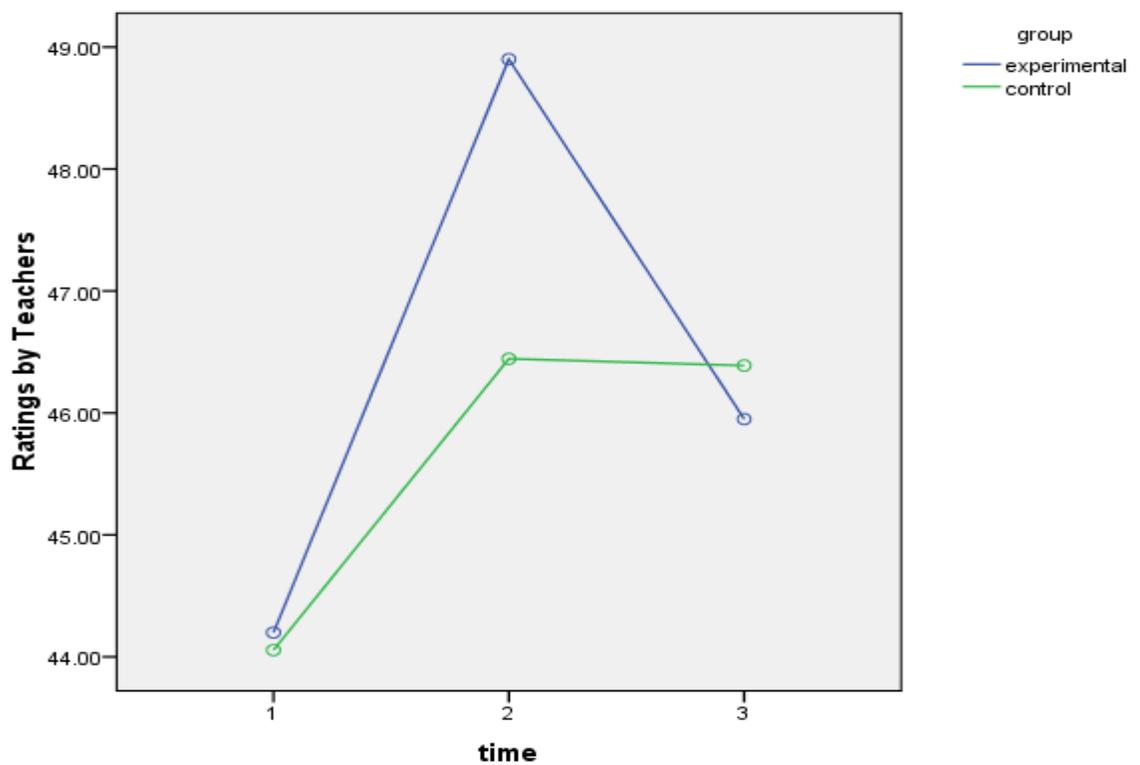
Figure 10. Mean scores on teachers' questionnaire, over time.

Follow up tests using non-parametric related samples Wilcoxon signed ranks test showed significant differences between the scores at Time 1 and Time 2 (Z = 3.25, p = .001). However difference between Time 2 and Time 3 was non-significant (Z = .807, p = .420).

Further follow up tests using mixed ANOVA with one related samples factor of Time (Teachers' ratings at T1, and at T2) and one between samples factor of Group (Control vs Experimental) as the two factors, revealed a significant effect of Time $F(1,21)$ = 13.329, p = .001, and non significant effect of group $F(1,21)$ = 0.192, p = .665 as well as non-significant interaction Time*Group $F(1,21)$ = 1.847, p = .188. This confirms that the improvement in the baseline phase between Time 1 and Time 2, was not different between the experimental and control groups.

Several children in both groups were rated improved at T2, but scores dropped again at T3, suggesting variability in the ratings of children's difficulties at different times, possibly related to demands made on them at different times in the school year rather than any DA intervention effect.

Items that described areas of difficulty for numerous children were 'problem solving or working through a task', and 'explaining what they are thinking or feeling'. Eighteen of the 24 children were described by their teachers as taking 'quite a lot' or 'a great deal' of care over their work.

Correlation between the teachers' ratings of children's difficulties at T1 and their performance on the CELF at T1 was non-significant (r_s = .32, p = .128), suggesting that teachers' ratings of the children's functional difficulties were based on different, probably non-linguistic criteria. Only one child whose scores on the CELF increased

substantially over time was also rated by his teacher to improve considerably over time. There was also no correlation between the teachers' ratings at T1 and the DA scores at T1, (r_s = .062, p= .775), again suggesting that the two measures were assessing different skills, and that the DA was not sensitive to the functional difficulties detected by the teachers nor was the teachers' scale sensitive to the language learning criteria assessed by the DA.

Correlations were performed between the teachers' ratings at T1 and the T1-T4 change scores, using Total Raw scores on the CELF. Correlations were non-significant (r_s = .380, p = .147). It can be seen that the children scoring well on the teachers' ratings scale did not go on to improve more than those for whom the teachers identified functional difficulties.

In summary, the teachers' ratings did not appear to be reliable indicators of change in the children as they were variable over time. The rating scale devised for teachers was not able to predict change in children's language scores over time as measured by the CELF-3(UK).

The only measure adopted to have demonstrated differences between the control and experimental groups in their progress in therapy after addition of data from the DA was the Response to Mediation Scale which suggested that the experimental group made significant progress over Time that reduced the difference between their rating and that of the Control group that was evident at the start of the study. The next possibility to explore would be whether the intervention was actually changed by the information supplied and significantly different programmes of intervention were indeed implemented.

7.3. Was the information derived from the DA used/useful in planning intervention?

7.3.1. Therapists' Feedback

A qualitative questionnaire (see Appendix XIII) was used to find out whether the participating SLTs found the information contained in the children's reports after the DA, useful. Responses were received from eight out of ten SLTs, from six of the seven schools involved in the study. Responses to questions are summarized in Table 19.

Table 19.   Feedback from SLTs

	Do you think the information supplied to you by the experimenter about the children, after they had had the DA, was useful?	Were you able to make use of this information in your planning of intervention for the children?	Do you think it made any difference to the outcomes of your therapy?
No of SLTs (n=8)	Yes - 6 No - 2	Yes- 5 Partly - 1 No - 2	Yes – 4 Uncertain – 3 No - 1

The SLTs were not certain, however, that the information and planning influenced the outcomes of their intervention. Open questions elicited the following observations about changes noted in the participants:

- feeling more positive
- asking more questions,
- understanding their own targets better

- being more motivated
- generalising strategies

A compilation of all the responses received from SLTs is contained in Appendix XIV.

7.3.2 Target setting

The targets of intervention in each phase of the study were monitored by means of questionnaires given to the SLTs, in which they were required to state three targets of the intervention programmed for each child for the time period specified. Return rate for these questionnaires at each stage of intervention was 100%

Each target was determined to be addressing one out of 11 areas of language, compiled by the experimenter, as described in Chapter 6. The number of targets in each linguistic category was totalled by the experimenter separately for the experimental and control groups in therapy phases 1 and 2, as summarized in Table 20.

Table 20. Areas of Language targeted for intervention by SLTs

| | Therapy Phase 1 | | | Therapy Phase 2 | | |
| | Group | | | Group | | |
	Control	Expt'l	Total	Control	Expt'l	Total
Attention & Listening	2	3	5	1	3	4
Morphology	2	1	3	1	1	2
Narrative	1	1	2	2	1	3
Phonological Processing	0	1	1	0	1	1

Pragmatics	8	12	20	3	6	9
Semantic Organization	8	10	18	7	7	14
Sentence Comprehension	2	3	5	4	1	5
Sentence Construction	4	5	9	9	12	21
Speech	1	0	1	1	0	1
Vocabulary	5	3	8	4	1	5
Other				1	4	5
Memory					2	2

Inspection of the table of raw data revealed that there was a change from phase 1 to phase 2 that spanned both groups, with a decrease in the number of pragmatic targets, and an increase in the number of sentence construction targets. This corresponds with the nature of information supplied by the experimenter in the reports of the DA which had a focus on syntactic and sentence construction aspects, arising out of the task used in the DA.

Categories of language were simplified into four areas, namely
1. Grammatical – comprising morphology, sentence comprehension and sentence construction
2. Domain general – comprising attention and listening, memory and 'other' designations
3. Sound level – comprising speech and phonological targets
4. Semantics, Pragmatics and Narrative

McNemar analysis was conducted on the change in the four categories from Phase 1 to Phase 2 for the whole cohort and found to be significant (χ^2 = 15.8, p = .01), suggesting that there was

overall change in the targets set for both groups in the second period of intervention. The two groups were then examined separately to ascertain if the changes in targets were specifically related to the recommendations made for children in the experimental group. The findings are summarized in Table 21.

Table 21. Areas of Language targeted for intervention in each group

| | Therapy Phase 1 | | | Therapy Phase 2 | | |
| | Group | | | Group | | |
	Control	Expt'l	Total	Control	Expt'l	Total
Grammar	8	9	17	14	14	28
Domain general	2	3	5	2	9	11
Sound level	1	1	2	1	1	2
Semantics, Pragmatics & Narrative	22	26	48	16	15	31

When considered separately, the change appeared in both the experimental and control groups. However, McNemar analysis revealed that the change over time was marginally significant in the control group, (χ^2 = 7.57, p = .056), but was significant in the experimental group (χ^2 = 9.256, p = .026). The SLTs therefore reported changing the nature of their intervention targets to a greater extent in the group for whom additional grammatical, metalinguistic and metacognitive information was made available to inform the second phase of intervention.

The changes in both groups may reflect the influence of the information supplied by the experimenter to the SLT, which was

generalised to all participants. A small number of the children were actually managed in groups, with experimental and control group children working together. Thus intervention targets were applied to control group participants as well as experimental group participants, and inspection of the raw data supplied by the SLTs shows four instances of identical targets for more than one child.

Identification of the linguistic categories of targets was independently rated by an experienced paediatric SLT, who examined the response sheets for 6 participants at Phase 1, and all 24 at phase 2, blind to group allocation. At Phase 2, item by item agreement between the independent ratings and those of the experimenter for phase 2 was 79% exact agreement. The total numbers were affected by 4 targets left blank by the independent rater when she was unsure which category to select, and in these instances, the ratings were counted as non-agreement. A summary of the categorization by the two raters is shown in Table 22.

Table 22. Categorization of intervention targets by two independent raters.

	Experimenter rating	Independent rater
Grammar	28	24
Domain general	11	12
Sound level	2	2
Semantics, Pragmatics & Narrative	31	30

7.3.3 Methods of intervention

Responses to the SLTs' questionnaires were also used to determine whether targets and activities reflected 'skills acquisition', 'metalinguistic activities', 'meta-cognitive activities' or combinations of these, adapted from Law et al (2008). Ratings of the questionnaires were carried out for each group in phase 1 and again in phase 2. Results can be seen in Table 23.

Table 23. Categorization of Intervention methods in each group, across phases of intervention.

	Therapy Phase 1 Group			Therapy Phase 2 Group		
	Control	Expt'l	Total	Control	Expt'l	Total
Skills	16	21	37	15	10	25
Metalinguistic knowledge	16	14	30	14	20	34
Metacognitive awareness	1	4	5	4	8	12

Inspection of the table shows that the total number of skills only based interventions has decreased substantially in the second period of intervention, in the experimental group. There is a corresponding increase in the metalinguistic activities in the second period of intervention in the experimental group.

A McNemar analysis was carried out on the change in the three methods of intervention from Phase 1 to Phase 2 for the whole cohort, and found to be just short of significance ($p = .051$). Splitting the groups, however, revealed a non-significant difference

in the control group ($\chi^2 = 2, p = .572$), but a significant difference in the experimental group ($\chi^2 = 8.33, p = .04$).

There was, therefore a reported difference in the nature of the intervention activities planned by the SLTs for the experimental group children in the second phase of therapy that was not significant for the children in the control group. The results however should be interpreted with caution as the information provided by the SLTs proved difficult to consistently align with the rating criteria, and may not be strongly reliable.

7.3.4. Delivery of therapy

Changes to the intervention programmes implemented by SLTs were further monitored by a sample of videos (n=3 x 2 phases) as described in section 6.4.3. These videos verified the degree of fidelity of the SLTs' accounts of their intervention, and also demonstrated whether there was evidence of mediational techniques in the therapy in either phase.

Videos were rated for instances in which the following specific mediations were used; Intentionality, Meaning, Transcendence and Feelings of competence. The definitions of these were made available to the SLTs in the manual provided to SLTs with the reports of the DAs, at the start of the experimental interventions (see Appendix IX). However there was little use either before instruction or after, of specifically mediational techniques such as making clear to the child what is important about what he is learning (mediation of meaning), that the therapist intends to help him grasp a specific skill (intentionality), and how that skill would be useful in other contexts (transcendence).

Mediation to the child of a feeling of his own competence, defined in the manual as ''targeting praise so that the child learns what he has done well' was used more than 6 times by one SLT in Phase 1 and twice by two of the SLTs in phase 2. Instances counted were those specifically containing praise attached to a particular behaviour such as 'you've done very well with remembering your words'.

In addition, the presence of specific behaviours described as mediational in the manual were recorded and totalled, as shown in Table 24. The behaviours were clearly defined in the manual in cognitive terms, and although the instances identified in the recordings were more related to the linguistic tasks, this was found to be an objective process of counting and recording, and no checking by an independent rater was considered necessary.

Table 24. Incidence of occurrence of mediational interventions by SLTs in sample videos of therapy.

	Phase 1			Phase 2		
	Sample 1	2	3	1	2	3
Process Questions	0	0	0	2	2	4
Bridging to different applications	0	1	0	0	0	0
Challenging Answers	2	3	1	0	4	0
Teaching Rules	2	1	1	0	0	2
Emphasizing Order	4	5	1	1	0	1
Creating Motivation	0	1	0	0	0	0

It can be seen that there is some individual difference between SLTs, but there is little change from the first phase of intervention to the

second. All SLTs were inclined to emphasize order, most did this by introducing the session with a plan or visual timetable for the session. SLTs in both intervention periods, challenged the children to justify their responses which is an intrinsically mediational behaviour (Haywood 1993). Questions phrased as 'How do you know..?' were identified as Process questions, and these were used only by the SLTs in phase 2. Questions such as 'Does it sound right?' Or 'Does that make sense?' in which the children were only called upon to confirm their grammatical judgements, were rated as 'Challenging answers' and were used by all three SLTs in phase 1, and one in phase 2.

The use of Process questions by all the SLTs in phase 2, and none in phase 1, suggests some uptake of the recommendation to mediate using this type of challenge. The questions, however, were met with content based responses, rather than metacognitive reflections by the children, and the sessions contained little that would be described as facilitating children's awareness of their own problem solving.

In summary, it would seem that the SLTs involved in the study found the information provided by the Dynamic Assessments useful, and incorporated it into their planning of intervention, by adjusting their therapeutic activities. There was a slight shift in all the children towards more focus on sentence construction targets, but only the therapy for the children in the experimental group became more metalinguistic.

Despite the focus on more explicit linguistic interventions and self knowledge rather than practice of language structures, SLTs did not implement specifically mediational interventions, recommended to improve transfer of knowledge. It may be that while knowledge of

metalinguistics is within the remit of SLT, and the recommendation to focus on the child's explicit knowledge is taken up, the specifically mediational style of intervention was not sufficiently informed by the manual supplied to the SLTs, and they may in fact benefit from extended training such as that offered in Dynamic Assessment training courses. It may also emerge that it is the implementation of mediational intervention that is the crucial factor in eliciting significantly improved outcomes from intervention. This issue will be further explored in the Discussion.

7.4 <u>What was the effect of individual variation on treatment efficacy?</u>

The small number of participants resulted in the considerable variation within the groups affecting group data, and reducing generalizability of results. The two main sources of variation are the children themselves, and their educational placements.

7.4.1 Variation in the Participants

With regard to the participants, a considerable range of severity of linguistic impairment was apparent in the CELF-3(UK) raw scores (see Table 1, Section 4.2.3). Furthermore, eligibility testing included evaluation on the Ravens CPM (see Table 1, Section 4.2.3), and results of this revealed a considerable range in non-verbal reasoning skills. Ranges of ability were further apparent in the DA, the responses to mediation and the teachers' ratings. These individual differences will be explored further in individual case studies in section 7.5 and in Chapter 8.

Individual variation was reflected not only in the static and DA scores at the start of the study, but also in the response of each child to the regular ongoing therapy offered in the baseline phase. Pre-post therapy change controlling for baseline (T1 to T2) change was

investigated in the entire group of children using ANCOVA on test scores immediately pre-intervention (T2 for experimental group (n=13) and T3 for control group (n=9) and immediately post intervention (T3 for experimental group and T4 for control). Within group analysis with the whole cohort shows a significant effect of therapy ($F(1,20) = 8.96$, $p = .007$) once baseline development during traditional therapy has been controlled for (T2-T1).

A further analysis of individual variation was also carried out. Splitting the whole group into those who improved in the baseline period and those who stayed approximately the same was carried out using an arbitrarily selected cut-off score of +10 points change in the Total CELF-3 raw score which enabled the effects to be demonstrated another way. This resulted in one group (n=14) of 'improvers' at baseline, and one group (n=8) who did not improve by at least 10 points on the CELF-3 Total Raw score.

The results of a mixed design ANOVA with one within subjects factor, Time of measurement (CELF at T1, CELF Pre-experimental intervention, and CELF Post-experimental intervention), and one between subjects factor, Group (Improvers and non-improvers at baseline), indicated significant main effects of Time, ($F(2,40) = 26.123$, $p < .001$) and Group, ($F(1,20) = 5.539$, $p = .029$) and significant interaction ($F(2,40) = 6.381$, $p = .004$). The interaction is shown in Figure 11.
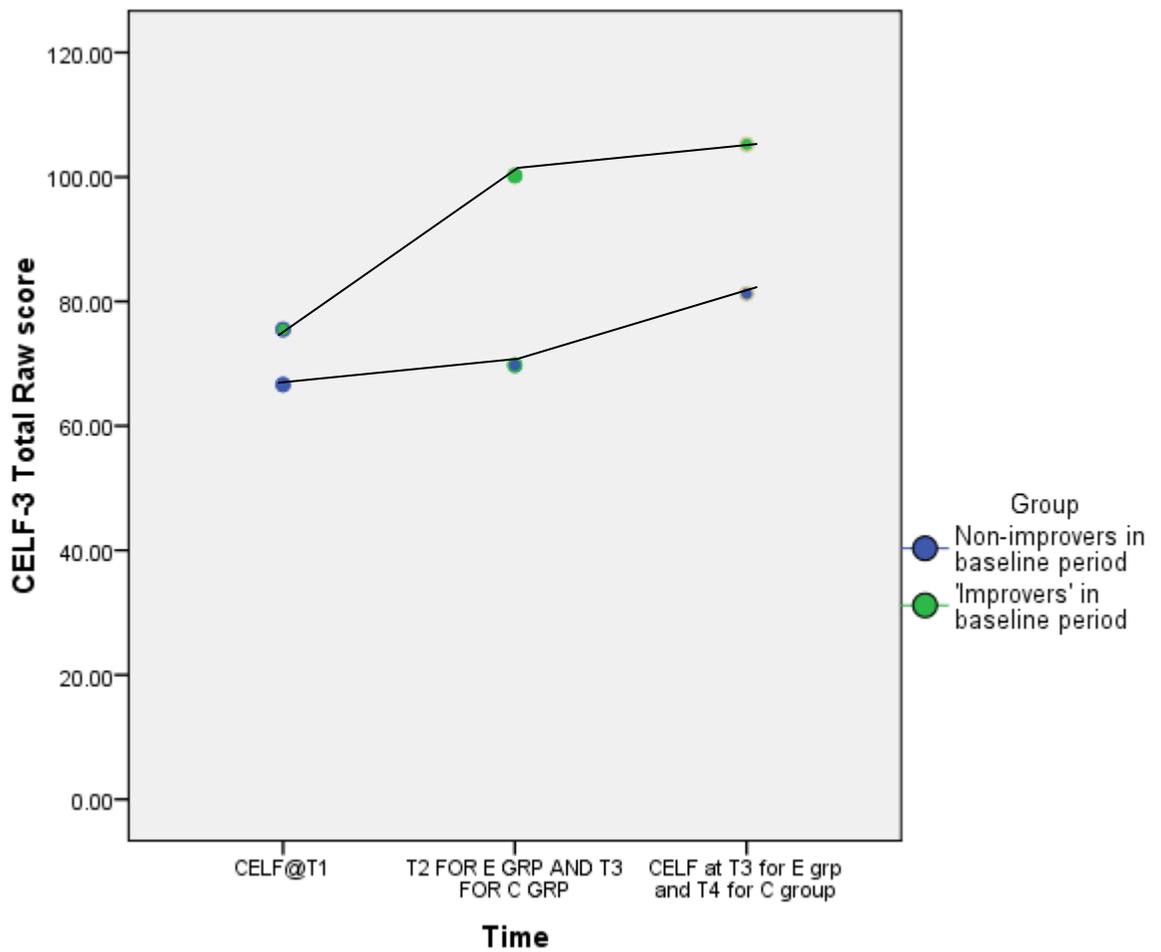
Figure 11. Raw Scores on CELF-3 of Improvers and Non-improvers according to time of test.

The graph illustrates that the group of children designated 'Improvers' who improved more than ten points in the baseline period of ongoing intervention, did not benefit significantly in the period of experimental intervention ($t(13) = 1.159$, $p = .267$). This suggests that they were improving in their original therapy, and changing the targets or style of the intervention was not beneficial. Those 'non-improvers' who benefited less from their ongoing intervention, made significant gains in the period of modified teaching in the experimental intervention ($t(7) = 5.555$, $p = .001$). This group included several of the children with the greatest

amounts of difficulty, according to the CELF-3, but also at least one high scoring child. The results also suggest that the significant effects of the modified intervention are visible within the first four months of intervention, and the development of strategies for language learning is not only beneficial in the long term.

7.4.2 Variation in educational placement

Intervention in the current study was carried out by ten SLTs in seven schools. There was one special school for children with speech, language and communication needs (SLCN). The remainder were Language Units and Resource Bases, four from Hertfordshire, and two from inner-London schools. Mean scores on the CELF-3(UK) obtained at the special school, suburban and city language bases over four time points were examined to determine if results obtained were affected by the school attended by the children. Mixed design 2-way ANOVA, with one within subjects factor, Time of measurement (T1, T2, T3, T4) and one between subjects factor, School (Special school, suburban, Inner London) was carried out.

Mauchly's test showed that the assumption of sphericity had been violated for the main effect of Time (Mauchly's $W(2) = .688$, $p < .05$); therefore, degrees of freedom were corrected using the Greenhouse-Geisser estimates of sphericity for the main effect of Time ($\varepsilon = .762$). Using this adjustment, the analysis revealed a significant main effect of Time ($F(1.5,32.02) = 28.9$, $p < .001$ $\eta_p^2 = .579$). The main effect of school was not significant, ($F(2,21) = .688$, $p = .514$), and Time*School interaction was also not significant ($F(3.05, 32.02) = .956$, $p = .426$).

These results reflect the changes in CELF-3 score over time, previously identified, but suggest that there was no difference according to the school attended by the children.
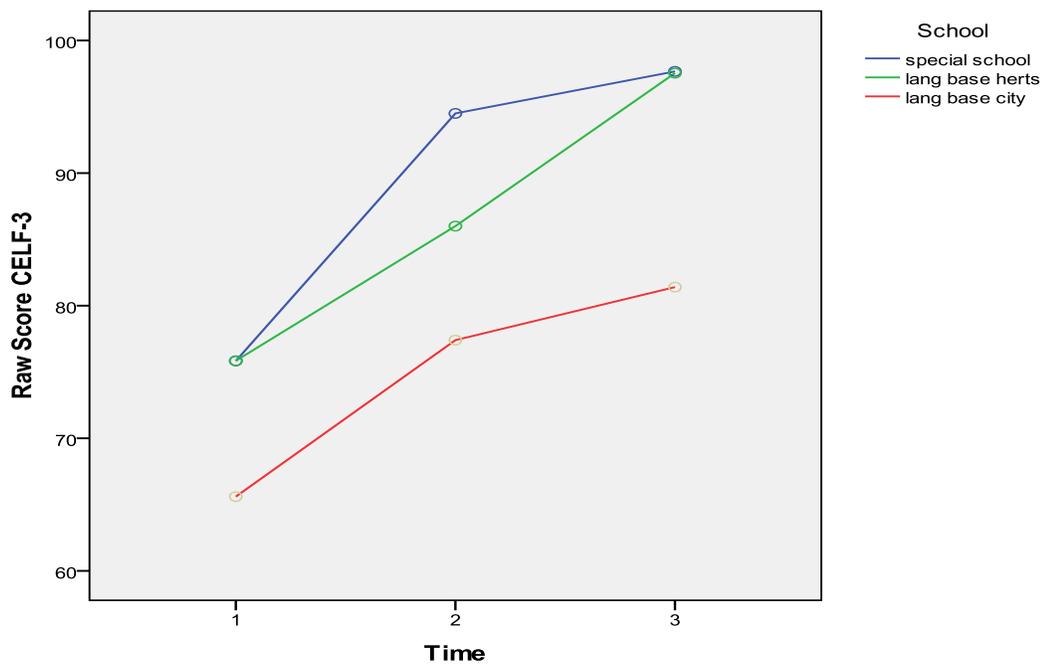
Figure 12. Scores on CELF-3 by participants in different schools, over time.

Another factor of note (see section 6.4.3) was that there was variation in the length and frequency of intervention sessions between schools and participants in the study, although dosage of intervention was kept consistent from phase 1 to phase 2 of therapy for each individual. As some SLTs worked with the children in pairs or small groups, and these groups sometimes found children from the experimental and control groups working together on similar or common targets, there was a possibility that control group children would have also worked on modified targets and/or with modified strategies of intervention.

7.5 Case Studies

Examination of individual case data revealed that some small groups of participants were linked by similar profiles of test scores, and these profiles have been used in this section, to organize the findings.

Individuals whose data differed from these patterns were apparent in each instance, and some of these will be examined in greater detail in this section. For ease of reference and to preserve anonymity of the participants, masculine pronouns will be used to refer to all individual participants.

'Gainers' High (poor) score in DA at T1

Four children, D1, CP2, BH9 and R1 scored poorly on DA at T1, and poorly on CELF. However, at T2 their DA scores had markedly improved, and there were steady, if not substantial improvements in CELF raw scores over time. All had Ravens ratings in the high range. Response to Mediation ratings and teachers' ratings were variable, consistent with the experimenter's comments relating to confidence and engagement with the task, but all were concerned about their performance and anxious to succeed in the task.

Two other children in the cohort differed in specific ways from this profile.

D4 was markedly inconsistent in test performance. At T1, he was 8;10, and when retested at T2 at age 9;2, so as a result, slightly different subtests of the CELF-3(UK) were used. This did not however, appear to be the only reason for D4's poorer performance. At T2 he was noted to be very tired and struggling to concentrate and engage with the tasks, and scored less well on both CELF and DA measures than at T1. He was also a poor reader and relied mostly on working memory in the DA task, rather than being able to make use of the printed words. Tiredness and poor attention would affect his performance on the task. This was also reflected in the Response to Mediation rating that decreased at T2, showing poorer understanding and interest in the task, poorer attention and problem

solving, and less inclination to seek help from the examiner. His ratings, however, recovered at T3.

Teacher ratings, based on more general performance over a longer period of time were unaffected, and pragmatic and interpersonal skills were strong, along with a high non-verbal reasoning ability according to the Ravens CPM. Whilst his receptive language scores were raised by his performance on the Word classes subtest, especially at Time1 and Time 3 and 4, Concepts and Directions remained poor throughout. All expressive subtests improved by small amounts steadily over time.

The DA revealed some substantial weaknesses in cognitive processes and metacognitive awareness, which led to poor task intrinsic motivation, a tendency to employ strategies that are not useful e.g. counting words, and failure to consistently use his knowledge. Recommendations included reflecting on semantic content of sentences, and new semantic relationships, in order to help him construct ideas, and increasing the variety of his sentence structures, without increasing sentence length which overloads his memory (See Appendix XV for the report supplied to SLT following the DA of D4).

R3 followed a similar pattern to the group of 'gainers' but differed in that, unusually, his expressive language total score was above that of his receptive language. This was accounted for by high scores in the Recalling Sentences subtest. R3 was able to retain and repeat long sentences, well above the level of complexity that he was able to generate spontaneously. R3 also differed from the profile of the group in respect of his Ravens CPM score which was in the low range, 25-50th percentile. Problem solving abilities were weak, as recorded both in the Response to Mediation scale, and the teacher's questionnaire, although both noted a slight improvement later in the

study period. There was a lack of attention to detail noted in the DA that accounted for morphological and syntactic inaccuracies and this observation was supported by an apparent lack of concern and anxiety about his own performance.

Recommendations were made for ongoing SLT intervention to address both receptive and expressive language difficulties, focusing on structural features of language and morphology, as well as attention to detail and accuracy of structures. It was further recommended that intervention might make use of explicit rule teaching and application, and explicit tasks addressing precision and accuracy in the gathering of information, following of directions, decoding of information, and selection and encoding of expressive language, to decrease the reliance on memory. It was seen that R3 was typical of the group of participants who benefited from experimental therapy when baseline gain was low.

'Non-gainers' High (poor)score in DA at T1

Three children, D3, TF1 and BH6, scored poorly on the DA at T1 and did not improve more than 5 points on subsequent re-test. Similarly their CELF-3 scores were low, and showed little sustained improvement over the therapy periods. D3 and TF1, both aged 8 through therapy phases 1 and 2, were also identified on the Response to Mediation scale as having significant difficulties with problem solving, comprehension of the task, and displaying poor response to challenge by giving up easily, rather than persevering to solve the task. Although BH6 who was rated higher, was more able to understand the task, persevere and talk through it, he showed weak strategic problem solving. Supporting data came from the teachers, who identified significant difficulties on almost all criteria for D3 and TF1, but fewer difficulties in BH6. They did, however, note that BH6 had significant problems in essential skills such as

'problem solving or working through a task', 'remembering or retaining what he has learnt', and 'explaining what he is thinking or feeling'. All three children scores were in the lower range of normal on the Ravens CPM.

Recommendations for all of these participants included general learning about following rules and recognizing patterns, as applied to language, communication and other behaviours; sequencing and sorting; and the Colourful Semantics approach that imposes structural regularity and attaches labels to parts of the sentence. Gains in language as a result of these interventions would be likely to become apparent after a longer time period than the duration of this study. It may be worthy of note that all three of these participants were randomly allocated to the experimental group.

Low/medium scores in DA at T1
The remainder of the cohort (n=15) scored less than 60 in the DA at Time 1. Although they improved over time on the CELF-3, there was little change in their DA scores over time, and little to be gained from repeating the DA. The profiles of several children, however, were unique in specific ways.

Both BH3 and BH4 scored within normal limits on the CELF-3 at Time 2. BH3 improved on all subtests, most markedly on Word classes and Recalling Sentences. Although scores on 5 of the 6 subtests were within normal limits, the scores on Sentence Recall were strikingly high, with BH3 achieving a standard score of 11. It is notable that ability in sentence recall is unusual in children with language impairment. The Response to Mediation scale, did not identify any specific weaknesses, other than little interest in the activity at T1 and T3. The questionnaire from BH3's teacher identified difficulties with BH3 explaining what he is thinking or

feeling, and also that his behavioural response to situations or challenges was not very appropriate. Similarly, the SDQ scale carried out at the time of participant selection identified BH3 as borderline in the Hyperactivity score, the Emotional Symptoms score and the Peer problems score. As each of the scales was borderline and there were no significantly high scores, it was decided to accept BH3 as a participant in the study. In each of the individual DA sessions, BH3's behaviour was good, however at Time 3, BH3's teacher rated difficulties in 10 out of 15 criteria, and it would appear that emotional and behavioural difficulties were present. The targets set for BH3 were mostly in the areas of attention and pragmatic skills, which would seem to be appropriate in the light of the high CELF-3 scores.

BH4, however, scored erratically on the CELF over time, and his scores at T3 were comparable with those at T1. Peaks in scores on Concepts and Directions, Word Classes and Formulated Sentences at T2 were subsequently reversed. The experimenter noted that he had a good facility with manipulating grammatical structures, but did not always apply his knowledge well when confronted with a task. Strategic problem solving was rated better over time.

Although BH4 was noted by the experimenter to have some inappropriate social and pragmatic skills, ratings on the SDQ, response to mediation and teacher rating did not highlight any significant concerns. Goals set by the SLT, however, targeted attention and listening, and pragmatic skills, and recommendations made on the basis of the DA at T2, included more work on reflection and increased self-awareness of when he is not applying his knowledge, or responding impulsively as well as work on more complex sentence structures involving conjunctions and adverbials. Strategies for problem solving, such as planning, making hypotheses

and checking, could be used, and generalised to contexts other than language.

Child CP1, aged 9;2 at the start of the study, and in year 5 at school, scored in the mid-range on the DA measure at T1, improving on subsequent retests. He made significant progress in language as measured by the CELF-3 over the study period, rated as severe language impairment at T1, moderate impairment at T2 and 3, and mild at T4. During the last term of the study, CP1 was transferred from the Language Base to a mainstream class in his school, although he was simultaneously given a diagnosis of ASD.

While the improvement in language measures would suggest that CP1 benefited from the management in place in the Language Base, consideration of the profile of assessments highlights discrepancies that support the diagnosis of ASD. The SDQ scale score carried out at the time of participant selection, was outside the normal range (18, when normal <15). CP1 was accepted for the study on the basis that his Hyperactivity criteria on the SDQ were normal, so that while his total score was 3 points over the threshold for 'Normal,' behavioural criteria were unlikely to affect the required mediation in the DA. The raised score was, however due to an abnormally high score on Emotional Symptoms, and borderline scores on Peer relations and Prosocial behaviours, and may in retrospect have been indicative of the subsequent ASD identification. There was also a considerable discrepancy between performance in the 1st percentile in language according to the CELF-3(UK), but performance in the 75-90th percentile on non-verbal reasoning according to the Ravens CPM.

Further, while the Response to Mediation scale total score was unremarkable, there was no difficulty in problem solving or understanding the task, but absence of self-talk, and little interest in

the activity. The experimenter noted that CP1 was quiet, and reported him to be 'serious throughout the session'. CP1's teacher, however, identified significant difficulties in almost all of the functional criteria on the questionnaire, including a rating of 'inappropriate' on behavioural responses to situations or challenges. Considerable improvement was recorded by the teacher at Time 2. It would appear, therefore, that the performance of CP1 throughout the study period may have been related to social and emotional factors as well as language. The individual nature of the DA, however, would make it equally applicable to a child with mild ASD, and the recommendations arising out of the DA should have been valid and applicable. The SLT set targets relating to sentence level comprehension and expression as well as semantic organization, and multiple meaning of words. This was supported by the recommendations arising from the DA, which also recommended raising metalinguistic awareness. Targets in the subsequent phase, however, addressed more vocabulary and semantic organization, but with greater reflection and metacognition involved.

Individual data such as these indicate the impact of variation that is compounded by the small group sizes, and suggest that analysis of a series of case studies may be more informative than group data. Further discussion regarding individual performances on the DA and test battery, and in response to the intervention offered, is contained in Chapter 8. In that review certain influential factors were identified, and the discussion arranged around these themes. High non-verbal scores and substantial discrepancy between verbal and non-verbal ability was associated with low DA scores, suggesting good learning potential, in two children. Both were seen to make gains in language following recommendations for metacognitive intervention. A substantial influence of social and emotional factors also emerged as affecting the outcomes of two further participants with good learning

potential, and differentiated the responsiveness of two children who differed slightly in learning potential, but were managed together. Finally the progression in DA scores emerged as the most effective way to record the progress of two further participants, whose progress in therapy was not otherwise apparent. Details of the outcomes of these participants throughout the phases of intervention are contained in section 8.1.4.4.

7.6 <u>Summary of findings</u>

No simple, straightforward effects of DA-based intervention were found in this study either using static or DA measures. This did not appear due to the sensitivity of measures which were well constructed (see Chapter 5), and indeed the analysis of different sub-measures also did not reveal change. However, two factors that may in part contribute to the main finding were the application of DA information to therapy, and the individual progress of the children. Specifically, whilst therapists implemented significant shifts in therapy goals, there was less evidence in video monitoring data, that strategies of mediational intervention were adopted, suggesting a gap in the information giving process. Finally, analysis controlling for baseline change revealed that the DA might usefully inform intervention for a proportion of children with SLI whose progress in regular ongoing therapy was slow. In addition, detailed individual case studies demonstrated the valuable contribution of additional data obtained from the DA to the understanding of individual case profiles.

These findings will be further explored in Chapter 8.

CHAPTER 8

IMPLICATIONS OF DYNAMIC ASSESSMENT FOR INTERVENTION: DISCUSSION

CHAPTER 8   IMPLICATIONS OF DYNAMIC ASSESSMENT FOR INTERVENTION: DISCUSSION

The current study, designed to explore the effect of supplying supplementary information derived from the DA to intervention programmes, was not intended as an efficacy study in which variables of therapeutic implementation would be tightly controlled. Instead, the study examined effectiveness and the population of children with language impairments studied was heterogeneous. It was predicted that different participants would be likely to benefit in different ways, and to a varying extent, from different aspects of the information supplied, and the recommendations derived, from this particular DA.

The investigation of the impact of the findings of the DA of language on intervention outcomes was carried out via a number of different avenues of investigation. All the children in the cohort improved significantly during their language therapy at each stage of the project, although a significant difference between the outcomes of their regular therapy and modified intervention was not shown on any of the measures used. However, a significant change to the targets set by the SLTs for the children involved in the modified intervention signalled a positive uptake of recommendations for intervention, which was accompanied by a positive response about the usefulness of the information, obtained from the SLTs. A significant effect of the modified intervention in those participants who were shown to have minimal improvement in their ongoing intervention programmes indicated that the DA might usefully inform intervention for a proportion of children with SLI whose progress in therapy was slow. The role of individual differences within the group of children was also discussed in a series of case studies. The findings will be discussed in more detail in this section.

8.1 Discussion of Results

8.1.1 Effectiveness of the DA information in improving therapy

8.1.1.1 Effectiveness of the therapy recommended by the DA in facilitating improvements in language, as measured by the CELF-3.

The gains in language made by the participants over the time of the study were measured primarily by repeated administration of the CELF-3(UK). Although there were some limitations to this measure, which will be discussed further later, the use of raw versus standard scores on the CELF-3 facilitated the detection of small improvements, as well as poorer performance by some children in both groups, on later retests, supporting the earlier reported inconsistent performance of children with SLI on standardised tests (Hasson and Botting 2010).

Previous studies have also published equivocal findings from intervention studies with a range of explanations for the failure to demonstrate positive effects of particular interventions being suggested. Few intervention studies have demonstrated positive intervention outcomes for expressive syntax, the meta-analysis by Law, Garrett and Nye (2004) finding positive outcomes for those with expressive impairments only in the absence of receptive impairments. Law, Garrett and Nye (2004) also noted that a minimum of 8 weeks of intervention was crucial in eliciting outcomes from intervention, and McCartney's cohort study (2004, 2011) suggested that a period of 16 weeks, in which the number of sessions could vary widely, was insufficient to reflect gains in therapy. The dosage of intervention was not controlled in the current effectiveness study, and the time periods over which measurements were taken were relatively short and may not have been long enough to show differentiated effects. Boyle et al's (2009) RCT

identified dosage as a crucial but unresolved issue in intervention studies. Similar conclusions were reached by Gallagher and Chiat (2009) who achieved positive outcomes in children with receptive language difficulties after intensive intervention, although outcomes with pre-school and nursery aged children tend to be better than with older, school aged children.

8.1.1.2 Effectiveness of the therapy recommended by the DA in facilitating improvements in language and learning as measured by the DA.

In a clinical context, the DA would only be administered once to inform subsequent intervention for a child who had a language impairment, although it may be used on a second occasion to evaluate maintenance of the mediated learning over time. The results of the current study suggest that while the DA is indeed useful to evaluate outcomes after a period of intervention, it is less useful with repeated administrations, as further learning from the mediation implemented on the second and third occasions raised scores by an amount that is not statistically significant for a group of children.

One possible explanation for this lies in the heterogeneity of the SLI sample, for which group mean scores may have obscured the wide range of variation within the cohort, with scores of the individual children changing to a different extent at each time point. In the second DA, some of the strategies for solving the particular task may have been repeated, or slightly different ones mediated, depending on the individual's response to each item. Some strategies may have been taken up by the child and retained after the first DA, other children may have grasped concepts only after the second trial. Some children may require yet more explicit or different intervention in order to be able to progress. Strategy use also interacts with the

level of difficulty of the particular linguistic construction for the individual child, and he may not be able to use what he has learnt in the face of a complex or poorly understood grammatical construction. This variation, which was exemplified in greater detail in the case studies, is precisely the observation made by Vygotsky that led to his development of the concept of the Zone of Proximal Development (ZPD) (Vygotsky 1986).

8.1.1.3   Effectiveness of the therapy recommended by the DA in facilitating transfer of learning

Scores on the Sentence Assembly (SA) subtest were examined in particular to determine whether intervention utilising information derived from the DA facilitated increased gains in the experimental group children on a task of near transfer. Modified intervention would have been expected to facilitate strategic learning that was transferred to a parallel task in an unsupported environment. This finding was not supported and the gains on SA in both groups mirrored the findings of the total CELF-3. This suggests that there was little transfer of learning strategies to the static task although there were gains on repeated trials of the DA itself, and qualitative evidence of transfer of learning both between items on one test, and between subsequent tests. Amount of transfer of learning did not emerge as a discriminating feature between groups.

In the light of this lack of near transfer, it would have been unlikely to detect greater gains over time on a test of far transfer, in this instance the Formulated Sentences (FS) subtest. Mediation during the DA, to a variable degree across participants, addressed aspects of sentence construction, such as that sentences usually start with the person, followed by the action, and this information may have led to improved performance on the FS subtest of the CELF-3. Similarly, recommendations for intervention to address aspects of

sentence construction might have resulted in measurable gains in the FS subtest for children in the experimental group. As in the SA subtest, the gains on FS mirrored the findings of the total CELF-3, and no particular acceleration of learning was detected on this measure.

8.1.1.4 Effectiveness of the therapy recommended by the DA in facilitating improvements in other behaviours, e.g. functional communicative behaviours

Following the assessments on the CELF and the DA, individual recommendations for intervention were made by the experimenter, many of which were for addressing basic cognitive processes such as sequential organization and following rules. Change in these skills would not necessarily be reflected in assessed language performance in the short term, but may have begun to impact on the functional communicative skills of the children, as examined via the RtM and the teachers' reports.

i. Response to Mediation scale

The Response to Mediation scale (RtM Lidz 2003) set out to measure the behavioural responsiveness of children to the mediational intervention carried out during the DA. The gains measured by the RtM were found in both the first and second phases of therapy, and showed that the experimental group became more responsive, in comparison to the control group, but the improvement cannot be attributed specifically to the differentiated intervention linked to the DA.

It is also not clear whether the gains in scores on the RtM are indicative of a change in behaviour that would necessarily facilitate better performance on a language task, or whether the behavioural aspects do not have an impact on the language tasks, as the trends

uncovered in language tasks are not the same as those shown by the RtM. Behavioural variations would, of course, affect a child's performance on any formal or informal assessment and might account for some of the variability in DA and CELF scores. The flexible nature of Dynamic Assessments accommodates behavioural variation to some degree, allowing for example, the tester to repeat cues, and the child to ask questions. Such interaction enables the examiner and the child to build a relationship based on trust and familiarity that would affect responsiveness, as measured by the RtM. It is proposed that this enables more accurate assessment of the individual. In the clinical context, the particular profile of each participant on the RtM scale should be considered alongside his performance on any other assessment (Lidz 2003).

The executive functioning required in 'Strategic problem solving' which reflected whether participants actively planned their responses was particularly poor, which was noted to be typical for a group of children with LI (Marton 2008). Also predictable from a group of children with LI was an overall rating of minimal interest in a decontextualised, specifically linguistic task, and the lowest overall rating for 'Evidence of self-talk when working on a challenging task'. Many of the children did not verbalize their own problem solving, as might be expected in a group of children with LI (Leonard 1998) although this behaviour was modelled and mediated by the examiner during the DA.

In summary it would appear that while the Response to Mediation scale was sensitive to an interaction that reflected a trend towards greater improvement in the experimental group, only some items in the Response to Mediation scale were sensitive to the changes in learning potential shown by the children, and the whole scale may be less useful than the selective use of the five criteria most

applicable to the participants satisfying inclusion criteria for the current study. The lack of differential results during the phase in which the intervention provided to the groups was altered, adds little to the determination of whether value is added to the intervention by the addition of information derived from the DA.

ii. Teachers' reports

The rating scale devised for completion by the teachers of all the children emerged as the least useful assessment component of the study. Variations in scoring could have been attributed to differing demands of the curriculum at different times, and behavioural fluctuations within children in a class, over the school year. Familiarity of the teachers with the children may also affect their ratings later in the year and similarly, teachers' expectations of children vary over time. Further, a certain amount of reluctance was expressed by some of the teachers when they were asked to complete the rating scales for the second and third time, and while they were reassured by the experimenter that completing the questionnaire quickly, based on an overall impression of the child would suffice, this may not have led to the most reliable ratings. A small number of questionnaires were not returned.

The lack of correspondence between the outcomes of the teachers' scale and the other assessments may be because the teachers based their assessments on their knowledge of the children gleaned from daily contact over an extended period of time, while all the other assessments, including the RtM were based on a particular session of assessment. In retrospect, due to the differences in timing and the context from which they were taken, the two sources of behavioural data were not entirely comparable. Further, no attempt was made to follow up on whether information from the teachers' ratings, or in fact from the behavioural aspects of the DA and the RtM was

specifically used by the SLTs in their planning. It was thought later that the teachers' ratings might have been predictive of the children's progress over the study period, their sensitivity to functional skills in the classroom context being related to the children's gains in language, but this did not emerge as a significant finding.

In summary, like the measurements on the formal standardised test, changes in the behaviour or functional communication of the children in the study were not found to be related to the differentiated intervention delivered at different stages of the study, although there was an identified difference between the two groups at the start of the study. Other scales, with greater sensitivity to changes, or greater reliability may have detected significant differences. From a functional point of view, the therapy informed by the DA should aim to produce an improvement in the performance of children in the experimental group that was noticeable to their teachers, and would have justified the assessment and the modification to intervention. This may turn out to be the case for a number of individual children, and will be explored in greater detail in the case studies section. There are several possible reasons for the lack of significance in the differentiation between groups, attributable to the measures used, the specificity of the intervention carried out by the Speech and Language Therapists involved in the study, or variation within the cohort of child participants, and these alternatives will be explored in greater detail in the following sections.

8.1.2 Factors related to measurement

8.1.2.1   Factors related to the instruments used

The measures used to gauge progress in intervention assessed only a limited range of linguistic skills, in a population in whom a wide variety of communicative impairments exist. As previously discussed,

the recommendations for intervention addressed a variety of pragmatic and social targets, as well as vocabulary and narratives, which were not tested in the evaluation of gains in intervention.

One solution for future research may be to use a battery of tests, including for example, the BPVS (Dunn et al 1997), ERRNI (Bishop 2004), the Children's Communication Checklist (Bishop 1998), the Pragmatics profile of everyday communication skills (Dewart and Summers 1988) or the Social Use of Language Programme (Rinaldi 1995). However, practical time limitations often restrict the testing sessions spent in standardised testing, especially when an additional DA session for each child at each time of testing is included. It was most logical to select an assessment of grammar, rather than any other single measure, and the CELF-3(UK) contains a range of subtests accessing receptive and expressive components at sentence level, as well as some word level semantic organization in the Semantic Relationships subtest. Further, there were obvious links to the Sentence Assembly subtest, and the CELF-3(UK) is widely used in clinics and schools in the UK, and has applicable norms for the population for which it was being used.

Another alternative approach may be to use a sample of spontaneous language, and analyse it comprehensively using LARSP (Crystal, Fletcher and Garman 1976) or STASS (Armstrong and Ainley 2007), but also scrutinize the data for choice of vocabulary, conversational appropriateness, and functional discourse skills. This would capture a greater range of potential improvements that could arise from the broader information derived from the DA being fed into the intervention programme. However, results from such measures are not entirely objectively quantifiable and would have limited quantitative and statistical evaluation of gains and comparison of groups. Nevertheless, Joffe (2008 p.79) also

advocated the use of informal assessments without standardised scoring to measure change.

Functional evaluation of outcomes may have used school or curriculum based measures, as advocated by Joffe (2008), who used the school SATS scores as a measure. Again these are not without problems, as SATS are not completed in full every year, and because children in Language Resource bases and units, and particularly in special schools, such as the cohort in the current study, frequently have 'special arrangements' made for them, to accommodate their particular needs, and these may limit the comparison of test outcomes from one child to another. Other curriculum based measures used in the schools were, like the tests, specific to particular content criteria, and unique to the schools, thus not suitable for group comparisons.

One issue across much intervention research is the fact that tests do not cover a wide range of ages, limiting longitudinal investigation. The selection of the CELF-3(UK) was for example problematic, primarily in the arrangement of subtests for different age groups, which directly affected the age group under observation. However while this issue with the age range of tests made the current DA validation difficult, the development of DA itself should facilitate measurement of skill across a wide range of contexts including age and culture.

Despite this difficulty, the Sentence Assembly subtest was administered to all children, regardless of age, at each stage of testing, and the static standardised form of the test served as a baseline against which subsequent progress in the same task could be monitored. It was apparent that the majority of children (in both groups) did make gains in the Sentence Assembly, and when their

scores were included in the total raw scores after their 9th birthday, the children already had experience in this subtest, and scored higher than would have been anticipated by a first administration of a new subtest.

These findings are apparent in the data from participants aged 8 at the start of the study (see Appendix XVI). While participants made gains in subtests that were repeated, their total raw scores may have dropped because of the switch in subtests, for example from 'Sentence Structure' to 'Semantic Relationships'. Gains in the Sentence Assembly test, as a result of transfer from the mediation in the DA, result in overall scores that are not truly comparable with that of a child tested in a standardised way on the CELF-3(UK) such as those in the normative sample. Standard scores were not used in the present study, but raw scores were also subject to measurement errors, and for this reason the study ought to have been limited to participants already aged nine at the start of the study.

In addition to the test selection limitations and measurement errors, the experimenter was conscious of a theoretical inconsistency in the reliance on a static standardised test to measure outcomes in this study. The difficulty rests on the assumptions presented at the outset, that static administration of standardised tests does not reliably gauge the optimal performance of children with LI. These children are frequently unreliable in test performance, indeed in many instances, children scored less well on subsequent retests of any or all of the subtests, which should not in theory be possible. The subsequent lower score, or failure on an item that was previously correct, in spite of having had practice on that item suggests that children may have guessed items correctly, and as they did not receive feedback, were apt to guess again, incorrectly on a subsequent retest. This is a poor indication of their ability on a

task. While this type of performance is accounted for in the standardisation sample of a test, it assumes disproportionate importance in within-subject retests, such as when children with difficulties are re-evaluated to measure progress.

Further, while fluctuations in mood, motivation, anxiety and emotional responses to testing were detected and noted, and sometimes compensated for in the Dynamic test procedure, they were likely to affect scores on the standardised test. The experimenter was unable to provide hints or cues to help the child, nor even repeat items, in accordance with the rules for standardised administration of the test, and it was felt at times that failure of an item was not representative of ability. The conditions of administration did little to foster the rapport and trust that was built up in Dynamic Assessment sessions and is a valuable component in intervention sessions. The experimenter felt that principles for representative assessment were compromised by the use of static testing in the study.

8.1.3 Factors related to the programming of intervention

8.1.3.1 Recommendations supplied to the SLTs after the DA

Ten Speech and Language Therapists from seven schools participated in the study. Test results and feedback from the DA were supplied as described in section 6.4.3, and SLTs were contacted by the experimenter. The experimenter acknowledged that she was not providing the SLT with knowledge about the child that would be entirely new to her, but rather asked the SLT, whose knowledge was based on longer experience of the child, to confirm the findings. What did become clear was that a great deal of information and insight about the child had been gained by the experimenter after only two testing sessions with the child. The experimenter was satisfied that all the SLTs had had the opportunity

to ask questions about the recommendations for modifications to their intervention presented in the manual, and were clear that they could contact the experimenter for further discussion. This offer was taken up on a small number of occasions. SLTs were again asked to send their targets and examples of activities to the experimenter. These were used to explore changes made to the intervention programmes after the DA reports were supplied. The results are discussed in section 8.1.1.3 below.

Closer examination of the recommendations included in the report for each child that were given to the SLTs after the DAs, revealed that the experimenter did not advocate changes to syntactic targets for all of the children. A summary of all recommendations showed that the experimenter mentioned pragmatic interventions for seven of the children, and identified grammatical structures as targets for eleven participants. The greater proportion of the recommendations made for the nature of intervention, concerned metalinguistic awareness work for 19 of the 24 participants. In accordance with the principles of Feuerstein's clinical approach to DA and MLE, intervention targeting metacognitive awareness of thinking and learning skills was advocated for 22 children, and these strategies for improving skills were not specifically related to language outcomes, but generalizable to other contexts. For example, recommendations on reports included the following;

'*it would be recommended that this be approached from the perspective of problem solving, and that strategies for problem solving, such as planning, making hypotheses and checking, be used*' and

'*He may benefit from explicit self regulation and reflection activities, accompanied by learning of metacognitive vocabulary*

such as 'remember, plan, check, compare, explain' (see Appendix XVII).

Thus recommendations for metacognitive interventions were made to address planning and checking of responses, application of rules, accuracy and attention to detail, as well as self monitoring and justification of responses. These are consistent with the cognitive functions identified by Feuerstein (1979), as features of the learner in the interaction between learner, assessor and task (see section 1.2.2). However, Feuerstein set out to evaluate these cognitive functions, or identify deficiencies in them, as the outcome of the LPAD. In contrast, the current study was more content based, identifying linguistic knowledge as well as problem solving, and being less systematic in the rating of cognitive functions. Weaknesses have, however, been identified in some of these 'executive functions' in children with LI (Marton 2008, Henry 2010) so it is not surprising that these recommendations have emerged from the current study.

Further, the identification of targets for intervention, be they pragmatic, grammatical or cognitive, were inter-related with the nature of the intervention. Recommendations for addressing linguistic targets were in most cases via awareness of linguistic structures, by explicit naming, explanation, or techniques such as Colourful Semantics (Bryan 1997) or Shape coding (Ebbels 2007). Similarly, for several children, pragmatic awareness was to be facilitated by 'barrier games' which highlight explicitly the information needs of both partners in a conversation. More detail about *how* these notions should be implemented, was not, however, specified in the reports. The experimenter noted only that metalinguistic awareness should be increased, and which concepts might benefit the child, giving no details of specific strategies. SLTs

had to determine their own intervention approach, with support from the manual that presented mediational strategies.

It might be seen therefore that the distinction previously made between targets and methods, is an artificial one, and that metalinguistic and metacognitive awareness are also in fact *targets* of intervention, rather than *methods*. The recommended method of intervention would be *mediation*, but as the SLTs were not trained in mediation, fewer instances of implementation were expected and indeed observed (see 7.3.4). The natural blurring of boundaries between targets and methods reflects the state of the art of language therapy in which few intervention studies have been able to empirically demonstrate significant effects. Studies devoted to methods of therapy, such as those outlined in Chapter 2, focussed on the *methods*, that were then assumed to be applicable to a range of linguistic targets, but this assumption was not comprehensively tested. For example, the outcomes of intervention using Conversational recast techniques are frequently based on general measures of language growth, such as MLU (Camarata and Nelson 2006), when in fact, Camarata and Nelson (1992, cited by Camarata and Nelson 2006) noted that specific grammatical targets may be better suited to recasts than other linguistic structures. Further, Camarata and Nelson (2006) noted that the recast technique is frequently included in other broader interventions again blurring targets with techniques of intervention. SLTs are inclined to define the goals or targets of their intervention, and adopt eclectic approaches to therapy (Law et al 2008). The rationales for intervention given by SLTs in that study were a combination of descriptions of deficits with a focus on therapy. The overlap between planning targets and methods might have contributed to the poor reliability of ratings by independent SLTs of the therapy data in the current study.

8.1.3.2 Transfer of information

In the planning of the project, it was assumed that reports of the DA would be most useful to SLTs if they were structured in a predictable way, and restricted in length to 1-2 pages, in order that they would be readable and accessible. These guidelines were adhered to by the experimenter. SLTs were not asked to view videos of the DA of their children as it was thought that this would place an unreasonable demand on the time of the SLT. On the questionnaire at the conclusion of the study, however, all of the therapists stated that videos would have been useful to add clarity to the DA findings, and all stated that they would have been willing to spend the time watching videos of the children's performance in the DA procedure. Haywood and Lidz (2007 p.220) also noted that observers of a DA session gain more insight than they can through written reports or oral feedback. Delclos, Burns and Kulewicz (1987 cited by Delclos, Burns and Vye 1993) found that teachers given the opportunity to observe videos of DA of young children altered their expectations of those children, and viewed them as generally more competent than they assumed after the children were observed in static testing. Subsequent research by Delclos, Burns and Vye (1993) however, concluded that reports written on the basis of DAs were also able to alter the expectations of teachers relative to their assumptions based on static assessment reports.

In retrospect, therefore, videos could have been used alongside reports for the purposes of assisting SLTs to plan intervention. SLTs were not asked to rate the usefulness of the reports themselves, the focus on feedback being on the information about the participants but Bosma and Resing (2010) reported that the teachers they worked with had preferences for reports containing useful strategies for instruction, and had difficulties with the unfamiliar format of DA

reports. Bosma and Resing further reported that there was no significant evidence of changes to the teaching practice of teachers supplied with reports of DAs. In the current study, care was taken to familiarize SLTs with the structure they could expect to see in the reports and to avoid or define unfamiliar terms. The next section discusses the findings taking into account the information supplied in the reports and the SLTs' intervention planning.

8.1.3.3 Intervention targets and methods

Inspection of the raw data from the SLTs report sheets revealed that the SLTs planned more sentence construction targets for all of the children in the second phase of therapy, after results of the DA for half of their children had been seen. This may have been expected for the children in the experimental group, whose syntactic abilities and knowledge of sentence construction had been exemplified in reports given to the SLT, but appeared to have generalised to planning for the children in the control group as well. This may be explained by a heightened attention to syntactic targets in the SLTs planning, or the fact that some of the children were seen in groups with targets common to all.

SLTs' responses to the questionnaires were also used to determine whether targets and activities reflected 'skills acquisition', 'metalinguistic activities', 'meta-cognitive activities' or combinations of these, using definitions adapted from Law et al (2008). Despite the definitions, it was difficult to determine the intentions of the SLTs in this regard, from the information given on the sheets. It might have been possible to add accuracy to these interpretations by consulting the SLTs about their exact methods and how they viewed the nature of the interventions done, but it was thought that this may have elicited subjective responses.   Results should be interpreted with caution due to the poor reliability of the ratings.

Nevertheless, there was thought to be sufficient difference between the two groups for some assumptions to be made. Unlike the targets, there was a clear difference between the nature of the intervention activities planned for the experimental and control groups, and cross-over effects were not apparent in the reported activities. There was a significant shift from skill based to more metalinguistic interventions for the experimental group after the information from the DA was received. It may be that as intended, the SLTs used the assessment of metalinguistic knowledge reported after the DA to plan activities that included explicit metalinguistic explanation of the language that the children were reported to have difficulty with. The information was specific to each child, and not generalized to others.

Thus although there was little difference between the groups in the choice of the language area for intervention, there was a difference in the nature of the intervention. This may have been attributable to the information supplied following DA that focussed on the metalinguistic knowledge and metacognitive awareness of participants, and the recommendations made for process based targets and mediational intervention. Mediational intervention implies directing the participant towards explicit self directed problem solving, which would require a degree of metalinguistic and metacognitive knowledge and reflection. Bosma and Resing (2010) reported studies in which teachers found it difficult to plan interventions based on psychological reports following DA, and only those trained in mediated learning and process-oriented approaches were able to generate interventions from the reports.

Similarly, Delclos, Burns and Vye (1993) found that teachers trained in mediated learning were more able to appreciate the programme planning recommendations in both DA and static assessment reports,

than were teachers with more conventional training in direct instruction. The training in mediated learning enabled teachers to focus on the processes used by the learner, and infer these from the information contained in all reports. SLTs in the current study had less difficulty inferring implications for intervention from the findings of the DA, but whether this is due to the professional training of the SLTs or any feature of the reports written in the study, cannot be determined. Haywood and Lidz (2007), however, noted that stimulability, a concept closely linked to DA, is one that is familiar to speech and language professionals.

8.1.3.4 Delivery of therapy

A small sample of intervention sessions from phases 1 and 2 of therapy carried out by SLTs with children involved in the study, were videoed and the nature of the therapy monitored by the experimenter. Observation of the video tapes revealed that several of the SLTs made use of behaviours that would be described by Haywood and Lidz (2007) as mediational, in both phases. These behaviours, for example helping the child to understand the structure of his own therapy session by presenting him with a visual timetable at the start, and specific contingent reinforcement of desired behaviours, are typical of the usual practice of SLTs and were present without instruction from the experimenter. Similarly, SLTs frequently facilitate self monitoring in children by asking them to make grammatical judgements or justify answers, and these prompts were seen to have been used by the SLTs in the baseline phase, with fewer in the second phase, possibly replaced by phrasing these prompts in the more process based format of 'How do you know that..?' which was suggested in the manual.

Specifically mediational metacognitive techniques such as emphasizing meaning of the session to the child, or ensuring his

awareness of the role of the therapist in helping him learn a skill, and how that skill would be useful in other contexts, all of which are required components for an interaction to be deemed mediational, were not evident. Thus it would seem that the therapeutic skill of trained and experienced SLTs includes some mediational behaviours, but because they had not been specifically trained in mediation, some other essential features of that style were not present. In the experience of the examiner, training in mediation requires a modification of well known and well used facilitative skills, and characterisation of these into a different theoretical framework with similar skills but different emphases, rather than learning of an entirely new repertoire of behaviours. Kester, Peña and Gillam, (2002) investigated the interaction between materials used in intervention and procedures used, and concluded that it was the use of mediational practices rather than the materials, that contributed to the greater gains in the teaching stage of a DA. Although their study was on preschool children from low SES and CLD, but not LI backgrounds, the principles of employing mediational strategies that focused for example on '*why labelling is important',* (p.46) may also be applicable in the current context.

In summary, because comparative studies of transfer of training during DA found that the mediational strategies used by Feuerstein and his colleagues were most effective, it might be recommended that SLTs receive more explicit training in mediation if they are to derive the maximum benefit from carrying out Dynamic Assessments and related intervention programmes. There is support for the need for training SLTs in the implementation of strongly mediational interventions. In the current study, SLTs found the information useful and were willing to use it to make changes to the children's therapy, but they were not altogether effective in carrying this out – targets and methods were changed, but style of intervention was not.

These findings also suggest that it may not have been the information elicited from the DA that was inadequate to inform intervention with improved outcomes, but an inadequate preparation by the experimenter of the SLTs involved in the study, with regard to the intervention techniques that would be expected of them.

The experimenter hoped to demonstrate that additional training in mediation would not be required in order for the current procedure to be used by experienced practicing SLTs. While the first recommendation would be to share the video material with the SLTs to make the nature of the children's learning more explicit, and supply more detailed reports and recommendations, the results also suggest that more explicit training in mediational techniques may be needed after all if improved outcomes are to be linked to findings of a DA procedure.

8.1.4 Individual Variation

8.1.4.1 Effect of individual variation at entry to study

The population of children with SLI is a heterogeneous one, and the current study recruited a broad range of children with differing abilities and needs, from different backgrounds who were placed in different educational systems. This, along with the small number of participants in each group, resulted in considerable variation within the groups, and the need to consider individual profiles of performance. Nettelbladt et al (1989) earlier concluded that detailed case studies are the only way to elucidate crucial individual differences in language disordered children, and Delclos et al (1992) made use of extended DA to describe the learning processes, and identify the optimal instructional techniques for individual students.

Variability was evident in the range of severity of linguistic impairment, and the range of non-verbal reasoning skills, as well as

in the nature of the presenting communication impairment. Some children for example, demonstrated milder expressive language difficulties in the area of syntax, but may have had greater problems in the areas of social communication and pragmatics. There is evidence that SLTs had selected targets of this nature in the first phase of therapy. Evaluation of generalizable cognitive functions such as planning of output, analogical learning from examples, following of rules and checking responses was possible from the DA, but the context of recommendations was that of sentence formulation, and examples of therapy targets and activities given in the manual for constructing intervention (Appendix IX) were similarly based on syntactic elements. Many of the participants were engaged in word level intervention, learning vocabulary and semantic organization, as well as pragmatics, and integration of recommendations into such programmes would have been more challenging for SLTs.

Thus only a proportion of participants in the study were in a position to benefit directly from the specifically linguistic assessment that constituted the DA, while the majority benefited from the recommendations regarding more domain general skills.   Further, the recommendation for increased focus on metalinguistic activities and reflection on language structures limited the applicability to participants with sufficient levels of language, and the severity of the manifesting language impairment would affect the uptake of metalinguistic syntactic interventions (Benelli et al 2006; Magnusson and Naucler 1993).

8.1.4.2 Progress in baseline intervention

In the light of indications that individual variation affected the outcomes from groups, an attempt was made to account for the progress made by the cohort in their ongoing therapy, against which their gains in the 'modified' therapy could be measured. Gains for the whole cohort in the period of modified therapy (informed by the DA) were significant, when gains in the baseline period were controlled for. Further statistical analysis revealed that the group of children who were progressing substantially during their ongoing therapy, made significantly less change in the period of modified therapy, than the children who had, up to that point, been making little measurable progress. The 'improvers' group was composed of both high and low achievers on the CELF, the Raven's CPM, the teacher's ratings, the DA and the RtM. Thus regardless of the child's profile of scores, the therapy programme devised by their SLT was effective in eliciting positive linguistic outcomes, measurable by the CELF-3, and these children did not benefit much more from the altered intervention after the DA. The smaller number of 'non-improvers', about one third of the cohort, however, were not making substantial gains in their therapy, as measured by the CELF-3, and this group were shown to benefit to a greater degree from the altered intervention after the DA. Although they may have been making gains in other areas of language, recommendations from the DA facilitated improvements in grammatical skills of language as measured by the CELF-3.

This complex interaction suggests that aside from the range of different approaches to intervention that individual children may require, there is a clinical need for SLTs to closely monitor the outcomes of their intervention programmes after short periods of management, and by the use of measures besides the criterion based targets of therapy. This may initially seem controversial, as the pragmatic and functional progress being made would not be

detected by the CELF-3, but the findings of this analysis suggest that if there is little detectable change on the total of the six subtest areas, the child may benefit from changes to the nature of the intervention programme. The changes may not involve implementation of any more grammatical based targets than previously, but rather fundamental cognitive organizational and problem solving strategies that would have an overall benefit on the child's learning. These learning needs may be identified by the process of the Dynamic Assessment, as feeding the findings of the DA back to the therapists planning therapy, did result in better outcomes for children.

8.1.4.3 Environmental variation

Intervention in the current study was carried out within seven schools with different organizational structures. Children in the Special school were taught in small groups with a great deal of collaboration between SLTs and teachers at all times. In one of the units the children were taught entirely within the mainstream, the remainder were placed for a proportion of the day in the resource base, where teachers and SLTs worked more closely together with a smaller group of children. Comparison of mean scores obtained at the special school, suburban and city language bases revealed no significant effects of type of school on the gain in CELF scores at T3. Outcomes obtained from the different schools were not significantly different in spite of differences in SES, amount of collaboration between SLTs and teachers, contact time in therapy, and therapists' experience, any or all of which were variables that could have impacted upon the rate of progress of participants (Kester Peña and Gillam 2002; Bosma and Resing 2010). As it would have been impossible to determine which of these factors were implicated in the outcomes of therapy, and as the numbers of children in each school were too small for meaningful statistics to be extracted, no

more detailed analysis of the outcomes of intervention by school was carried out.

The SLTs participating in the study were asked only to maintain a regular timetable of intervention across both phases of therapy, in order that progress in the second phase would be comparable with that in the first, baseline phase, and both intervention phases spanned an equivalent length school term, although events in the school calendar disrupt regular therapy sessions. Accurate data about the frequency of SLT intervention or the length of sessions were not gathered from each SLT and there was some variation in the dosage of therapy. Further, some children had individual therapy sessions while others worked with their SLT in pairs or small groups and in some instances control group children would have also worked on modified targets and/or with modified strategies of intervention. None of these variables were shown to have affected the outcomes significantly between types of schools, but individual differences may have been obscured by the group data.

Further to the case study data presented in 7.5, insights into the effects of individual scores and performances on group data may be gained by closer inspection of individual outcomes of intervention in the next section.

8.1.4.4 Case Studies

Interesting individual profiles emerged from close examination of the progress in therapy, across the different stages of the intervention in the project. Small groups of children, who were associated by a similar profile of assessment results, nonetheless showed differentiated responses to their intervention, and a number of influential factors have been used to organize the discussion.

Influence of high non-verbal scores

Two boys who scored highly on the Raven's CPM were on the first percentile of the CELF-3(UK) suggesting a very large discrepancy in abilities. Both boys had low scores on the DA suggesting good learning potential confirmed by the good progress they made during the baseline phase of therapy, with improved scores at T2. The first child benefited from word-level interventions, addressing synonyms and opposites, in skills based therapy in the first phase of intervention. After DA, his targets changed to more syntactic ones, with the approach to intervention making greater use of metalinguistic awareness, and leading to further progress at T3. Further moves towards greater metacognitive awareness in the third phase of therapy resulted in further progress at T4, by which time he had climbed 12 percentile points on the CELF-3. Strategic planning enabled him to make use of the linguistic knowledge that he had, and some progress was made towards functional improvement in terms of information selected (see Appendix XVIII for the targets and recommendations for CH1).

The other boy also made good progress on the CELF-3 with his original targets that were semantically based, but showed even greater gains in phase two with a change to syntactic targets although feedback from the DA had not yet been supplied. In phase three, with information from the DA, his SLT implemented metacognitive strategies of self-monitoring and planning, and although further gains were not evident on the CELF-3, improvements were detected by the DA reflecting improved learning strategies. Thus while increased scores on the CELF-3 were not seen to be related to the intervention informed by the DA, the potential for improvement was identified and predicted by the DA scores, in spite of low scores on the CELF-3 at T1, and gains were facilitated by different approaches to management.

In contrast, two other participants with a similar profile of scores reflecting substantial verbal:non verbal discrepancy and good learning potential, showed very modest gains in intervention. R2 presented as a very intelligent child, with low total language scores that obscured the fact that four out of six subtests were scored within normal limits (standard scores of 7 or above). Recalling sentences was a significant weakness and remained so throughout the study period. In response to findings of the DA, the targets for R2 were changed in the second intervention phase from semantic and pragmatic ones to more strategic and syntactic ones. However, the methods used had previously been focused on some metalinguistic concepts, with the use of colour and shape coding (Ebbels 2007) and this was maintained, as it was consistent with some of the recommendations of the experimenter (see Appendix XIX). On reflection, however, it was seen that R2 made only small gains in the baseline therapy, and some greater changes to the programme may have been necessary to improve outcomes. Stronger recommendations in favour of metacognitive awareness might have been made, and greater changes in this direction may have facilitated more significant gains. In particular, because of the intransigent weakness in sentence recall, R2 may be taught compensatory strategies, more reliant on reasoning than memory, to support linguistic performance.

A further barrier to R2's progress would appear to be social and behavioural difficulties which were apparent to the examiner during the DA, identified on the RtM and by the teacher on her rating scale. Aspects of interaction and expression of his feelings, as well as appropriateness of responses to a situation were rated as low. Certainly unusual intonation patterns and anxiety in the test situation were apparent to the examiner, although test performance

was not affected. Social competence and pragmatic skills were included in the goals of the SLT. CP3, similarly demonstrated little progress in therapy, even after input from DA, when one of his targets became metalinguistic, and gains remained modest. The examiner recommended more work on metacognitive awareness, but like R2, a considerable lack of confidence and social anxiety could have been a barrier to progress.

Profiles of gains in this group of participants suggest that progress in metalinguistic and metacognitive areas may be facilitated by high scores in nonverbal reasoning abilities. These finding concur with the conclusions reached by Magnusson (1993), who found, as reported in section 2.1.6, that children with grammatical difficulties may be at risk for metalinguistic difficulties, that may in turn be mediated by strong nonverbal abilities. The data further suggest however, that this progress may be disrupted by anxiety, lack of confidence and weak social interaction skills.

Influence of Social-emotional factors
Social-emotional factors and pragmatic skills were again potentially implicated in the discrepancies in the profiles of another two participants who also scored in the low range on the DA at T1. They both improved in subsequent DA trials and also both showed some gains in CELF-3 scores over time, and especially in the Sentence Assembly subtest, reflecting uptake of the mediation offered in the DA procedure, and good learning potential. The experimenter noted, however, that TF4 had some difficulties with semantically logical sentences, poor pragmatic skills and inconsistent interpersonal skills. These observations were borne out by the ratings on the Response to Mediation scale, and the examiner struggled to build up a comfortable rapport with the child. The teacher, however, rated TF4 as having almost no significant difficulties at all. The targets of the

SLT at T1 contained two pragmatic targets, suggesting that she was also aware of interpersonal, pragmatic difficulties. After the information from the DA in phase 2, the SLT incorporated 'Thinking for Learning' and Colourful Semantics' approaches into two of her targets. Modest gain was seen at T2, and less at T3, although this scoring is difficult to interpret as TF4 turned 9 during phase 2, and thus was retested on different subtests which are not comparable. On the repeated subtests, only Formulated Sentences improved slightly.

TF5, in contrast, was thought by the experimenter to have good interpersonal skills at T1 and T2, he appeared cheerful and confident, built up good rapport, and Response to Mediation scales identified no significant difficulties. At T3, an incident prior to the testing session appeared to have upset TF5 and considerable emotional and behavioural difficulties were apparent. These were presumably known to his teacher who rated him as having significant difficulties on almost all of the functional skills on the questionnaire at each of the test times. As TF5 was allocated to the Control group, no additional information was provided to his SLT for phase 2, and the therapy goals and methods were unchanged. At T3, results of the DA were supplied, but there are no records of the intervention planned for that period owing to a fire at the school. The fire also necessitated relocation to new premises, and TF5 was tested at Time 4 in sub-optimal conditions, in a noisy place, with continuous interruption, and testing of the last subtest was not completed. Surprisingly, in spite of the disruption, TF5 achieved extremely well, and showed improved performances on all five of the repeated subtests. For the purposes of statistical analysis, he was given the same score in the last remaining subtest as he had previously gained, and even with that score, his overall total was considerably (24 raw score points) higher. It cannot be assumed that intervention targets

were altered, but TF5 demonstrated good gains on the CELF-3 as predicted by the low score on the DA at T1. Final outcomes of therapy as the end of the study period appear to be related to a greater degree to the examiner's perception of the pragmatic and interpersonal skills of the participants than to their test scores.

Two children in the same school, one of whom was allocated to the experimental group and one to the control group were managed together in paired intervention. Both children were working on word level targets and listening skills in the first phase, and when the focus on syntactic targets and metalinguistic intervention was fed back to the SLT with regard to one child, two syntactic targets with metalinguistic awareness as a method of facilitation were included for both children. TA1 scored low (48) on the first DA, suggesting good learning potential, but made modest gains in intervention, in phase 1 and phase 2, with gains in Sentence Assembly in phase 2, accounting for more than half of his raw score gain. At Time 4, TA1 suddenly achieved 11 correct items in Semantic Relationships, having managed only 0 or 1 in previous trials. This seems to be unrelated to specific therapy and cannot be explained except by assuming that TA1 no longer assumed this subtest to be too difficult, and attended more carefully to the items. This assumption was made on the basis of observed anxiety in TA1, with an obvious fear of failure that led to TA1 becoming anxious, sometimes tearful and uncertain throughout all assessment sessions. His self-monitoring was poor and he was quick to give up on a task, waiting for help from the examiner. Formal testing resulted in unreliable and less than optimum performances from TA1, but the DA elicited good outcomes, with improvement over time, and transfer of learning to the standardised Sentence Assembly subtest. TA1 has subsequently been referred for management of excessive anxiety.

In contrast, TA2 scored more moderately (57) on the DA at T1, and benefited slightly from both from baseline intervention and in phase 2 when targets were aligned with those of TA1, but made large gains in the final phase of intervention, in all three expressive language subtests. It may be that strategies such as self monitoring, attention to detail and justifying his responses had a greater effect over a slightly longer period of intervention, but there are no direct links between intervention and outcomes.

What emerges as an interesting finding is the role of emotional and behavioural characteristics of the participants, and the impact that these have on test results. Impairments in emotional and behavioural responses are typical of these children, (Botting and Conti-Ramsden 2000) and interact with their knowledge in affecting test performance. Feuerstein et al (2002 p.141) noted that affective factors impact substantially on cognitive performance in standardised tests. Assessors carrying out DA on a single occasion are as likely to perceive a 'snapshot' of a child's ability, impacted by his emotional state at the time, as those carrying out a standardised test, but in some cases the lesser emphasis on achievement, and the opportunity to build up a greater rapport enabled good learning potential scores to be demonstrated on the DA, many of which were subsequently shown to have predicted gains in language as measured on the CELF-3. The DA ought to include detailed notes of the emotional behaviour of the child, as recommended by Lidz (2003), Haywood and Lidz (2007) and Tzuriel (1991).

Influence of progress in DA scores
Two further case studies illustrate particular benefits of repeated administrations of the DA procedure that do not specifically link to initial test results or outcomes of intervention. D4 was presented in section 7.5 as a child with inconsistent responses, and poor cognitive

and metacognitive skills. Despite being in the experimental group, with these findings presented to his SLT at T2, the goals set in the following phase of therapy were not substantially altered, nor were more metacognitive strategies introduced, and this may have been justified by the results from T2 which were substantially poorer on both the CELF and the DA, than at T1. However, despite performing extremely poorly on the tests at T2, D4 recovered substantially at T3, particularly on the DA, while the CELF total scores were still lower than at T1. It was thought that these scores do not reflect the effects of intervention informed by the DA, as much as inconsistency in test performance. However, both standard scores and total raw scores on the CELF-3(UK) obscure the details of D4's performance, and there was in fact a steady improvement in scores on expressive subtests, most notably Formulated Sentences and Sentence Assembly at T4 which are tasks related to the DA task. Similarly, there was improvement on the DA scores with each successive retest. Gains may possibly be in response to the mediation received during the DA and it would appear that D4 may have benefited from the mediational strategies used in the DA, and implementation of these in his intervention may have resulted in further improved outcomes. The present author would also make use of metacognitive strategies to support poor working memory in this child who was not reading more than simple single words, and therefore reliant on working memory, yet had very poor performance on the Recalling Sentences subtests of the CELF-3. Mediation should however, be addressed to the substantial weaknesses in receptive language that are substantially below D4's non-verbal reasoning scores according to the Ravens CPM, which should support concept learning (Magnusson and Naucler 1993).

Child D1 was randomly allocated to the Control group, and findings from the DA were not passed on to his SLT until the third phase of

therapy, but it was seen that some changes to the targets of intervention were altered in each phase of therapy. Results measured by the CELF showed D1 to improve from T1 to T2 and T3 in all the Receptive subtests, and in Formulated Sentences, Recalling Sentences and Sentence Assembly. Between T3 and T4, with modified intervention, D1 made further gains in the expressive Subtests, but not Receptive. It would appear that D1 was responsive to therapy and improved in whichever areas were specifically targeted, gaining a few raw score points, but insufficient to alter his standard scores from the lowest rankings. In summary, the findings of the DA at T1 did not identify a good learning potential, and specific recommendations did not lead to more improved outcomes from therapy, but at T2 and again at T3, the substantial improvement in DA scores signalled responsiveness to previous mediation and retention of learning strategies, and thus the value of DA in this child lay in the measurement of incremental change and like in the previous case, in the identification of useful learning strategies.

In summary, single case data demonstrated that group scores used in statistical analysis obscure the variable outcomes of individual participants. A small number of substantial gains in CELF-3 scores affect the group means that are based on small numbers of participants. The majority of children make modest changes from one retest to the next. The value of the Dynamic Assessment in identification of those with good learning potential was verified by case data, although the prediction of substantial gains in intervention is largely dependent on the emotional and behavioural readiness of the child to respond to learning. Clear links between the nature of the intervention implemented and the outcomes cannot be established, and in many cases, although recommendations were

made, it would have been inappropriate to alter a successful intervention programme.

8.2 General Discussion

8.2.1 Aims and Outcomes of DA

The aim of a Dynamic Assessment is to elicit information about an individual's potential to learn, and although this concept is bound to the construct of intelligence in complex ways, proponents of DA (e.g. Jeltova et al 2007, Hessels 2009) have shifted the focus of assessment away from intelligence, and into more content specific areas. This approach was adopted in the current study; it was thought that the assessment of potential to improve *in language* would facilitate greater insights into children with language impairments, and potentially enable formulation of more effective interventions. Indeed this was demonstrated for some children on a case-by-case level, specifically those children not benefitting from regular ongoing intervention. The task enabled the examiner to understand which aspects of sentence formulation were most difficult for the participants, and which structures were more amenable to instruction. In the end, however, the examiner produced recommendations for intervention related not to the structural complexities of language, but to the domain general cognitive processes implicated in language learning.

Detailed analysis and description of the responses to particular grammatical structures included in the DA, revealed difficulties that were supported by previous accounts in the literature of difficulties for children with LI. Dative sentences, and those containing non-canonical word orders emerged as the most difficult items, and the facility to reverse semantic items in simple argument structures but not in more complex ones, supported the findings of Bishop (1997) and van der Lely and Harris (1990). However, the SLTs were aware

of the grammatical weaknesses of the children with LI, and included them in their longer term if not immediate, plans for therapy. The experimenter did not find it necessary to recommend specific grammatical structures for intervention. Rather the focus was on the problem solving skills of the children, including their ability to label parts of the sentence, and explain the rules for sentence construction in explicit terms. In addition, the assessment (and subsequent recommendations) addressed the children's cognitive skills of gathering and examining data, formulating plans, checking their responses and justifying their decisions. Weaknesses identified in this area were consistent with the findings of impaired executive functioning described by Marton (2008) and Henry (2010).

In effect, the DA successfully and usefully complemented the results of standardised language tests by providing additional information pertaining to more domain general cognitive skills. This is consistent with the position of Haywood and Lidz (2007), who see DA and standardised tests as complementary. Further, Feuerstein (1979) incorporated subtests utilising different modalities of processing, into his 'Cognitive Map' and into his battery of LPAD tasks. Cognitive strengths or 'deficiencies' (Feuerstein's term) elicited by the verbal 16-Word Memory Test may be compared with those elicited by the Positional Learning Test (Feuerstein et al 2002) which relies on visual stimuli, and the Trimodal Analogies Test (Feuerstein et al 2002 p208) specifically compares analogical reasoning in the three modalities of verbal, pictorial and symbolic stimuli. Thus the information about cognitive functions elicited by the verbal/linguistic demands of the current procedure might usefully supplement skills demonstrated in other modalities. The direction of this data is suggested by the discrepancy between the results of the CELF-3 and the non-verbal Ravens CPM, but in DA terms, the current test should be supplemented by the dynamic LPAD Set Variations, (Feuerstein et

al 2002 p209) which is a dynamic measure, based on the structure of the Ravens CPM, probing the strategies used and the learning potential elicited in non-verbal analogical reasoning.

The reports resulting from the study were also comparable with the outcomes of the case studies of DA of children referred to Speech and Language Therapy, reported by Peña and Gillam (2000). The following combination of linguistic and metalinguistic /metacognitive recommendations for intervention was produced by the authors:
'Provide F with support for learning the names and functions of things..help her use her new vocabulary during everyday activities…use MLE strategies to focus her attention and help her understand the importance of the activities selected. …'(p 553).

In this and further examples which may be seen in Appendix XVII, the similarity between the recommendations produced by Peña and Gillam (2000) and those produced by the examiner in the current study are apparent. It can be seen that the targets for intervention are couched in similarly broad terms to those produced in the current study, except that more specific mediational techniques are identified. The methods of DA used differed from those of Peña, who used a more traditional test-mediate-retest format, but the application of mediational intervention with language specific content was similar, and the recommendations that ensued represent similar language based targets.

The procedure was based to a large extent on the Graduated Prompts approach to DA, influenced by the work of Campione and Brown (1987) and of Resing (1997). The measurement of learning potential was operationalized as an inverse of the number of 'hints' or prompts an individual needed to reach a specific criterion. Again the construction of the training procedure was designed to include

both teaching of task specific strategies and metacognitive executive control skills (Campione et al 1982, cited by Resing 1997 p.74). This was effected by having the first prompts in the *'weak-to-strong series'* (Campione 1984 cited by Resing 1987 p75), as ones that tapped into general metacognitive strategies such as remembering what was done before, while the final, most specific hint ensured that the child sees how the problem is solved. The current procedure was shown in Chapter 5, to be able to differentiate between children on the basis of their need for prompting, and the needs for teaching of both task specific material and metacognitive strategies emerged. Resing (1997 and 2009) also noted that the advantage of the DA was to deliver qualitative information about the quality of the answers and the reasoning strategies that children employ when solving the task, and the reports written about the children demonstrate that this type of information was indeed elicited. The effectiveness of the Graduated Prompts procedure in the current protocol validated its selection as a method, and the outcomes achieved parallel those recorded by researchers who have similarly used the procedure such as Guthke (1997) and Resing (1993).

8.2.2 Outcomes of the exploration of the implications of DA for intervention

Fey, Long and Finestack (2003) formulated ten principles for facilitation of grammar in children with language impairments, four devoted to target selection, and six to intervention methods. They note that some of those related to intervention procedures are supported by empirical evidence, while the first four principles of goal selection are *'crucial intervention elements that are not considered controversial'* (p.5). These include the fourth principle, which proposes a concept embodied by DA, namely that the specific goals set for grammatical intervention for a particular child must be

based in part on the child's "functional readiness" to achieve the target (p.7). This appears as a rewording of Vygotsky's ZPD, yet the assumptions on which this 'readiness' is based, are not clear. Fey, Long and Finestack assume children to be ready to master forms which they have already partially mastered and tend to use inconsistently, and this principle recurs in developmental studies of the acquisition of both phonological and morphological forms. However they note that Nelson et al (1996) thought that partially mastered targets may continue to develop to full mastery even without intervention. This was, however, disputed by Connell and Stone (1992, cited by Fey, Long and Finestack 2003), who further advocated that 'readiness' might be determined even when there is no evidence of partial mastery by a child, purely on the basis of developmental appropriateness. Although justified by theoretical descriptions of language acquisition in children with LI, there appears to the current author to be a confused and poorly substantiated practical implementation of a sound principle.

Determination of readiness to learn structures, defined by stimulability to lower level prompts or minimal mediation in a dynamic assessment, would seem to provide the evidence to support the implementation of specific goals as recommended by Fey, Long and Finestack. The intention of the stated principle is to increase intervention effects by selecting appropriate goals (Fey, Long and Finestack p.7) which is consistent with the aims of the current study, and it would be recommended that the contribution of DA be incorporated into general principles for management of children with language difficulties. It may be recommended, however, that reports of the findings of DA such as those produced in the current study, express the knowledge gained about an individual's use of grammatical structures in terms of their readiness for intervention.

The remainder of Fey, Long and Finestack's (2003) principles for goal selection are also compatible with the theoretical basis for DA. It seems eminently sensible to argue that goals should facilitate wider functional improvements, which can in part be implemented by 'bridging' (Feuerstein's 'Mediation of Transcendence' 1991; Lidz, 1991) or extending learning into other contexts. The intervention procedures, however, are different to those advocated by the current cognitive approach to intervention, and relate to evidence for positive outcomes of intervention arising from studies described in section 2.3.1 as' Traditional'. Some of these studies were able to provide empirical evidence for the positive effects of specific interventions, for example Nelson et al (1996) demonstrated significantly higher rates of learning of a variety of grammatical targets treated via conversational recasts and by imitation procedures, in comparison to untreated control structures. Instances of spontaneous use of structures were higher for those treated by recasting than imitation, in both children with language impairments and typically developing children. Outcomes in this study were measured by counts of spontaneous use of structures in clinic or home settings, both during and after the intervention programme. However Friedman and Friedman (1980) achieved overall gains in language, but did not elicit positive treatment effects after a programme of eight months of intervention. They noted that the main effect statistic did not take into account individual differences in performance, despite group sizes of 25 and 16. The current study went further than Friedman and Friedman in investigating these individual differences at group and case-study level. The findings indeed confirmed Friedman and Friedman's view that intervention effects may be masked by this heterogeneity. Furthermore, the instrument of measurement in the study was the DSS (Developmental Sentence Score, Lee 1974), and the authors noted that the scale itself may not have captured the syntactic growth that

occurred in some of the children. Again, the current study was able to explore the issue of measurement more fully, but in this respect did not find particular cause to conclude that outcome measurement issues were at the heart of the results. Weismer and Murray-Branch (1989) similarly found treatment effects, but failed to find significant differences between the two treatments compared, using responses to 'generalisation probes' as a measure of outcomes.   This may be due to the ways in which these probes were actually used, since the present study identified a gap in application of the DA knowledge to therapeutic practice.

Ebbels (2008) summarized the findings of intervention studies concluding that there are few randomized control trials targeting school age children with language impairment that are able to determine the superiority of one method over another for achieving language gains. One of the reasons for this is that studies tend towards a model of examining whether the intervention works in a majority way.   However, findings that different children and different targets appeared to respond to different approaches emphasises the need for more in depth studies that can tease apart the characteristics of children who will benefit most. Studies are also difficult to compare because of the widely differing lengths of the intervention periods and the different methods of measurement of outcomes. Similarly, considering metalinguistic approaches, Ebbels noted that trials of Colourful Semantics (Bryan 1997) did not include control groups and therefore effects of intervention could not be definitively determined. The efficacy of Shape Coding (Ebbels 2007), however, was supported by a randomized control trial, but Ebbels (2008) noted that only a limited number of areas of language had been investigated, and further studies are necessary to determine which children and which targets benefit from this intervention.   In the light of this summary, it is not wholly unexpected that definitive

overall intervention effects were not demonstrated in the current study. Group sizes were small, exaggerating the effects of individual variations, and heterogeneity of language impairments in terms of both severity and nature of impairment was great. Few treatments can be shown to benefit all children, and the baseline therapy to which the 'modified' therapy was compared, was itself not actually defined, but a mixture and combination of targets and methods, many of which continued and overlapped into the 'modified' intervention, so that the comparison was not made between two clearly defined and mutually exclusive programmes. However the use of children as their own baseline control meant that it was possible to adjust for non-DA intervention progress.

In effect, in this analysis more stringent controls over individual variation were in place, with repeated measures of the same children being compared, and variables of dosage, SLT skill, and school were also controlled implicitly by using a 'within group' covariate. Further, overlapping targets and interventions between baseline and modified intervention were subsumed with only the changes made differentiating the two phases of therapy, and this highly controlled analysis did indeed reveal intervention effects.

8.2.3 Aptitude x Treatment

Recent texts have emphasized the need to return to multidimensional assessments and insights into clients, in management of communication disorders of all types. Joffe, Cruice and Chiat (2008) for example, pointed out that we need a greater understanding of interactions between different levels of the language system, verbal and non-verbal processing skills, language and other abilities, as well as 'specific vulnerabilities' and interactions with the environment. They also emphasize the need to understand how the person, his environment, *language, ability, activity,*

participation and life context work together' (p xviii). Further, in intervention it is important to consider the interaction between variables of the client, clinician, task and context. Given the number and range of these variables, it would seem almost impossible to plan any intervention that would be applicable for more than one individual, and in fact to plan management for an individual in advance, as all criteria are continuously interacting and changing.

The findings of static assessments are representative of the status of the multiple variables at the time of testing only, and it is well known to practising SLTs that children are reported to perform and achieve differently in SLT sessions, at home and at school. Dynamic Assessments permit manipulations of some clinician variables, identified by Dodd (2008) as affecting intervention, for example manner and style of interaction, and support offered to the child. 'Online' manipulation of explanations, scaffolds, and demands, may result in findings with greater applicability. For example a rapid, unsupported demand for responses may elicit anxiety in a child that affects test performance, while encouragement to 'have a go', and hints towards a solution may elicit a more representative indication of ability. Similarly, Joffe (2008) identified variables related to the task, and as previously described, DA enables manipulation of modality of presentation of materials, as well as number, type and complexity of stimuli, to determine the effect that these differences have on the performance of an individual.

Thus DA sets the clinician up with a broader knowledge of variables affecting the individual client's performance, and arguably a better starting point for the planning of intervention. The methods also elicit some evidence of what works for the client in terms of prompts or mediations that are required in order to facilitate the child achieving the solution of a problem. Again this information is limited

to the solving of a particular problem at a particular time and in a particular place, but facilitating generalization at every stage of learning is an important component of effective management. Similarly continuous monitoring of criterion learning should enable modifications to be made to the intervention programme as and when required. Findings of the present study also indicated that lack of measurable progress in intervention signals that changes to the programme should be considered.

Early studies that identified aptitude x treatment effects produced some interesting but inconsistent group findings. Methods of intervention for children with language impairments were linked to measures of IQ, language and visuo-motor achievement (Friedman and Friedman 1980) and there was some apparent logic to the finding that lower functioning children were more responsive to more didactic approaches. However, this finding was not supported by Cole and Dale (1986, cited by Leonard 1998) an issue attributed to the lower non-verbal IQs of the children in the study. Similarly Yoder, Kaiser and Alpert (1991) who in fact studied pre-school children with learning difficulties, also found imitation based approaches to favour the higher functioning children in their study, and it was thought that the actual range of IQ might be the determining factor. However, it may be seen that numerous other variables pertaining to the different populations must have made the findings of these two studies incomparable.

A discussion of this issue by Leonard (1998) concluded that there is no clear evidence of which children benefit from which therapy approach, or which target structures may be more amenable to teaching via which method. The present author supports this view and contends that interventions must be individualised for every child. Mediation, however, is a technique that is theoretically

grounded in Vygotsky's socio-cultural theory, and has no fixed methodology other than the transmission of learning by a mediator that as Feuerstein explains, stands between the stimulus and the learner (1991) and makes the task accessible to the learner, in as far as it is within his zone of proximal development. To facilitate retention and transfer of this learning the individual needs to actively participate in the learning, and for this reason metacognitive monitoring is required. However, the content and level of metacognitive awareness that can be achieved is individually variable, and may be related to linguistic skills, and non-verbal processing skills, and thus even the metacognitive requirements must be individually tailored and mediated to the individual. Moreover, whilst individual static tests may not predict progress in therapy well, the use of DAs such as the one used here are likely to increase predictive validity at this low-functioning end of the language spectrum.

8.2.4 Mediation and SLT

The fundamental principles of how to make a session mediational were made explicit to the SLTs participating in the study via the manual (Appendix IX). The seven functions of mediation described by Haywood (1993) were included, along with simple, practical examples of the kind of questions, challenges and activities that could be included. It was thought that this would be sufficient to assist SLTs in modifying their therapy into a more mediational style, consistent with the theoretical underpinnings of the study. However, examination of a sample of videotapes of sessions showed that there was little uptake of specifically mediational behaviours by SLTs in their sessions. There was a slight substitution of challenging questions, which require self monitoring of linguistic productions by the children, by process based questions which require them to justify their decision making. However, informal observation of the

responses to these questions revealed that the children produced linguistic based responses rather than metacognitive ones. For example, children asked 'How do you know that's right?' answered 'Because it starts with 'Is'/Because it's a question/ Because it makes sense' rather than 'Because I have checked it / because that's how I did it before'. Further prompting and probing by SLTs to elicit these reflective responses was not evident.

Despite the concrete descriptions of mediational behaviours provided, the present author believes there are more fundamental differences between traditional SLT and mediational intervention. Traditional therapies for language impairment include modelling, recasting, which is a form of modelling 'correct' structures, and imitation. These supply the 'answer' to language formulations at the outset of the intervention. Mediational methods, in contrast, consider providing a model and requesting imitation to be the highest level of mediation, which should be used only as a last resort, when less directive prompts have failed to elicit the required response from a participant. This framework is represented in the 'Required Mediational Intervention' (RMI) scoring framework used by Feuerstein (Feuerstein et al 2002 p.533 see Appendix XX). Less directive cues should be used first, with more and more specific assistance delivered as required. The intervention is adaptive, contingent upon the needs of the learner, rather than prescriptive, using predetermined models.

An example of recasting, with a request for imitation contrasted with an imagined script for mediation of the structure may be seen in Box 1. In this example, the clinician chose to mediate rule governed behaviour to the child as a strategy for managing irregular past tenses. Although this is consistent with the dual process mechanism for acquisition (Pinker 2000), the content was selected for the

accessibility of the strategy for learning, and clinicians may select whatever strategy is consistent with their theoretical interpretation of problems, as well as any that they think would assist the learning of the individual child with whom they are working. What is essential however, is that the child is guided to the principle of problem solution and the strategy that he may use in future examples, to facilitate transfer. The session would also include a discussion of other life examples in which rules do not apply, in order to reinforce the metacognitive concept. Examples from the study include the child who was able to say that she knew she had to make a question 'because you said'. Discussion proceeded to clarify that doing what she had previously been shown by a teacher was a good way to try and solve a new example.

Box 1. Contrasting scripts for Intervention

---

1. Recasting with request for imitation

C: Boy runned* to the bus

T: Yes, the boy **ran** to the bus……………..(recast)

T: Can you say the boy **ran** to the bus?.............(request for imitation)

2. Mediation of the structure

C: Boy runned* to the bus

T: Yes, you are right that's what happened, but did you say that right?

C: yes

T: How do we say something that happened yesterday, or in the past? (Make sure that the child knows the concept of past before addressing past tense morpheme)

C: Runned*

T: Can you think of how we might say it with walk instead of run?

C: walked

T: Right, you know how to make the word. Can you tell me what you have to do to make the word right for the past

C: put on –ed

T: Excellent, that's the rule isn't it? Can you tell me the rule?

C; Put –ed for the past

T: Good, you know there are rules for changing some words that we have to remember to use. But do the rules work for all words?

C; I don't know

T: Can you think of any words where the rule doesn't work?

C: No

T: what about 'Go'? Do we say goed*?

C: No

T; What do we say?

C: (Not sure)

T: We say 'went' don't we?

C: yes

T: So some words don't work. What about 'run'? Could it be one that doesn't work with the rule?

C: yes

…etc. Proceed to teaching.

Reinforcement of target behaviours in mediational sessions is arguably more specific and directed than that delivered in traditional interventions, although the nature of feedback used by clinicians is very dependent on the individual clinician. Frequent use of the 'Good sitting/ good listening/good talking' framework is not sufficiently specific for a child to grasp what he has gained by listening and why it is important that he listen to input from others, and the current

writer would argue that this does not constitute differentiated feedback. Further, attention in SLT sessions is frequently maintained by rewards extrinsic to the task itself, whereas mediational intervention recommends the creation of motivation intrinsic to the task by demonstrating how, when and why the skills learnt are useful and important to the learner.

Although there are some similarities between the fundamental principles of language therapy and mediation, there are also differences that may only become apparent through more study of mediation, and training in mediational techniques may be necessary in order for SLTs to be able to adopt the principles of mediational intervention. Whether SLTs would be able to adopt Dynamic Assessments without further training is a question that has not been investigated. Deutsch and Reynolds (2000) found that educational psychologists in the UK expressing an interest in DA had not been able to take up the practice primarily because of lack of training, as well as materials and support, and also because of restrictions on their time, and the constraints imposed by local authority provision. However, Miller, Gillam and Peña (2001) have published a manualised Dynamic Test of Narrative that includes recommendations for mediated intervention sessions as part of their test-teach-retest format. The assumption is that trained speech and language therapists would be able to adopt the procedure from the information contained in the manual, without further training.

Subsequently, Peña, Gillam and Miller (2003) produced revised scripts for the intervention linked with the DA of narratives. They noted that it was discovered that children made greater gains when the focus of intervention was on episode elements and story structure, rather than individual aspects of narration as described in the original manual. Two scripts for mediation of these aspects were

produced, and are explicit regarding the mediational intervention to be provided, giving prompts to use as well as cues to the clinician about what concepts to expand. This presumably supports the clinician untrained in mediation to a greater extent, and enables them to complete a DA incorporating mediational intervention. However, it would not be possible to script mediational interventions for the range of structures and targets recommended for the children in the current experimental study, nor for the extended periods of intervention that were not part of the assessment procedure, but rather management following assessment. Under these conditions, SLTs would have to be trained in principles of mediational intervention in order that they could devise their own sessions.

8.2.5 Emotional and Behavioural Issues in LI

A recurrent trend that emerged from the case study data, was the effect of emotional and particularly anxiety states on the performance of children during both the static and dynamic assessments. While formal assessment procedures do not usually take account of such factors, other than the opportunity for testers to make observational comments after scoring the test, the DA and in particular the RtM scale offered opportunities to the examiner in the current study to record behavioural responses. It is worth noting however that this was the part of the DA that was least amenable to psychometric scrutiny. According to the LPAD theory (Tzuriel 1991), tasks in the dynamic model of assessment are presented with simpler, more manageable examples first, with the level of difficulty gradually increased in order to help testees overcome their fears of failure. Further, a component of the DA should, according to Tzuriel (1991) try to change the motivational and affective features during the assessment, and record the type and degree of mediation required to effect that change in a similar fashion to the recording of

mediation required to effect changes in intellectual or in this instance language, functioning.

Botting and Conti-Ramsden (2000) reported that children with language impairments are often reported to have emotional, social or behavioural difficulties, and investigated these features in a population of children aged 6.6-7.9, taken from language units in the UK. Like the cohort of children in the present study who were screened with the SDQ (Goodman 1997) and did not show any extreme behavioural abnormalities, the average behavioural scores of children in Botting and Conti-Ramsden's study were within a normal range. However, a high proportion of children with expressive and receptive or complex language impairments were found to have a clinical level of emotional and behavioural difficulties. At the later stage of the study, when the participants were 7.5-8.9, the difficulties emerged as more emotional than anti-social, except in children whose language difficulties included pragmatic components, who were shown to have more peer relationship problems.

When the same young people were 16 years of age, Conti-Ramsden and Botting (2008) again reported a higher incidence of anxiety and depression in adolescents with LI than in typically developing adolescents. A review of studies by Benner, Nelson and Epstein (2002) found that on average, clinical diagnoses of EBD occurred in 57% of children with language impairments, and Beitchman et al (2001, cited by Conti-Ramsden and Botting 2008) reported that in later years, children with SLI were at greater risk of anxiety disorders. However, Conti-Ramsden and Botting's review of studies (2008) pointed out that there was little consensus between all studies, regarding emotional health of individuals with language impairments. Furthermore, the latest follow up of the Conti-Ramsden cohort has shown that the difference between those with LI and their

TD peers had disappeared once they left formal education at 17 (Wadman, in press), suggesting that the pathways of emotional difficulties are complex in this population.

Although the relationship between behavioural problems and language impairment may be complex, with causality versus co-morbidity difficult to determine, Botting and Conti-Ramsden noted that behavioural problems may lead to slower progress of children in language therapy. While this was not investigated directly in their study, the authors noted that the relationship between different subtypes of LI and the patterns of behavioural difficulties associated with them suggest that different strategies of intervention may need to be considered. Schery (1985) found social-emotional factors to be a significant predictor of improvement in language in young children with LI enrolled in special school programmes. Although the social-emotional factors were not a significant predictor of pre-school language levels, they did have an impact on remediation of those problems. Similarly, Lauchlan and Elliott (1997) found that the low self esteem of children with learning difficulties which they attributed to the children having experienced long periods of academic failure, resulted in difficulty for the assessors to mediate to the children. Tzuriel, (1991) noted that some children with poor motivation, low self-esteem or resistance to support were not accessible to mediation. The recommendation for mediators was to intervene in specific ways to interpret the emotional response, and emphasize meaning in an attempt to increase the child's acceptance of mediation. More detailed assessment of behavioural features in the present study may have resulted in differentiated targets for intervention to accommodate such variables, resulting in different gains.

The behavioural ratings used in the Botting and Conti-Ramsden study, while more comprehensive and rigorous than the measures used in the present study were similarly based on the perceptions of teachers, which in the present study were shown to correlate poorly with language tests and to have poor retest reliability over time. However, in the current study there were discrepancies between the ratings by teachers and the perceptions of SLTs and the experimenter in 1:1 sessions, who may have different experiences with the children, and progress in individual intervention may not be the same as that achieved in the classroom setting. Comprehensive monitoring of performance during assessments in the classroom and in individual SLT, and likewise of performance in teaching or 1:1 intervention sessions, would add to the knowledge about a child, and the impact of environmental factors on his performance.

Haywood and Lidz (2007 p.237-256) presented a detailed case study of a child with language impairments alongside some other academic failures and attention difficulties, who also manifested anxiety and behavioural features such as resistance to treatment and defensiveness in intervention sessions. Detailed Dynamic Assessment demonstrated better performance and less anxiety when tests were not administered under timed conditions, but recognition of the examiner's efforts to identify her needs and help her, made the child more anxious, inattentive and resistant. Cognitive deficiencies in attention, and memory, self-regulation and flexibility were determined. Interventional recommendations were for educational support as well as for regulation of attention, problem solving, and of her emotional reactions, some of which were to be dealt with by cognitive strategy coaching. The authors concluded that without Dynamic Assessment, the academic difficulties would have been described and possibly overestimated, while the underlying executive control problems may not have emerged, and

ultimately recommendations for useful strategies were provided. The case demonstrated, however, that it was only after a series of lengthy and frequently modified individualised Dynamic Assessments that the nature of the child's difficulties could be determined, and the examiners were not deterred by poor performance, anxiety or resistance, but continued to modify mediations until the child's true abilities were determined.

In the current study, observations were made regarding anxiety and resistance that the experimenter was aware were affecting a child's performance, but the protocols of the study did not permit the extended probing and additional testing required to fully assess the extent of the children's difficulties. This would require a return to entirely individualised assessment using a battery of assessments in different modalities. It is apparent, however, that proceeding with traditional language therapy without considering the underlying cognitive and behavioural difficulties may not achieve optimal outcomes.

8.3 Limitations of the Study

8.3.1 Dosage of Intervention

Numerous variables in the intervention programmes were not controlled in the current effectiveness study, but the design of the study as a case study series allowed for between child variables to be controlled by comparison of change in a child with the change in his/her own baseline period of intervention. The issue of dosage was not, however, sufficiently controlled because of the range of number of sessions of therapy made available to the children that in some cases was too few for therapy gains to be likely to be measurable. Although SLTs were not actually asked about the length and frequency of their sessions, some children may have attended only one SLT session per week, which with interruptions for absences,

half terms, school outings, events, and timetable changes, may have amounted to approximately eight sessions of intervention between assessments. In some instances, the child's SLT session was given to the experimenter for assessment, and it was also in the experimenter's experience that children were unavailable for testing on several occasions so that the experimenter returned to the school two or three times in order to complete the assessment.

8.3.2 Quality of feedback to SLTs and the use of video

It was previously noted that the video recordings of the Dynamic Assessment sessions might have been provided to the SLTs of the children in order to supplement the information provided in the reports with live and concrete examples of the children's achievements in the DA. It might also have proved useful to make the videos available to teachers in the language bases and schools who were in most cases extremely involved in the SLT programmes and working in close co-operation with the SLTs. Research by Haywood and Lidz (2007 p.220) and Delclos, Burns and Vye (1993) showed benefits to observers of DA sessions who gain more insight into the children in their care, and make more optimistic predictions about what they may achieve. Reports were necessarily kept short and accessible, but it may have been useful to phrase the findings about grammatical knowledge and transfer skills in terms of the demonstrated readiness to learn structures. This may have made the information about transfer more transparent and more closely linked to explicit recommendations for intervention.

There is a considerable amount written about the reports written by educational psychologists (EPs) after they have conducted DAs. Freeman and Miller (2001) found that despite the content being unfamiliar, SENCos rated the information contained in reports of DAs useful, especially with regard to the use of strategies by teachers

and pupils. The DA was also helpful in understanding children's abilities and needs and suggested an appreciation by the SENCos that these could be altered by intervention. Similar findings were reported by Bosma and Resing (2010) in whose study the teachers also found the reports of DAs interesting and useful, but not significantly more so than reports of static assessments. Lauchlan and Elliott (1997) similarly reported that teachers of children with learning difficulties found the reports of DAs informative and accurate, but that they did not impact upon the work carried out in classrooms. The professional relationship between the EPs who carry out assessments of children and make recommendations and those educators who implement the recommendations and programmes in the schools, is not, however, replicated by speech and language therapists who both assess and treat, and for whom the processes are more intertwined. It was considered an informative exercise to expose experienced SLTs who were unfamiliar with DA to reports arising out of DAs related to children with whom they were familiar. However the expectation placed on the SLTs that they would be able to modify their intervention practices on the basis of a brief information leaflet about mediation was too great, and additional training would need to be supplied to SLTs.

In the ideal study, with fewer practical and time constraints, a wider battery of tests would be used to monitor progress after intervention, and the length of intervention blocks would be lengthened, increasing the interval between tests. SLTs would be provided with more detailed outcomes of the DA, accompanied by the video of the DA, that they might watch in the company of the experimenter who could provide an interpretative commentary if necessary. Training of the SLTs in mediational techniques, and more detailed planning or scripting of interventions would have changed the nature of the

study altogether, and are aspects for future studies to be considered in the next section.

8.4 Future directions for Research

The following research questions posed by Hasson and Joffe (2007) were identified by Lidz and Peña (2009 p.129) as relevant questions not only for SLTs but for all those trying to integrate DA into school based management.

1. Can we differentiate modifiability/stimulability/learning potential in groups of children so as to determine children in the population who will require specific SLT intervention in order to progress?

2. Can we then standardise and possibly quantify, methods for measuring responsiveness to language learning in the individual child?

3. How can we use DA methods to identify which methods of intervention, and how much intervention will benefit a particular child?

These questions have to some extent been addressed in the current study which devised a method of assessment that was able to differentiate responsiveness within a group of children with diagnosed language impairments. The method was quantifiable and standardised enough to be replicable, yet accessed individualised information about intervention for a particular child. The method has potential for extension into a reliable determiner of the needs of an individual in order to progress.

However, the real validation of the DA would be a demonstration that it does in fact lead to improved outcomes from intervention, and

this has not emerged convincingly from the current study. It would be consistent with the theory of DA adopted in the current study, that of modifiability by mediational learning (Feuerstein 1991) if intervention utilising mediational methods could be shown to be informed by the DA and effective in achieving maximal outcomes. To this end, further studies are required.

It may have been possible within the protocols of the current study to plan the individual interventions for the children in detail in collaboration with the SLTs. A framework was devised, but not implemented, that permitted individual goals to be fitted into a sequential programme of mediated learning. Each stage was defined in broad terms to ensure that the child

a) was involved in the setting of his own targets, and understanding why they were important;

b) knew or learnt the relevant vocabulary for his task including where necessary, metalinguistic vocabulary;

c) monitored his own learning at every stage;

d) generalized learning to new examples within language and in other domains.

A further study implementing such a programme would be important to establish firstly whether improved outcomes can be elicited by the use of such a programme, and secondly whether it could be manualised or passed on to SLTs without the need for further training. A series of case studies exploring the efficacy of interventions closely linked to the findings of DA would add to the body of knowledge in this area.

Further research that is needed is to train SLTs in techniques of mediation, and then determine whether outcomes of intervention are altered, with and without the information provided by DA. SLTs could

be trained in the same way as EPs, but without the tools for intellectual assessment (usually the LPAD in training offered in the UK). Training in mediation might also make the reports of DA or indeed the independent practice of DA, accessible to SLTs.

An investigation of the current practices of SLTs regarding DA, and or their awareness of DA as an alternative assessment, would prepare the ground and establish the need for and interest in the training of SLTs.

In summary, there is considerable scope for extension of the current study, both to extend the test itself by modifications to the materials and application to a wider population, and to explore further the role that DA might have in informing intervention. The outcomes that may be achieved by training SLTs in DA and in mediated intervention, and by using alternative methods of DA have not yet been addressed, and a battery of Dynamic Assessments addressing various levels of language may be applied to the study of both typical and impaired language development.

8.5 **Summary**

This chapter has addressed some of the major issues arising from a research project that both developed a novel assessment of language, and explored its use in a clinical context. Multiple individual results have been analysed to identify the strengths of the procedure as an assessment tool, and its value in contributing to the devising of individualised interventions. Factors influencing the effectiveness of intervention in the context of school based programmes were explored. The next chapter summarizes the conclusions that were drawn.

CHAPTER 9

SUMMARY AND CONCLUSIONS

CHAPTER 9   SUMMARY AND CONCLUSIONS

There were two main components to the present research study. The first devised a novel methodology for the assessment of language in children with language impairments, and the second investigated whether feeding back the results of the assessment to the Speech and Language Therapists working with the children, would affect their outcomes of therapy. A number of interesting findings emerged, and have been presented and discussed in the foregoing chapters.

Although DA has been in use for many years, psychologists working in the field have acknowledged the minimal uptake of the methods, and the limited use being made of the methods. Nevertheless, according to Haywood and Lidz (2007) the expansion in the field is recent, and their recommended approach is to offer guidelines to encourage a range of professionals to include DA in their practice, particularly to '*bridge the gap between assessment and treatment*' (p.20). The current project responds to the call to ask appropriate assessment questions and use enhanced information to inform clinical intervention. The application is novel in a number of specific ways. Primarily a combination of theoretical models and methods has been used in order to meet the need for detailed individualised information about clients to be extracted in a short period of time and in a way that enables comparison of individuals in a clinically informative way. Predictions that facilitate planning of services are relevant in the current economic climate. Secondly, the findings refer to several issues current in the language impairment literature that relate to the role of working memory and executive function, and to the intervention issues of dosage and outcome measurement.

With regard to the procedural aspects of the Dynamic Assessment, the following conclusions were reached:

- The combination of Graduated Prompts and mediational strategies was shown to be effective in eliciting and scoring the responsiveness of children with LI to cues for managing sentence construction.

- The content of the test items was appropriately selected for elicitation of structures reported in the literature to be specifically difficult for children with language impairments.

- The procedure was manageable within a short period of time and the information elicited by the procedure justified the investment of time spent in administration and analysis of findings.

- The procedure was validated by moderate correlations with other standardised test results, but shown to add a significant amount of information to that gained from standardised assessments of language.

- The procedure was shown to be sensitive to individual variation and to changes over time.

- Inter-rater reliability of the scoring of mediational levels was high. The internal consistency of the test was high although a small number of individual items in some versions of the test were not reliable, and additional grammatical structures could have been included for investigation, so some slight modifications to the procedure would be recommended.

- The findings of the DA might best be shared with other professionals by making videos of the DA session available to them.

The clinical utility of the test procedure for the informing of service delivery was considered to be an important rationale for the development and trial of the procedure. The improved predictive validity of Dynamic Assessments over static

assessments has been reported, and was similarly demonstrated in the current study. Highlights of the findings were as follows:

- All participants shown to have low DA scores, suggesting good learning potential, were seen to make significant gains in their intervention in the duration of the study period. This suggests that the DA may be included in a battery of assessments as an indicator of prognosis for improvement in the short term.

- Information elicited from the DA pertained to metalinguistic knowledge and metacognitive awareness as well as to abilities in language and facilitated the formulation of recommendations for management in all of these areas.

- Findings of the DA were found to be useful and informative by Speech and Language Therapists involved with management of the children.

- Information about the behavioural and performance variations of the individual children that become apparent to professionals and teachers involved with a child during a course of therapy, was ascertained by the examiner within one or two sessions, enabling individualised planning of intervention from an earlier stage.

- Children shown to be making minimal progress in their regular ongoing programmes of Speech and Language Therapy benefited most from changes made to their intervention following Dynamic Assessment.

Significantly improved or changed outcomes of intervention following modifications informed by the DA were not found when the group was examined as a whole. Several explanations for this finding were proposed, and avenues for further research were recommended. Key findings of this aspect of the research study include the following:

- Sensitivity of outcome measures was explored but not convincingly shown to be a factor in the lack of simple intervention effects

- Heterogeneity of the group of children with language impairments, multiplicity of the recommendations made for remediation, and lack of standardisation of the interventions were all implicated in increasing variability in the cohort study. Resulting group data was not significant and obscured individual changes that were apparent in individual case studies.

- Previous literature suggests that the implementation of intervention ought to be mediational in style. Mediational intervention would be more consistent with the theoretical basis of the assessment and recommendations arising out of the assessment would be better implemented by this mode of intervention. However video monitoring suggested that this was not occurring to maximum effect in the experimental intervention group. Therefore, SLTs may need to be specifically trained in techniques of mediational intervention.

- Emotional and behavioural components of a child's performance emerged as important variables. Regardless of the nature of the assessment carried out and the intervention undertaken, anxiety, resistance and fear of failure affect performance and need to be taken into account in evaluating performance and planning facilitations.

- Dosage of intervention is an unresolved issue that requires further research to ascertain the length and frequency of sessions and the duration of a programme of intervention that is effective in producing positive outcomes.

- Time scale of outcome measurement is also an unknown quantity – several researchers argue that longer time

periods of follow up are needed to truly assess the nature
of effect of the DA on intervention.

In order to situate the findings of the intervention component of
the current study in the context of existing literature about
outcomes of intervention for language impairments, an extensive
literature search was conducted. The procedures of the search
mirror to a large extent, those of Gillam and Gillam (2006), who
in their search for evidence to underpin intervention, similarly
summarize the findings of a small number of studies specifically
addressing interventions with school age children.
Implementation of evidence based procedures then needs to be
balanced against individual and family variables, and service
delivery constraints to select appropriate management strategies
in the Gillam and Gillam report, or to evaluate research outcomes
in the current study. These latter issues of individual variables
contribute substantially to the investigations of *effectiveness* of
intervention methods in the real world context.

Throughout the last decade several studies have attempted to
show whether speech and language therapy is demonstrably
effective, and it is unsurprising, given the range of variables, that
outcomes from intervention studies have at best, been mixed.
Evidence based intervention has assumed centre stage, yet the
evidence for effective language interventions is sparse, especially
with regard to school age children, and data from both EBP
reports and meta-analyses have concluded that individual
variation is crucial. Studies typically involve small numbers of
children, or groups have to be subdivided to reflect different
areas of language, targets of therapy and different baseline
abilities. In this context, findings of measurable improvements on
the CELF-3 (UK) resulting from intervention informed by a DA,
targeting a range of grammatical structures for a group of 8-10
year old children with mixed expressive and receptive language

impairments, who were not improving a great deal in their previous therapy programmes, is encouraging.

With regard to assessment, the procedures used to assess outcomes of intervention, like the procedures used to assess children for diagnostic and prescriptive purposes, produce variable results, despite being individually reliable (Dockrell and Law 2007). These authors also recommend that individual profiles of assessment and informal measures of progress in intervention are utilised as well as group data and standardised tests. In that respect, the DA developed in the current study, when carried out alongside a standardised test, as recommended, meets both needs simultaneously. Previous discussion regarding the needs for more individual dimensional assessments further supports the value of DA on account of the range of information that is elicited from the procedure.

In summary, the inclusion of a Dynamic Assessment of expressive language in a format like that presented in the present thesis would be a valuable addition to the assessment battery used by clinicians in practice with school age children. The information elicited would enable insight into the processing and strategies used by the children, and makes an important contribution to prognosis and intervention planning for children on an individual basis in the clinical setting.

REFERENCES

REFERENCES

ADAMS, C., BYERS BROWN, B. and EDWARDS, M., 1997. *Developmental Disorders of Language.* 2nd edn. London: Whurr.

ADAMS, C., COKE, R., CRUTCHLEY, A., HESKETH, A. and REEVES, D., 2001. *Assessment of Comprehension and Expression 6-11.* London: NFER-Nelson.

ALONY, S. and KOZULIN, A., 2007. Dynamic assessment of receptive language in children with Down syndrome. *International Journal of Speech-Language Pathology,* 9(4), pp. 323-331.

ARAM, D.M., MORRIS, R. and HALL, N.E., 1992. The validity of discrepancy criteria for identifying children with developmental language disorders. *Journal of learning disabilities,* 25(9), pp. 549-554.

ARAM, D.M. and NATION, J.E., 1975. Patterns of language behavior in children with developmental language disorders. *Journal of speech and hearing research,* 18(2), pp. 229-241.

ARCHIBALD, L.M.D. and GATHERCOLE, S.E., 2006. Short-term and working memory in specific language impairment. *International Journal of Language & Communication Disorders,* 41(6), pp. 675-693.

ARMSTRONG, S. and AINLEY, M., 2007. *The South Tyneside Assessment of Syntactic Structures 2.* 2nd Ed. UK: Stass Publications.

ASHMAN, A.F., 1992. Process-based Instruction: Integrating Assessment and Instruction. In: H.C. HAYWOOD and D. TZURIEL, Eds, *Interactive Assessment.* New York: Springer-Verlag, pp. 375-396.

BADDELEY, A., 2000. The episodic buffer: a new component of working memory? *Trends in cognitive sciences,* 4(11), pp. 417-423.

BADDELEY, A., 1996. Exploring the central executive. *The Quarterly Journal of Experimental Psychology Section A,* 49(1), pp. 5-28.

BAIN, B.A. and OLSWANG, L.B., 1995. Examining readiness for learning two-word utterances by children with specific expressive language impairment: Dynamic assessment validation. *American Journal of Speech-Language Pathology,* 4(1), pp. 81-91.

BALTHAZAR, C.H. and SCOTT, C.M., 2007. Syntax-Morphology. In: A.G. KAMHI, J.J. MASTERSON and K. APEL, Eds, *Clinical Decision Making in Developmental Language Disorders.* Baltimore: Paul H Brookes, pp. 143-163.

BANNERT, M. and MENGELKAMP, C., 2008. Assessment of metacognitive skills by means of instruction to think aloud and reflect when prompted. Does the verbalisation method affect learning? *Metacognition and Learning,* 3(1), pp. 39-58.

BECKMANN, J.F., 2006. Superiority: Always and everywhere? On some misconceptions in the validation of dynamic testing. *Educational and Child Psychology,* 23(3), pp. 35-49.

BENELLI, B., BELACCHI, C., GINI, G. and LUCANGELI, D., 2006. 'To define means to say what you know about things': the development of definitional skills as metalinguistic acquisition. *Journal of child language,* 33(1), pp. 71-97.

BENNER, G.J., NELSON, J.R. and EPSTEIN, M.H., 2002. Language Skills of Children with EBD: A Literature Review. *Journal of Emotional and Behavioral Disorders,* 10(1), pp. 43-59.

BERGER, J.L., 2004. Assessment of Learning Potential: Construction and First Evaluation of the Psychometric Characteristics of an Analogical Reasoning Test. *Journal of Cognitive Education and Psychology,* 4(1), pp. 144-145.

BISHOP, D.V.M., 2008. Specific language impairment, dyslexia, and autism: Using genetics to unravel their relationship. In: C.F.NORBURY, J.B. TOMBLIN AND D.V.M. BISHOP Eds, Understanding Developmental Language Disorders. Hove: Psychology Press, pp. 67-78.

BISHOP, D.V.M., 2004. *Expression, Reception and Recall of Narrative Instrument.* UK: Pearson Assessment.

BISHOP, D.V.M., 1998. Development of the Children's Communication Checklist (CCC): A method for assessing qualitative aspects of communicative impairment in children. *Journal of Child Psychology and Psychiatry,* 39(6), pp. 879-891.

BISHOP, D.V.M., 1997. *Uncommon Understanding. Development and Disorders of Language Comprehension in Children.* Hove: Psychology Press.

BISHOP, D.V.M., 1982. Comprehension of spoken, written and signed sentences in childhood language disorders. *Journal of Child Psychology and Psychiatry,* 23(1), pp. 1-20.

BISHOP, D., ADAMS, C. and ROSEN, S., 2006. Resistance of grammatical impairment to computerized comprehension training in children with specific and non-specific language impairments. *International Journal of Language & Communication Disorders,* 41(1), pp. 19-40.

BLEILE, K.M. and HAND, L., 1995. Metalinguistics. *Clinical Linguistics & Phonetics,* 9(1), pp. 25-28.

BODROVA, E. and LEONG, D.J., 2003. Learning and development of preschool children from the Vygotskian perspective. In: A. KOZULIN, B. GINDIS, V.S. AGEYEV and S.M. MILLER, Eds, *Vygotsky's Educational Theory in Cultural Context.* New York: Cambridge Univ Press, pp. 156-176.

BOSMA, T. and RESING, W., 2010. Teacher's Appraisal of Dynamic Assessment Outcomes: Recommendations for Weak Mathematics-Performers. *Journal of Cognitive Education and Psychology,* 9(2), pp. 91-115.

BOTTING, N., 2005. Non-verbal cognitive development and language impairment. *Journal of Child Psychology and Psychiatry and Allied Disciplines,* 46(3), pp. 317-326.

BOTTING, N. and CONTI-RAMSDEN, G., 2000. Social and behavioural difficulties in children with language impairment. *Child Language Teaching and Therapy,* 16(2), pp. 105-120.

BOYLE, J.M., MCCARTNEY, E., O'HARE, A. and FORBES, J., 2009. Direct versus indirect and individual versus group modes of language therapy for children with primary language impairment: principal outcomes from a randomized controlled trial and economic evaluation. *International Journal of Language & Communication Disorders,* 44(6), pp. 826-846.

BRANSFORD, JD, DELCLOS, VR, VYE, NJ, BURNS, MS, HASSELBRING, TS, 1987. State of the art and future directions. . In: C.S. LIDZ, Ed, *Dynamic Assessment. An Interactional Approach to Evaluating Learning Potential.* New York: Guilford Press, pp. 479-496.

BRODY, N., 2000. History of theories and measurements of intelligence. In: R.J. STERNBERG, Ed, *Handbook of Intelligence.* Cambridge Univ Pr, pp. 16-33.

BROWN, A.L. and FERRARA, R.F., 1985. Diagnosing Zones of Proximal Development. In: J.V. WERTSCH, ed, *Culture, Communication and Cognition: Vygotskian Perspectives.* New York: Cambridge University Press, pp. 273-305.

BRYAN, A., 1997. Colourful Semantics. In: S. CHIAT, J. LAW and J. MARSHALL, Eds, *Language Disorders in children and adults: Psycholinguistic approaches to therapy.* London: Whurr, pp. 143-161.

BUDOFF, M., 1987(a). The validity of learning potential assessment. In: C.S. LIDZ, Ed, *Dynamic Assessment. An Interactional Approach to Evaluating Learning Potential.* New York: Guilford Press, pp. 52-81.

BUDOFF, M., 1987 (b). Measures for assessing learning potential. In: C.S. LIDZ, ed, *Dynamic Assessment. An Interactional Approach to Evaluating Learning Potential.* New York: Guilford Press, pp. 173-195.

BURDEN, R., 1998. How can we best help children become effective thinkers and learners? The case for and against thinking skills programmes. In: R.L. BURDEN and M. WILLIAMS, Eds, *Thinking through the Curriculum.* New York: Routledge, pp. 1-27.

CAESAR, L.G. and KOHLER, P.D., 2009. Tools Clinicians Use: A Survey of Language Assessment Procedures Used by School-Based Speech-Language Pathologists. *Communication Disorders Quarterly,* 30(4), pp. 226-236.

CAMARATA, S. and NELSON, K.E., 2002. Measurement and the Diagnosis and Treatment of Language Disorders in Children. *Peabody Journal of Education,* 77(2), pp. 106-116.

CAMARATA, S., NELSON, K.E., GILLUM, H. and CAMARATA, M., 2009. Incidental receptive language growth associated with expressive grammar intervention in SLI. *First Language,* 29(1), pp. 51-63.

CAMARATA, S.M. and NELSON, K.E., 2006. Conversational recast intervention with preschool and older children. In: R.J. MCCAULEY and M.E. FEY, Eds, *Treatment of Language Disorders in Children.* Paul H Brookes, pp. 237-264.

CAMILLERI, B. and BOTTING, N., In press. Beyond static assessment of children's receptive vocabulary: A dynamic assessment of word learning ability.

CAMILLERI, B. and LAW, J., 2007. Assessing children referred to speech and language therapy: Static and dynamic assessment of receptive vocabulary. *International Journal of Speech-Language Pathology,* 9(4), pp. 312-322.

CAMPBELL, T., DOLLAGHAN, C., NEEDLEMAN, H. and JANOSKY, J., 1997. Reducing bias in language assessment: Processing-dependent measures. *Journal of Speech, Language, and Hearing Research,* 40(3), pp. 519-525.

CAMPIONE, J.C., 1989. Assisted assessment: A taxonomy of approaches and an outline of strengths and weaknesses. *Journal of learning disabilities,* 22(3), pp. 151-165.

CAMPIONE, J.C. and BROWN, A.L., 1987. Linking dynamic assessment with school achievement. In: C.S. LIDZ, Ed, *Dynamic Assessment. An Interactional Approach to Evaluating Learning Potential.* New York: Guilford Press, pp. 82-115.

CAMPIONE, J.C., BROWN, A.L. and FERRARA, R.A., 1982. Mental retardation and intelligence. In: R.J.STERNBERG, Ed, Handbook of Human Intelligence. New York: Cambridge Univ Press, pp. 392-490.

CARLSON, J.S. and WIEDL, K.H., 1992. Principles of dynamic assessment: The application of a specific model. *Learning and Individual Differences,* 4, pp. 153–166.

CARLSON, J.S. and WIEDL, K.H., 1978. Use of testing-the-limits procedures in the assessment of intellectual capabilities in children with learning difficulties. *American Journal of Mental Deficiency,* 82(6), pp. 559-564.

CASALIS, S. and COLÉ, P., 2009. On the relationship between morphological and phonological awareness: Effects of training in kindergarten and in first-grade reading. *First Language,* 29(1), pp. 113-142.

CHAIKLIN, S., 2003. The zone of proximal development in Vygotsky's analysis of learning and instruction. In: A. KOZULIN, B. GINDIS, V.S.

AGEYEV and S.M. MILLER, Eds, New York: Cambridge University Press, pp. 39–64.

CIRRIN, F.M. and GILLAM, R.B., 2008. Language intervention practices for school-age children with spoken language disorders: A systematic review. *Language, Speech, and Hearing Services in Schools,* 39(1), pp. S110-S137.

CIRRIN, F.M., SCHOOLING, T.L., NELSON, N.W., DIEHL, S.F., FLYNN, P.F., STASKOWSKI, M., TORREY, T.Z. and ADAMCZYK, D.F., 2010. Evidence-Based Systematic Review: Effects of Different Service Delivery Models on Communication Outcomes for Elementary School-Age Children. *Language, Speech, and Hearing Services in Schools,* 41(3), pp. 233-264.

COHEN, W., HODSON, A., O'HARE, A., BOYLE, J., DURRANI, T., MCCARTNEY, E., MATTEY, M., NAFTALIN, L. and WATSON, J., 2005. Effects of Computer-Based Intervention through Acoustically Modified Speech (Fast ForWord) in Severe Mixed Receptive-Expressive Language Impairment: Outcomes From a Randomized Controlled Trial. *Journal of Speech, Language, and Hearing Research,* 48(3), pp. 715-729.

CONTI-RAMSDEN, G. and BOTTING, N., 1999. Classification of children with specific language impairment: Longitudinal considerations. *Journal of Speech, Language, and Hearing Research,* 42(5), pp. 1195-1204.

CONTI-RAMSDEN, G. and BOTTING, N., 2008. Emotional health in adolescents with and without a history of specific language impairment (SLI). *Journal of Child Psychology and Psychiatry,* 49(5), pp. 516-525.

CONTI-RAMSDEN, G., BOTTING, N. and FARAGHER, B., 2001. Psycholinguistic markers for specific language impairment (SLI). *The Journal of Child Psychology and Psychiatry and Allied Disciplines,* 42(06), pp. 741-748.

CONTI-RAMSDEN, G., CRUTCHLEY, A. and BOTTING, N., 1997. The extent to which psychometric tests differentiate subgroups of children with SLI. *Journal of Speech, Language, and Hearing Research,* 40(4), pp. 765-777.

CRYSTAL, D., 1982. *Profiling Linguistic Disability.* London: Edward Arnold.

CRYSTAL, D., FLETCHER, P. and GARMAN, M., 1976. *The grammatical analysis of language disability: A procedure for assessment and remediation.* London*:* Edward Arnold.

DAS, J.P. and CONWAY, R.N.F., 1992. Reflections on remediation and transfer: A Vygotskian perspective. In: H.C.HAYWOOD and D.TZURIEL, Eds, Interactive Assessment. New York: Springer-Verlag, pp. 94-115.

DAS, J.P., NAGLIERI, J.A. and KIRBY, J.R., 1994. *Assessment of cognitive processes: The PASS theory of intelligence.* Allyn & Bacon.

DAWES, P. and BISHOP, D.V.M., 2010. Psychometric profile of children with auditory processing disorder and children with dyslexia. *Archives Dis Child,* 95(6), pp. 432-436.

DEAN, E., HOWELL, J., HILL, A. and WATERS, D., 1990. *Metaphon resource pack.* NFER-Nelson.

DEAN, E.C., HOWELL, J., WATERS, D. and REID, J., 1995. Metaphon: A metalinguistic approach to the treatment of phonological disorder in children. *Clinical Linguistics & Phonetics,* 9(1), pp. 1-19.

DELCLOS, V.R., BURNS, M.S. and VYE, N.J., 1993. A comparison of teachers' responses to dynamic and traditional assessment reports. *Journal of Psychoeducational Assessment,* 11(1), pp. 46-55.

DELCLOS, V.R., VYE, N.J., BURNS, M.S., BRANSFORD, J.D. and HASSELBRING, T.S., 1992. Improving the quality of instruction: Roles for dynamic assessment. In: H.C.HAYWOOD and D.TZURIEL, Eds, Interactive Assessment. New York: Springer-Verlag pp. 317-331.

DEPARTMENT FOR EDUCATION, May 2006-last update. http://education.gov.uk/publications/standard/publicationDetail/Page1/SESCO6-18 [Accessed January 26, 2011].

DEUTSCH, R. and REYNOLDS, Y., 2000. The use of dynamic assessment by educational psychologists in the UK. *Educational Psychology in Practice,* 16(3), pp. 311-331.

DEWART, H. and SUMMERS, S., 1988. *The pragmatics profile of early communication skills.* Windsor: Nfer-Nelson.

DICK, F., RICHARDSON, F. and SACCUMAN, M.C., 2008. Using magnetic resonance imaging to investigate developmental language disorders. In: C.F.NORBURY, J.B. TOMBLIN AND D.V.M. BISHOP Eds, Understanding Developmental Language Disorders. Hove: Psychology Press, pp. 53-66.

DOCKRELL, J.E., 2001. Assessing Language Skills in Preschool Children. *Child Psychology and Psychiatry Review,* 6(2), pp. 74-84.

DOCKRELL, J.E. and LAW, J., 2007. Measuring and understanding patterns of change in intervention studies with children: Implications for evidence-based practice. *Evidence-Based Communication Assessment and Intervention,* 1(2), pp. 86-97.

DODD, B., 2008. Speech-language therapy and evidence based practice. In: V. JOFFE, M. CRUICE and S. CHIAT, Eds, *Language Disorders in Children and adults: New issues in research and practice.* UK: Wiley, pp. 54-67.

DONALDSON, A.L. and OLSWANG, L.B., 2007. Investigating requests for information in children with autism spectrum disorders: Static versus dynamic assessment. *International Journal of Speech and Language Pathology,* 9(4), pp. 297-311.

DUNN, L.M., DUNN, D.M., WHETTON, C. and & BURLEY, J., 1997. *The British Picture Vocabulary Scale II.* 2nd Edition. Berkshire UK: NFER-Nelson.

EBBELS, S., 2008. Improving grammatical skill in children with specific language impairment. In: C.F. NORBURY, J.B. TOMBLIN and D.V.M. BISHOP, Eds, *Understanding Developmental Language Disorders. From Theory to Practice.* UK: Psychology Pr, pp. 149-174.

EBBELS, S., 2007. Teaching grammar to school-aged children with specific language impairment using shape coding. *Child Language Teaching and Therapy,* 23(1), pp. 67-93.

EBBELS, S. and VAN DER LELY, H., 2001. Meta-syntactic therapy using visual coding for children with severe persistent SLI. *International Journal of Language & Communication Disorders,* 36(s1), pp. 345-350.

EBBELS, S.H., VAN DER LELY, H.K.J. and DOCKRELL, J.E., 2007. Intervention for verb argument structure in children with persistent SLI: a randomized control trial. *Journal of Speech, Language, and Hearing Research,* 50(5), pp. 1330-1349.

ELLIOTT, J., 2003. Dynamic assessment in educational settings: Realising potential. *Educational Review,* 55(1), pp. 15-32.

EMBRETSON, S.E., 1987a. Toward development of a psychometric approach. In: C.S. LIDZ, Ed, *Dynamic Assessment. An Interactional Approach to Evaluating Learning Potential.* New York: Guilford Press, pp. 141–170.

EMBRETSON, S.E., 1987b. Improving the measurement of spatial aptitude by dynamic testing. *Intelligence,* 11(4), pp. 333-358.

FABIO, R.A., 2005. Dynamic assessment of intelligence is a better reply to adaptive behavior and cognitive plasticity. *The Journal of general psychology,* 132(1), pp. 41-66.

FEUERSTEIN, R., FEUERSTEIN, R.S., FALIK, L.H. and RAND, Y., 2002. *The dynamic assessment of cognitive modifiability: The learning propensity assessment device: Theory, instruments and techniques.* Jerusalem: ICELP Press

FEUERSTEIN, R. and FEUERSTEIN, S., 1991. Mediated learning experience: A theoretical review. In: R. FEUERSTEIN, P.S. KLEIN and A.J. TANNENBAUM, Eds, *Mediated learning experience (MLE) : theoretical, psychosocial and learning implications.* Freund Publishing House, pp. 3-51.

FEUERSTEIN, R., KLEIN, P.S. and TANNENBAUM, A.J., Eds, 1991. *Mediated learning experience (MLE): Theoretical, psychosocial and learning implications.* Tel Aviv: Freund Publishing House Ltd.

FEUERSTEIN, R., RAND, Y., HAYWOOD, H.C., KYRAM, L. and HOFFMAN, M.B., 2003. Learning Propensity Assessment Device Manual. Jerusalem: ICELP Press.

FEUERSTEIN, R., RAND, Y. and HOFFMAN, M.B., 1981. The dynamic assessment of retarded performers: The learning potential assessment device, theory, instruments and techniques. *International Journal of Rehabilitation Research,* 4(3), pp. 465-466.

FEUERSTEIN, R., RAND, Y. and HOFFMAN, M.B., 1980. *Instrumental Enrichment: An intervention programme for cognitive modifiability.* Baltimore: University Park Press.

FEUERSTEIN, R., RAND, Y. and HOFFMAN, M.B., 1979. *The dynamic assessment of retarded performers : the learning potential assessment device, theory, instruments, and techniques*. Baltimore*:* University Park Press.

FEY, M.E., 1986. Determining the Content and Sequence of Specific goals:Developmental versus Remedial Logic. In: M.E. FEY, Ed, *Language Intervention with Young Children.* NJ: Alleyn and Bacon, pp. 101-108.

FEY, M.E., CLEAVE, P.L., LONG, S.H. and HUGHES, D.L., 1993. Two approaches to the facilitation of grammar in children with language impairment: An experimental evaluation. *Journal of speech and hearing research,* 36(1), pp. 141-157.

FEY, M.E., LONG, S.H. and FINESTACK, L.H., 2003. Ten principles of grammar facilitation for children with specific language impairments. *American Journal of Speech-Language Pathology,* 12(1), pp. 3-15.

FINESTACK, L.H. and FEY, M.E., 2009. Evaluation of a Deductive Procedure to Teach Grammatical Inflections to Children With Language Impairment. *American Journal of Speech-Language Pathology,* 18(3), pp. 289-302.

FONTENEAU, E. and VAN DER LELY, H.K.J., 2008. Electrical Brain Responses in Language-Impaired Children Reveal Grammar-Specific Deficits. PLoS ONE 3(3): e1832. doi:10.1371/journal.pone.0001832

FREEMAN, L. and MILLER, A., 2001. Norm-referenced, Criterion-referenced, and Dynamic Assessment: what exactly is the point? *Educational Psychology in Practice,* 17(1), pp. 3-16.

FRIBERG, J.C., 2010. Considerations for test selection: How do validity and reliability impact diagnostic decisions? *Child Language Teaching and Therapy,* 26(1), pp. 77-92.

FRIEDMAN, P. and FRIEDMAN, K.A., 1980. Accounting for individual differences when comparing the effectiveness of remedial language teaching methods. *Applied Psycholinguistics,* 1(2), pp. 151-170.

FUCHS, D. and FUCHS, L.S., 2006. Introduction to Response to Intervention: What, why, and how valid is it? *Reading Research Quarterly,* 41(1), pp. 93-99.

GALLAGHER, A.L. and CHIAT, S., 2009. Evaluation of speech and language therapy interventions for pre-school children with specific language impairment: a comparison of outcomes following specialist

intensive, nursery-based and no intervention. *International journal of language & communication disorders,* 44(5), pp. 616-638.

GARCIA, M. and CONTE, V., 2004. Improvement of cognitive functions in children with Down syndrome through the Bright Start Program. *Journal of Cognitive Education and Psychology,* 4(1), pp. 65-86.

GARDNER, H., FROUD, K., MCCLELLAND, A. and VAN DER LELY, H.K.J., 2006. Development of the Grammar and Phonology Screening (GAPS) test to assess key markers of specific language and literacy difficulties in young children. *International Journal of Language & Communication Disorders,* 41(5), pp. 513-540.

GATHERCOLE, S.E., (2010) Working memory, language and classroom learning. Conference presentation. Child Language Seminar. City University, London,

GATHERCOLE, S.E. and BADDELEY, A.D., 1993. *Working memory and language.* Hove: Psychology Press.

GATHERCOLE, S.E. and BADDELEY, A.D., 1990. Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and language,* 29(3), pp. 336-360.

GILLAM, R. and MCFADDEN, T.U., 1994. Redefining assessment as a holistic discovery process. *Communication Disorders Quarterly,* 16(1), pp. 36-40.

GILLAM, R.B., LOEB, D.F., HOFFMAN, L.V.M., BOHMAN, T., CHAMPLIN, C.A., THIBODEAU, L., WIDEN, J., BRANDEL, J. and FRIEL-PATTI, S., 2008. The efficacy of Fast ForWord Language intervention in school-age children with language impairment: A randomized controlled trial. *Journal of Speech, Language, and Hearing Research,* 51(1), pp. 97-119.

GILLAM, S.L. and GILLAM, R.B., 2006. Making evidence-based decisions about child language intervention in schools. *Language, speech, and hearing services in schools,* 37(4), pp. 304-315.

GLASPEY, A. and STOEL-GAMMON, C., 2007. A dynamic approach to phonological assessment. *International Journal of Speech-Language Pathology,* 9(4), pp. 286-296.

GOMBERT, J.E., 1992. *Metalinguistic development.* University of Chicago Press.

GOODMAN, R., 1997. The Strengths and Difficulties Questionnaire: A Research Note. *Journal of Child Psychology and Psychiatry,* 38, pp. 581-586.

GOPNIK, M., 1990. Feature blindness: A case study. *Language Acquisition,* 1(2), pp. 139-164.

GOPNIK, M. and CRAGO, M.B., 1991. Familial aggregation of a developmental language disorder. *Cognition,* 39(1), pp. 1-50.

GRAHAM, S. and PERIN, D., 2007. A Meta-Analysis of Writing Instruction for Adolescent Students. *Journal of educational psychology,* 99(3), pp. 445-476.

GREDLER, M.E. and SHIELDS, C.C., 2008. *Vygotsky's legacy: a foundation for research and practice.* New York: The Guilford Press.

GRIGORENKO, E.L. and STERNBERG, R.J., 1998. Dynamic assessment. *Psychological bulletin,* 124, pp. 75-111.

GRUNDY, K., 1995. Metaphon: Unique and effective? *Clinical Linguistics & Phonetics,* 9(1), pp. 20-24.

GUMMERSALL, D.M. and STRONG, C.J., 1999. Assessment of complex sentence production in a narrative context. *Language, Speech, and Hearing Services in Schools,* 30(2), pp. 152-164.

GUTERMAN, E., 2002. Toward dynamic assessment of reading: applying metacognitive awareness guidance to reading assessment tasks. *Journal of Research in Reading,* 25(3), pp. 283-298.

GUTHKE, J., 1993. Developments in learning potential assessment. In: J.H.M. HAMERS, K. SIJTSMA and A.J.J.M. RUIJSSENAARS, Eds, *Learning Potential Assessment. Theoretical, Methodological and Practical Issues.* Amsterdam: Swets and Zeitlinger, pp. 43-68.

GUTHKE, J., BECKMANN, J.F. and DOBAT, H., 1997. Dynamic testing-problems, uses, trends and evidence of validity. *Educational and Child Psychology,* 14(4), pp. 17-32.

GUTHKE, J. and WINGENFELD, S., 1992. The learning test concept: Origins, state of the art, and trends. pp. 64-93.

GUTIERREZ-CLELLEN, V.F. and PEÑA, E., 2001. Dynamic assessment of diverse children: A tutorial. *Language, Speech, and Hearing Services in Schools,* 32(4), pp. 212-224.

HASSON, N. and BOTTING, N., 2010. Dynamic assessment of children with language impairments: A pilot study. *Child Language Teaching and Therapy,* 26(3), pp. 249-272.

HASSON, N. and JOFFE, V., 2007. The case for dynamic assessment in speech and language therapy. *Child Language Teaching and Therapy,* 23(1), pp. 9-25.

HAYWOOD, H.C., 1993. A Mediational Teaching Style. *International Journal of Cognitive Education and Mediated Learning,* 3(1), pp. 27-38.

HAYWOOD, H.C., BROOKS, P. and BURNS, S., 1992. *Bright start: Cognitive curriculum for young children.* Charlesbridge Pub.

HAYWOOD, H.C. and LIDZ, C.S., 2007. *Dynamic Assessment in Practice. Clinical and Educational Applications.* New York: Cambridge University Press.

HENRY, L., 2010. Executive *Functioning in children with Specific Language Impairment.* Seminar Presentation. City University, London.

HESSELS, M.G.P., 2009. Estimation of the Predictive Validity of the HART by Means of a Dynamic Test of Geography. *Journal of Cognitive Education and Psychology,* 8(1), pp. 5-21.

HESSELS, M.G.P., BERGER, J.L. and BOSSON, M., 2008. Group Assessment of Learning Potential of Pupils in Mainstream Primary Education and Special Education Classes. *Journal of Cognitive Education and Psychology,* 7(1), pp. 43-69.

HICKMANN, M.E., 1985. The implications of discourse skills in Vygotsky's developmental theory. In: J.V. WERTSCH, ed, *Culture, communication and cognition: Vygotskian perspectives.* New York: Cambridge University Press, pp. 236-257.

HIGGINS, S., HALL, E., BAUMFIELD, V. and MOSELEY, D., 2005. A meta-analysis of the impact of the implementation of thinking skills approaches on pupils. Project Report. EPPI-Centre, Social Science Research Unit, Institute of Education, University of London, London.
http://eppi.ioe.ac.uk/cms/Default.aspx?tabid=338

HUDDLESTON, R.D., 1984. *Introduction to the Grammar of English.* Cambridge UK: Cambridge University Press.

HULME, C. and SNOWLING, M., 2009. *Developmental disorders of language learning and cognition.* Chichester: Wiley-Blackwell.

JELTOVA, I., BIRNEY, D., FREDINE, N., JARVIN, L., STERNBERG, R.J. and GRIGORENKO, E.L., 2007. Dynamic assessment as a process-oriented assessment in educational settings. *International Journal of Speech and Language Pathology,* 9(4), pp. 273-285.

JITENDRA, A.K. and KAMEENUI, E.J., 1993. Dynamic assessment as a compensatory assessment approach: A description and analysis. *Remedial and Special Education,* 14(5), pp. 6-18.

JOFFE, V., 2008. Some (but still insufficient) evidence exists for language interventions for school-age children with developmental speech and language impairments. *Evidence-Based Communication Assessment and Intervention,* 2(2), pp. 76-80.

JOFFE, V., CRUICE, M. and CHIAT, S., Eds, 2008. *Language disorders in children and adults : new issues in research and practice.* Chichester: Wiley-Blackwell.

JUSTICE, L.M. and EZELL, H.K., 2004. Print referencing: An emergent literacy enhancement strategy and its clinical applications. *Language, Speech, and Hearing Services in Schools,* 35(2), pp. 185-193.

KAHN, R.J. and KING, S.R., 1997. Dynamic Procedures for Assessing Children's Cognitive and Emotional Strengths and Needs. *Journal of Cognitive Education,* 6(2), pp. 101-114.

KAMHI, A.G., 2006. Treatment decisions for children with speech-sound disorders. *Language, Speech, and Hearing Services in Schools,* 37(4), pp. 271-279.

KANIEL, S., 2009. Domain Specific vs Domain General: Implications for Dynamic Assessment. Conference Presentation IACEP XII International Conference, University of Osnabruck, Germany.

KANIEL, S., 2001. Not all testing is Dynamic Testing. Overcoming the Fuzzy Identity of Dynamic Assessment. *Issues in Education,* 7(2), pp. 211-223.

KARMILOFF-SMITH, A., 2009. Nativism versus Neuroconstructivism: Rethinking the study of Developmental Disorders. *Developmental Psychology,* 45(1), pp. 56-63.

KARPOV, Y., 2008. Do All Dynamic Assessment Techniques Assess Learning Potential? *Journal of Cognitive Education and Psychology,* 7(3), pp. 411-418.

KARPOV, Y., 2003. Vygotsky's Doctrine of Scientific Concepts: Its role for Contemporary Education. In: A. KOZULIN, B. GINDIS, V.S. AGEYEV and S.M. MILLER, Eds, *Vygotsky's Educational Theory in Cultural Context.* New York: Cambridge University Press, pp. 65-82.

KEANE, K.J., 1987. Assessing Deaf Children. In: C.S. LIDZ, Ed, *Dynamic Assessment. An Interactional Approach to Evaluating Learning Potential.* New York: Guilford Press, pp. 360-377.

KESTER, E.B., PEÑA, E.D. and GILLAM, R.B., 2002. Outcomes of Dynamic Assessment with Culturally and Linguistically Diverse Students. *Journal of Cognitive Education and Psychology,* 2(1), pp. 37-54.

KLIMACKA, L., 1995. Managing disordered phonological development with the Metaphon approach. *Clinical Linguistics & Phonetics,* 9(1), pp. 36-42.

KOHNERT, K. and DANAHY, K., 2007. Young L2 learners' performance on a novel morpheme task. *Clinical Linguistics and Phonetics,* 21(7), pp. 557-569.

KOZULIN, A., 2003. Psychological tools and mediated learning. In: A. KOZULIN, B. GINDIS, V.S. AGEYEV and S.M. MILLER, eds, *Vygotsky's Educational Theory in Cultural Context.* New York: Cambridge Univ Pr, pp. 15-38.

KOZULIN, A., 1990. *Vygotsky's psychology: A biography of ideas.* UK: Harvester Wheatsheaf.

KOZULIN, A., GINDIS, B., AGEYEV V.S. and MILLER S.M., Eds, 2003. *Vygotsky's educational theory in cultural context.* New York: Cambridge Univ Press.

KOZULIN, A., LEBEER, J., MADELLA-NOJA, A., GONZALEZ, F., JEFFREY, I., ROSENTHAL, N. and KOSLOWSKY, M., 2010. Cognitive modifiability

of children with developmental disabilities: A multicentre study using Feuerstein's Instrumental Enrichment-Basic program. *Research in developmental disabilities,* 31(2), pp. 551-559.

LAHEY, M., 1990. Who shall be called language disordered? Some reflections and one perspective. *Journal of Speech and Hearing Disorders,* 55(4), pp. 612-620.

LAING, S.P. and KAMHI, A., 2003. Alternative assessment of language and literacy in culturally and linguistically diverse populations. *Language Speech and Hearing Services in Schools,* 34(1), pp. 44-55.

LARSEN, J.A. and NIPPOLD, M.A., 2007. Morphological analysis in school-age children: Dynamic assessment of a word learning strategy. *Language, Speech, and Hearing Services in Schools,* 38(3), pp. 201-212.

LAUCHLAN, F. and ELLIOTT, J., 1997. Using dynamic assessment materials as a tool for providing cognitive intervention to children with complex learning difficulties. *Educational and Child Psychology,* 14(4), pp. 137-148.

LAW, J., BOYLE, J., HARRIS, F., HARKNESS, A. and NYE, C., 1998. Screening for speech and language delay: a systematic review of the literature. *Health Technology Assessment,* 2(9).

LAW, J. and CAMILLERI, B., 2007. Dynamic assessment and its application to children with speech and language learning difficulties. *International Journal of Speech-Language Pathology,* 9(4), pp. 271-272.

LAW, J., CAMPBELL, C., ROULSTONE, S., ADAMS, C. and BOYLE, J., 2008. Mapping practice onto theory: the speech and language practitioner's construction of receptive language impairment. *International Journal of Language & Communication Disorders,* 43(3), pp. 245-263.

LAW, J., GARRETT, Z. and NYE, C., 2004. The efficacy of treatment for children with developmental speech and language delay/disorder: A meta-analysis. *Journal of Speech, Language, and Hearing Research,* 47(4), pp. 924-943.

LAW, J., 2008. Making new connections - where to next? In: V. JOFFE, M. CRUICE and S. CHIAT, Eds, *Language Disorders in Children and Adults.* UK: Wiley-Blackwell, pp. 159-175.

LEONARD, L.B., 2009. Is Expressive Language Disorder an Accurate Diagnostic Category? *American Journal of Speech-Language Pathology,* 18, pp. 115-123.

LEONARD, L.B., 1998. *Children with Specific Language Impairment.* Cambridge MA: MIT press.

LEONARD, L.B., 1995. Functional categories in the grammars of children with specific language impairment. *Journal of speech and hearing research,* 38(6), pp. 1270-1283.

LEONARD, L.B., CAMARATA, S.M., PAWLOWSKA, M., BROWN, B. and CAMARATA, M.N., 2008. The acquisition of tense and agreement morphemes by children with specific language impairment during intervention: Phase 3. *Journal of Speech, Language, and Hearing Research,* 51(1), pp. 120-125.

LEVY, H. and FRIEDMANN, N., 2009. Treatment of syntactic movement in syntactic SLI: A case study. *First Language,* 29(1), pp. 15-50.

LIDZ, C.S., 2003. *Early childhood assessment.* New Jersey:Wiley.

LIDZ, C.S., 2000. The Application of Cognitive Functions Scale (ACFS): An example of curriculum-based dynamic assessment. In: C.S. LIDZ and J. ELLIOTT, Eds, *Dynamic Assessment: Prevailing models and applications.* Amsterdam: Elsevier, pp. 407-439.

LIDZ, C.S., 1991. *Practitioner's guide to dynamic assessment.* New York: The Guilford Press.

LIDZ, C.S., 1987. *Dynamic assessment: An interactional approach to evaluating learning potential.* New York: Guilford Press

LIDZ, C.S. and ELLIOTT, J., 2000. *Dynamic assessment: Prevailing models and applications.* Amsterdam: Elsevier Science.

LIDZ, C.S. and GINDIS, B., 2003. Dynamic assessment of the evolving cognitive functions in children. In: A. KOZULIN, B. GINDIS, V.S. AGEYEV and S.M. MILLER, eds, *Vygotsky's Educational Theory in Cultural Context.* New York: Cambridge Univ Pr, pp. 99–115.

LIDZ, C.S., JEPSEN, R.H. and MILLER, M.B., 1997. Relationships between cognitive processes and academic achievement: Application of a group dynamic assessment procedure with multiply handicapped adolescents. *Educational and Child Psychology,* 14(4), pp. 56-67.

LIDZ, C.S. and PEÑA, E.D., 2009. Response to intervention and dynamic assessment: do we just appear to be speaking the same language? *Seminars in speech and language,* 30(2), pp. 121-133.

LLOYD, J.W. and BLANDFORD, B.J., 1991. Assessment for instruction. In: H.L. SWANSON, Ed, *Handbook on the Assessment of Learning Disabilities.* Texas: Pro-Ed, pp. 45-58.

LOEB, D.F. and LEONARD, L.B., 1991. Subject case marking and verb morphology in normally developing and specifically language-impaired children. *Journal of speech and hearing research,* 34(2), pp. 340-346.

MAGNUSSON, E. and NAUCLÉR, K., 1993. The development of linguistic awareness in language-disordered children. *First language,* 13(37), pp. 93-111.

MARINELLIE, S.A. and JOHNSON, C.J., 2002. Definitional skill in school-age children with specific language impairment. *Journal of communication disorders,* 35(3), pp. 241-259.

MARTON, K., 2008. Visuo-spatial processing and executive functions in children with specific language impairment. *International Journal of Language & Communication Disorders,* 43(2), pp. 181-200.

MARTON, K. and SCHWARTZ, R.G., 2003. Working memory capacity and language processes in children with specific language impairment. *Journal of Speech, Language, and Hearing Research,* 46(5), pp. 1138-1154.

MASTERSON, J.J., EVANS, L.H. and ALOIA, M., 1993. Verbal analogical reasoning in children with language-learning disabilities. *Journal of speech and hearing research,* 36(1), pp. 76-82.

MASTERSON, J.J. and PERREY, C.D., 1999. Training analogical reasoning skills in children with language disorders. *American Journal of Speech-Language Pathology,* 8(1), pp. 53-61.

MATHENY, N. and PANAGOS, J.M., 1978. Comparing the effects of articulation and syntax programs on syntax and articulation improvement. *Language, Speech, and Hearing Services in Schools,* 9(1), pp. 57-61.

MCCARTNEY, E., 2004. *A Survey and Cohort Intervention Using Indirect Speech and Language Therapy for Children with Primary Language Impairment in Schools.* West of Scotland Primary Care NHS Trust.

MCCARTNEY, E., BOYLE, J., ELLIS, S., BANNATYNE, S. and TURNBULL, M., 2011. Indirect language therapy for children with persistent language impairment in mainstream primary schools: outcomes from a cohort intervention. *International Journal of Language & Communication Disorders,* 46(1), pp. 74-82.

MCCAULEY, R.J. and FEY, M.E., 2006. *Treatment of language disorders in children.* Baltimore: Paul H. Brookes Pub.

MCCAULEY, R.J. and SWISHER, L., 1984. Psychometric review of language and articulation tests for preschool children. *Journal of Speech and Hearing Disorders,* 49(1), pp. 34-42.

MCFADDEN, T.U., 1996. Creating language impairments in typically achieving children: The pitfalls of "normal" normative sampling. *Language, Speech, and Hearing Services in Schools,* 27(1), pp. 3-9.

MELTZER, L. and REID, D.K., 1994. New directions in the assessment of students with special needs: The shift toward a constructivist perspective. *The journal of special education,* 28(3), pp. 338-355.

MERRITT, D.D. and CULATTA, B., 1998. Dynamic assessment, language processes, and curricular content. In: D.D. MERRITT and B. CULATTA, Eds, *Language Intervention in the Classroom.* Singular Pub Group, pp. 99–142.

MICCIO, A.W., 1995. Metaphon: factors contributing to treatment outcomes. *Clinical Linguistics & Phonetics,* 9(1), pp. 28-36.

MIKUCKI, B.A. and LARRIVEE, L., 2006. Validity and reliability of twelve child language tests, *Unpublished poster session presented at the American Speech–Language–Hearing Association's national convention, Miami, FL,* 2006, http://www.eshow2000.com/asha/2006/handouts.cfm.

MILLER, L., GILLAM, R.B. and PEÑA, E.D., 2001. *Dynamic assessment and intervention: Improving children's narrative abilities.* Austin Texas: Pro-Ed.

MISLEVY, R.J. and YIN, C., 2009. If Language Is a Complex Adaptive System, What Is Language Assessment? *Language Learning,* 59(1), pp. 249-267.

MISSIUNA, C. and SAMUELS, M., 1988. Dynamic assessment: Review and critique. *Special Services in the Schools,* 5(1-2), pp. 1-22.

MONTGOMERY, J.W., 2002. Information processing and language comprehension in children with specific language impairment. *Topics in Language Disorders,* 22(3), pp. 62-84.

MUMA, J.R., 1984. Semel and Wiig's CELF: Construct Validity? *Journal of Speech and Hearing Disorders,* 49, pp. 101-103.

NELSON, K.E., CAMARATA, S.M., WELSH, J., BUTKOVSKY, L. and CAMARATA, M., 1996. Effects of imitative and conversational recasting treatment on the acquisition of grammar in children with specific language impairment and younger language-normal children. *Journal of speech and hearing research,* 39(4), pp. 850-859.

NETTELBLADT, U., 1995. The Metaphon approach to phonological therapy from a Swedish perspective. *Clinical Linguistics & Phonetics,* 9(1), pp. 42-49.

NETTELBLADT, U., SAHLÉN, B., ORS, M. and JOHANNESSON, P., 1989. A multidisciplinary assessment of children with severe language disorder. *Clinical Linguistics & Phonetics,* 3(4), pp. 313-346.

NEWBURY, D.F. and MONACO, A.P., 2008. The application of molecular genetics to the study of language impairments. In: C.F.NORBURY, J.B. TOMBLIN AND D.V.M. BISHOP Eds, Understanding Developmental Language Disorders. Hove: Psychology Press, pp. 79-91.

NIPPOLD, M.A., 1998. *Later language development: The school-age and adolescent years.* Austin: Pro-Ed.

NORBURY, C.F., TOMBLIN, J.B. and BISHOP, D.V.M., 2008. *Understanding developmental language disorders: from theory to practice.* Hove: Psychology Press.

OLSWANG, L., BAIN, B. and JOHNSON, G., 1992. Using dynamic assessment with children with language disorders. In: S. WARREN and J. REICHLE, Eds, *Causes and effects in communication and language intervention.* Baltimore: Paul H Brookes, pp. 187-215.

OLSWANG, L.B. and BAIN, B.A., 1996. Assessment information for predicting upcoming change in language production. *Journal of speech and hearing research,* 39(2), pp. 414-423.

OWENS, R.E., 2001. *Language development: An introduction.* 5th Edn. Boston: Allyn and Bacon.

PALINCSAR, A.S., BROWN, A.L. and CAMPIONE, J.C., 1991. Dynamic assessment. In: H.L. Swanson, Ed., *Handbook on the assessment of learning disabilities: theory, research, and practice.* Austin: Pro-Ed, pp. 75-94.

PEÑA, E. and IGLESIAS, A., 1992. The Application of Dynamic Methods to Language Assessment: A non-biased procedure. *The Journal of Special Education,* 26(3), pp. 269-280.

PEÑA, E., IGLESIAS, A. and LIDZ, C.S., 2001. Reducing test bias through dynamic assessment of children's word learning ability. *American Journal of Speech-Language Pathology,* 10(2), pp. 138-154.

PEÑA, E.D., 2001. Assessment of semantic knowledge: Use of feedback and clinical interviewing. *Seminars in speech and language,* 22(1), pp. 51-63.

PEÑA, E.D., 2000. Measurement of modifiability in children from culturally and linguistically diverse backgrounds. *Communication Disorders Quarterly,* 21(2), pp. 87-97.

PEÑA, E.D., GILLAM, R. and MILLER, L., 2003. *Dynamic Assessment of Narratives. Revised Scripts for Assessment Mediation.* Austin: Pro-Ed.

PEÑA, E.D., GILLAM, R.B., MALEK, M., RUIZ-FELTER, R., RESENDIZ, M., FIESTAS, C. and SABEL, T., 2006. Dynamic assessment of school-age children's narrative ability: An experimental investigation of classification accuracy. *Journal of Speech, Language, and Hearing Research,* 49(5), pp. 1037-1056.

PEÑA, E.D., RESÉNDIZ, M. and GILLAM, R.B., 2007. The role of clinical judgements of modifiability in the diagnosis of language impairment. *International Journal of Speech-Language Pathology,* 9(4), pp. 332-345.

PEÑA, E.D. and GILLAM, R.B., 2000. Dynamic Assessment of Children referred for Speech and Language Evaluations. In: C.S. LIDZ and J. ELLIOTT, Eds, *Dynamic Assessment: Prevailing Models and Applications.* Amsterdam:Elsevier science, pp. 543-575.

PENNINGTON, B.F. and BISHOP, D.V.M., 2009. Relations among speech, language, and reading disorders. *Annual Review of Psychology,* 60, pp. 283-306.

PINKER, S., 2000. *Words and rules: The ingredients of language.* Harper Perennial.

POEHNER, M.E., 2008. *Dynamic assessment : a Vygotskian approach to understanding and promoting L2 development.* New York: Springer.

PRING, T., 2005. *Research methods in communication disorders.* London: Wiley.

RAPIN, I. and ALLEN, D.A., 1987. Developmental dysphasia and autism in preschool children: Characteristics and subtypes, *Proceedings of the first international symposium on specific speech and language disorders in children,* 1987, pp. 20–35.

RAVEN, J.C., COURT, J.H. and RAVEN, J., 1990. *Manual for Raven's Progressive Matrices and Vocabulary Scales - Section 2: Coloured Progressive Matrices.* Oxford: Oxford Psychologists Press.

RAVID, D. and GEIGER, V., 2009. Promoting morphological awareness in Hebrew-speaking grade-schoolers: An intervention study using linguistic humor. *First Language,* 29(1), pp. 81-112.

RAVID, D. and HORA, A., 2009. From implicit to explicit language knowledge in intervention: Introduction to the Special Issue on intervention and metalanguage. *First Language,* 29(1), pp. 5-14.

RENFREW, C., 1988. *The Renfrew Action Picture Test.* UK: Winslow Press.

RESING, W.C.M., 2001. Beyond Binet. All Testing should be Dynamic Testing. *Issues in Education,* 7(2), pp. 225-235.

RESING, W.C.M., 1997. Learning potential assessment: the alternative for measuring intelligence? *Educational and Child Psychology,* 14, pp. 68-82.

RESING, W.C.M., 1993. Measuring inductive reasoning skills: The construction of a learning potential test. In: J.H.M. HAMERS, K. SIJTSMA and A.J.J.M. RUIJSSENAARS, Eds, *Learning Potential Assessment.* Amsterdam: Swets & Zeitlinger, pp. 219-242.

RESING, W.C.M., DE JONG, F.M., BOSMA, T. and TUNTELER, E., 2009. Learning During Dynamic Testing: Variability in Strategy Use by Indigenous and Ethnic Minority Children. *Journal of Cognitive Education and Psychology,* 8(1), pp. 22-37.

RICE, M.L. AND WEXLER K., 2001. *Test of Early Grammatical Impairment. Examiner's Manual.* USA: The Psychological Corporation.

RICE, M.L. and WEXLER, K., 1996. Toward tense as a clinical marker of specific language impairment in English-speaking children. *Journal of speech and hearing research,* 39(6), pp. 1239-1257.

RICE, M.L., WEXLER, K. and CLEAVE, P.L., 1995. Specific language impairment as a period of extended optional infinitive. *Journal of speech and hearing research,* 38(4), pp. 850-863.

RICE, M.L., 2007. Children with Specific Language Impairment: Bridging the Genetic and Developmental Perspectives. In: E. HOFF and M. SHATZ, Eds, *Blackwell handbook of language development.* Malden Mass Oxford: Blackwell, pp. 411-431.

RINALDI, W., 1995. *The social use of language programme (Primary and Pre-School Teaching Pack).* Windsor: NFER-Nelson.

RUTTER, M., 2008. Diagnostic concepts and risk factors. In: C.F.NORBURY, J.B. TOMBLIN AND D.V.M. BISHOP Eds, Understanding Developmental Language Disorders. Hove: Psychology Press, pp. 205-215.

SCHERY, T.K., 1985. Correlates of language development in language-disordered children. *Journal of Speech and Hearing Disorders,* 50(1), pp. 73-83.

SEMEL, E., WIIG, E.H. and SECORD, W.A., 1987. *Clinical Evaluation of Language Fundamentals.* 3rd Edition(UK). UK: The Psychological Corporation.

SMITH-LOCK, K.M., 1995. Morphological usage and awareness in children with and without specific language impairment. *Annals of Dyslexia,* 45(1), pp. 161-185.

SPAULDING, T.J., PLANTE, E. and FARINELLA, K.A., 2006. Eligibility Criteria for Language Impairment: Is the Low End of Normal Always Appropriate? *Language, Speech, and Hearing Services in Schools,* 37(1), pp. 61-72.

SPECTOR, J.E., 1992. Predicting progress in beginning reading: Dynamic assessment of phonemic awareness. *Journal of educational psychology,* 84(3), pp. 353-363.

SPEKMAN, N.J. and ROTH, F.P., 1984. Clinical Evaluation of Language Functions (CELF) Diagnostic Battery: An Analysis and Critique. *Journal of Speech and Hearing Disorders,* 49, pp. 97-100.

SPOONER, L., 2002. Addressing expressive language disorder in children who also have severe receptive language disorder: A psycholinguistic approach. *Child Language Teaching and Therapy,* 18, pp. 289-313.

STACKHOUSE, J. and WELLS, B., 1997. *Children's speech and literacy difficulties: A psycholinguistic framework.* London:Whurr Publishers.

STERNBERG, R.J., 2000. *Handbook of intelligence.* New York:Cambridge Univ Press.

STERNBERG, R.J. and GRIGORENKO, E.L., 2001. All testing is dynamic testing. *Issues in Education,* 7(2), pp. 137-170.

SWANSON, H.L., 1996. *Swanson-Cognitive Processing test.* Austin, Texas: Pro-Ed.

SWANSON, H.L., 1995. Using the Cognitive Processing Test to assess ability: Development of a dynamic assessment measure. *School Psychology Review,* 24(4), pp. 672-693.

SWANSON, H.L. and HOWARD, C.B., 2005. Children with Reading Disabilities: Does Dynamic Assessment Help in the Classification? *Learning Disability Quarterly,* 28(1), pp. 17-35.

SWANSON, H.L. and LUSSIER, C.M., 2001. A selective synthesis of the experimental literature on dynamic assessment. *Review of Educational Research,* 71(2), pp. 321-349.

SWISHER, L., RESTREPO, M.A., PLANTE, E. and LOWELL, S., 1995. Effect of implicit and explicit "rule" presentation on bound-morpheme generalization in specific language impairment. *Journal of speech and hearing research,* 38(1), pp. 168-173.

TALLAL, P., 2000. Experimental studies of language learning impairments: From research to remediation. In: D.V.M. BISHOP, and L.B. LEONARD, Eds.,*Speech and language impairments in children: Causes, characteristics, intervention and outcome,* Hove: Psychology Press, pp. 131–155.

TALLAL, P., 1976. Rapid auditory processing in normal and disordered language development. *Journal of speech and hearing research,* 19(3), pp. 561-571.

TALLAL, P., MERZENICH, M.M., MILLER, S. and JENKINS, W., 1998. Language learning impairments: integrating basic science, technology, and remediation. *Experimental Brain Research,* 123(1), pp. 210-219.

TALLAL, P., MILLER, S.L., BEDI, G., BYMA, G., WANG, X., NAGARAJAN, S.S., SCHREINER, C., JENKINS, W.M. and MERZENICH, M.M., 1996. Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science,* 271, pp. 81-84.

TALLAL, P. and PIERCY, M., 1973. Defects of non-verbal auditory perception in children with developmental aphasia. *Nature,* 241, pp. 468-469.

TIEKSTRA, M., HESSELS, M.G. and MINNAERT, A.E., 2009. Learning capacity in adolescents with mild intellectual disabilities. *Psychological reports,* 105(3 Pt 1), pp. 804-814.

TOMBLIN, J.B., 2009. Co-morbidity of Autism and SLI: Kinds, Kin or Confounds? *IJLCD Winter Lecture*, City University, London.

TOMBLIN, J.B., RECORDS, N.L., BUCKWALTER, P., ZHANG, X., SMITH, E. and O'BRIEN, M., 1997. Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language and Hearing Research,* 40(6), pp. 1245-1260.

TOMBLIN, J.B., RECORDS, N.L. and ZHANG, X., 1996. A system for the diagnosis of specific language impairment in kindergarten children. *Journal of Speech, Language and Hearing Research,* 39(6), pp. 1284-1294.

TOMBLIN, J.B. and ZHANG, X., 2006. The dimensionality of language ability in school-age children. *Journal of Speech, Language, and Hearing Research,* 49(6), pp. 1193-1208.

TYLER, A.A., LEWIS, K.E., HASKILL, A. and TOLBERT, L.C., 2002. Efficacy and cross-domain effects of a morphosyntax and a phonology intervention. *Language, Speech, and Hearing Services in Schools,* 33(1), pp. 52-66.

TZURIEL D. AND SAMUELS M.T, 2000. Dynamic Assessment of Learning Potential: Inter-rater reliability of deficient cognitive functions, types of mediation and non-intellective factors. *Journal of Cognitive Education and Psychology,* 1(1), pp. 40-64.

TZURIEL, D., 2000. Dynamic assessment of young children: Educational and intervention perspectives. *Educational Psychology Review,* 12(4), pp. 385-435.

TZURIEL, D., 1997. A novel dynamic assessment approach for young children: Major dimensions and current research. *Educational and Child Psychology,* 14, pp. 83-108.

TZURIEL, D., 1991. Cognitive modifiability, Mediated Learning experience and Affective-motivational processes: A Transactional Approach. In: R. FEUERSTEIN, P.S. KLEIN and A.J. TANNENBAUM, Eds, *Mediated learning experience (MLE) : theoretical, psychosocial and learning implications.* Tel Aviv: Freund Publishing House, pp. 95-120.

TZURIEL, D., KANIEL, S., KANNER, E. and HAYWOOD, H.C., 1999. Effects of the "Bright Start" program in kindergarten on transfer and academic achievement. *Early Childhood Research Quarterly,* 14(1), pp. 111-141.

VAN DER LELY, H.K.J., (2010) 20 years on: What have studies of children with Grammatical-SLI and typically developing children taught us about language and the brain?'. *Seminar Presentation*, City University, London.

VAN DER LELY, H.K.J., 2005. Domain-specific cognitive systems: insight from Grammatical-SLI. *Trends in cognitive sciences,* 9(2), pp. 53-59.

VAN DER LELY, H.K.J. and BATTELL, J., 2003. Wh-movement in children with grammatical SLI: A test of the RDDR hypothesis. *Language*, 79(1), pp. 153-181.

VAN DER LELY, H.K.J. and HARRIS, M., 1990. Comprehension of reversible sentences in specifically language-impaired children. *Journal of Speech and Hearing Disorders,* 55(1), pp. 101-117.

VAN DER LELY, H.K.J., JONES, M. and MARSHALL, C.R., 2011. Who did Buzz see someone? Grammaticality judgement of wh-questions in typically developing children and children with Grammatical-SLI. *Lingua,* 121(3), pp. 408-422.

VERNES, S.C., NEWBURY, D.F., ABRAHAMS, B.S., WINCHESTER, L., NICOD, J., GROSZER, M., ALARCÓN, M., OLIVER, P.L., DAVIES, K.E. and GESCHWIND, D.H., 2008. A functional genetic link between distinct developmental language disorders. *New England Journal of Medicine,* 359(22), pp. 2337-2345.

VYE, N.J., BURNS, M.S., DELCLOS, V.R. and BRANSFORD, J.D., 1987. A comprehensive approach to assessing intellectually handicapped children. In: C.S. LIDZ, Ed, *Dynamic Assessment. An Interactional Approach to Evaluating Learning Potential.* New York: Guilford Press, pp. 327-359.

VYGOTSKY, L.S., 1986. *Thought and language.* Cambridge MA:MIT Press.

WADMAN, R., BOTTING, N., DURKIN, K. and CONTI-RAMSDEN, G., In press. Changes in Emotional Health Symptoms in Adolescents with Specific Language Impairment. *International Journal of Language and Communication Disorders*.

WATKINS, R.V. and RICE, M.L., 1991. Verb particle and preposition acquisition in language-impaired preschoolers. *Journal of speech and hearing research,* 34(5), pp. 1130-1141.

WEAVER, P.A., 1979. Improving reading comprehension: Effects of sentence organization instruction. *Reading Research Quarterly,* 15(1), pp. 129-146.

WEISMER, S.E., 1991. Hypothesis-testing abilities of language-impaired children. *Journal of speech and hearing research,* 34(6), pp. 1329-1338.

WEISMER, S.E. and MURRAY-BRANCH, J., 1989. Modeling versus modeling plus evoked production training: A comparison of two language intervention methods. *Journal of Speech and Hearing Disorders,* 54(2), pp. 269-281.

WILLIAMS, M. and BURDEN, R.L., 1997. *Psychology for Language Teachers: a Social Constructivist Approach.* Cambridge: Cambridge University Press.

World Health Organization, International Classification of Diseases, Last updated 2007. http://www.who.int/classifications/icd/en/

http://www.dynamicassessment.com/ [Accessed: December 8, 2009].

YODER, P.J., KAISER, A.P. and ALPERT, C.L., 1991. An exploratory study of the interaction between language teaching methods and child characteristics. *Journal of speech and hearing research,* 34(1), pp. 155-167.

YUILL, N., 2009. The relation between ambiguity understanding and metalinguistic discussion of joking riddles in good and poor comprehenders: Potential for intervention and possible processes of change. *First Language,* 29(1), pp. 65-79.

APPENDIX

APPENDIX 1

Feuerstein's Deficient Cognitive Functions
(LPAD Manual Feuerstein et al 1995)

Deficient cognitive functions at the input phase:

1. Blurred and sweeping perception.
2. Unplanned, impulsive, and unsystematic exploratory behavior
3. Lack of, or impaired receptive verbal tools that affect discrimination (e.g. objects, events and relationships are not appropriately labeled).
4. Lack of, or impaired spatial orientation and lack of stable systems of reference by which to establish organization of space.
5. Lack of, or impaired temporal concepts.
6. Lack of, or impaired conservation of constancies (e.g. size, shape, quantity, color, orientation) across variations in one or more dimensions.
7. Lack of, or deficient need for precision and accuracy in data gathering.
8. Lack of capacity for considering two or more sources of information at once. This is reflected in dealing with data in a piecemeal fashion rather than as a unit of facts that are organized.

Deficient cognitive functions at the elaboration phase:

1. Inadequacy in the perception of the existence of a problem and its definition.
2. Inability to select relevant as opposed to irrelevant cues in defining a problem.
3. Lack of spontaneous comparative behavior or the limitation of its application by a restricted need system.
4. Narrowness of the mental field.
5. Episodic grasp of reality.
6. Lack of need for the education or establishment of relationships
7. Lack of need for and/or exercise of summative behavior.
8. Lack of, or impaired need for pursuing logical evidence.
9. Lack of, or impaired inferential hypothetical ("if") thinking.
10. Lack of, or impaired strategies for hypothesis testing.
11. Lack of, or impaired planning behavior.
12. Lack of, or impaired interiorization.
13. Non-elaboration of certain cognitive categories because the verbal concepts are not a part of the individual verbal inventory or a receptive level, or because they are not mobilized at the expressive level.

Deficient cognitive functions at the output phase:
1. Egocentric communication modalities.
2. Difficulty in projecting virtual relationships.
3. Blocking.
4. Trial and error responses.
5. Lack of, or impaired verbal or other tools for communicating adequately elaborated responses.
6. Lack of, or impaired need for precision and accuracy in the communication of one's responses.
7. Deficiency in visual transport.
8. Impulsive, random, unplanned behavior.

APPENDIX II

The Cognitive Map
Feuerstein, Rand and Hoffman (1979) (p.122-125)

The cognitive map includes seven parameters by which a mental act can be analyzed, categorized and ordered – content, modality, phase, operations, level of complexity, level of abstraction, and level of efficiency – and enable the use of a process-oriented approach.

Content. Each mental act can be described according to the subject matter with which it deals and the universe of content on which it operates. Experiential and educational background and culturally determined saliency of a content contribute to differential levels of competence among individuals.

Modality. The mental act is presented in a variety of languages: verbal, pictorial, numerical, figural, or a combination of these and other codes, which range from mimicry and metalinguistic communication to conventional signs that are totally detached from the content they signify. The efficiency in use of specific modalities may differ in various socioeconomic, ethnic, or cultural groups, as well as in individuals.

The modality in which the tasks are presented deserves careful consideration, for a quasi-total failure may be converted into a correct response by shifting the modality of presentation of the task and the expression of its solution. One cannot decide that an operation, sui generis, in inaccessible to a child simply on the basis of his inability to perform it in a specific modality. On the other hand, the difficulty involved in using a particular modality must be understood in order to be challenged.

Phase. A specific mental act can be divided into three basic phases: input, elaboration and output. The identification of a phase is neither necessary nor possible when the response is appropriate; however with failure, it is necessary to isolate the phase responsible and to assign a differential weight to it. The individual's response may have been inadequate because of incomplete, imprecise or inappropriate gathered data, which even if elaborated properly would lead, ipso facto, to a failure in the output phase. Failure may occur despite proper input and elaboration if the examinee is unable to communicate the response adequately because of egocentricity or the lack of verbal tools.

Operations. A mental act may be analyzed according to the operations that are required for its accomplishment. An operation may ne understood as a strategy or a set of rules, in terms of which information derived from internal or external sources is organized, transformed, manipulated and acted upon. In defining the nature of the operation, it is important to identify

the prerequisites necessary for its generation and application. Operations may be relatively simple or complex. Classification, seriation, logical multiplication, or analogical, syllogistic, or inferential thinking are obviously more complex than recognition or comparison.

Level of Complexity. The level of complexity of a mental act may be understood as the quantity and quality of units of information it contains. The quality of the information is a function of its degree of novelty. The more familiar the units, even if they are multiple, the less complex the act; the less familiar, the more complex the mental act is. To determine the complexity of a task for an examinee, then, requires a differentiated count that considers simultaneously both the number of items and their degree of familiarity. Either teaching the examinee how to break a task into its component parts and/or familiarizing him with them, thereby making them accessible to him, may help us to view the failure differently and ascribe a different meaning to it.

Level of Abstraction. The level of abstraction defines the distance between the given mental act and the object or event upon which it operates. Thus a mental act may involve operations on the objects themselves, such as sorting, or it may involve relationships between purely hypothetical propositions without reference to real or imagined objects or events.

Level of Efficiency. Efficiency can be perceived as both qualitatively and quantitatively different from the other six parameters although it may be determined or affected by one or more of them, either singly or in combination. For instance, a high level of complexity attributable to a lack of familiarity may lead to a relatively inefficient handling of a task. The inability to isolate efficiency from capacity is an important source of error in the assessment of an examinee's true capacity and repertoire of information and skills. It results in faulty labelling and an erroneous prognosis. The lack of efficiency, defined by slowness, reduced production, or imprecision, may be totally irrelevant to the capacity of the individual to grasp and elaborate a particular problem.

Inefficiency may be caused by a variety of task-intrinsic and/or task-extrinsic factors. Fatigue, anxiety, lack of motivation, and the amount of required investment may all affect the individual in his performance of a task. The recency of acquisition of a pattern of behaviour must also be considered, inasmuch as a behaviour that is neither automatic nor crystallized is more vulnerable to the impact of interfering factors. The more established and crystallized the patterns, the less it will be disrupted by emotional or extrinsic factors. Conventional test scores actually reflect efficiency in terms of rapidity and the number of correct responses, without taking into account any of the other parameters of the mental act.

APPENDIX III				1	2	3	4	5	
RESPONSE TO MEDIATION SCALE		**LIDZ 2003**							
Child:			**Date:**	**Score**					
				1	2	3	4	5	
A SELF REGULATION OF ATTENTION									
	1	Unable to maintain attn to task							
	2	Fleeting attention to task even with input from adult							
	3	Maintains with significant input from adult							
	4	Maintains with occasional input from adult							
	5	Maintains with no input from adult							
		Does not apply							
B SELF REGULATION OF MOTOR ACTIVITY									
	1	Impulsive to point of disruption							
	2	Impulsiveness needs significant restraint from adult							
	3	Impulsive control needs moderate restraint from adult							
	4	Impulsive control needs minimal restraint from adult							
	5	No evidence of difficulty with impulse control							
		Does not apply							
C SELF REGULATION OF EMOTIONS									
	1	Extreme emotional lability; difficulty self-calming							
	2	Significant emotional lability; difficulty self-calming							
	3	Minimal emotional lability; able to self-calm							
	4	Rare emotional lability; able to self-calm							
	5	No evidence of emotional lability							

D STRATEGIC PROBLEM SOLVING						
1	Does not engage in any organized manner with task					
2	Engages but uses trial and error approach					
3	Pauses for seeming momentary reflection before proceeding					
4	Some evidence of planful, organized task involvement					
5	Clearly planful and well organized approach					
	Does not apply					
E EVIDENCE OF SELF-TALK WHEN WORKING ON CHALLENGING TASK						
1	No evidence					
2	Makes noises, but these express effort, not task					
3	Verbalizes, but content is not task related					
4	Makes task related comments					
5	Task-related comments guide efforts at task solution (inc. muttered, unclear comments)					
	Does not apply					
F INTERACTIVITY WITH MEDIATOR						
1	Does not engage in turntaking communications					
2	Minimal engagement in turntaking communications					
3	Moderate engagement in turntaking communications					
4	Comfortable, frequent engagement in turntaking communications					
5	Initiates and responds appropriately and expansively in several chains of conversational interactions					
	Does not apply					
G RESPONSIVENESS TO INITIATIONS OF MEDIATOR						
1	Resisitive to mediator's initiatives					
2	Passive noncompliant					
3	Passive minimally responsive					
4	Consistently responsive					
5	Enthusiastic and responsive					
	Does not apply					

H COMPREHENSION OF THE TASK						
1 No evidence of task comprehension						
2 Willing imitator, but needs model, demonstration, or move through						
3 Slow to comprehend, but does eventually get it						
4 Average comprehension of task						
5 Quick to comprehend task						
Does not apply						
I RESPONSE TO CHALLENGE						
1 Refuses, cries, or tantrums in response to challenge						
2 Begins but quickly gives up						
3 Persists, but with significant encouragement from adult						
4 Persists and completes task, with minimal adult encouragement						
5 Energized by challenge; enjoys the challenge						
Does not apply						
J USE OF ADULT AS A RESOURCE WHEN CHILD NEEDS HELP						
1 Does not refer to adult						
2 Nonverbally, passively signals need for help						
3 Nonverbally actively seeks help						
4 Verbally asks for help						
5 Actively seeks help and seems to appreciate help provided						
Does not apply						
K INTEREST IN ACTIVITY MATERIALS						
1 Shows dislike of materials						
2 Neutral reaction to materials						
3 Minimal interest in materials						
4 Fluctuating interest in materials						
5 Consistently strong interest in materials						
Does not apply						

APPENDIX IV

Grammatical structure of Test items

Item **Content structure**
No

1   SVO   Declarative with auxiliary plus main verb + PossPron NP
2   SVO   Declarative with auxiliary plus main verb + future tense VP
3   SVcSV   Declarative with coordination
4   SVOcSVO   Declarative with coordination
5   SVOdOi   Dative N and N in Subject position
6   SVOiOd   Dative N and N in Obj position
7   SVA   Declarative with copula verb and Adj NP
8   SVA   Declarative with copula verb, + contracted neg
9   SVA   Declarative with modal auxiliary plus main verb and prep phrase
10 SVOA Declarative with modal auxiliary plus main verb, Object, and prep Adverbial phrase
11 SVOsSVO Declarative with subordinate clause (Use of anaphoric pronoun)
12 SVsSV   Declarative with subordinate clause

APPENDIX V   EQUIVALENT VERSIONS OF TEST MATERIALS

Version	Test Items
A	The man is painting my wall
	Is the man painting my wall?
B	The boy is washing my car
C	The dog is eating my bread
D	The girl is reading my book
A	Billy is going to score a goal
	Is Billy going to score a goal?
B	Mary is going to draw a picture
C	Jon is going to fly a kite
D	Sue is going to see a film
A	Mum is eating and Dad is drinking
	Mum is drinking and Dad is eating (etc)
B	Sally is walking and Sue is running
C	Jos is winning and Charlie is losing
D	Sara is sitting and Jeni is standing
A	Mum is picking the flowers and Dad is cutting the grass
	Mum is cutting the grass and Dad is picking the flowers (etc)
	NB that 'picking the grass'* is unacceptable
B	
C	Mum is driving the car and Dad is riding the bike
D	Mark is reading the newspaper and Tom is writing a story
	Phil is pouring the tea and Mum is cutting the cake
A	The girl and the boy showed the monkey a banana
	The boy and the girl showed the monkey a banana
B	The brother and the sister fed the baby a bottle
C	The Mum and the Dad sent the boy a letter
D	The girl and the boy gave the dog a bone
A	The girl gave the boy a drink and a biscuit
	The boy gave the girl a biscuit and a drink (etc)
B	The man sent the lady a card and a present
C	The girl lent the boy a pencil and a rubber
D	The man showed the boy a bat and a ball
A	The big boy is hungry
	Is the big boy hungry?
B	The little girl is happy
C	The old man is tired
D	The clever boy is naughty
A	The bird's cage isn't broken
	Isn't the bird's cage broken?

APPENDIX V   EQUIVALENT VERSIONS OF TEST MATERIALS

B   The girl's room isn't messy
C   The dog's paw isn't hurt
D   The cat's fur isn't dirty
A   Susie will hide under the table
           Will Susie hide under the table?

B   David will sleep on the sofa
C   James will sit in the car
D   Marc will look under the bed
A   You can hang your coat on the hook
           Can you hang your coat on the hook?

B   You can stick your picture on the fridge
C   He can put his keys on the rack
D   I can leave my note on the table
A   Joe brushes his teeth before he goes to bed
           Before he goes to bed, Joe brushes his teeth (etc)
           NB He goes to bed before Joe brushes his teeth*

B   Adam locks the door before he starts the car
C   Miriam goes to school after she combs her hair
D   Dan eats his dinner after he washes his hands
A   Debby cried because the window broke
           Because the window broke, Debby cried
           NB the window broke because Debby cried*

B   Hannah laughed because the clown slipped
C   Neil sneezed because the pepper spilt
D   Cassie screamed because the door banged

APPENDIX VI

Score Sheet (for use during the DA session)

Child……….   Time ……

Item No	Content structure	Cue level 1	2	3	4	5	No of cues
1a	Declarative with auxiliary plus main verb + PossPron NP						
1b	Interrogative with aux reversal						
2a	Declarative with auxiliary plus main verb + future tense VP						
2b	Interrogative with aux reversal						
3a	Declarative with coordination SVcSV						
3b	Reversible content						
4a	Declarative with coordination SVOcSVO						
4b	Reversible content with some semantic constraints						
5a	SVOdOi Nand N in Subject position						
5b	Reversible content						
6a	Declarative SVOiOd N and N in Obj position						
6b	Reversible content						
7a	Declarative with copula verb and Adj NP						
7b	Interrogative with verb reversal						
8a	Declarative with copula verb, +contracted neg						
8b	Interrogative with verb reversal						
9a	Declarative with modal auxiliary plus main verb and prep phrase SVPrepP						
9b	Interrogative with aux reversal						
10a	Declarative with modal auxiliary plus main verb, Object, and prep phrase SVOPrepP						

10b	Interrogative with aux reversal
11a	Declarative with (temporal) subordinate clause SVOsSVO (Use of anaphoric pronoun)
11b	cSVOSVO reversal of subordinating conjunction
12a	Declarative with (causative) subordinate clause SVsSV (Semantic constraints)
12b	cSVSV reversal of subordinating conjunction

Total No of cues:

_____

APPENDIX VII
Information sheet for parents

Dynamic assessment and Intervention for Children with Specific Language Impairment.

Information for Parents

Invitation

Your child is being invited to take part in a project. Before you decide, it is important for you to understand why the project is being done, and what it will involve. Please take time to read the information carefully, and if anything is not clear to you, or you would like any more information, please ask me any questions you have. Take time to decide whether or not you wish your child to take part.

This sheet tells you the purpose of this study, what will happen to your child if he/she takes part, and gives you more detailed information about the conduct of the study.

What is the purpose of the project?

The project involves testing your child's language skills thoroughly, before and after a period of intervention. After that, an additional test procedure will be carried out, that involves assisting your child through a task, and evaluating the amount of prompting he/she needs to succeed. This is called a **dynamic** test. The information gained from this procedure will be incorporated into a further period of intervention, and again the outcomes will be checked after a period of therapy. The aim is to see whether using the dynamic test procedure gives information that is more helpful in planning therapy for your child, than the usual tests.

Why has my child been chosen?

A small number of children with language difficulties, known to Speech and Language Therapists, have been chosen as I believe that the dynamic testing may help them, and I may be able to make detailed recommendations for their further progress, based on the procedure. The children were selected on the basis of their age and the nature of their language difficulties. In addition your child's Speech and Language Therapist has agreed to participate in the study.

Does he/she have to take part?

It is up to you to decide. I will describe the study and go through this information sheet, which I will then give to you to keep. I will then ask you to sign a consent form to show you have agreed for your child to take part. You are free to withdraw your child at any time, without giving a reason. This would not affect the standard of care he or she receives.

What will happen to him/her if he/she takes part?

Your child will be seen by me, on his/her own at his school for approximately an hour of language assessment at the start of the Autumn term, instead of being tested by his regular Speech and Language Therapist.

Then he/she will continue in his/her regular Language therapy for the next 2 months. At the end of that time, (towards the end of the Autumn term), I will do some more language tests to find out how much he/she has improved. I will also take an additional hour for the dynamic assessment, but that will not feel like a test, and your child may think they're getting some extra help. The activities will aim to find out more about the nature of your child's language difficulties, and what strategies may help him/her to improve.

Then, if your child is in the experimental group, I and your child's regular Speech and Language Therapist will use the test results to plan a further 2 months of therapy, in the Spring Term, that will be carried out by his/her therapist at the school in his/her regular therapy times. At the end I will repeat all the tests to find out how much he/she has improved.

If your child is allocated to the control group, I will hold all the results for two months, and then give all the information to his/her Speech and Language Therapist and teacher before the end of the Spring Term. If there have been positive results that suggest that the changes made in the therapy for the experimental group have been useful, we will plan improved therapy for your child in the same way. If there has been no difference, your child's regular therapy programme will continue.

All of that should be carried out during one school year, starting in September 2008, and finishing by the summer of 2009. I may come back for one hour during the following term to test him/her one more time to find out whether he/she has retained everything he/she learnt.

Your child will be getting exactly the same amount of therapy that he/she usually gets, and it will be carried out by the same therapist, but there will be extra testing to see how much he/she is learning,

and whether we have been able to help him/her learn more easily or quickly.

I will need to videotape the tests so that I can analyse them in more detail. I will keep all the recordings on my personal computer, and they will not have your child's name attached to them. No information about your child will be passed on to anyone else, all the results will only have a number on them. I will need your permission to tape your child. I will also ask your for permission to use the tapes for lectures or conference presentations, if the findings of the study will be useful for training other Speech and Language Therapists.

What will happen if he does not take part, or drops out?

Your child will continue to receive his/her regular language therapy from his /her Speech and Language Therapist, as before.

What happens when the research study stops?

Your child will continue to receive the language therapy that he/she needs from his/her Speech and Language Therapist

What are the benefits of taking part?

Your child will be thoroughly tested and receive a period of therapy to help him/her with his/her difficulties. You will receive a report on what I have found out about your child, and ideas for future therapy, using methods that have
been shown to help him/her. Your child's school teachers and his speech and language therapist will receive a copy of this report.

Will my taking part be kept confidential?

Yes. We will follow ethical and legal practice and all information which is collected about your child during the project will be kept strictly confidential. His/her name will not appear on the videotapes, in any published materials, or any lecture or conference presentations.

Who has reviewed the study?

All research in the NHS is looked at by independent group of people, called a Research Ethics Committee to protect your safety, rights, wellbeing and dignity. This study has been reviewed and given favourable opinion by Redbridge and Waltham Forest Research Ethics Committee.

What will happen to the results of the study?

You, the school, and your child's speech and language therapist will receive a report on what I have found out about your child, and my recommendations for his/her further management. If the methods used are found to be useful, they may be reported in a professional journal for Speech and Language therapists. In addition, the findings may be presented at conferences or lectures for the purpose of training other Speech and Language Therapists. A paper detailing the study will be presented to City University as part of my further degree. These reports will not identify the children taking part.

What if there is a problem?

If you have a concern about any aspect of this study, you should ask to speak to the researcher (on 07711 649550) who will do their best to answer your questions. If you remain unhappy and wish to complain formally, you can do this through the NHS Complaints Procedure or City University.

Any Other Questions?
Contact: Natalie Hasson, Department of Language and Communication Science, City University, 020 7040 8280, or 07711 649550 n.k.hasson@city.ac.uk

Appendix VIII

Transcription of mediation

<u>Child X</u>   Time 2,   Item 5   T= Therapist

T: What's this one about?

X: Dad and…..Dad and….. Mum and Dad..

T: Mum and Dad, good, good place to start. What did they do in this

        one?   Remember we've got a letter, we've got boy, and we've got

        sent   (*reminder   for non-reading child*).

        So what did they do?

X: sent a…..

T: sent a ? What did they send?

X: sent a…. sent a……

T: letter

C: a letter

T: they sent a letter

C: to boy, to the boy!

T: Very good, that's the idea. We haven't got all those words. Can we

         try Mum and Dad sent….

C;   a

T: ..sent the boy..

C: the boy, a letter!

T   Well done. Can you make another one?

C: the boy sent a letter to Mum and Dad

T: Good. The boy sent Mum and Dad <u>a letter</u>

C: <u>a letter }</u> *simultaneously*

<u>Child Y</u>   Time 2   Item 8

 Y: The dog isn't hurt paw   (*shakes head, No*)

T: Nearly there

C: The dog……isn't hurt

T: Very close. What do you see here (*point to apostrophe 's*)

C: dog**s**

T: dog**s**. What does that 's mean?

C;   More than one

T: It could do, but not in this one. It means something belongs to the

        dog. You know, if I said this is Y's jumper, it belongs to you

C: Right

T: Right, I would say Y's with a **s**, that means it belongs to you. So

        here, look, something belongs to the dog. What belongs to the dog?

C: paw

T;   Yes, that's right. So now can you make the sentence?

C: dog's paw hurt isn't the   (*shakes head, no*)

T: Hang on….the dog's.. what did you say belongs to the dog?

C: Paw

T: The dog's paw..

C: The dog's paw isn't hurt   (*shakes head, No*)

T: (repeats back) The dog's paw isn't hurt   (*C nods head, Yes*)

APPENDIX IX

Manual for the SLTs delivering the experimental intervention

1. **Introduction**
2. **Targets of the Intervention**
3. **Strategies and Methods of Intervention**
4. **Application**

1. **Introduction**

The experimental intervention phase, like the first intervention, will be delivered over a period of one school term, within the usual timetabled SLT sessions.

There are two main differences in the intervention programmes;

i)   that the experimental intervention is informed by the results of the Dynamic assessment,

ii)   that the theoretical basis of the experimental intervention is consistent with that of the dynamic assessment, and derives from the theory of Feuerstein.

The information derived from the Dynamic Assessment consists of:

- learning needs, ie whether the individual requires metacognitive monitoring, strategy training or item specific application of knowledge
- the individual's learning needs in terms of amount of input required from examiner

In addition:

6. Detail of language structures that the child has difficulty with, that is additional to that obtained from the standardised tests
7. the effect of amount of content and nature of semantic content on the child's construction of linguistic structures
8. the child's ability to transfer, or generalise learning or strategies ie. item - to item transfer
9. the child's *metalinguistic* knowledge, ability to label, explain and manipulate linguistic concepts
10. the child's *metacognitive* ability ie awareness of the processes and strategies that are used to solve the given task

The DA will also contribute information about the child's:

- attention /activity/ emotion while engaged in the presented task
- motivation / attitude to learning / interest / response to input, while engaged in the presented task
- use of strategies, including reliance on others for help

This information will influence both the targets set for the intervention, and the methods and strategies used in the delivery of intervention. In addition, Feuerstein's theory of Mediated Learning Experience, (Feuerstein 1980) informs the targets that address cognitive awareness of the tasks, reflective thinking and problem solving, and also the nature of the intervention, that is mediational in style.

2. Targets of the Intervention

Like the first intervention, the programme will address language and communication targets.
The SLT will set 3-4 SMART targets for the period of intervention
The targets will be derived from-

1. Ongoing intervention - considering the previous targets of intervention, and the progress made towards achieving them

2. CELF - specific targets arising from subtests

3. DA information

4. Identification of common or recurrent difficulties

The SLT and the experimenter, will extract features of the child's language processing difficulties that impact on more than one application (as in 4 above), and devise process based intervention to address these difficulties, as well as metalinguistic knowledge of the structures and metacognitive awareness of the processes.

Eg:
• difficulty with reversible sentences, relates to verb arguments
Therapy needs to identify verbs, then thematic roles, and use these as a means to sequence and structure elements of a sentence

• difficulty with the use of auxiliary verbs in declarative sentences to carry tense/aspect marking as well as to form questions.
In therapy, identify verbs, tenses, verb endings. Then auxiliary verbs and the way they mark tense/aspect. Formulate rules for constructing questions

• difficulty with sentence length
Formulate strategies for grouping, or chunking phrases, and then combining

• if child does not benefit from previous examples - does not transfer
Make examples explicit, identify similarities and analogies

Finally, the SLT should identify extended application of the target structures, eg into a variety of linguistic contexts, such as the construction of narratives, and include functional use of linguistic constructions.

Eg. AMY

Specific difficulty on sentence - 'Joe has a bath before he goes to bed'
Required prompt to group words into clauses 'has a bath' / 'goes to bed'
before being able to determine the use of 'before'

Target would be *grouping* words into Verb-Object or Verb-Complement or
even Verb-Adverbial units.

Plan to look at a wide range of verbs, both semantically specific eg
'brushes' vs general 'has' - to capture very large number of activities.

Address sequencing activities according to daily life and experience, then
use of both 'before', 'after' and also 'then' to arrange sentences.

Then extend to use in narrative.

3. Strategies and Methods of Intervention

All intervention will be delivered by SLTs as before, and will be informed
by:

1) The Therapist's own experience and expertise in managing the
individual
2)   The information elicited from the Dynamic Assessment, (as above)
3) The theoretical basis for the intervention, which is to adopt the mediated
learning framework (Feuerstein, Lidz 1991, Haywood 1993)

The SLT should consider the information made available from the DA and
response to mediation scale, and adapt therapy as required.

All activities and materials to be used are at the discretion of the SLT.

Theoretical Basis
Mediational teaching is the intervention method recommended in the current
study. Derived from Feuerstein's theory, it is the application of principles
thought to be essential for adequate cognitive development in children
(Haywood 1993).

In the current application, mediation has the function of helping children to
understand the structure and meaning of language, and that following rules
of language will enable expression and further understanding of ideas, that
in turn enable further learning to take place.

The intervention should incorporate the essential components of mediated
intervention. These essential components, (Lidz 1991) which will be
explained in more detail later, are:
- mediation of intentionality
- mediation of meaning
- mediation of transcendence
- mediation of a feeling of competence

Mediation of meaning and transcendence imply explicit, metacognitive
teaching, making sure at each stage that the child grasps the principle that he
is learning, its importance and application to the task and wider, functional
use. This is consistent with the targets that address cognitive awareness of
the tasks, and reflective thinking that were identified earlier.

The method shifts the emphasis of therapy away from modelling and
towards a more problem-solving approach. Facilitation is minimal, and
introduced only if and when required to enable the child to formulate a
strategy for problem solving.

4. Application

What makes the session *mediational*?

Mediated Learning Experience is defined by the presence of a number of mediating behaviours. The most essential of these, have been adapted from Feuerstein by Carol Lidz, (1991 and 2003 P63) and presented as follows:

Mediation of intentionality – conveying to the child that you intend to help him improve,
 'This includes communication to the child of the purpose for the interaction, as well as attempts by the mediator to maintain the child's involvement in the interaction.' (P63)

Mediation of meaning – sharing the purpose of the activity, 'moving the content from neutral to a position of value and importance' (P63)

Mediation of transcendence – linking the activity to other contexts in which the skill can be used, 'promotion of cognitive bridges between the task or activity and related but not currently present experiences of the child; these may refer to the past or may anticipate the future' (P64)

Mediation of a feeling of competence – targeting praise so that the child learns what he has done well, learns that the tester has confidence in him, and gains confidence in his own ability.

What do I do?

According to Haywood (1993 P31) mediators have the following functions:

1. Supply the information that may be needed to learn relationships or find solutions
2. Ask questions, ie elicit rather than give answers
3. Guide learning by arranging and directing sequences of experiences in a
   developmental fashion
4. Bring about induction of rules by calling attention to similarities among events or
    examples
5. Facilitate application of rules
6. Build confidence of children
7. Maintain a metacognitive emphasis, ie focus attention on the child's own thinking
    processes and encourage them to do the same

How do I Mediate? (Haywood 1993)

Ask process questions - usually containing 'How?'
Eg
How did you know?
How else could you do that?
How can you find out?

Bridge to different applications
Elicit ideas from the children - rather than telling them
Eg..
When is another time you could do that?

Challenge the child to justify his answers
Challenge both right and wrong responses
E.g.
That's right, but how else could you do it?
Why is that way better than this?

Teach about Rules
E.g.
If we have that and that - what rule can we make? Does it apply to this?
Would it help to have a rule here?
Would a rule help us to know what to do?

Emphasize order, system, sequence and strategy
Try and facilitate a predictable approach and enable planning
E.g.
Use a timetable for the day and/or the session
Reduce trial and error behaviour, guessing and random answers.

Create Task -Intrinsic Motivation (Feuerstein)
Help the child to appreciate that the task is meaningful and motivate him to
complete a task by emphasizing the achievement

APPENDIX X

Questionnaire for Teachers   Date:

School:………   Child's Initials:……..   Child Code:….

Please circle the number of the descriptor you feel is most applicable.
Please feel free to write in comments if you want to clarify or add
anything

1. How motivated and engaged is X in an average language lesson in
   the classroom?

Not at all	Not very	Reasonably	Well	Very well
1	2	3	4	5

2. How well does X manage to keep his/her attention on a language or
   literacy task?

Not at all	Not very	Reasonably	Well	Very well
1	2	3	4	5

3. How well does X manage to keep his/her attention on a practical
   task?

Not at all	Not very	Reasonably	Well	Very well
1	2	3	4	5

4. How much difficulty does X have in problem solving or working
   through a task?

None at all	Not much	Some	Quite a lot	A great deal
1	2	3	4	5

5. How well does X follow rules? (eg school rules, rules of a game, or
   of a class activity)

Not at all	Not very	Reasonably	Well	Very well
1	2	3	4	5

6. How well does X follow verbal instructions?

Not at all	Not very	Reasonably	Well	Very well
1	2	3	4	5

7. How frequently do you repeat instructions specifically for X ?

Not at all	Occasionally	Sometimes	Often	All the time
1	2	3	4	5

8. How frequently or consistently does X ask for clarification or help?

Not at all	Occasionally	Sometimes	Often	All the time
1	2	3	4	5

9. How willing is X to take up and learn from directions or feedback?

Not at all	Not very	Reasonably	Well	Very well
1	2	3	4	5

10. How well does X remember or retain what he/she has learnt?

Not at all	Not very	Reasonably	Well	Very well
1	2	3	4	5

11. How well can X get his/her own message over in spoken language?

Not at all	Not very	Reasonably	Well	Very well
1	2	3	4	5

12. How well can X relate a story or incident in order that you can clearly understand what took place?

Not at all	Not very	Reasonably	Well	Very well
1	2	3	4	5

13. How much difficulty does X have explaining what he/she is thinking or feeling?

None at all	Not much	Some	Quite a lot	A great deal
1	2	3	4	5

14. How much care does X take over his/her work?

None at all	Not much	Some	Quite a lot	A great deal
1	2	3	4	5

15.   How would you rate the appropriateness of X's behavioural responses to situations or challenges?

Very appropriate	Usually appropriate	Reasonably	Not very	Not at all
1	2	3	4	5

Thank you for your help.

APPENDIX XI

Questionnaire for SLTs

Monitoring of Intervention:

Child's Initials…………   School……………….   Date ....

1. Outline 3 targets that you are currently working towards with the child:

i)

ii)

iii)

2. 'Outline 3 activities that you have recently or currently engaged in with this child' (Law et al 2008 P249)

i)

ii)

iii)

For each activity, specify at least one method of facilitation used, and how change was measured.

Thank you,

Natalie

APPENDIX XII

Sample report

Findings of Dynamic assessment

Name…X   **Date ….Time 3 April 2009**

1. Detail of language structures that the child has difficulty with, that is additional to that obtained from the standardised tests

Good, stable use of SVO
Difficulty with SV OdOi
Unsure that 'Isn't' has same function as 'is'
Does not recognize possessive '-'s'
Difficulty arranging elements of sentence using 'after' – confused temporal sequence

2. The effect of amount of content and nature of semantic content on the child's construction of linguistic structures

 Confused by semantic use of 'after'
Inclined to attempt grammatically correct but semantically illogical sentences

3 The child's ability to transfer, or generalise learning or strategies ie. item - to item transfer

Transferred strategy – formulated Q spontaneously on following item
-Unable to transfer OdOi structure
X retains and repeats strategies and ideas previously learnt in class or SLT, but is not always able to apply them to new examples or situations presented

4 The child's metalinguistic knowledge, ability to label, explain and manipulate linguistic concepts

- Aware of task requirements and concept of sentence' 'You have to make a sentence then another sentence'
- Can formulate a question when prompted - knew she should start with 'Is-'
- Cannot identify action (or 'doing word') in sentence
- Recognizes plural '-s' but unable to explain the concept

5 The child's metacognitive ability ie awareness of the processes and strategies that are used to solve the given task

- Aware of task demands –can apply concept of 'rules' of task – eg using all of the given words

- Able to identify difficulties – I've left that one out
- Identified strategy –'mix' the words around, swopped the words
- Identified that strategy was used previously
- Identified that it is difficult to arrange function words eg 'the' and 'and' easy to arrange content words

Sometimes uncertain of judgment of grammaticality

Behavioural Factors:

- *attention /activity/ emotion while engaged in the presented task*

X attended to the task throughout with no input from the SLT

- *motivation / attitude to learning / interest / response to input, while engaged in the presented task*

X appeared really interested in, and motivated by the task, and was keen to persevere and succeed, reinforced by verbal praise and feedback from the tester. She responded well to all input from the tester and retained content of what was said to her.

- *use of strategies, including reliance on others for help*

X used strategies for problem solving spontaneously, but was aware of the tester's willingness to help her, and used support only when she had attempted the task herself, and been unable to solve it.

Summary:
- learning needs, ie whether the individual requires metacognitive monitoring, strategy training or item specific application of knowledge
- the individual's learning needs in terms of amount of input required from examiner

X has structural regularity in her language and can formulate basic sentences reliably. She still has considerable difficulty manipulating grammatical morphemes, and formulating more varied sentence structures accurately, and this is exacerbated by uncertainty in making judgements of correctness. She demonstrates metalinguistic awareness and knowledge, and attempts to use strategic problem solving spontaneously.

X has good interpersonal communication skills and willingness to engage and to learn.

X may require ongoing SLT intervention to address both receptive and expressive language difficulties, focusing on structural features of language and morphology, but her inclination to try and transfer and generalize what she has learnt may facilitate learning of systematic rule governed aspects. Intervention might therefore, make use of explicit rule teaching and application.

As X is uncertain about her own judgments of correct grammar, it may be useful to carry out activities requiring judgment as well as justification for her choices, based on rules she has learnt. This could be applied to contexts wider than language, and may facilitate increased confidence.

APPENDIX XIII

 Feedback questions for SLTs

Feedback from Speech and Language Therapists involved in the DA study

Do you think the information supplied to you by the experimenter about the children, after they had had the DA, was useful?

Were you able to make use of this information in your planning of intervention for the children?
How?

Do you think it made any difference to the outcomes of your therapy?

Did it make any difference to you, in your thinking about the child or their therapy?

Do you think it would have been useful to see the videos taken of the DA procedure?

Would you have been willing to watch a 30-40 minute video for each child, as well as read the report?

How do you think the procedure used or the information supplied to you, might have been improved?

APPENDIX XIV

Compilation of feedback from SLTs

Feedback from Speech and Language Therapists involved in the DA study

Do you think the information supplied to you by the experimenter about the children, after they had had the DA, was useful?

1 Yes, both in terms of the children's metalinguistic potential and the actual target setting.

2 Yes it was. However it would have been even more useful if all the therapists concerned had been able to attend the training about the Dynamic Assessment beforehand, local training restrictions limited the number of therapists that could be released.

3 Yes

4 Yes. It provided a further dimension to the child's insight into their difficulties and their problem solving abilities that I had not fully explored.
5 Yes, very much so. It was useful to have my thinking redirected towards the children's *process* of learning and how well they could explain rules etc when they were attempting language intervention tasks.

6 Yes, very!

7 The test results were useful for the reviews. A more in depth qualitative summary would have been more informative

8 The scores were useful, however the children I work with have mainly social and pragmatic needs, with several of their scores within normal limits.

Were you able to make use of this information in your planning of intervention for the children? How?

1 Yes, Planning now includes explicit introduction, i.e. talking about the target itself, with the child reflecting on the particular skill or area of difficulty, then discussing how they can best deal with it.
2 Yes, I was able to introduce more metalinguistic targets into my planning and asked the child more about what he had done that had helped him to succeed.
3 Yes. I reflected more on the child's meta-cognitive / meta-linguistic skills and this influenced some of my targets.

4 Yes. I used the information to encourage greater levels of problem-solving and generalising of strategies. I also ensured that at the end of each session each child reflected more thoroughly on what they had done that had helped them with that task.

5 Yes, I used the information from the DA to develop new targets that slotted in nicely to interventions that I was already running. For example in verbal reasoning activities I adapted the targets so that the children had to explain how they knew that was the answer.

6 *It reinforced some of the ideas we were already using in therapy, which was reassuring! It was also helpful to see which of our Resource children met the criteria for the project – I've started working differently with those not found to meet the criteria for good non-verbal skills.*

7 No – due to limited information and the fact that in this setting we have the luxury of getting to know the children's, strengths and difficulties very well.

8 As we know the children and their needs in depth in this setting, the information was useful, however a lot of the information incorporated things we were already doing.

Do you think it made any difference to the outcomes of your therapy?

1 Yes. The children feel positive and smile 'knowingly' during activities. I feel they take more responsibility for achieving the targets.

2 As yet there is no evidence from the quantitative data, but qualitatively the child does ask more questions when he is unsure and is more confident to "have a try" rather than sit silently. This means that functional communication has improved. This may also be due to the fact that he has settled into the base routine however.

3 I think the children developed a better understanding of their own targets and progress and for some this improved motivation and hence functional progress.

4 I think it helped with the generalisation of strategies so that skills and strategies were more likely to be applied to tasks outside of the therapy session.

5 It's hard to tell – the children have made progress towards their targets over the past few terms, but they also have significant language impairments so their progress is usually small.

*6 Difficult to say – our therapy has been so disrupted due to ongoing
problems in agreeing interventions with the school, so would be hard to say
whether effects seen are due to our therapy/recommendations out of DA
project/maturation etc.*

7 No due to the reasons above.

8 Very difficult to say as we would be doing it anyway.

**Did it make any difference to you, in your thinking about the child or
their therapy?**

1 Yes. For a long time, as a therapist, I have called upon the children's
ability to reflect on their own experience and performance. However, I have
never actually probed the extent to which each child is able to participate in
a discussion about their language.

2 I think I have always tended to work in this way, but what it did do
is consolidate for me that there was a reason behind why it was a
useful way to work.

3 I put more emphasis on developing the child's awareness of what we
were trying to achieve and why

4 Yes. As stated above it made me think more metalinguistically
and metacognitively rather than just teaching and practising
skills. It also highlighted the different ways different children
approach tasks and how they learn and reflect on their learning,

5 Yes – for one child in particular it has helped me to understand
why he 'guesses' answers so often (the DA revealed he poor ability
to apply or explain language rules). It has also encouraged me to
think about rationales for my interventions more, and to consider the
learning process for a child when introducing new therapy goals.

6 Yes, see response to Q2 above.

7 No – not enough information was received.

8 It was good to hear that the approach taken with this children seemed to
the be the correct one.

**Do you think it would have been useful to see the videos taken of the DA
procedure?**

1 Yes, very much so. It was only when I saw an extract of one at a training
session that I realised how to actually guide a child in this kind of discovery.
2 Yes, very!

3 yes
4 Definitely

5 Yes – I am very interested in DA and would like to be able to use some of the techniques in my ongoing assessment and therapy, and feel it would be useful to see how Natalie used these.

6 Yes, that would have been very interesting.

7 Yes, very

8 Yes – very. I feel I would have understood it all a lot better. Email contact is difficult as we are so busy.

Would you have been willing to watch a 30-40 minute video for each child, as well as read the report?

1   Yes.
2   Personally yes. I feel it would have been worth doing in my own time if necessary, however others may not feel the same way.
3 yes
4 Yes
5 Yes
6 In theory yes – finding time to do extra things is always difficult of course, but it would have been interesting to see.
7 Yes
8 Yes

How do you think the procedure used or the information supplied to you, might have been improved?

1 I would have liked to actually watch the session or a video, as said above. I cannot comment on the procedure as this was for the purpose of the research, which I do not know enough about.

2. By being able to see the video and having more knowledge of the technique beforehand, some of the references given were difficult to get hold of "in the field"

3. Knowing more about the assessment procedure would have enhanced my understanding of the information in the reports. Watching the videos would probably have been the best way to achieve this.

4 I'm not sure that it could have been different bearing in mind it was important to the study that we were not fully aware of what you were doing. However, attending the training afterwards was a very good way of fully understanding and identifying how I could use these techniques in assessment and therapy.

5 I felt the information that went to parents (apart from the final letter) was not particularly accessible – it made sense to me as a clinician and I was able to interpret and apply it, but don't think it would have meant much to parents without a detailed accompanying explanation.

6 I think it was fine really, no suggestions here.

7 A meeting to discuss the child and the specific techniques used in the assessment that were affective.

8 Meetings and video footage would have been useful – again, contact via email can make it harder to fully understand the project.

APPENDIX XV

Report of Findings of Dynamic assessment

Name…D4   **Date ….T2 January 2009**

1. Detail of language structures that the child has difficulty with, that is additional to that obtained from the standardised tests

Can formulate question given start with auxiliary verb 'Is..' 'Isn't' 'Will' 'Can'
Difficulty formulating question with copula 'Is' – produced <u>'is the man *is* tired'</u>
Knows 'is –ing' present progressive structure. Searches for 'is' in order to make sentence. Difficulty with alternative sentence structures.
Does not recognize relationship between 'is' and 'isn't'
Difficulty using possessive 's (the dog's paw)
Difficulty formulating indirect object – 'give *the boy* a pencil'

2. The effect of amount of content and nature of semantic content on the child's construction of linguistic structures

Basic agent-action concept
Working memory limitation – not able to read all words and cannot retain which words he has used, so longer sentences more difficult for him
Does not always consider semantic/logical organization of sentences

3. The child's ability to transfer, or generalise learning or strategies ie. item - to item transfer

Attention to detail – changes 'a' to 'the'
Good transfer from one sentence to another with reversible content
Poor transfer of unfamiliar sentence structure eg 'sent *the boy* a letter'
Transfers strategy of making a question to later items

4. *The child's metalinguistic knowledge, ability to label, explain and manipulate linguistic concepts*

Awareness of question construction using 'Did..' Is.. and Isn't..
Can identify 'doing word' (verb) in sentence
'A lot of words' – used counting of words as indication of difficulty
Knows that sentence starts with 'person' (though doesn't always spontaneously use this as strategy)
Can identify 'he' as person
Aware of reversibility – 'just swopped A and B'. Knows which items may be reversed 'can change what the people are doing'

5. *the child's metacognitive ability ie awareness of the processes and strategies that are used to solve the given task*

Poor awareness of process of judging grammaticality – does not understand 'How do you know it was right?'
Poor awareness of judging correctness – does not follow rules
Own rating of difficulty does not relate to amount of scaffolding required to solve the task. Cannot give reasons for difficulty.

Good awareness of purpose of session – 'learnt to make sentences and questions'. Difficult to elaborate on strategies for sentence construction.

Behavioural Factors:

- *attention /activity/ emotion while engaged in the presented task*

Good attention, on 1:1 level with support to redirect to task, for first part of session. Became increasingly unwilling to attend.
No fidgeting or distractible behaviour, but tired, wanting to put head down and avoid.

- *motivation / attitude to learning / interest / response to input, while engaged in the presented task*

Poor motivation to do task or to learn. Avoidance – does not want to continue when task is difficult. Passive response to input – responds when required to. Listened and some response to input. Positive but unenthusiastic response to praise.

- *use of strategies, including reliance on others for help*

Willing to seek help from adult and responsive to encouragement.

Summary:
- learning needs, ie whether the individual requires metacognitive monitoring, strategy training or item specific application of knowledge
- the individual's learning needs in terms of amount of input required from examiner

D4 demonstrated a consistent knowledge of some basic grammatical structures, and his spontaneous language is grammatical, if simple. He has a limited number of sentence constructions that he is familiar with, and he tries to recreate these. He struggles with transfer to even quite closely related structures like is/isn't .

Some structures are easily prompted, (eg questions with aux reversal) and D4 can formulate sentences when given the starting item, but cannot

generate them himself. Unfamiliar structures (eg indirect and direct object) require more repetition before D4 can retain them.

D4 could potentially benefit from increasing the variety of his sentence structures, without increasing sentence length which overloads his memory.

D4 has some strategies that he employs that are not useful – eg counting words, and other knowledge that he does not make use of consistently, eg knowing that a sentence starts with 'a person'. It may be useful for him to increase cognitive awareness of strategy use. Similarly, he could benefit from reflecting on semantic content of sentences, and new semantic relationships, in order to help him construct ideas.

He recognizes and can judge grammatical correctness, but there is little if any awareness of this process. He is, however, aware of his own difficulty and avoids tasks that are too difficult, with little motivation to persevere with the task.

D4 has extremely strong interpersonal and pragmatic skills, and continues task to please the adult, rather than with any interest in the task or his own learning.

TIME	NAME	AGE	GRP	CD	WC	SS	SR	FS	RS	WS	SA	REC.	EXPR.	TOTAL	%ile	total	change
				Subtests - Standard scores								TOTAL STD SCORE				Raw sco	Raw score
T1	D1	8;2	C	3	5	5		3	3	3	3	66	66	64	1	60	
T2	D1	8;6	C	4	8	4		3	3	3	5	73	66	64	1	75	15
T3	D1	8;9	C	5	9	7		3	3	4	6	83	66	64	1	95	20
T4	D1	9;4	C	3	8		4	3	3		7	71	71	65	1	68	-27
T1	D3	8.5	E	3	3	3		3	3	3	4	66	66	64	1	51	
T2	D3	8;9	E	3	4	3		3	3	3	6	66	66	64	1	58	7
T3	D3	8;11	E	3	6	4		3	3	3	2	66	66	64	1	75	17
T4	D3	9;6	E	3	6		4	3	3		5	65	65	65	1	63	-12
T1	D4	8;10	E	4	8	6		3	3	4	4	73	66	64	1	83	
T2	D4	9;2	E	4	5		3	3	3		5	65	65	65	1	63	-20
T3	D4	9;5	E	4	8		5	3	3		5	71	65	65	1	77	14
T4	D4	9;11	E	3	9		3	4	3		8	71	71	65	1	84	7
T1	CP2	8,9	C	3	7	8		3	3	3	5	73	66	64	1	81	
T2	CP2	9;0	C	3	5		4	3	4		4	65	65	65	1	67	-14
T3	CP2	9;3	C	3	8		4	3	3		5	71	65	65	1	66	-1
T4	CP2	9;10	C	3	8		4	3	3		9	71	71	65	1	80	14
T1	R1	8;4	E	3	5	10		3	4	3	3	73	66	64	1	87	
T2	R1	8;8	E	5	6	8		4	4	5	3	76	66	64	1	107	20
T3	R1	8;11	E	4	6	10		4	4	4	2	79	66	64	1	103	-4
T4	R1	9;6	E	6	5		3	3	7		5	71	71	65	1	92	-11
T1	TF1	8;2	E	3	3	3		3	3	3	3	66	66	64	1	28	
T2	TF1	8;6	E	3	4	3		3	3	3	3	66	66	64	1	37	9
T3	TF1	8;10	E	3	3	3		3	3	3	3	66	66	64	1	45	8
T1	TF4	8;7	E	4	5	5		3	3	4	7	66	66	64	1	77	
T2	TF4	8;11	E	3	7	4		3	3	4	7	66	66	64	1	88	11
T3	TF4	9;1	E	3	5		4	3	3		10	65	71	65	1	74	-14

APPENDIX XVII

EXCERPTS FROM RECOMMENDATIONS SECTIONS OF ALL PARTICIPANTS' REPORTS

Key:
Recommendations for Metalinguistic targets are highlighted in green
Recommendations for Metacognitive targets are highlighted in blue
Recommendations for grammatical targets are highlighted in pink
Recommendations for social or pragmatic targets are highlighted in red

BH3
further facilitate X's language and communication, via mediation of metalinguistic knowledge and reflection.
colourful semantics to help X to recognize, label and manipulate sentence components, and give him tools for reflection on language structure.
further reflection on processes such as planning, relating new information to previous knowledge, hypothetical thinking, and justifying his thoughts, as well as maximising his apparent motivation to problem-solve (at least in a 1:1 situation)

BH4
benefit from more work in reflective thinking,
increased awareness of when he is not applying his knowledge, or responding impulsively, which he is sometimes inclined to do.

from assessment of X's inferencing ability.
work on more complex sentence structures involving conjunctions and adverbials, but it would be recommended that this be approached from the perspective of problem solving, and that strategies for problem solving, such as planning, making hypotheses and checking, be used.

BH6
It would be recommended that X focus on language structures in intervention – and may benefit from a programme such as 'Colourful semantics' or Shape Coding, wherein she learns the content of various semantic 'slots' in a sentence.
H might also benefit from general cognitive skills such as sequencing, sorting, and attention to detail,
then later noting the tenses of verbs, or arranging events into sequence, and later linking with temporal conjunctions.

BH7

X requires ongoing SLT intervention to address both receptive and expressive language difficulties, as well as social skills and emotional control. He could benefit from activities addressing precision and accuracy, making explicit the need for these, and monitoring of accuracy in himself and others. He may benefit from explicit self regulation and reflection

activities, accompanied by learning of metacognitive vocabulary such as 'remember, plan, check, compare, explain'.

BH9

X requires ongoing SLT intervention to address both receptive and expressive language difficulties, as well as pragmatic, social and indirect language.
 He may benefit from the use of pragmatic tasks such as barrier games that highlight his own needs for information, as well as that of others, with explicit discussion of information formulation, and appropriate, accurate answers to questions. He may also benefit from explicit metalinguistic explanation – separating the concept of word and sentence from the meaning of the words in the sentence themselves.

X is able to generate examples from a model – he cannot however, explain that he has done this, and it is not clear whether he is aware of the process of comparison or matching. It may be useful to explicitly teach this concept, to enable him to identify similarities and differences, and begin to reflect on using this as a strategy for learning and problem solving.

BH11

X requires ongoing SLT intervention to address both receptive and expressive language difficulties, as well as pragmatic, and social language.
 He may benefit from the use of pragmatic tasks such as barrier games that highlight his own needs for information, and processing of incoming information. X was able to use concepts such as the temporal 'before' and the causal 'because', but does not use semantic processing accurately in all tasks. He may benefit from explicit self monitoring activities, accompanied by learning of metacognitive vocabulary such as 'don't know' vs 'can't remember' and to plan, check, match and compare.

CH1

In summary, it would appear that X's typical performance does not reflect the extent of his knowledge. He would benefit from improvement of his cognitive functioning – increasing awareness and control over behaviours such as careful gathering of information, attention to detail, checking, selecting and planning his responses, and reflecting on his performance. He demonstrated responsiveness and understanding of some of these concepts during the DA. These generalisable skills may enable him to perform better in language tasks, and in class.

CP1

X managed the task fairly well, demonstrating good linguistic and metalinguistic knowledge. He was also aware of his own learning and reflective about his own skills, apparently motivated to improve. He interacted well with the tester, and asked for assistance, but not without trying himself first.

X would benefit from ongoing metalinguistic intervention. He understood the content of input and mediation, but took time to be able to use new

material himself, and potentially needs a good deal of help in this. However, he is self reflective, and the fluency with linguistic concepts may give him confidence in his own work.

CP2
X may require ongoing SLT intervention to address both receptive and expressive language difficulties, focusing on concepts such as spatial organization, conditionals (if..then..) and indirect or non-literal language.

 Intervention might make use of explicit rule teaching and application, following of directions, decoding of direct and indirect information, and selection and encoding of expressive language. He may benefit from the use of pragmatic tasks such as barrier games that highlight his own needs for information, as well as that of others, with explicit discussion of intentions and information formulation.

CP3
It would appear that X would benefit from metalinguistic approaches and being encouraged to verbalize about the language himself. He is benefiting from shape coding, and should continue with that, but using the concepts more independently and being required to express the process himself, as well as self monitor his use.

X could also benefit from metacognitive awareness – eg identifying what he knows, and the processes he uses, including metacognitive vocabulary such as remembering. knowing, copying, comparing etc.

D1
X may require ongoing SLT intervention to address both receptive and expressive language difficulties, focusing on structural features of language and morphology, but her inclination to try and transfer and generalize what she has learnt may facilitate learning of systematic rule governed aspects. Intervention might therefore, make use of explicit rule teaching and application.

As X is uncertain about her own judgments of correct grammar, it may be useful to carry out activities requiring judgment as well as justification for her choices, based on rules she has learnt. This could be applied to contexts wider than language, and may facilitate increased confidence.

D2
X, however, has little metacognitive awareness, and a very passive approach to learning. He did not check his responses, and self correct, and intervention might address these skills of self monitoring, reflection and self correction. Although X retains what is taught to him, he does not initiate learning, and does not appear motivated to problem solve. It may be beneficial to address problem solving strategies in a context other than language or school, to try and increase his spontaneous efforts.

D3

X has little structural regularity in his language and virtually no metalinguistic knowledge.

He may benefit from general learning about following rules and recognizing patterns, as applied to language, communication and other behaviours.

He may also benefit from the Colourful Semantics approach that imposes structural regularity and attaches labels to parts of the sentence.

There was some indication of phonological processing difficulty, and auditory phonological awareness training may be recommended as a precursor to literacy.

X has good interpersonal communication skills and may be able to benefit from pragmatic tasks such as barrier games demonstrating communicative effectiveness, to increase accuracy of encoding of messages.

X may require intensive, long term and SLT directed intervention to address both receptive and expressive language difficulties.

D4

X could potentially benefit from increasing the variety of his sentence structures, without increasing sentence length which overloads his memory.

X has some strategies that he employs that are not useful – eg counting words, and other knowledge that he does not make use of consistently, eg knowing that a sentence starts with 'a person'. It may be useful for him to increase cognitive awareness of strategy use. Similarly, he could benefit from reflecting on semantic content of sentences, and new semantic relationships, in order to help him construct ideas.

He recognizes and can judge grammatical correctness, but there is little if any awareness of this process. He is, however, aware of his own difficulty and avoids tasks that are too difficult, with little motivation to persevere with the task.

D5

X has little metacognitive awareness, and did not check her responses, or self correct. Intervention might address these skills of self monitoring, reflection and self correction, as well as skills of explicit explanation and analysis. These tasks might accompany ongoing intervention for language structures, or may be applied to other contexts in order that T practice accounting for the sequence of everyday activities, the planning or preparation of everyday tasks, etc

R1

Little metalinguistic and metacognitive reflection was elicited from X, though it would seem that he understood the concepts, and there was some retention in the summary at the end of the session. Nevertheless, it would be recommended that X be encouraged to consolidate his linguistic knowledge by metalinguistic reflection, perhaps addressing the semantic and

morphological details of items, and he be enabled to adopt a problem-solving approach to language tasks that he would meet in the National Curriculum. The challenge may be to sufficiently motivate and engage X in these tasks.

R2

There was a considerable discrepancy between X's responses to the task items and his response to reflective questions. He struggled to explain the task, could not define a word he said he knew, and could not easily identify the types of words or their role in a sentence, despite his referring back to Colourful semantics'. In addition, he showed very poor knowledge of metacognitive processes, poor awareness of his own knowledge skill, or problem solving. He did not use any terms such as know, remember, practice, was very unsure when making grammatical judgements, and could not justify these judgements or any of his answers.

It would be recommended that intervention with X address these concepts, as the ability to explain and justify his responses may increase X's confidence in producing answers. He could benefit from metalinguistic terminology and explanation, as well as metacognitive vocabulary and explanation. This might begin with aspects he is good at, such as semantic relationships between words, and thematic roles.

R3

X may require ongoing SLT intervention to address both receptive and expressive language difficulties, focusing on structural features of language and morphology, as well as attention to detail and accuracy of structures. Intervention might therefore, make use of explicit rule teaching and application, and explicit tasks addressing precision and accuracy in the gathering of information, following of directions, decoding of information, and selection and encoding of expressive language.

R4

X may require ongoing SLT intervention to address more complex, and longer, recursive receptive and expressive language concepts. Much of the intervention could make use of his excellent self awareness and motivation to problem solve with explicit teaching of rules, and explicit explanations of word meanings and their use. Continued expectations of self –monitoring would be beneficial as X is a reliable judge of correctness, but has a tendency towards impulsive responding.

TA1

It is thought that X may benefit from intervention focused on labelling and explaining language structures, in order that he has tools to reflect on the language, and problem solve. This may help him to explain and justify what he has done, and may enable him to gain some confidence in his ability, and reduce anxiety. He retained some of the concepts discussed in the session, such as 'starting with a person', and switching elements around. He had less grasp of the metacognitive processes presented, and although he acknowledged the need to check, he had little idea how to go about it. X

might, therefore, benefit from some content-free thinking skills training – removing some of the anxiety associated with getting the right answer, and focusing more on the skills of problem solving, and self monitoring. In addition, he might benefit from some identification and labelling of feelings and emotions, although this vocabulary was not investigated.

TA2
X demonstrated some metalinguistic knowledge, and task awareness. His self monitoring was reasonable, he made accurate judgements of grammatical accuracy, and was fairly well aware of his own performance. He may need to devote more attention to the learning of morphological elements, and this may be facilitated by metalinguistic awareness of the structures, ie explicit teaching of which to use, and how, as well as mediation of his own need to be more precise and accurate in his selection of items. He could benefit from greater facility with metalinguistic terms and their use, in order to justify and explain his understanding and use of language, eg. X knew that he had reversed the agents in the sentence, but was unable to explain that he had switched words around, and which words he had used.

X requires ongoing SLT intervention to address his expressive language difficulties, primarily. He could benefit from activities targeting formulation of sentences, and explicit explanation of these, as well as self monitoring activities to increase awareness, and maximise strengths in grammatical judgement.

TF1
Recommendations for intervention for X would include structural organization of language, using categorization of items, colourful semantics and explicit teaching of concepts, much of which is already included in his SLT targets. He may benefit from general learning about following rules and recognizing patterns, as applied to language and communication behaviours, but there is little evidence of readiness for executive or cognitive control over his own learning.

TF4
Although it was not assessed, it was observed that X has some poor pragmatic skills, and he was slow to grasp semantically inappropriate utterances. It may also be useful to investigate X's knowledge of inference.

Although he demonstrated knowledge of metalinguistic concepts, X struggled to formulate explanations, and seemed unsure of some vocabulary for expressing processes. Similarly he was aware of cognitive processes, but unable to explain clearly. It would seem that X may benefit from increasing metalinguistic explanation alongside language learning, ie. providing explicit labelling and description of the tasks he is required to do. Colourful semantics; the introduction of metalinguistic vocabulary such as 'verb' and 'adjective'; and focus on accuracy of word endings, may be useful,

alongside planning and checking strategies. This would enable him to reason out his language learning, and support the development of more complex structures. Verbal and cognitive justification may also give X more confidence in his performance and improve his interpersonal skills.

TF5

X demonstrated some metalinguistic knowledge, and task awareness. Knowledge of more complex linguistic concepts is weak, and may be addressed in ongoing intervention. His self monitoring was reasonable, but he may need to devote more attention to the learning of morphological elements, and this may be facilitated by metalinguistic awareness of the structures, ie explicit teaching of which to use, and how, as well as mediation of his own need to be more precise and accurate in his selection of items.

X requires ongoing SLT intervention to address his receptive and expressive language difficulties. He could benefit from activities targeting formulation of sentences, and explicit explanation of these, as well as self monitoring activities to increase awareness, and accuracy in sentence formulation

APPENDIX XVIII

SUMMARY AND RECOMMENDATIONS FOR PARTICIPANT CH1

Summary:

X was able to arrange most of the sentences independently, and seemed to find the structured task easier than having to formulate expressive language spontaneously. His language is characterized by long and rambling constructions and a lack of precision in getting his meaning across. There is a lack of detail and accuracy, resulting in confusion of some structures eg articles and pronouns.

X attempted to impose order on his responses and explain or justify what he had produced, however these explanations were imprecise and X did not seem to have the vocabulary and concepts to explain himself. His metalinguistic knowledge is implicit, he indicated that he was aware of the manipulations, but unable to express them clearly. There is a need for X to increase his metalinguistic vocabulary alongside syntactic expression to reason linguistically and develop higher level language structures. X has a good semantic understanding and appreciates absurdity and humour, although the structural details are not grasped.

In the first CELF-3 test in October, X's expressive language score was higher than his receptive language, and although this may be an artefact of testing, X's attention to detail and careful gathering of information may be impaired so that he does not fully process incoming language. He has difficulty following instructions and checking that he has planned his responses carefully. His teacher identified several functional difficulties related to attending to and processing verbal information, instructions and rules.

In summary, it would appear that X's typical performance does not reflect the extent of his knowledge. He would benefit from improvement of his cognitive functioning – increasing awareness and control over behaviours such as careful gathering of information, attention to detail, checking, selecting and planning his responses, and reflecting on his performance. He demonstrated responsiveness and understanding of some of these concepts during the DA. These generalisable skills may enable him to perform better in language tasks, and in class.

APPENDIX XIX

SUMMARY AND RECOMMENDATIONS FOR PARTICIPANT R2

Summary:

X completed the linguistic task easily – demonstrating familiarity with sentence construction and formulation, using a variety of grammatical structures. He is accurate in his use of the given words, and performs more successfully arranging the words than when asked to generate his own sentences. He uses exaggerated and slightly inappropriate intonation that was observed on occasions during standardised testing as well.

There was a considerable discrepancy between X's responses to the task items and his response to reflective questions. He struggled to explain the task, could not define a word he said he knew, and could not easily identify the types of words or their role in a sentence, despite his referring back to 'Colourful semantics'. In addition, he showed very poor knowledge of metacognitive processes, poor awareness of his own knowledge skill, or problem solving. He did not use any terms such as know, remember, practice, was very unsure when making grammatical judgements, and could not justify these judgements or any of his answers.

It would be recommended that intervention with X address these concepts, as the ability to explain and justify his responses may increase X's confidence in producing answers. He could benefit from metalinguistic terminology and explanation, as well as metacognitive vocabulary and explanation. This might begin with aspects he is good at, such as semantic relationships between words, and thematic roles.

Appendix XX

REQUIRED MEDIATIONAL INTERVENTION

Distance Level	Examiner	Examinee
0	Produces response via direct imposition on examinee	Passive, conforms to pressure of examiner to reproduce model
1	Models act to be copied, encourages imitation, withdraws as examinee starts to respond	Initiates partially successful representation of model
2	Points out specific examples of rules, concepts, attributes of the problem, identifies constant and changing elements	Spontaneously responds to task, attends to mediation
3	Identifies general class characteristics	Encouraged to apply response to new situation
4	Refers to previously identified strategies	Acts on previous mediation, applies and repeats, no rules formulated
5	Selects/encourages strategies based on insight and rules	Chooses adequate strategies based on derived insight
6	Point out previously used strategies using transcending verbal and metalinguistic rules	Applies previously used strategies, reflects awareness of rules and operations
7	Focuses examinee attention on problem anticipatory, and pre-response mediation, to provide initial regulation of response	Formulates specific rules, strategies, attitudes, meanings. Self regulatory
8	Alerts to metacognitive elements, directs mediation to structural change, challenges for resistance	Elements of structural change present
9	Passive presence in elicitation of responses	Mediation is internalized, self regulation

Ref: Feuerstein R., Feuerstein R.S., Falik, L. and Rand Y. 2002 The Dynamic Assessment of Cognitive Modifiability. ICELP Press.