Dynamic assessment of word learning skills of preschool children with primary language impairment

BERNARD CAMILLERI[1], and JAMES LAW[2]

1. City University London, England

2. Newcastle University, Newcastle, England

Correspondence: Bernard Camilleri, Division of Language and Communication Science, City University, London EC1V 0HB, England. E-mail: Bernard.camilleri.1@city.ac.uk.

Key words: dynamic, assessment, child, speech-language pathology

Running head: Dynamic assessment and word learning

Abstract

Dynamic assessment has been shown to have considerable theoretical and clinical significance in the assessment of socially disadvantaged and culturally and linguistically diverse children. In this study it is used to enhance assessment of preschool children with primary language impairment.

The purpose of the study was to determine whether a dynamic assessment (DA) has the potential to enhance the predictive capacity of a static measure of receptive vocabulary in preschool children.

Forty preschool children were assessed using the static British Picture Vocabulary Scale (BPVS), a DA of word learning potential and an assessment of non-verbal cognitive ability. Thirty seven children were followed up six months later and re-assessed using the BPV**S.** Although the predictive capacity of the static measure was found to be substantial, the DA increased this significantly especially for children with static scores below the 25th centile. The DA of children's word learning has the potential to add value to the static assessment of the child with low language skills, to predict subsequent receptive vocabulary skills and to increase the chance of correctly identifying children in need of ongoing support.

Introduction

One of the greatest outstanding challenges in paediatric speech-language pathology is the difficulty in predicting patterns of emerging language development in the preschool years. There has been considerable discussion over the reasons for this (Desmarais, Sylvestre, Myer, Bairati, & Rouleau, 2008; Desmarais, Sylvestre, Miller, Bairati, & Rouleau, 2008; Reilly et al., 2008) but the fact remains that although many children with speech and language difficulties at school entry are likely to continue to experience difficulties into adolescence and beyond (Tomblin, 2008), a substantial proportion of those who are considered to be "early language delayed" (Dale, Price, Bishop, & Plomin, 2003, p. 544) at the age of two to three, appear to have resolved any language difficulties by the time they start formal schooling aged five.

Central to this debate is the nature of the assessment process. Typically, when concerns about an individual's language development are raised, standardised, norm-referenced tests are used to compare the individual child's performance to that of his/her peers in order to determine whether level of functioning is appropriate for his/her age. It is important to obtain systematic measures of a child's strengths and weaknesses on different domains of language. However, in the preschool years, the severity of the language delay (i.e. the extent to which a child deviates from the expected norms/patterns on standardised assessments) is actually not a sufficient indicator of whether the delay constitutes a clinically significant, potentially persistent language impairment (Rutter, 2008). It is within this context that dynamic assessment (DA) generally, and dynamic assessment of receptive vocabulary in particular are being considered. In addition to the need for assessing different domains of language (e.g., receptive vocabulary, morpho-syntactic ability, phonological skills), it is also important to carry out assessments which reveal additional levels of information regarding a given domain.

Dynamic assessment (DA) may constitute a useful approach, providing insight on the child's response to interaction/intervention and potential to learn. The defining feature of dynamic assessment is that "instruction and feedback are built into the testing process and are differentiated on the basis of an individual's performance" (Elliott, 2003, pp. 16-17).

Both static (norm-referenced) and dynamic assessments have been developed in the field of psychology to measure cognitive functioning or "intelligence". Dynamic assessment was synonymous with cognitive assessment in the former Soviet Union, where DA was first developed under the influence of Vygotsky's sociocultural theory. Outside of the former Soviet Union, extensive interest in dynamic assessment began in the 1960s during a time of increased criticism of static testing which also coincided with the publication for the first time of Vygotsky's works in English (Sternberg & Grigorenko, 2002).

One of the identified problems was that "whatever their predictive validity in a group or correlational sense (which might be different with respect to ethnic group membership), standardised and normative tests have less success when applied to prediction in individual cases." (Haywood et al., 1992, p. 40). This may be especially true when dealing with children from minority groups or diverse sociocultural backgrounds, and the problem is not resolved by making sure that the minority groups are proportionally represented in normative samples in which it is assumed that it is possible to capture the characteristics of the "average" child. Such practices are likely to lead to the misclassification of some subgroups (Haywood, Tzuriel, & Vaught, 1992). Brown and Ferrara (1985) have argued that static assessment measures are particularly poor, whether used as diagnostic tools or as predictors of later outcome, when used with younger, preschool-age children. This is known to also be true of standardised assessments of language with young children (Rutter, 2008).

Practitioners and researchers have long realized the need for assessments which could be used with populations form diverse linguistic, cultural, social and educational backgrounds with whom the use of static assessments was highly problematic (Feuerstein, 2003; Sternberg & Grigorenko, 2002). They have also appreciated the value of estimating individuals' learning potential as opposed to the product of what had been learned, irrespective of linguistic, cultural, social or educational background (Brown & Ferrara, 1985). If a child's learning potential or ability to benefit from an instructional interaction is the construct to be considered, assessment tools which allow the possibility of examining the effects of deliberately induced change are required (Haywood & Wingenfeld, 1992).

Although DA was originally developed to extend cognitive assessments it has started to be used in the field of language development and specifically language impairment over the past two decades (Law & Camilleri, 2007). For example, one of the earliest studies in this field used DA to predict the potential for change in language production of preschool children with specific expressive language impairment (SELI) (Bain & Olswang, 1995; Olswang & Bain, 1996). The authors reported a consistent relationship between dynamic scores and children's progress and concluded that DA captures children's potential for immediate change based on their responsiveness to their DA procedure. They also considered that their weighted DA score was closely associated with the discrepancy between the child's receptive and expressive language scores thereby highlighting the potential role played by receptive language skills, especially those associated with vocabulary development which are amongst the most sensitive to environmental modification (Harris, 1992).

DA has also been employed to distinguish between “language difficulties” and “language differences” in bilingual populations using a “test-teach-retest” approach (Fagundes, Haynes, Haak, & Moran, 1998; Gutierrez-Clellen & Peña, 2001; Lidz & Peña, 1996; Peña, Quinn, & Iglesias, 1992; Peña & Quinn, 1997; Peña & Gillam, 2000; Peña, Iglesias, & Lidz, 2001; Peña et al., 2006; Peña, Resendiz, & Gillam, 2007). Children from diverse cultural and linguistic backgrounds who achieved low scores on a static test of a language area (e.g., narrative) were grouped at the outset into one of two categories: children with language difficulties (or low language ability) and children without possible language impairment (or typically developing children/language difference). Peña et al. (2001) suggest that DA has considerable clinical applications for reducing both over-referral if depressed static performance does not reflect true learning potential (Peña & Quinn, 1997) and under-referral when real language difficulties are explained away as language differences (Gutierrez-Clellen & Peña, 2001).

To date, DA has tended to be focused on the child’s expressive language skills. Given the significance of receptive language in predicting longer term outcomes (Olswang, Rodriguez, & Timler, 1998) and the problems that have been noted about intervention in this area (Law, Garrett, & Nye, 2003). This is an important omission. It is clear that children’s lexical levels are susceptible to environmental influences, and specifically that children with language impairments have difficulty establishing lexical representations (Rice, Buhr, & Nemeth, 1990; Rice, Buhr, & Oetting, 1992; Rice, Oetting, Marquis, Bode, & Pae, 1994). This makes the process of learning to both understand and express new words an ideal candidate for a DA procedure. “Fast mapping" is the term used to refer to the ability to establish an initial link between the word and its referent, based on limited exposure (Carey, 1978). Research on fast mapping has shown that, when typically developing preschool children encounter a new word, they are quickly able to draw upon linguistic and contextual information to develop an initial, if incomplete, understanding of the word (Carey, 1978; Dollaghan, 1985; Heibeck & Markman, 1987). Through their fast-mapping abilities children are able to establish a large number of lexical entries in a relatively short time (Clark, 1993). A DA of children's word learning ability could therefore investigate the child's "fast mapping" abilities (Dollaghan, 1985; Dollaghan, 1987; Rice et al., 1990) within an interactive situation.

In a recent study by the present authors, a DA procedure for word learning for use with preschool children with language impairments was developed. Children were provided with a hierarchy of cues towards initially matching word and referent (picture card) and subsequently given opportunities to demonstrate their retention of the word-referent match for both receptive and expressive purposes (Camilleri & Law, 2007). Although referred monolingual English-speaking children were found to achieve significantly higher scores on the static British Picture Vocabulary Scales (BPVS) (Dunn, Dunn, Whetton, & Burley, 1997) than children with English as an additional language (EAL), their DA scores were comparable. Of particular significance was the fact that dynamic scores were highly correlated with BPVS scores when the whole group of referred children was considered, while the same was not true for the group of children with the lower static scores. Among these lower scoring children, the variability in DA scores could not be predicted by their BPVS scores. The obvious extension of this analysis is to examine the difference between the lower scoring children and the referred group as a whole, on static and dynamic assessments across time. This is the focus of the present paper. The children identified in Camilleri and Law's (2007) study were followed up six months after the initial assessment with a view to addressing the following questions.

1) To what extent do the scores on the static and dynamic assessment of vocabulary predict vocabulary performance across time?
2) To what extent does the pattern of prediction differ for low scoring children, as opposed to referred children generally?

Method

Participants

The forty children involved in the study were randomly selected from among children referred to an inner-city speech-language pathology (SLP) service. Referrals were obtained from a range of sources, including general practitioners, health visitors, paediatricians and parents. At the time of referral, the SLP service completed a form indicating the primary cause for concern and whether any additional known difficulties were present. Children were selected among those for whom speech and/or language was indicated as the primary cause for concern. Following convention those with language difficulties secondary to other conditions, such as hearing impairment, cerebral palsy and autism were excluded. Children's non-verbal cognitive ability was not used as a criterion for inclusion/exclusion. Children from a diverse cultural/linguistic background were included, but English was required to be one of the home languages. The specific languages spoken by the children were also indicated on the referral forms.

The children were aged between 41 and 60 months when they were first assessed using a combination of static and dynamic measures (see below). Following initial assessment, the group of 40 children was subdivided into two groups of higher and lower scoring children. The 25th percentile on the BPVS (administered as the pre-test) was adopted as a cut-off to distinguish between these two groups. Different cut-offs have been adopted when distinguishing between typically developing children and those with a language delay. This liberal cut-off was adopted here to separate children for whom receptive vocabulary was clearly within the normal range (14 higher scoring children) from a group of lower scoring (LS) children for whom vocabulary was a possible cause for concern (26 children). The children had no prior experience of speech and language therapy services.

The children were followed up after a six month waiting period and re-assessed using the static BPVS. During this time the children were on a waiting list, which meant that they did not receive speech and language therapy services as reported by parents at the time of re-assessment. This meant that any changes in language scores on the BPVS were not a result of specific language interventions.

Ethical approval

Ethical approval was obtained in November 2004 from the local NHS Ethics Committee, prior to the start of data collection.

Procedure

The procedure carried out at time 1 consisted of a dynamic assessment consisting of a pre-test phase, a dynamic or interactive phase and a post-test phase, carried out in a single session lasting a maximum of forty-five minutes. The “block building” and “picture similarities” sub-tests of the British Ability Scales (Elliott, 1996) were also included as measures of non-verbal ability. While non-verbal ability was not the main focus of the current research, measures of non-verbal ability were included to explore the possible relationship between this variable and children’s language outcomes at follow-up. The “block building” test was carried out as an ice-breaker which did not require any language from the child. This was followed by the BPVS. This constituted the pre-test phase and provided a baseline measure of the child's receptive vocabulary. It also allowed the assessor to identify words which the child was able to identify correctly and others which the child was unable to identify in the BPVS.

The static BPVS was followed by the interactive phase, during which six of the vocabulary items (three nouns and three verbs) that the child was unable to identify correctly were targeted. The three nouns were targeted first, followed by the three verbs. The interactive phase was carried out in the form of a game which involved posting picture cards into a letter box. The focus of the interactive phase was to establish how easily the children could identify the targeted words by using graduated levels of assistance (leading to a dynamic score). Additionally the interactive phase involved establishing how well the children established a representation for the words for both receptive and expressive purposes, as measured by a first "Recall" task and two Expressive tasks (see below). The focus of the post-test was to establish whether the children had retained the targeted items for receptive purposes, as measured by a second Recall task (see below). Each of the tasks and phases will now be described (further details can be found in the paper by Camilleri and Law, 2007).

Interactive phase: Graduated assistance/support

The following description will illustrate how the three nouns were targeted. Verbs were subsequently targeted in a similar fashion. For each of the six targeted vocabulary items (three nouns and three verbs), the child was presented with three cards, one of which was a targeted item. The other two pictures consisted of 'easier' pictures of the same grammatical category, which the child had identified correctly during the BPVS. Children were given three levels of assistance (from least to most assistive), towards identifying the target item from among the distractors.

Independent Identification

Children were initially asked to find the target and given the possibility of adopting independent problem solving skills in order to match the word with the target picture, without specifically focussing the children's attention towards the appropriate word-picture match. This constituted the first level of (minimal) support and the child's response is referred to as "independent identification" (Camilleri & Law, 2007, p. 316).

Implicit Identification

If the child pointed to the wrong picture, the assessor gave the child feedback that he/she had not found the 'hard' word. The assessor then asked the child to find the two easier distractors in turn, before asking the child for a second time to find the target item. The word-referent match is implicit at this stage, although the child must still attend to the auditory information and visually scan the pictures before choosing the correct card. This constituted the second level of assistance and the child's response constituted 'implicit identification'.

Explicit Identification

If the child still failed to identify the correct card for the target word, the assessor provided a third and final level of support, making the link between the word and the referent explicit. On this occasion, when the child correctly identified each of the two distractors they were turned face down, thus effectively eliminating the cards as a potential response. This final level of support leads to a response which may be referred to as ‘explicit identification’.

During this phase, children were assigned scores by using a weighted scoring system, similar to that adopted by Bain and Olswang (1995). This was based on the amount of assistance they required during the interactive phase of the DA. The first level of assistance was assigned three points, the second level was assigned two points and the third level was assigned one. A child carrying out ‘independent identification’ was assigned the three points associated with the first level as well as another three points (2+1) associated with the lower levels. A child carrying out ‘implicit identification’ was assigned the two points assigned to this level, plus one additional point. A child carrying out ‘explicit identification’ was given one point. The total number of points for the six vocabulary items was added up to derive the child’s dynamic weighted score (DWS). Each child could obtain a range of scores from six (one for each word) to 36 (six for each word).

Interactive phase: Expressive tasks and (first) Recall task

Once the child successfully identified the three cards on the table (one target and two distractors), the child was prompted to post them into a letter box. The child was asked which card he/she wanted to post until all three cards were posted. This gave the assessor an opportunity to find out whether the child was able to use the target noun expressively. A correct response at this stage would contribute towards the child’s score on the First Expressive task (Expressive 1) within the interactive phase. The process described above was repeated with the other two targeted nouns. Once all three targeted nouns were identified with one or other degree of support, the child was assessed for comprehension of these three newly established items. The three targeted picture cards were presented, and the child asked to point to each named card. This part of the interactive phase enabled the assessor to establish whether the child had retained the words for receptive use, very shortly after having been exposed to the word-referent match. The child's responses contributed (together with their subsequent responses for verbs) to their score on the Immediate Recall task (Recall 1). The child was subsequently given a second opportunity to demonstrate that they could say the three targeted nouns. This was again carried out in the form of a posting game, where the child was asked to name the item he/she wished to post first, second and third. This constituted the final measure of children's expressive responses with the first grammatical category (nouns) and contributed to their score on the Second Expressive task (Expressive 2). Once the process was completed with nouns, it was repeated with verbs. The total number of words produced/identified correctly across grammatical categories (i.e. nouns and verbs) constituted their Expressive 1, Recall 1 and Expressive 2 scores respectively. Given that six words were targeted, children could achieve a score from zero to six for Expressive 1, Expressive 2 and Recall 1.

Picture Similarities Task, followed by Post-Test

The interactive phase was followed by the Picture Similarities sub-test of the BAS. This second sub-test allowed for a time-lag between the interactive phase and the post-test, in addition to providing information on the child's non-verbal cognitive ability.

The final phase of the dynamic assessment consisted of the receptive vocabulary post-test. The focus of this post-test was to establish how many of the targeted words had been retained by the child (for receptive purposes), several minutes after they had been exposed to them in the Interactive phase. All six targeted items were presented to the child and he/she was asked to point to each item named randomly by the assessor. The child was not given feedback after each response but was given feedback after they had attempted to point to all six named items. Children achieved a score from zero to six on this receptive post-test. This second receptive task will be referred to as “Recall 2”.

Analysis

Once the children were re-assessed with the BPVS (at time 2) it was possible to combine data from the initial assessment with the follow-up scores. This allowed for an analysis of the change in BPVS scores from time 1 to time 2, as well as of the correlation between the various measures taken at time 1 and the BPVS scores at time 2.

Finally, in order to determine which of the measures derived at time 1 were predictive of subsequent vocabulary outcomes, multiple regression analysis was adopted (see Table 4). This was carried out first with the whole group of children who were followed up, in order to evaluate whether the measures at time 1 could be used with children achieving static BPVS scores across the full range. Subsequently the analysis was repeated with the sub-group of low scoring children, the group for whom further assessment information is most likely to be required. The dependent or outcome variable was the BPVS score at Time 2. The first predictor variable which was entered in both multiple regression analyses was the BPVS score at Time 1. This variable was chosen as scores on a given standardised assessment might reasonably be expected to correlate highly with subsequent or previous scores on the same measure, although the authors of the BPVS themselves point out that “although the original

BPVS has been used extensively in Britain, original reports showing direct statistical evidence for its concurrent and predictive validity are few" (Dunn et al., 1997, p. 35). The second variable to be entered hierarchically into the multiple regression consisted of BAS scores, representing children's non-verbal cognitive abilities. Correlations between BAS and BPVS scores at time 1 were significant although they were lower with children achieving scores below the 25th percentile on the BPVS (0.43) than for the entire group of referred children (0.50). Inclusion at the second step of the multiple regression allowed the possibility to investigate the extent to which non-verbal ability is related to later vocabulary outcome.

Finally, the different measures obtained from the dynamic component of the assessment were entered into the multiple regression analysis in a forward stepwise manner, as appropriate for exploratory research (Wright, 1997).

Results

The distribution of the scores for the 40 referred children on both the static and dynamic assessments carried out at the point of referral (time 1) are provided in table 1. As expected, their percentile scores on the BPVS were, on average considerably lower than on the BAS.

Table 1 about here

Thirty-seven of the 40 children were seen at time 2 and re-assessed with the static BPVS. Of these, 23 were in the low scoring (LS) group of children who had scored below the 25th percentile on the BPVS at time 1. On average, children achieved higher percentile scores on the BPVS at time 2 than at time 1, with this difference achieving statistical significance (see Table 2).

Table 2 about here

In spite of the average increases in scores between Time 1 and Time 2, BPVS percentile scores were found to be stable across time for some children – five children achieved the same percentile score at both Time 1 and Time 2. These were children achieving consistently low scores. However, as many as 24 children achieved higher scores at Time 2, while eight children achieved lower percentile scores. On average, the increase in scores was slightly greater for the group as a whole (9 percentile points) than for the LS sub-group (7 percentile points).

As Table 3 shows, the correlation between BPVS scores at Time 1 and at Time 2 for this group was greater (.90) than for the LS sub-group (.56). While both of the correlations in question were highly significant, the figures meant that BPVS scores at Time 1 accounted for a considerably greater proportion of the variance for the group as a whole than for the LS sub-group.

Table 3 about here.

The assessment procedure undertaken at Time 1 provided the static BPVS and BAS scores, as well as the dynamic measures of dynamic weighted score (DWS), two expressive vocabulary measures (Expressive 1 and Expressive 2) and two receptive vocabulary measures (Recall 1 and Recall 2). The dynamic measures were designed to reflect different aspects of children's processing abilities in establishing, retaining and using representations for new word-referent matches. Each of these dynamic measures was found to correlate significantly with BPVS scores at time 2 when the whole group of referred children was considered (see Table 3). However, when the low scoring subgroup was considered, only one dynamic score – that derived from the second expressive task was found to be significantly correlated with the BPVS score.

Table 4 about here

When the multiple regression analyses were carried out, the static BPVS and BAS scores were the first two variables entered. When the measures derived from the dynamic assessment were entered in a forward stepwise manner, only one additional variable met the statistical criteria for inclusion in the model - both with the whole group of children and with regards to the subgroup of LS children. This was the score on the second round of the expressive task (Expressive 2).

As was to be expected, given the high correlation between the two measures on the BPVS, the assessment's scores at Time 1 were highly predictive of scores at Time 2, throughout the three steps of the multiple regression analysis with the whole group of referred children. When the BPVS score at Time 1 was the only variable considered, it accounted for 81% of variance in BPVS scores at Time 2. The question of interest was whether other predictor variables could account for a significant proportion of the remaining variance. When both BAS scores and scores on the second expressive task were included in the model, a further 5% of variance in BPVS scores at Time 2 was accounted for (see Table 4). Although both variables (BAS and Expressive 2) were separately found to be highly correlated with the outcome variable (BPVS Time 2), the BAS variable could only account for an additional 2% and the Expressive 2 an additional 3%. This was partly because the BPVS score at Time 1 had accounted for such a large percentage of variance at Time 2, and partly because the other predictors were also quite highly correlated with the first predictor variable – BPVS at Time 1.

As can be seen in Table 5, the situation was reversed with regards to the sub-group of LS children. At step one of the multiple regression, BPVS scores at Time 1 accounted for a relatively low 32% of variance in BPVS scores at Time 2. In addition, the measures achieved through the DA procedure were not highly correlated with the first predictor variable. BAS scores, on the other hand were highly correlated with BPVS scores at Time 1, though not to the same extent as for the group as a whole.

Table 5 about here

BAS scores were added to the model at the second step. This variable was found to account for an additional 6% of the variance in BPVS scores at Time 2. In spite of accounting for a higher percentage of variance with LS children (than with the group as a whole), this contribution did not achieve statistical significance. Children's performance on the second round of the expressive task accounted for a further 20% of the variance in BPVS scores at Time 2.

The majority of referred children who scored below the 25^{th} percentile on the BPVS at Time 1 achieved scores at Time 2 which were very close to those predicted on the basis of their original BPVS scores, their performance on the second expressive task and, to a lesser extent, their BAS scores (see Table 6).

Table 6 about here

Discussion

Both static and dynamic assessment scores have utility in predicting performance over time. It might be argued that the fact that the static assessment is quicker to use gives it the edge in terms of its applicability for clinicians and researchers. Closer examination of the data suggests that the dynamic element enhances the value of the static assessment in terms of the level of information that it provides and we would argue that this type of information is critical in the context of assessing the performance of the individual child. DA is therefore particularly valid for the practitioner who wants to know how to read a child's performance and discuss their potential with teachers, family etc. The risk of overreliance on statistical criteria from individual standardised assessment is that it can make interpretation across time difficult. Children move in and out of clinical categories based on those criteria as has been clearly demonstrated in large-scale longitudinal studies (Silva, McGee, & Williams, 1983). The fact that the impact of the dynamic assessment was found to be more relevant at the lower end of the distribution makes this especially relevant to the practitioner who, after all, by definition is focussing on this group.

Among low-scoring referred children, BPVS scores at Time 1 could account for 32% of variance in scores at Time 2. While 32% is a considerable proportion of the variance, it is lower than one might expect when dealing with repeated measures of the same assessment. It is clear that for LS children in particular, static scores on the BPVS taken at a particular point in time are relatively poor predictors of subsequent outcomes on the same measure. It would therefore be inappropriate to define children's language status, in this case regarding their receptive vocabulary abilities, solely in terms of their static performance on a standardised assessment. While the weighted dynamic score was not found to be predictive of children's subsequent receptive vocabulary, the crucial finding was that at least one of the other scores derived from this DA of receptive vocabulary could be usefully adopted together with the static assessment in defining children's abilities in terms of both current performance as well as ability to establish and retain new word-referent representations. The fact that scores on Expressive 2 were predictive of subsequent outcome, whereas the other dynamic measures were not, can be interpreted in terms of the requirements of the tasks. In order for a child to carry out the expressive tasks, he/she would have had to not only match the word and referent but, through the learning opportunities, have established and retained in memory a sufficiently strong mental representation of the word to be able to produce the word expressively. By comparison, the receptive tasks could have been carried out successfully without having established this strong mental representation. A partial representation involving parts of the phonological sequence and an ability to eliminate wrong responses and/or guess the right response may have led to a child identifying several items correctly, with or without a strong underlying lexical representation. This is consistent with findings from research on fast mapping which has suggested that children with language impairments are able to identify a novel word within a stream of words and establish an initial link between word and referent (Dollaghan, 1985, Rice, Buhr & Nemeth, 1990). However, establishing and retaining the phonetic information, particularly for subsequent expressive use, was found to be the most vulnerable aspect of the fast mapping process (Dollaghan, 1985).

The dynamic weighted score provided a measure of children's ability to problem solve in matching an unfamiliar word with a new referent, but did not reflect ability to establish and retain the underlying representations for the new word-referent matches. The findings of this research confirm the view that an ability to match an unfamiliar word with a new referent is a necessary skill for word learning, but is not sufficient in itself.

Children's performance on the expressive tasks, on the other hand, provided a reflection of their ability to establish all the components of the underlying lexical representation, retain them and access them for expressive purposes. The second expressive task proved to be the more valid predictor of the two (expressive tasks), perhaps due to the fact that the learning opportunities accrued by the second round were necessary for the lexical representation to be established sufficiently.

The model which emerged from the multiple regression analysis suggests that Expressive 2 scores and, to a lesser degree, scores on non-verbal components of the BAS may be usefully adopted, together with BPVS scores at Time 1, to predict vocabulary outcomes six months later, as measured by the BPVS. Specifically, each correctly named item, with a maximum of six, (during the second expressive task) contributed nearly 3.5 percentile points to children's predicted percentile score at time 2 (cf. Table 5). This model was derived from a group of children within a relatively wide age range (41 to 60 months at time 1) and included children from different language backgrounds. Post-hoc analyses were carried out to investigate whether age or language background was related to children's progress over the duration of the study. Neither variable was found to be related to children's progress, whether considered independently or in combination with children's BPVS, BAS or expressive scores. This confirms that the DA can be used effectively with low scoring children across the age range, irrespective of whether English was the only language or an additional one.

Study limitations

Given their initial low scores, one possible explanation for the change in BPVS scores with the LS group was that this constituted regression to the mean (Zhang and Tomblin, 2003). However, the increase in BPVS scores occurred across both the LS subgroup and the group of referred children as a whole, even with children whose initial scores were not particularly low. In spite of the fact that the recommended six month time lag between first and second presentations had been observed, the children's performance at Time 2 may have been influenced by their experience at time 1, both of the BPVS itself and of the dynamic component of the assessment. This may be considered an undesirable influence by those wishing to use BPVS scores to make purely norm-based judgements, but is less of a cause for concern insofar as children's differential responses at Time 1 were being explored in terms of their predictive value of different outcomes at Time 2. The extent to which the activities incorporated in the dynamic component of the assessment have a direct influence on children's growth in receptive vocabulary would need to be explored by carrying out a separate controlled intervention study.

Implications for practice

Our findings would suggest that, although the age of the children involved in this study was rather lower than that conventionally engaged in the majority of DA research, this approach may be especially useful in assessing language at this age which is of critical importance to parents, early educators and speech and language pathologists. DA has the potential to add clarity to the decision making of speech and language pathologists.

Although DA has been described in this paper as a type of assessment, one could equally well describe it as a kind of brief intervention which has the effect of contributing to a classification or diagnosis. Indeed it could usefully be seen as an element of the type of diagnostic therapeutics advocated in the development of clinical decision making (Hoben, Varley, & Cox, 2010). Equally, it is important to note that children's responsiveness to the sort of dyadic interactions involved in dynamic assessment can be mediated by a range of other factors, such as executive functioning, joint attention and memory. This needs to be taken into account when interpreting individual children's responses within a dynamic assessment.

Specifically in the context of the dynamic assessment of word learning, the potential role played by the monitoring of generalisation across words is of particular interest. It is important to point out that children who achieved higher scores on the BPVS on the six-month follow-up identified several new items beyond the ones directly targeted during the dynamic component of the assessment at Time 1. This can be interpreted in terms of a genuine increase in receptive vocabulary and/or an increased ability in carrying out the task involving matching a word with a pictorial referent under assessment conditions. It cannot be interpreted as a reflection of the child merely retaining the vocabulary items targeted during the DA. In fact, several children identified new items which had not been targeted, while failing to identify some of the targeted items. Our sample was too small to explore this further but it is clear that this capacity to extend even such a brief learning experience may be of considerable clinical value. The clinical value of the very procedure used in this study and described in the initial paper (Camilleri & Law, 2007) is being explored in different settings, with different age groups and in different countries. A revised version of the DA used in this study has in fact been included in a battery of dynamic assessments aimed at assessing preschool children with English as an additional language (Hasson, Camilleri, Jones, Smith & Dodd, 2013). Yet another version, which incorporates a conversational approach to exposing vocabulary to children, has being explored for its potential to uncover greater qualitative information about children's word-learning skills in naturalistic contexts (Camilleri & Botting, 2013).

Conclusions

Dynamic assessment, originally developed to tap into a broad range of cognitive skills, clearly has the potential to enhance the assessment of both expressive and receptive language skills. Indeed it could be argued that the assessment of language skills in the early years may be particularly prone to both developmental "noise" and to contextual factors and thus, we would argue, should be seen as a useful test case for dynamic assessment. The emphasis in this paper on receptive language and word learning skills is important because these skills underpin all other aspects of language learning and can be seen as the precursors of the skills which play such a large part in the way that the child integrates into school and starts to read. Our finding that DA would appear to be especially salient for the child with low language skills has important implications for the way that assessments are used with young children. While static assessments are clearly important they are only part of the story and it is these children's capacity to react to a controlled learning environment which is critical to their subsequent performance.

Reference List

Bain, B. A. & 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,* 81-91.

Brown, A. L. & Ferrara, R. A. (1985). Diagnosing zones of proximal development. In J.W.Wertsch (Ed.), *Culture, communication and cognition: Vygotskian perspectives* (pp. 273-305). Cambridge: Cambridge University Press.

Camilleri, B. & Botting, N. (2013), Beyond static assessment of children's receptive vocabulary: The dynamic assessment of word learning (DAWL). *International Journal of Language and Communication Disorders.*

Camilleri, B. & 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,* 312-322.

Carey, S. (1978). The child as a word learner. In M.Halle, G. Miller, & J. Bresnan (Eds.), *Linguistic theory and psychological reality* (pp. 264-293). Cambridge, MA: MIT Press.

Clark, E. V. (1993). *The lexicon in acquisition*. Cambridge: Cambridge University Press.

Dale, P. S., Price, T. S., Bishop, D. V. M., & Plomin, R. (2003). Outcomes of early language delay: 1. Predicting persistent and transient delay at 3 and 4 years. *Journal of Speech, Language, and Hearing Research, 46,* 544-560.

Desmarais, C., Sylvestre, A., Miller, F., Bairati, I., & Rouleau, N. (2008). Reply to S. Reilly and colleagues. *International Journal of Language and Communication Disorders, 43,* 476-477.

Desmarais, C., Sylvestre, A., Myer, F., Bairati, I., & Rouleau, N. (2008). Systematic review of the literature on characteristics of late-talking toddlers. *International Journal of Language and Communication Disorders, 42,* 361-389.

Dollaghan, C. (1985). Child meets word: "Fast mapping" in preschool children. *Journal of Speech and Hearing Research, 28,* 449-454.

Dollaghan, C. (1987). Fast mapping in normal and language impaired children. *Journal of Speech and Hearing Disorders, 52,* 218-222.

Dunn, L., Dunn, L., Whetton, C., & Burley, J. (1997). *The British Picture Vocabulary Scale*. Windsor: NFER-Nelson.

Elliott, C. D. (1996). *British Ability Scales II*. Windsor: NFER-Nelson.

Fagundes, D. D., Haynes, W. O., Haak, N. J., & Moran, M. J. (1998). Task variability effects on the language test performance of Southern lower socioeconomic class African American and Caucasian five-year-olds. *Language, Speech, and Hearing Services in Schools, 29,* 148-157.

Feuerstein, R. (2003). The theory of structural cognitive modifiability and mediated learning experience. In *Feuerstein's theory and applied systems* (pp. 17-49). Jerusalem: The International Centre for the Enhancement of Learning Potential.

Gutierrez-Clellen, V. F. & Peña, E. (2001). Dynamic assessment of diverse children: A tutorial. *Language, Speech, and Hearing Services in Schools, 32,* 212-224.

Harris, M. (1992). *Language experience and early language development: From input to uptake*. Hove, East Sussex: Psychology Press.

Hasson, N., Camilleri, B., Jones, C., Smith, J., & Dodd, B. (2013), Discriminating disorder from difference using dynamic assessment with bilingual children. *Child Language Teaching and Therapy*, *29*, 57-75.

Haywood, C. H., Tzuriel, D., & Vaught, S. (1992). Psychoeducational assessment from a transactional perspective. In C.H.Haywood & D. Tzuriel (Eds.), *Interactive assessment* (pp. 38-63). New York: Springer-Verlag.

Heibeck, T. H. & Markman, E. M. (1987). Word learning in children: An examination of fast mapping. *Child Development, 58,* 1021-1034.

Hoben, K., Varley, R., & Cox, R. (2010). Clinical reasoning of speech and language therapy students. *International Journal of Language and Communication Disorders, 42,* 123-135.

Law, J. & Camilleri, B. (2007). Dynamic assessment and its applications to children with speech and language learning difficulties. *International Journal of Speech-Language Pathology, 9,* 271-272.

Law, J., Garrett, Z., & Nye, C. (2003). Speech and language therapy interventions for children with primary speech and language delay or disorder. *Cochrane Database of Systematic Reviews, 3*.

Lidz, C. S. & Peña, E. (1996). Dynamic assessment: The model, its relevance as a nonbiased approach, and its application to Latino American preschool children. *Language, Speech, and Hearing Services in Schools, 27,* 367-372.

Olswang, L. B. & Bain, B. A. (1996). Assessment information for predicting upcoming change in language production. *Journal of Speech and Hearing Research, 39,* 414-423.

Olswang, L. B., Rodriguez, B., & Timler, G. (1998). Recommending intervention for toddlers with specific language learning difficulties: We may not have all the answers, but we know a lot. *American Journal of Speech-Language Pathology, 7,* 23-32.

Peña, E. & Gillam, R. B. (2000). Dynamic assessment of children referred for speech and language evaluations. In C. S. Lidz (Ed.), *Dynamic assessment: Prevailing models and applications* (pp. 543-575). New York, Elsevier Science.

Peña, E., Gillam, R. B., Malek, M., Ruiz-Felter, R., Resendiz, M., Fiestas, C. et al. (2006). Dynamic assessment of school-age children's narrative ability: An experimental investigation of classification accuracy. *Journal of Speech, Language, and Hearing Research, 49,* 1037-1057.

Peña, E., Iglesias, A., & Lidz, C. S. (2001). Reducing test bias through dynamic assessment of children's word learning ability. *American Journal of Speech-Language Pathology, 10,* 138-154.

Peña, E. & Quinn, R. (1997). Task familiarity: Effects on the test performance of Puerto Rican and African American children. *Language, Speech, and Hearing Services in Schools, 28,* 323-332.

Peña, E., Quinn, R., & Iglesias, A. (1992). The application of dynamic methods to language assessment: A nonbiased procedure. *The Journal of Special Education, 26,* 269-280.

Peña, E., Resendiz, M., & Gillam, R. B. (2007). The role of clinical judgement of modifiability in the diagnosis of language impairment. *International Journal of Speech-Language Pathology, 9,* 332-345.

Reilly, S., Wake, M., Bavin, E., Eadie, P., Bretherton, P., & Prior, M. R. (2008). Letter regarding 'A systematic review of the literature on characteristics of late-talking toddlers, by Desmarais et al. *International Journal of Language and Communication Disorders, 43,* 473-475.

Rice, M. L., Buhr, J. C., & Nemeth, M. (1990). Fast mapping word-learning abilities of language-delayed preschoolers. *Journal of Speech and Hearing Research, 55,* 33-42.

Rice, M. L., Buhr, J. C., & Oetting, J. B. (1992). Specific-language-impaired children's quick incidental learning of words: The effect of a pause. *Journal of Speech and Hearing Research, 35,* 1040-1048.

Rice, M. L., Oetting, J. B., Marquis, J., Bode, J., & Pae, S. (1994). Frequency of input effects on word comprehension of children with specific language impairment. *Journal of Speech and Hearing Research, 37,* 106-122.

Rutter, M. (2008). Diagnostic concepts and risk processes. In C. Frazier Norbury, J. B. Tomblin, & D. V. M. Bishop (Eds.), *Understanding developmental language disorders: From theory to practice* (pp. 205-215). New York: Psychology Press.

Silva, P. A., McGee, R., & Williams, S. M. (1983). Developmental language delay from three to seven years and its significance for low intelligence and reading difficulties at age seven. *Developmental Medicine and Child Neurology, 25,* 783-793.

Sternberg, R. J. & Grigorenko, E. L. (2002). *Dynamic testing: The nature and measurement of learning potential*. Cambridge: Cambridge University Press.

Tomblin, J. B. (2008). Validating diagnostic standards for specific language impairment using adolescent outcomes. In C. Frazier Norbury, J. B. Tomblin, & D. V. M. Bishop (Eds.), *Understanding developmental language disorders* (pp. 93-114). New York: Psychology Press.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Wright, D. B. (1997). *Understanding statistics: An introduction for the social sciences*. London: Sage.

Tables

Table 1 Scores on static and dynamic measures

	Referred (whole group; n=40)	
	Mean	*SD*
BAS percentile score	40.94	23.22
BPVS percentile score	25.80	26.33
Dynamic weighted score	31.83	3.40
Expressive 1	2.48	1.43
Expressive 2	3.28	1.55
Recall 1	5.33	1.05
Recall 2	4.08	1.61

Table 2 Comparison of scores on BPVS at Time 1 and Time 2

	BPVS Time 1		BPVS Time 2				
	Mean	SD	Mean	SD	Mean diff.	t	df
Referred (n=37)	26.84	27.10	36.03	30.15	-9.19	-4.21****	36
Referred <25 (n=23)	9.13	7.26	16.17	10.89	-7.04	-3.72**	22

**p<.01, **** p <.0001

Table 3 Correlations between BPVS scores at time 2 and scores derived at time 1 with the whole group of referred children and the low scoring group

	Referred (whole group)	LS Referred
	BPVS Time 2	BPVS Time 2
	n=37	n=23
BPVS Time 1	.898**	.561**
BAS	.557**	.458*
Dynamic weighted score.	.492**	.249
Expressive 1	.498**	.184
Expressive 2	.627**	.444*
Recall 1	.435**	.015
Recall 2	.549**	.005

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

Table 4 Multiple Regression predicting vocabulary levels (BPVS scores) on follow-up (Time 2) of whole group of referred children

	B	SE B	β
Step1			
Constant	9.23	3.14	
BPVS Time 1	.99	.08	.90***
Step 2			
Constant	3.73	4.41	
BPVS Time 1	.92	.09	.83***
BAS	.18	.11	.14
Step 3			
Constant	-9.45	6.30	
BPVS Time 1	.75	.11	.67***
BAS	.25	.10	.20*
Expressive 2	4.57	1.68	.23*

Dependent variable: BPVS Time 2
* = $p < 0.05$, *** = $p < .001$
R^2 = .81 for Step 1; ΔR^2 = .02 for Step 2; ΔR^2 = .03 for Step 3

Table 5 Multiple Regression for prediction of vocabulary levels (BPVS scores) on follow-up (Time 2) of LS referred children

	B	SE B	B
Step1			
Constant	8.49	3.14	
BPVS Time 1	.84	.271	.56**
Step 2			
Constant	5.96	3.58	
BPVS Time 1	.67	.29	.45*
BAS	.12	.09	.27
Step 3			
Constant	-3.62	4.42	
BPVS Time 1	.53	.25	.35
BAS	.17	.08	.37*
Expressive 2	3.49	1.17	.46**

Dependent variable: BPVS Time 2

* = $p < 0.05$, ** = $p < .01$

$R^2 = .32$ for Step 1; $\Delta R^2 = .06$ for Step 2; $\Delta R^2 = .20$ for Step 3

Table 6 Casewise diagnostics

Case number	Std. Residual	BPVS Time 2	Predicted Value	Residual
1	-1.46	14	25.2	-11.2
3	1.18	34	25.0	9
4	-.57	1	5.4	-4.4
7	.19	18	16.6	1.4
8	-.05	12	12.4	-.4
9	-.55	8	12.2	-4.2
11	1.84	24	9.9	14.1
13	-1.39	1	11.6	-10.6
14	-.54	1	5.1	-4.1
17	-.06	28	28.4	-.4
18	.08	28	27.4	.6
21	.21	22	20.4	1.6
22	-.21	7	8.6	-1.6
26	1.17	26	17.1	8.9
27	.43	20	16.7	3.3
28	.47	37	33.4	3.6
29	.04	11	10.7	.3
30	.51	14	10.1	3.9
32	-1.07	6	14.2	-8.2
34	-.58	4	8.5	-4.5
36	-1.56	8	19.9	-11.9
38	1.78	20	6.4	13.6
39	.13	28	27.0	1