Child and Adolescent Mental Health
a journal for all professionals working with children and young people

Child and Adolescent Mental Health Volume 13, No. 4, 2008, pp. 198–206 doi: 10.1111/j.1475-3588.2008.00503.x

Parental Report of Infant Language Skills: A Review of the Development and Application of the Communicative Development Inventories

James Law[1] & Penny Roy[2]

[1]*Centre for Integrated Healthcare Research, Queen Margaret University, Edinburgh EH21 6UU, Scotland. E-mail: jlaw@qmu.ac.uk*
[2]*Department of Language and Communication Science, City University, Northampton Square, London EC1V 0HB, UK*

The Communicative Development Inventories (CDIs) are parent report measures of vocabulary and other aspects of language development in very young children. They have evolved over the past 20 years to be one of the most well recognised assessments of infant language. Of particular significance is the fact that the CDIs are the first measures of their kind to be widely translated and adapted for use in many different languages. The inventories have served a variety of functions including measuring early language acquisition, deriving normative data on language acquisition, and both identifying and describing children whose early language is significantly delayed. This review describes the development of the CDIs, summarises the volume of research that has been generated in a range of applications of the measures, and evaluates their current standing both as a research tool and as a clinical measure. Issues around the sensitivity and predictive value of the CDIs are also considered.

Key Practitioner Message:

- CDIs are one of the most commonly used research tools in the study of early language development
- Despite their wide use there are concerns about their validity and reliability
- Although they are versatile, efficient and valid, they should not be considered a panacea for child language assessment and particularly for predicting persistent language delay.

Keywords: Child language; measurement; communicative development inventories (COI); literature review

Introduction

There are three broad approaches to the assessment of language in young children: standardised norm referenced measures, analysis of language samples, and parent/carer reports of language performance or concern about language skills. None of these approaches is problem free. It is generally agreed that no one measure is sufficient to capture the multidimensional nature of language (Dockrell, 2001) and the optimal method for assessing language in infants and toddlers remains a matter of intense debate (Feldman et al., 2005). The number of standardised language measures available for use with children under 2.6 is limited, moreover accurate assessment of skills depends on the child's engagement with the task. A significant minority of young children fail to cooperate with strangers and, unsurprisingly, the younger the child, the less likely they are to comply (Chiat & Roy, 2007). Similar problems may arise when eliciting spontaneous language samples from young children in the clinical setting, and these are time consuming to analyse. These problems are particularly acute in the assessment of children with very low levels of both language comprehension and production, such as some children with autistic spectrum disorder (ASD), for whom many formal language tests may be inappropriate (Charman et al., 2003).

There are a number of pros and cons related to parental report. On the one hand, they have the advantage that they are based on the parent's extensive knowledge of their child's language skills in a wide variety of naturalistic settings, and as a result they are likely to be able to capture genuine performance in a way that a formal test of the young child commonly does not. Furthermore, they do not require highly trained clinicians to administer them. On the other hand, parents may lack an accurate frame of reference or be prone to bias. Parental intelligence or education may influence assessments (Feldman et al., 2000). In certain clinical or educational settings parents may have a vested interest in the outcome and either overestimate or underestimate their child's abilities (Oliver et al., 2003). Key to the whole endeavour, and common to all

Published by Blackwell Publishing, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA

three approaches, is the need to provide reliable and valid measures that can be used for both research and clinical purposes (Dockrell, 2001).

Dale and colleagues (1989) observed that the validity and reliability of parent reporting is dependent on three factors: the information called for is current and not retrospective; the language skills are emergent; and, lastly, skills are identified by recognition and not recall. These factors were incorporated into the design of one of the most widely used parental questionnaires, the MacArthur Communicative Development Inventories (MCDI) (Dale et al., 2001; Fenson et al., 1993, 1994) or as they have recently been renamed the MacArthur-Bates Communication Development Inventories (Fenson et al., 2007). Because the scales have been adapted in a number of ways in a range of studies, the measure is now most commonly referred to as the Communicative Development Inventories (CDIs or the CDI). The CDIs were developed as cost-efficient instruments for assessing communicative skills in infants and toddlers. They have been used in two main ways: first, to describe and chart the acquisition of early language in typically developing children and provide normative data; and second, to identify and describe children whose early language is significantly delayed and/or differs from normal development. The aim of this paper is to describe the development of the CDIs in the last 20 years, summarise the volume of research they have generated, and evaluate their current standing both as a research tool and a clinical measure.

The development of the CDIs

The CDIs comprise three complementary scales: the CDI Infant form (Words and Gestures: CDI-WG) that was standardised on 659 children aged 0.8 to 1.4; the CDI Toddler form (Words and Sentences: CDI-WS) standardised on 1130 children aged 1.4 to 2.6; and the more recent upward extension for children aged 2.6 to 3.6, the CDI-III. The CDI-WG is divided into two parts, one concerned with receptive and expressive vocabulary consisting of a 396-item vocabulary checklist, and the other with the intentional and symbolic development of gestures. The CDI-WS is designed to measure language production. It comprises a 680-word vocabulary production checklist, and measures of morphological and syntactic development, including items on utterance length and grammatical complexity. The CDI-III includes a 100-item vocabulary checklist and measures of syntactic maturity and language use and, to date, has been much less widely used than the infant and toddler versions (Feldman et al., 2005).

Although the earlier infant and toddler versions were comprehensive, there was a demand for shorter versions which could be more readily adapted for use in large scale applications. Fenson, Pethick et al. (2000a) developed standardised short form versions (CDI-USSF) which provide quicker estimates of children's communicative skills that are less dependent on parents' literacy level than the original inventories: an infant version (0.8–1.6), comprising a 89- word checklist for vocabulary comprehension and production, and two parallel toddler versions (1.4–2.6) comprising a 100- word vocabulary production checklist and a question about word combinations. Less widely used than the long versions, the value of the short forms has been recognised in large population studies (Horwitz et al., 2003). Eley and colleagues (2001) adapted the toddler version for use in the UK (CDI-UKSF) in their large scale study of twins, the Twins Early Developmental Study (TEDS). This 100-item version of the vocabulary list (CDI-UKSF), and subsequently a shorter 50-item version, have been used in a national study monitoring the language levels of 2-year-olds from families living in socially and economically disadvantaged areas in the UK, areas that were involved in government sponsored Sure Start programmes (Harris, 2002; Harris et al., 2004 a, b; Harris, Law, & Roy, 2005). The distribution of word types in the 50 and 100-item lists was comparable, but words that had been found to be culturally biased in the 100 items were excluded. These were items that in previous administrations of the 100-item list had produced low frequency scores in minority bilingual groups. Normative data (percentile and standard scores) for both the 100-item and 50-item versions were derived from a large sample of 1300 parents across England. This sample was representative in terms of gender, regional variation, ethnicity and maternal educational level. Unlike the US standardisation, which used postal questionnaires, the UK CDI-SF was directly administered and speech and language therapists were on hand to help with completion if required (Roy, Kersley, & Law, 2004). All versions have developed separate norms for boys and girls.

Translations of the CDIs

One of the most striking aspects of the CDI in the last 20 years is the number of translations and adaptations that have emerged. Forty-two versions are currently available world wide according to figures drawn up by Conboy and recently revised by Martinez-Sussmann (see http://www.sci.sdsu.edu/cdi/adaptations_ol.htm), spanning Europe, Asia, Scandinavia, Africa, Russia, the Antipodes and the Far East. Sign languages are represented too: the American Sign version is currently available and standardisation of the British Sign version has recently been completed (Woolfe et al., in preparation).

The Psychometric Properties of the CDIs

Given the widespread use of the CDIs, it is important to examine the reliability and both the concurrent and predictive validity of the instruments before considering their applications. Research studies have been concerned to demonstrate the psychometric robustness of the different forms at a group level. However, the value of the CDI as a clinical tool rests on its sensitivity to the identification of cases at an individual level.

Group level: age sensitivity

Feldman et al. (2000) found parents were sensitive to the developmental changes in their children's language. The scores on all the scales of the CDI-WG and CDI-WS increased significantly with age. As far as the CDI-III is concerned, although the 'sentence' scale and to a lesser

extent the 'using language' scale showed significant age effects, the vocabulary scale did not (Feldman et al., 2005).

Reliability

Acceptable levels of internal consistency of the original CDI-WS and the New Zealand adaptation (NZ CDI-WS) have been reported as evidenced by high correlations between vocabulary and grammar sections (complexity and MLU) (r = .62–.76) (Fenson et al., 1994; Reese & Read, 2000). Reese and Read found correlations were slightly stronger at 25 months than at 19 months, which was likely to be due to greater variability in grammatical development at the older age. Long-term reliability across approximately 6 months was also high, particularly for vocabulary (r = .75–.81), with grammatical complexity somewhat lower (r = .59–.61) (Fenson et al., 1994; Reese & Read, 2000). However, these time intervals are longer than those typically used to assess test-retest reliability. In the UK, standardisation of the CDI-UKSF high levels of test – retest and inter-rater reliabilities were found after a 2 week interval (vocabulary α = .99, and 91% agreement on categories for combining words (Roy et al., 2004).

Concurrent validity

At a group level the concurrent validity of the CDI is calculated by correlating scores on the CDI with scores on standardised tests of language and cognition and measures of language conversation. We have not specified the gold standard measures against which the CDIs have been checked. The reader is referred to the cited papers for this information. Table 1 shows the concurrent validity of the CDIs for samples of typically developing children, for samples of older children with delayed language, with or without associated difficulties, and for a translated version.

As can be seen in Table 1, with some exceptions, reported correlations range from moderate to high,

Table 1. Summary of findings from a sample of CDIs' concurrent validity studies

Authors	Age (months)	Sample type	*n*	CDI measure	*r*
Dale (1991)	24	Typically developing	24	WS: Productive voc	.68–.73
				Complexity	.54–.79.
				M3L[1]	.48–.74
Feldman et al. (2005)	36	Typically developing	115	CDI-III vocabulary	.33–.53
				Sentences	.41–.56
				Using language	.26–52
Heilmann et al. (2005)	30	Late talkers	38	WS: Productive voc	.51–.63[2]
				Complexity	.43–.52
				M3L	.4–.6
Marchman & Martinez-Sussman (2002)	23–24	Typically developing English/Spanish	26	WS Eng: Product voc	.56–.79
				Spanish: Product. voc	.56–.78
				WS Eng: Grammar	.46–.64
				Spanish: Grammar	.60–.71
Miller, Sedey, & Miolo (1995)	11–26	Typically developing	46	WS: Productive voc	.70–.75
	16–68	Down syndrome	44	WS: Productive voc	.77–.82
Oliver et al. (2003)	36–39	Typically developing	85	CDI-III vocabulary	.42–.62
Pan et al. (2004)	24	Typically developing low income	105	CDI-SF	.5
Roberts, Burchinal, & Durham (1999)	18	African-American low SES	56	SF: Productive voc	.41–.43[3]
				Irregular Ns & Vs	.33–.34
				MSL	.34–.4
	24	African-American low SES	41	SF: Productive voc	.76–.85[3]
				Irregular Ns & Vs	.65
				MSL	.74
	30	African-American low SES	44	SF: Productive voc	.61–.76[3]
				Irregular Ns & Vs	.38–.63
				MSL[4]	.51–.60
Thal, DesJardin, & Eisenberg (2007)	32–86 majority 32–66	Profound hearing loss and cochlear implants		CDI-WG	.58–.93
				Phrases understood	.74–.85
				Words understood	.62–.89
				Words produced	.55–.88
				WS: Productive voc	.41–.78
				Complexity	.53–.81
				M3L	
Thal et al. (1999)	39–49	Late talkers	20	WS: Productive voc	.67–.82
				Complexity	.52–.77
				M3L	.59–.62
	24–32	Late talkers	12	WG: Productive voc	.66
				Receptive voc	ns
				Gesture	ns

[1]Mean of three longest utterances
[2]parent-child not examiner-child language samples
[3]excluding questionable CDI scores
[4]Maximum sentence length

based on a range of outcome/concurrent behavioural language measures, including standardised tests and language samples. On the whole, higher correlations have been found for production than for comprehension and grammar, which has led some researchers to see production as the 'best' measure (Thal et al., 1999). The figures for the concurrent validity of the CDI-WG comprehension and gesture scales are questionable, at least when assessing older language delayed children (Table 1). It is generally agreed that parents are poorer at judging their children's language comprehension compared with their productive skills (Eriksson, Westerlund, & Berglund, 2002; Tomasello & Mervis, 1994). By the second year of life the child's receptive ability is so extensive it is no longer possible for parents to monitor it accurately. The CDI-WS did not include an assessment of comprehension because of the poor validity of parental report in this sub domain for this age range (Feldman et al., 2005). Differences in the range of ages and/or language abilities of the children in any one sample may affect the size of correlations: the narrower the range, the lower the correlation. Factors that affect the values of concurrent validity need to be taken into consideration when using the CDI measures with different populations of children.

Predictive validity

Studies of the predictive validity of CDIs have found correlations around .5 between CDI measures and subsequent outcome measures for time periods between 8–21 months. Associations have been found to be higher for CDI vocabulary scores than grammar scores (CDI-WS, Miller, Sedey, & Miolo, 1995; CDI-SF, Pan et al., 2004; NZ CDI-WS, Reese & Read, 2000). However Pan et al. and Reese and Read found the predictive value in their samples varied in sub groups according to sampling characteristics. In the New Zealand study, mothers with lower education overestimated their children's vocabulary levels compared with their performance on standardised measures (Reese & Read, 2000).

Accuracy of parental reports

Other studies have called into question the accuracy of parental reports, particularly in minority groups, who were underrepresented in the original US standardisation sample. Some of the most serious questions about the measurement properties of the CDIs were raised by Feldman et al. (2000), with findings from their longitudinal study of a large socio-demographically diverse sample of 2156 children in the US. CDI scores were analysed in terms of age, sex, race and maternal education and health insurance as proxies for SES. For three out of four of the continuous scales on the CDI-WG (phrases understood, vocabulary comprehension and vocabulary production) in children aged 10–13 months and two out of five of the continuous scales of the CDI-WS (word forms irregular and word forms overgeneralised) in children aged 22–25 months they found that children with *less* favourable SES indicators had *higher* scores than those with more favourable indictors. In other words, the direction of results was opposite to that predicted. They recommended caution in the use of CDIs in research and clinical practice to: (i) identify individual children at risk for language deficits; (ii) compare children with different socio-demographic profiles; and (iii) evaluate the effects of intervention. In a more recent study the invalid subscales were excluded (Feldman et al., 2005).

On the basis of these studies it appears that parents with higher education are more likely to be accurate reporters of their children's language, whereas parents with less education tend to overestimate their children's language abilities. However, the picture is not so straightforward. Roberts, Burchinal and Durham (1999) found some evidence of *under* reporting of early vocabulary and grammatical development in their sample of 87 African-American children from primarily low income families at 30 months, but not at 18 or 24 months, compared with the CDI norming population. The authors presented a number of alternatives for this finding, including the possibility of systematic bias in reporting.

There is some evidence that completion of the CDIs by multiple reporters may be more valid than single reporters (DeHouwer, Bornstein, & Leach, 2005; Marchman & Martinez-Sussman, 2002). DeHouwer et al. (2005) found the middle to upper-class mothers in their sample of monolingual Dutch children tended to underestimate their children's vocabulary. However, this may have been a function of their familiarity with the children. The proportion of full-time working mothers was high. Only 3 out of the 30 mothers had no employment outside the home, and 1 of these 3 children was left in child care for 32 hours/week.

Applications of the CDIs

Studies of normal language acquisition

The original US standardisation sample (n = 1803) of the CDIs provided a rich source of normative data on the gestural, lexical and grammatical development of children aged 0.8–2.6 in the US. Studies have shown, for example, that within a single language, vocabulary size is a more powerful predictor of grammatical development than age or gender, contributing significant variance to measures of grammar after age and gender are controlled (Bates et al., 1994; Bates, Dale, & Thal, 1995; Bauer, Goldfield, & Reznick, 2002; Dale et al., 2000; Dale & Fenson, 1996; Fenson et al., 1994; Marchman & Bates, 1994). Whilst significant overall age trends emerged, one of the most striking findings was the wide variability across children at initial assessment and over the course of vocabulary development. Criticisms that this variability undermined the potential value of the CDIs (Feldman et al., 2000) were countered by the argument that it was due to authentic variations in early language development, and not to psychometric deficiencies of the CDIs (Fenson, Bates et al., 2000a).

As Robinson and Mervis (1999) pointed out, the CDI growth curves that plot vocabulary size against age are based on cross-sectional data and as such represent the group average, not individual differences. In a single case study, they assessed the longitudinal validity of the CDI by comparing the CDI growth trajectory with the growth trajectory based on daily diary entries. Overall, the CDI growth curve was validated, but the CDI underestimated the words in the diary study, and the discrepancy between the measures

increased with age. A group study of individual growth profiles of CDI-WG scores was also revealing. Bauer, Goldfield and Reznick (2002) found that, whilst the correspondence between rates for productive and receptive vocabulary development was strong, there were a few cases of clear dissociations. They argued that identification of these outliers could provide an important tool for exploring the mechanisms of language development.

Gender

Sex differences favouring girls were found in the original US samples, but typically accounted for a small amount of variance, and effects of SES and birth order were even smaller. Reese and Read (2000) found that gender accounted for a more substantial amount of unique variance (10%) in total vocabulary scores in their more SES diverse sample of New Zealand infants. Bauer et al. (2002) also found in their study of individual growth profiles that the lexical development of girls outpaced boys. They found distinctive 'fast' and 'slow' trajectories, favouring girls, but not exclusively so.

SES and education

Economic and social disadvantage has been seen as a risk factor in early and delayed language acquisition (Hart & Risley, 1995; Locke, Ginsborg, & Peers, 2002). Arriaga et al.'s (1998) study of low and middle income families found the CDI scores of children from very low income families were strikingly lower on three key measures: size of vocabulary, word combination, and grammatical complexity. They found a 30% downward shift in *all* scores, a substantially larger shift than previously reported for low SES samples. The size of the shift is partly a function of the unrepresentative composition of the US standardisation samples for both the long and short forms of the CDI.

As Arriaga et al. (1998) argued, the middle class biased norms can be seen as advantageous for studies of low SES samples in that they allow a direct comparison of the language status of a child from a low income/low education family with that of children from middle class families. However, biased norms could mean that inflated proportions of more SES diverse samples/populations fall below cut-offs for identifying low function (typically the 5th or 10th percentile). This may account, for example, for the high proportion (nearly 20%) of 'late talkers' (those performing at or below the 10th percentile on the CDI) identified at 24 months in the Early Language in Victoria Study (ELVS), a longitudinal study following the development of 1900 infants born in Melbourne from 0.8 to 4.0 (Reilly et al., 2007). Logistic regression identified lower levels of maternal education along with non-English speaking background and a family history of speech and language difficulty as risk factors for 'late talking'. This study found, as we did, that the risk of maternal education did not apply equally across the sample as a whole (Roy et al., 2004). In the UK standardisation sample the association between parental education level and lower productive vocabulary in the child only held for the parents with the lowest levels of education (Roy et al., 2004). Education level for SES advantaged populations did not emerge as a 'risk factor' (US: Fenson et al., 1994; UK: Hamilton, Plunkett, & Shafer, 2000) and explains little if any variance in children's early productive vocabularies.

Cross linguistic studies

The translations and adaptations of the CDI to other languages described above have opened up possibilities of cross-linguistic comparisons of communicative development (Bornstein, et al., 2004; Caselli et al., 1995; Caselli, Casadio, & Bates, 1999; Devescovi et al., 2005; Maital et al., 2000; Tardif, Gelman, & Xu, 1999; Tardif, Fletcher, & Liang, 2004). Use of the CDI with bilingual children affords a unique opportunity to investigate the parallel development of two emergent language systems and the relationships between them (Pearson et al., 1997) and to explore relationships between the development of grammar and vocabulary. Conboy and Thal (2006), using the US and Spanish version of the CDI, found the grammatical abilities in each separate language of the bilingual toddlers were paced by the lexical development primarily within the same language. They concluded that bilingual children acquire the grammar of each language separately.

Other uses of CDIs

The CDIs are being used increasingly to address theoretical issues such as estimating the relative contribution of genetics versus environmental factors in the rates of language development (Dionne et al., 2003; Price et al., 2000), and determining the prevalence and predictors of early language delay (Horwitz et al., 2003). CDIs have also been used as outcome measures to evaluate the role of hypothesised early predictors of later language development; to describe the language skills of children with developmental disorders; and in studies of 'at risk' populations.

CDIs as outcome measures

The CDIs have been adopted as outcome measures to assess the role of hypothesised early predictors of later language and communicative development (Heimann et al., 2006; Paavola, Kunnari, & Moilanen, 2005; Markus et al., 2000; Lyytinen et al., 1999; Colombo et al., 2004; Dixon & Smith, 2000). Tsao, Liu and Kuhl (2004), for example, found speech discrimination performance at 6 months predicted language at 2 years on the CDI, supporting the hypothesis that phonetic perception may play an important role in language development.

Children with developmental disorders

The CDIs have been used extensively as a research tool to investigate language skills in children with a range of developmental disabilities (Berglund, Eriksson, & Johansson, 2001; Bird et al., 2005; Caselli et al., 1998; Chilosi et al., 2001; Mayne, Yoshinago-Itano & Sedey, 2000a; Mayne, Yoshinago-Itano, Sedey, & Carey, 2000b; Mitchell et al., 2006; Scherer & Dantonio, 1995; Smith, Mirenda, & Zaidman-Zait, 2007; Stallings, Gao, & Svirsky, 2000; Thal et al., 2004; Thal, DesJardin, & Eisenberg, 2007). Some studies have used the CDIs to compare the early language development of children with different disabilities. For example, young children with Williams Syndrome were found to have superior linguistic skills and larger expressive scores than age

matched children with Down Syndrome (Harris et al., 1997; Mervis & Robinson, 2000). The Charman et al. (2003) study of children with autistic spectrum disorders illustrates the CDI's potential to distinguish between delayed and deviant language development, by comparing their language profiles with those of typically developing children.

'*At risk*' *populations*

The infant and toddler forms of the CDIs have also been used to assess the language development of children who are potentially 'at risk' of developing language impairment, including toddlers born preterm (Magill-Evans & Harrison, 1999); those exposed prenatally to stimulant drugs (Dixon et al., 1997); infants who suffer from otitis media (Feldman et al., 1999, 2003) and those at familial risk for dyslexia (Koster et al., 2005; Lyytinen & Lyytinen, 2004; Viholainen et al., 2002). The Dutch study, using the N-CDI standardised in Belgium, found differences in the *profiles* of acquisition of children whose production vocabulary was greater than 50 words. 'At risk' children produced fewer verbs and closed-class words compared with controls, but no differences were found in their production of nouns and predicates overall (Koster et al., 2005). Together with the evidence from Thal et al.'s longitudinal study, Early Identification of Risk for Language Impairment (Thal, 2005), Koster et al. argued that the *combined* risk of delay in lexical acquisition and deviance in its composition may differentiate transient and persistent problems in morphological development.

'*Late talkers*'

The issue of differentiating transient from persistent problems in morphological development is central to the study of 'late talkers' (LTs). One of the most common criteria for classifying children as LTs has been total productive vocabulary at or below the 10th percentile on the CDI-WS or the CDI-UKSF (Dale et al., 2003; Heilmann et al., 2005). LTs tend to be identified on average around 24 months (with a range of 18–30 months). At a group level, LTs are at a substantially elevated risk for continuing language impairment; however many low scoring children catch up in later preschool years (Dale et al., 2003; Thal, Tobias, & Morrison, 1991; Thal et al., 1999). For clinical purposes group level analysis is not sufficient. Clinicians want information on individual 'caseness'. They want to know the likelihood that children who do poorly on parent reports are the same children whose performance is low on later standardised or accepted measures that are used to define language or cognitive delay ('gold standards').

Westerlund, Berglund and Eriksson (2006) argued that if the prevalence of a disorder is low and the level of severity is high, as is the case with severe language disability, it is important to find as many cases as possible. Accordingly, priority should be given to the high sensitivity of a measure (proportion of true positives identified), even at the cost of low specificity (proportion of true negatives identified). There are only a handful of studies that have looked at the sensitivity of the CDI, and overall the findings are not encouraging. For example, Westerlund et al. (2006) used the Swedish Communication Screening (SCS18), a short screening version of the Swedish CDI (SECDI), in a population sample of toddlers. They found that productive vocabulary was the best predictor of the three SCS18 variables, but its sensitivity (50%) was too low to be clinically valid. They concluded that 18 months was too early to predict severe language impairment in 3-year-olds. However, 6 months later, the picture looks similar. Dale et al. (2003), using the CDI-SF in a large scale genetic study of 8386 2-year-old twins in the UK, found the predictive value of the CDI was adequate, but its sensitivity was poor. They concluded that 'vocabulary at 2 is a predictor of poor language at 3 and 4', but 'too poor to be of practical utility in discriminating persistent and transient difficulties' (p.555). Both the SCS18 and the CDI-SF are shortened versions of the CDIs, and as such are particularly suitable for use as population screens. But as Law et al. (2000) argued, although short screening measures are appealing options, the evidence to date is that they do not deliver the type of data that is really useful in a clinical context. One possibility is that the longer version, the CDI-WS, may prove more sensitive. In many respects, the results of a follow-up study at 30 months of 100 toddlers (38 LTs and 62 with a history of normal language development, according to reported productive vocabulary on the CDI-WS at 24 months) look more promising (Heilmann et al., 2005). However, despite excellent specificity (98%), sensitivity of the CDI-WS at the optimal cut-off (11th percentile) although substantially higher than the short versions, in clinical terms remained modest (68%). Although sensitivity is a key measure, tests rarely report these figures. A recent review of 43 tests of early language development revealed that measures of sensitivity and specificity were reported in 9, of which 5 only reached acceptable levels of sensitivity (Spaulding, Plante, & Farinella, 2006).

Conclusions

In summary, the CDIs are cost effective, easy to administer measures of early language and communicative development that have been very widely used in a variety of contexts to address a range of different questions. In their favour they have the distinct advantage of having been translated into many languages and used extensively in cross linguistic investigations. They have been shown to be versatile, efficient and valid measures of language development in young children both with and without developmental disabilities, and have been used effectively in both clinical and research settings. The checklists can be completed by parents prior to their child's evaluation, and clinicians can use this information to estimate child's language level and provide an opportunity for parents to become actively involved in the assessment process (Miller et al., 1995). The CDIs have provided a rich source of information on early communicative and language acquisition and revealed the wide range of individual differences in young children's early vocabulary acquisition and their rate of language development. As Devescovi et al. (2005) acknowledged, the CDIs should not be seen as substitutes for 'live' samples; nevertheless, they offer an efficient and relatively low cost means of gathering data from large samples in one or more languages and they have a significant role to play in generating working hypotheses.

On the more negative side, the CDIs should not be considered the panacea for all those wishing to assess language in very young children. The variability in developmental trajectories for early language development makes reliable predictions about individual outcomes of 'late talking' a challenge. Feldman et al.'s (2000) caution about using the CDIs to identify language delay in individual children is still relevant, but this problem is not unique to the CDIs. In common with other language measures, sensitivity remains an issue.

Researchers and clinicians should be aware of the potential value of different aspects of the CDIs and consider their psychometric properties when used in different populations. It would be helpful to know more about the potential impact of reporting bias, particularly in populations of vulnerable children. It is likely that the CDIs will continue to be widely used in the foreseeable future for all the pragmatic reasons to which reference has already been made. But it is important that these reservations are taken into consideration and that researchers and clinicians are aware of the strengths and limitations of the measure in its various forms.

References

Arriaga, R.I., Fenson, L., Cronan, T., & Pethick, S.J. (1998). Scores on the MacArthur communicative development inventory of children from low- and middle-income families. *Applied Psycholinguistics*, *19*, 209–223.

Bates, E., Dale, P., & Thal, D. (1995). Individual differences and their implications for theories of language development. In P. Fletcher & B. MacWhinney (Eds.), *Handbook of child language* (pp. 96–151). Oxford, UK: Basil Blackwell.

Bates, E., Marchman, V., Thal, D., Fenson, L., Dale, P., Reznick, J.S., Reilly, J., & Hartung, J. (1994). Developmental and stylistic variation in the composition of early vocabulary. *Journal of Child Language*, *21*, 85–123.

Bauer, D.J., Goldfield, B.A., & Reznick, J.S. (2002). Alternative approaches to analyzing individual differences in the rate of early vocabulary development. *Applied Psycholinguistics*, *23*, 313–335.

Berglund, E., Eriksson, M., & Johansson, I. (2001). Parental reports of spoken language in children with Down syndrome. *Journal of Speech, Language, and Hearing Research*, *44*, 179–191.

Bird, E.K.R., Cleave, P., Trudeau, N., Thordardottir, E., Sutton, A., & Thorpe, A. (2005). The language abilities of bilingual children with Down syndrome. *American Journal of Speech-Language Pathology*, *14*, 187–199.

Bornstein, M.H., Cote, L.R., Maital, S., Painter, K., Park, S.Y., Pascual, L., Pecheux, M.G., Ruel, J., Venuti, P., & Vyt, A. (2004). Cross-linguistic analysis of vocabulary in young children: Spanish, Dutch, French, Hebrew, Italian, Korean, and American English. *Child Development*, *75*, 1115–1139.

Caselli, C., Bates, E., Casadio, P., Fenson, J., Fenson, L., & Sanderl, L. (1995). A cross-linguistic study of early lexical development. *Cognitive Development*, *10*, 159–199.

Caselli, M.C., Casadio, P., & Bates, E. (1999). A comparison of the transition from first words to grammar in English and Italian. *Journal of Child Language*, *26*, 69–111.

Caselli, M.C., Vicari, S., Longobardi, E., Lami, L., Pizzoli, C., & Stella, G. (1998). Gestures and words in early development of children with Down syndrome. *Journal of Speech, Language, and Hearing Research*, *41*, 1125–1135.

Charman, T., Drew, A., Baird, C., & Baird, G. (2003). Measuring early language development in preschool children with autism spectrum disorder using the MacArthur Communicative Development Inventory (Infant form). *Journal of Child Language*, *30*, 213–236.

Chiat, S., & Roy, P. (2007). The preschool repetition test: An evaluation of performance in typically developing and clinically referred children. *Journal of Speech, Language, and Hearing Research*, *50*, 429–443.

Chilosi, A.M., Cipriani, P., Bertuccelli, B., Pfanner, L., & Cioni, G. (2001). Early cognitive and communication development in children with focal brain lesions. *Journal of Child Neurology*, *16*, 309–316.

Colombo, J., Shaddy, D.J., Richman, W.A., Maikranz, J.M., & Blaga, O.M. (2004). The developmental course of habituation in infancy and preschool outcome. *Infancy*, *5*, 1–38.

Conboy, B.T., & Thal, D.J. (2006). Ties between the lexicon and grammar: Cross-sectional and longitudinal studies of bilingual toddlers. *Child Development*, *77*, 712–735.

Dale, P. (1991). The validity of a parent report measure of vocabulary and syntax at 24 months. *Journal of Speech and Hearing Research*, *34*, 565–571.

Dale, P.S., Bates, E., Reznick, S., & Moriset, C. (1989). The validity of a parent report instrument of child language at 20 months. *Journal of Child Language*, *16*, 239–249.

Dale, P.S., Dionne, G., Eley, T.C., & Plomin, R. (2000). Lexical and grammatical development: A behavioural genetic perspective. *Journal of Child Language*, *27*, 619–642.

Dale, P.S., & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods Instruments and Computers*, *28*, 125–127.

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

Dale, P.S., Reznick, J.S., Thal, D., & Marchman, V.A. (2001). *A parent report measure of language development for three-year-olds*. Unpublished manuscript, University of Missouri-Columbia.

DeHouwer, A., Bornstein, M.H., & Leach, D.B. (2005). Assessing early communicative ability: A cross-reporter cumulative score for the MacArthur CDI. *Journal of Child Language*, *32*, 735–758.

Devescovi, A., Caselli, M.C., Marchione, D., Pasqualetti, P., Reilly, J., & Bates, E. (2005). A crosslinguistic study of the relationship between grammar and lexical development. *Journal of Child Language*, *32*, 759–786.

Dionne, G., Dale, P.S., Boivin, M., & Plomin, R. (2003). Genetic evidence for bidirectional effects of early lexical and grammatical development. *Child Development*, *74*, 394–412.

Dixon, W.E., & Smith, P.H. (2000). Links between early temperament and language acquisition. *Merrill-Palmer Quarterly-Journal of Developmental Psychology*, *46*, 417–440.

Dixon, S., Thal, D., Potrykus, J., Dickson, T.B., & Jacoby, J. (1997). Early language development in children with prenatal exposure to stimulant drugs. *Developmental Neuropsychology*, *13*, 371–396.

Dockrell, J.E. (2001). Assessing language skills in preschool children. *Child Psychology and Psychiatry Review*, *6*, 74–86.

Eley, T.C., Dale, P.S., Bishop, D.V.M., Price, T.S., & Plomin, R. (2001). Longitudinal analysis of components of cognitive delay: Examining the aetiology of verbal and performance aspects of cognitive delay. *Journal of Educational Psychology*, *93*, 698–707.

Eriksson, M., Westerlund, M., & Berglund, E. (2002). A screening version of the Swedish communicative development inventories designed for use with 18-month-old children. *Journal of Speech, Language, and Hearing Research*, *45*, 948–960.

Feldman, H.M., Dale, P.S., Campbell, T.F., Colborn, D.K., Kurs-Lasky, M., Rockette, H.E., & Paradise, J.L. (2005). Concurrent and predictive validity of parent reports of child language at ages 2 and 3 years. *Child Development*, *76*, 856–868.

Feldman, H.M., Dollaghan, C.A., Campbell, T.F., Colborn, D.K., Janosky, J., Kurs-Lasky, M., Rockette, H.E., Dale, P.S., & Paradise, J.L. (2003). Parent-reported language skills in relation to otitis media during the first 3 years of life. *Journal of Speech, Language, and Hearing Research*, *46*, 273–287.

Feldman, H.M., Dollaghan, C.A., Campbell, T.F., Colborn, D.K., Kurs-Lasky, M., Janosky, J., & Paradise, J.L. (1999). Parent-reported language and communication skills at one and two years of age in relation to otitis media in the first two years of life. *Pediatrics*, *104*, Art. No. e52.

Feldman, H.M., Dollaghan, C.A., Campbell, T.F., Kurs-Lasky, M., Janosky, J.E., & Paradise, J.L. (2000). Measurement properties of the MacArthur communicative development inventories at ages one and two years. *Child Development*, *71*, 310–322.

Fenson, L., Bates, E., Dale, P., Goodman, J., Reznick, J.S., & Thal, D. (2000a). Measuring variability in early child language: Don't shoot the messenger. *Child Development*, *71*, 323–328.

Fenson, L., Dale, P.S., Reznick, J.S., Bates, E., Thal, D.J., & Pethick, S.J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, *59*, 1–185.

Fenson, L., Dale, P.S., Reznick, J.S., Thal, D.J., Bates, E., Hartung, J.P., Pethick, S., & Reilly, J.S. (1993). *The MacArthur Communicative Development Inventories: Users*' guide and technical manual. San Diego, CA: Singular Publishing Group.

Fenson, L., Marchman, V., Thal, D.J., Dale, P.S., Reznick, J.S., & Bates, E. (2007). *Communicative Development Inventories*. Baltimore: Brookes.

Fenson, L., Pethick, S., Renda, C., Cox, J.L., Dale, P.S., & Reznick, J.S. (2000b). Short-form versions of the MacArthur Communicative Development Inventories. *Applied Psycholinguistics*, *21*, 95–115.

Hamilton, A., Plunkett, K., & Schafer, G. (2000). Infant vocabulary development assessed with a British communicative development inventory. *Journal of Child Language*, *27*, 689–705.

Harris, F. (2002). *The first implementation of the Sure Start Language Measure*. DfES. Available at: http://www.surestart.gov.uk/_doc/P0000128.pdf

Harris, F., Law, J., & Roy, P. (2004a). *The development of a 50 word adaptation of the UK short form of the MCDI for use with two years olds in Sure Start Programmes in England*. Paper prepared by the team responsible for this further development of the CDI in the UK for Professor Larry Fenson and the MCDI team in San Diego, USA. Available from Professor James Law, c/o Centre for Integrated Healthcare Research, Queen Margaret University, Edinbugh UK jlaw@qmu.ac.uk.

Harris, F., Law, J., & Roy, P. (2005). *The third implementation of the Sure Start Language Measure*. DfES. Available at: http://www.surestart.gov.uk/_doc/P0001802.pdf

Harris, F., Law, J., Roy, P., & Kermani, S. (2004b). *The second implementation of the Sure Start Language Measure*. DfES. Available at: http://www.surestart.gov.uk/doc/P0000809.pdf 2004

Harris, N.G.S., Bellugi, U., Bates, E., Jones, W., & Rossen, M. (1997). Contrasting profiles of language development in children with Williams and Down syndromes. *Developmental Neuropsychology*, *13*, 345–370.

Hart, B., & Risley, T.R. (1995). *Meaningful differences in the everyday experiences of young American children*. Baltimore: Paul Brooke Publishing Co.

Heilmann, J., Ellis Weismer, S., Evans, J., & Hollar, C. (2005). Utility of the MacArthur-Bates Communicative Development Inventory in identifying language abilities of late-talking and typically developing toddlers. *American Journal of Speech-Language Pathology*, *14*, 40–51.

Heimann, M., Strid, K., Smith, L., Tjus, T., Ulvund, S.E., & Meltzoff, A.N. (2006). Exploring the relation between memory, gestural communication, and the emergence of language in infancy: A longitudinal study. *Infant and Child Development*, *15*, 233–249.

Horwitz, S.M., Irwin, J.R., Briggs-Gowan, M.J., Heenan, J.M.B., Mendoza, J., & Carter, A.S. (2003). Language delay in a community cohort of young children. *Journal of the American Academy of Child and Adolescent Psychiatry*, *42*, 932–940.

Koster, C., Been, P.H., Krikhaar, E.M., Zwarts, F., Diepstra, H.D., & Van Leeuwen, T.H. (2005). Differences at 17 months: Productive language patterns in infants at familial risk for dyslexia and typically developing children. *Journal of Speech, Language, and Hearing Research*, *48*, 426–438.

Law, J., Boyle, J., Harris, F., Harkness, A., & Nye, C. (2000). The feasibility of universal screening for primary speech and language delay: Findings from a systematic review of the literature. *Developmental Medicine and Child Neurology*, *42*, 190–200.

Locke, A., Ginsborg, J., & Peers, I. (2002). Development and disadvantage: Implications for the early years and beyond. *International Journal of Language and Communication Disorders*, *37*, 3–15.

Lyytinen, P., Laakso, M.L., Poikkeus, A.M., & Rita, N. (1999). The development and predictive relations of play and language across the second year of life. *Scandinavian Journal of Psychology*, *40*, 177–186.

Lyytinen, P., & Lyytinen, H. (2004). Growth and predictive relations of vocabulary and inflectional morphology in children with and without familial risk for dyslexia. *Applied Psycholinguistics*, *25*, 397–411.

Magill-Evans, J., & Harrison, M.J. (1999). Parent-child interactions and development of toddlers born preterm. *Western Journal of Nursing Research*, *21*, 292–307.

Maital, S.L., Dromi, E., Sagi, A., & Bornstein, M.H. (2000). The Hebrew communicative development inventory: Language specific properties and cross-linguistic generalizations. *Journal of Child Language*, *27*, 43–67.

Marchman, V., & Bates, E. (1994). Continuity in lexical and morphological development: A test of the critical mass hypothesis. *Journal of Child Language*, *21*, 339–366.

Marchman, V.A., & Martinez-Sussmann, C. (2002). Concurrent validity of caregiver/parent report measures of language for children who are learning both English and Spanish. *Journal of Speech, Language, and Hearing Research*, *45*, 983–997.

Markus, J., Mundy, P., Morales, M., Delgado, C.E.F., & Yale, M. (2000). Individual differences in infant skills as predictors of child caregiver joint attention and language. *Social Development*, *9*, 302–315.

Mayne, A.M., Yoshinago-Itano, C., Sedey, A.L., & Carey, A. (2000a). Expressive vocabulary development of infants and toddlers who are deaf or hard of hearing. *Volta Review*, *100*, 1–28.

Mayne, A.M., Yoshinago-Itano, C., & Sedey, A.L. (2000b). Receptive vocabulary development of infants and toddlers who are deaf or hard of hearing. *Volta Review*, *100*, 29–52.

Mervis, C.B., & Robinson, B.F. (2000). Expressive vocabulary ability of toddlers with Williams syndrome or Down syndrome: A comparison. *Developmental Neuropsychology*, *17*, 111–126.

Miller, J.F., Sedey, A.L., & Miolo, G. (1995). Validity of parent report measures of vocabulary acquisition in children with Down-Syndrome. *Journal of Speech, Language, and Hearing Research*, *38*, 1037–1044.

Mitchell, S., Brian, J., Zwaigenbaum, L., Roberts, W., Szatmari, P., Smith, I., & Bryson, S. (2006). Early language and communication development of infants later diagnosed

with autism spectrum disorder. *Journal of Developmental and Behavioral Pediatrics*, *27*, S69–S78.

Oliver, B., Dale, P.S., Saudino, K.J., Petrill, S.A., Pike, A., & Plomin, R. (2003). The validity of a parent-based assessment of cognitive abilities in three-year olds. *Early Child Development and Care*, *172*, 337–348.

Paavola, L., Kunnari, S., & Moilanen, I. (2005). Maternal responsiveness and infant intentional communication: Implications for the early communicative and linguistic development. *Child Care Health and Development*, *31*, 727–735.

Pan, B.A., Rowe, M.L., Spier, E., & Tamis-Lemonda, C. (2004). Measuring productive vocabulary in toddlers in low-income families: Concurrent and predictive validity of three sources of data. *Journal of Child Language*, *31*, 587–608.

Pearson, B.Z., Fernandez, S.C., Lewedeg, V., & Oller, D.K. (1997). The relation of input factors to lexical learning by bilingual infants. *Applied Psycholinguistics*, *18*, 41–58.

Price, T.S., Eley, T.C., Dale, P.S., Stevenson, J., Saudino, K., & Plomin, R. (2000). Genetic and environmental covariation between verbal and nonverbal cognitive development in infancy. *Child Development*, *71*, 948–959.

Reese, E., & Read, S. (2000). Predictive validity of the New Zealand MacArthur Communicative Development Inventory: Words and sentences. *Journal of Child Language*, *27*, 255–266.

Reilly, S., Wake, M., Bavin, E.L., Prior, M., Williams, J., Bretherton, L., Eadie, P., Barrett, Y., & Ukoumunne, C.C. (2007). Predicting language at two years of age: Prospective community study. *Pediatrics*, *120*, 1441–1449.

Roberts, J.E., Burchinal, M., & Durham, M. (1999). Parents' report of vocabulary and grammatical development of African American preschoolers: Child and environmental associations. *Child Development*, *70*, 92–106.

Robinson, B.F., & Mervis, C.B. (1999). Comparing productive vocabulary measures from the CDI and a systematic diary study. *Journal of Child Language*, *26*, 177–185.

Roy, P., Kersley, H., & Law, J. (2004). *The Sure Start Language Measure Standardisation Study*. DfES. Available at: http://www.surestart.gov.uk/_doc/P0001797.pdf.

Scherer, N.J., & Dantonio, L. (1995). Parent questionnaire for screening early language-development in children with cleft palate. *Cleft-palate Craniofacial Journal*, *32*, 7–13.

Smith, V., Mirenda, P., & Zaidman-Zait, A. (2007). Predictors of expressive vocabulary growth in children with autism. *Journal of Speech, Language, and Hearing Research*, *50*, 149–160.

Spaulding, T.J., Plante, E., & 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*, 61–72.

Stallings, L.M., Gao, S., & Svirsky, M.A. (2000). Assessing the language abilities of pediatric cochlear implant users across a broad range of ages and performance abilities. *Volta Review*, *102*, 215–235.

Tardif, T., Fletcher, P., & Liang, W. (2004). First words in English and Chinese: Results from norming studies of the MacArthur Communicative Development Inventories. *International Journal of Psychology*, *39*, 515–515 Suppl.

Tardif, T., Gelman, S.A., & Xu, F. (1999). Putting the 'noun bias' in context: A comparison of English and Mandarin. *Child Development*, *70*, 620–635.

Thal, D. (2005). *Early detection of risk for language impairment: What are the best strategies?* Presented at the American Speech and Hearing Association Convention, Miami, USA.

Thal, D., DesJardin, J.L., & Eisenberg, L.S. (2007). Validity of the MacArthur-Bates Communicative Development Inventories for measuring language abilities in children with cochlear implants. *American Journal of Speech-Language Pathology*, *16*, 54–64.

Thal, D.J., O'Hanlon, L., Clemmons, M., & Fralin, L. (1999). Validity of a parent report measure of vocabulary and syntax for preschool children with language impairment. *Journal of Speech, Language, and Hearing Research*, *42*, 482–496.

Thal, D.J., Reilly, J., Seibert, L., Jeffries, R., & Fenson, J. (2004). Language development in children at risk of language impairment: Cross-population comparisons. *Brain and Language*, *88*, 167–179.

Thal, D., Tobias, S., & Morrison, D. (1991). Language and gesture in late talkers: A 1-year follow-up. *Journal of Speech and Hearing Research*, *34*, 604–612.

Tomasello, M., & Mervis, C.B. (1994). The instrument is great, but measuring comprehension is still a problem. *Monographs of the Society for Research in Child Development*, *59*, 174–179.

Tsao, F.M., Liu, H.M., & Kuhl, P.K. (2004). Speech perception in infancy predicts language development in the second year of life: A longitudinal study. *Child Development*, *75*, 1067–1084.

Viholainen, H., Ahonen, T., Cantell, M., Lyytinen, P., & Lyytinen, H. (2002). Development of early motor skills and language in children at risk for familial dyslexia. *Developmental Medicine and Child Neurology*, *44*, 761–769.

Westerlund, M., Berglund, E., & Eriksson, M. (2006). Can severely language delayed 3 year-olds be identified at 18 months? Evaluation of a screening version of the MacArthur-Bates communicative development inventories. *Journal of Speech, Language, and Hearing Research*, *49*, 237–247.

Woolfe, T., Woll, B., Herman, R., & Roy, P. (in preparation). *The BSL MacArthur-Bates Communicative Development Inventory*.