British Journal of Developmental Psychology (1998), 16, 349–363 *Printed in Great Britain*

The validity of parent-based assessment of the cognitive abilities of 2-year-olds

Kimberly J. Saudino
Boston University, Boston, MA 02215, USA

Philip S. Dale
University of Washington, Seattle, WA 98195, USA

Bonny Oliver, Stephen A. Petrill, Victoria Richardson, Michael Rutter and Emily Simonoff
Institute of Psychiatry, London SE5 8AF, UK

Jim Stevenson
University of Southampton, Southampton, UK

Robert Plomin*
Social, Genetic and Developmental Psychiatry Research Centre, Institute of Psychiatry, 113 Denmark Hill, London SE5 8AF

Parent reports have been widely used to assess child behaviours in the socio-emotional domain, but seldom have been used to assess behaviours within the cognitive domain. The present study examines the ability of parent reports and parent-administered tasks obtained through the post to assess non-verbal cognitive abilities in early childhood. In a sample of 107 2-year-olds, age-corrected scores on parent reports and parent-administered tasks assessing non-verbal reasoning significantly predicted performance on the Mental Development Index (MDI) of the Bayley Scales of Infant Development-II two weeks later ($r = .49$ and $r = .41$, $p < .0001$, respectively). The multiple correlation between the two components and the MDI was .55 ($p < .0001$). This ability of parental assessments to predict the MDI is comparable to the predictive power of standard tester-administered measures at this age, even though the parent measure specifically excludes verbal items that are included in the MDI. Adding parent reports of language development significantly improved the prediction of the MDI ($R = .66$, $p < .0001$). In addition, higher within-domain than cross-domain correlations reflect a significant ability of parents to discriminate verbal and non-verbal abilities.

* Requests for reprints.

Using parents to assess the cognitive abilities of 2-year-olds

Because parents know their children well and spend much time with them, they are a potentially rich source of information about their children's typical behaviours across many situations. Indeed, parent-report measures have been widely used in the social-emotional domain to assess child temperament, attachment, social skills and behaviour problems (e.g. Achenbach & Edelbrock, 1983; Goldsmith, 1996; Hogan, Scott & Bauer, 1992; Waters & Deane, 1985). In contrast, parent reports have seldom been used to assess behaviours within the cognitive domain.

Parent-report measures offer three potential advantages for studying cognitive development. First, as compared to traditional standardized tests, which typically involve a brief sample of behaviour in a specific test situation, parental assessments are based on more extensive behavioural sampling, and might therefore attenuate problems associated with situational influences. Second, parent reports are considerably more efficient and economical than in-person assessment. Third, because parental assessments can be administered via the post, they are particularly useful for behavioural research that requires large samples (e.g. behavioural genetic studies and studies that screen for exceptionality).

Of course, the utility of parent-report measures for assessing cognitive development depends on the ability of parents to make accurate judgments regarding behavioural indices of their children's cognitive skills. Although parent reports have been assumed to be prone to bias, forgetting and misinterpretation (as reviewed by Fenson, Dale, Reznick, Bates, Thal & Pethick, 1994), there is ample evidence to suggest that parents can provide reasonable estimates of their children's abilities. For example, a recent review of the literature found that across 23 studies, the average correlation between parental judgments, uncorrected for age, and professional assessment (i.e. standardized tests) of cognitive development of young children was .73 (Dinnebeil & Rule, 1994). Parental estimates do tend to yield higher means than scores based on standardized testing, a finding that has typically been interpreted as parental overestimation (Miller, 1986; Miller, Manhal & Mee, 1991); however, it can also be argued that the narrow test situation may actually *underestimate* children's abilities (Gradel, Thompson & Sheehan, 1981; Sheehan, 1988).

Although previous studies examining the validity of parental ratings within the cognitive domain have provided encouraging results, some questions still remain. In general, parental assessments of children's development have included a wide range of behaviours, only some of which are purely cognitive. For example, the Minnesota Child Development Inventory (Ireton & Thwing, 1974), a widely used parent-report screening measure, includes items designed to assess motor, language, social and self-help skills in addition to cognitive abilities. This is also true for more recent parent-report measures (e.g. Kent Infant Development Scale, Reuter & Bickett, 1985; Infant Monitoring Questionnaires, Bricker & Squires, 1989). Thus, rather than focusing specifically on cognitive abilities, these measures use parent ratings to assess general development.

An exception to the lack of cognitively oriented parent-report measures are the MacArthur Communicative Development Inventories (MCDI), a set of two parent-report measures designed to assess language development in infants and toddlers (Fenson *et al.*, 1994). Research findings on the MCDI clearly demonstrate that parents can provide valid

and reliable information about a wide variety of language competencies demonstrated by their children. Averaging across multiple studies, parent reports of children's vocabulary knowledge on the MCDI correlated .67 with laboratory measures of expressive language. Similarly, the average correlation between the MCDI parent report of sentence complexity and laboratory measures of observed mean length of utterance was .79. These validity correlations are of the same order of magnitude as the reliabilities of the criterion measures, thus predicting essentially all of the reliable variance in those criteria. Fenson *et al.* suggest that parent report is most likely to be accurate when assessment is limited to current behaviours and not retrospective accounts; when it is focused on emergent behaviours which the parent can still monitor; and when the format involves recognition rather than recall (e.g. asking if a child says 'green' rather than asking how many colour words the child says). Each of these conditions acts to place fewer demands on the respondent's memory.

Given the success of the MCDI as a parent-report measure of young children's verbal cognitive abilities, we were curious to see if parent reports could be used to provide a valid measure of non-verbal cognitive abilities in early childhood. In contrast to current parent-report measures in which cognitive abilities are defined primarily in terms of sensorimotor skills or abilities (e.g. 'holds and drinks from cup'), we were interested in problem-solving or reasoning abilities that had no distinct motoric component. Moreover, in addition to parent ratings of non-verbal cognitive abilities, we included a series of parent-administered non-verbal problem-solving tasks that could easily be administered within the home situation.

Based on the demonstrated validity of the MCDI and of parent reports of development more generally, we hypothesized that parents of 2-year-old children could provide valid assessments of their children's non-verbal reasoning skills. Specifically, we predicted that the parent-report and parent-administered measures of non-verbal cognitive ability, though correlated, would make independent contributions to the prediction of cognitive ability on the Mental Development Index of the Bayley Scales of Infant Development-II (BSID-II MDI; Bayley, 1993), a standard tester-administered measure of cognitive development.

The Bayley is a broadly ranged test that is meant to sample both language and non-language abilities. Although the structure of cognitive abilities at this early age is far from clear, of all the distinctions hypothesized thus far, the distinction between language and non-language ('performance' in the Wechsler sense) skills appears to be the best established, and probably earliest to emerge (Lewis, 1983; Sattler, 1988). Guided by this distinction, we examined two additional questions: First, does using a parent-report measure of language development in addition to our parental measure of non-verbal cognitive abilities improve our ability to predict performance on the BSID-II MDI? That is, because the MDI at this age includes many language items, we hypothesized that the prediction of the MDI by parental measures would be even better if the parent-report MacArthur Communicative Development Inventory is used in conjunction with parental assessments of non-verbal cognitive abilities. The second additional research question asked if parents can discriminate their children's abilities in language and non-language domains, as opposed to responding globally to their children's development. In keeping with the hierarchical model of the structure of cognitive abilities (Brody, 1992; Carroll, 1993), we assumed that verbal and non-verbal measures are correlated due to the pervasive

influence of general cognitive ability; however, we also expected that verbal and non-verbal measures would make unique contributions to the prediction of cognitive development.

Method

Sample

The sample included 62 females and 45 males with an average age at time of testing of 2.2 years (SD = .26 years). The children were members of 43 twin pairs and 7 triplets obtained through the Multiple Births Foundation of London, UK. Twins were specifically recruited because it is an eventual goal to use parental reports of cognitive ability in the Twins' Early Development Study (TEDS), a study of all twins born in England and Wales in 1994, 1995 and 1996. For this reason, it was necessary to develop a parent-report measure of non-verbal cognitive ability and determine how well it would work when parents administered the measure to their twin children. A particular concern was to ensure that the parent-report measure be manageable for parents with two toddlers. As a stringent test of whether the familial nature of the data inflated the probability of associations presented in this paper, analyses were first conducted separately for a single member of each family (i.e. for twins A and twins B singly). These singleton analyses yielded results similar in terms of effect, size and significance to that of the full sample. Therefore, the use of twin and triplet siblings does not appear to pose a serious problem for this data and, consequently, the data was pooled for subsequent analyses.

The sample was almost exclusively Caucasian (96 per cent of mothers and 92 per cent of fathers) and was predominantly middle to upper-middle class according to standard occupational classification levels in the UK (Office of Population Census and Surveys, 1991).

Measures

Parent Report of Children's Abilities (PARCA). The Parent Report of Children's Abilities (PARCA), a parental assessment measure of non-verbal cognitive ability, was developed for use in the present study to assess the non-verbal cognitive abilities of 2-year-old children. This measure consists of a parent-report component and a parent-administered component. PARCA test materials include an instruction/answer booklet for each child and a set of 10 plastic blocks. The total administration time for PARCA is approximately one hour.

Parent-report component. The parent-report component consists of 26 questions assessing the areas of quantitative skills, spatial abilities, symbolic play, planning and organizing, adaptive behaviours and memory. Items for this component included modified items from existing parent questionnaires (e.g. the Minnesota Child Development Inventory, Ireton & Thwing, 1974; and the Ages and Stages Questionnaires, Bricker, Squires & Mounts, 1995), as well as original items written specifically for the PARCA. Because the measure was designed to assess non-verbal abilities, an effort was made to ensure that the items included in this component did not directly or indirectly assess language-related skills. There is some evidence that the degree of congruence between parents' and professionals' estimates of children's ability is higher when parents are required to report whether or not their child can perform a specific task (Dinnebeil & Rule, 1994; Sattler, Feldman & Bohanan, 1985). Therefore, questions were phrased in terms of specific 'activities', and parents were asked to indicate whether or not they had seen their child perform the activity (e.g. 'Can your child put a simple piece, such as a square or an animal, into the correct piece on a puzzle board?' or 'Does your child recognize himself/herself when looking in the mirror?'). Parents were encouraged to try a task if they were not certain whether their child could perform the task. Each 'Yes' response was given a score of 1. 'No' or 'don't know' responses were scored as 0.A total score for the parent-report component was derived by summing the scores for each question. Internal consistency, as estimated by Cronbach's alpha coefficient, was .74.

Parent-administered component. The parent-administered component was based, in part, on non-verbal items drawn from existing measures of cognitive ability, such as the Bayley Scales of Infant Development

(Bayley, 1969) and the McCarthy Scales of Children's Abilities (McCarthy, 1972), on the basis of validity, reliability, ease of administration and potential for enjoyment for parent and child. Items were modified to permit parent administration. In most cases, this involved revising the wording of instructions to be as specific and clear as possible. There were 24 items within four categories of tasks: (1) *design drawing* which involved the child copying a design (scribble, circle, horizontal line, vertical line) that had been demonstrated by the parent; (2) *match-to-sample* where the child was shown a shape and then asked to find the same shape among a series of four alternatives; (3) *block building* which required the child to arrange plastic blocks so that they matched structures demonstrated by the parent (row, tower, bridge, wall); and (4) *imitative action* where the child was asked to copy a series of one, two, three and four actions that were demonstrated by the parent (actions included folding a piece of paper, opening and closing mouth, pulling earlobe, blinking eyes and patting cheek). Each task started off with simple items and became progressively harder.

Parents were asked to follow the instructions to administer each item and to indicate the child's response. This included circling the child's answer out of a series of four possible choices (i.e. match-to-sample); providing a numerical count (e.g. 'How many blocks did your child put together?'); or answering 'Yes' or 'No' to specific questions describing the child's performance (e.g. 'Does your child copy all three actions?'). With this information, we were able to use the parent's responses in the instruction/answer booklets to score the child's performance on each task according to decision rules similar to those employed in standard tester-administered measures of cognitive ability. A total score for the parent-administered component was obtained by summing across the child's scores on each task. Internal consistency for the parent-administered component was .83.

Feasibility testing. A preliminary version of the PARCA was pilot tested at the Pennsylvania State University to determine the feasibility of a parent-administered measure of non-verbal cognitive ability that could be sent through the mail. Twenty-two parents, who were solicited by mail, administered the PARCA to their 2-year-old children and returned the test materials along with their comments regarding the activities or any difficulties that they had administering the tasks or understanding the instructions. In addition, the parent-report component was pilot tested on a sample of 45 parents of 24- and 30-month-olds who were participating in a language development study at the University of Washington. Based on the data and comments from the participating families, the PARCA was then revised to its final form. This revision included further simplification of the instructions to parents and the deletion of items and tasks that showed little variability. The description of the PARCA above, the aforementioned estimates of internal consistency and all analyses presented in this paper are for the final version of the PARCA.

The MacArthur Communicative Development Inventories: UKSF adaptation. The parent report measure of language used in this study was an adaptation of the MacArthur Communicative Development Inventories: Words and Sentences (MCDI:WS), a questionnaire that has been demonstrated to have high internal consistency and validity for children between 16 and 30 months (Fenson *et al.*, 1994). The MCDI:WS contains a checklist of 680 words to assess expressive vocabulary, and a set of 37 forced-choice sentence pairs to assess grammatical development. Fenson, Pethick & Cox (1994) have developed a 100-item short form vocabulary checklist with excellent prediction to the full vocabulary list. That list, with a few minor changes to 'anglicize' items, comprised the vocabulary scale of the new language measure, called the MacArthur Communicative Development Inventories: UK Short-form Version (MCDI:UKSF). Parents check the words which they have heard their child speak, and the number of positive responses is totalled. The Sentence Complexity Scale of the MCDI:UKSF was constructed by selecting 12 items from the full set of 37 forced-choice sentence pair items on the MCDI:WS. Items were selected on the basis of appropriate level of difficulty, developmental variation and item–whole correlations. Parents are asked to choose the sentence of each pair which best reflects the child's present level; for example, *baby crying* vs. *baby is crying*. The sentence complexity score is the number of pairs for which the parent selected the more complex alternative. Children whose parents indicated that they were not yet combining words were given a score of zero on this scale. Internal consistencies for the Vocabulary and Sentence Complexity Scales of the MCDI:UKSF were .98 and .93, respectively.

Mental scale of BSID-II. The Mental Scale of the Bayley Scales of Infant Development-II (BSID-II; Bayley, 1993) was used as a standard tester-administered measure of cognitive development. Although not

identical to its predecessor, the BSID-II retains the basic qualities of the earlier version while improving reliability and validity (Bayley, 1993; Nellis & Gridley, 1994). The Mental Scale is designed to assess memory, habituation, number concepts, classification, generalization, problem solving, language and social skills. The Mental Development Index (MDI), derived from the Mental Scale, provides a measure of infant cognitive development. The MDI is a normalized standard score with a mean of 100 and a standard deviation of 15 in the standardization sample.

In addition to the MDI score, two subscales were constructed from Bayley performance to index language and non-language abilities, following the approach of Dale, Bates, Reznick & Morisset (1989) for the earlier version of the Bayley. All items on the BSID-II Mental Scale are classified as cognitive, language, social and/or motor (Bayley, 1993, p. 53). All items classified as 'language' were first selected. Then, those 15 'language' items which were also listed as 'cognitive' were examined to determine their most appropriate placement; as a result, seven items were moved from the language scale to the non-language scale. All items not assigned to the language scale were placed on the non-language scale. A list of language items is provided in the Appendix.

Procedure

Parents were contacted by mail through the Multiple Births Foundation of London and asked to take part in a study that would compare the PARCA and a standard tester-administered measure of cognitive development. Families who agreed to participate were sent the PARCA test materials and the MCDI:UKSF. Parents were asked to administer the PARCA at a time when the children were alert and content and when there would be few distractions. It was not necessary for the PARCA to be administered all at once, and parents were advised to move on to another task if they felt that their child was becoming frustrated or bored with an activity. Although parents were permitted to encourage their children to do the best that they could, the instructions stressed that parents should not give the child any help on the tasks.

Several precautions were taken to discourage parents from overestimating their children's abilities. First, parents were informed that the purpose of the study was to evaluate the measure and not their children's performance. It was also emphasized that the items on the PARCA were meant for a range of ages and that no 2-year-old child would be able to perform successfully on all tasks. Finally, parents were aware that their children's performance on the PARCA would be compared to a standard developmental measure administered by a professional tester.

Approximately two weeks after the PARCA was sent, families were visited at home by two female testers trained in infant assessment. At this visit, the completed PARCA test booklets and MCDI:UKSF forms were collected and each twin within a family was individually assessed by a different tester using the Mental Scale of the Bayley Scales of Infant Development-II (BSID-II; Bayley, 1993).

Results

Descriptive statistics

Table 1 presents the means and standard deviations for the MDI, the PARCA parent-report (PR) and parent-administered (PA) components and the MCDI:UKSF vocabulary and sentence complexity scores for the full sample. Despite the fact that our sample was skewed somewhat toward higher socio-economic status, the mean and standard deviation for the MDI (M = 97.3, SD = 12.8, range 50–124) compare favorably to the expected mean of 100 and standard deviation of 15, providing reassuring evidence that our sample was not atypical.

All measures were significantly correlated with age at time of testing. Because the MDI is standardized for age, the significant correlation between age and MDI score was somewhat unexpected. However, a relation between age and MDI has been reported elsewhere (e.g. Reznick, Robinson & Corley, 1997; Saudino & Plomin, 1997). A possible

explanation for this finding is that the age standardization of MDI is not complete. The standardization of the BSID-II for ages 12 to 30 months is based on 3-month age spans, with scores for intermediate ages being derived from interpolation (Bayley, 1993). Moreover, standard scores are provided for one-month age intervals. Given the rapid rate of development during infancy and early childhood, it is likely that significant developmental change takes place within normative age groups. Age-adjusted scores for each measure were created by regressing the obtained score on age and using the residuals as an age-adjusted score. All subsequent analyses were conducted on age-adjusted scores for each measure so that correlations among the measures do not merely reflect age. Nonetheless, we would note that the fact that unresidualized PARCA scores correlated with age, despite the relatively narrow age range of the sample, provides some initial evidence for the validity of the measure.

With the exception of MCDI:UKSF sentence complexity, in which girls showed a small advantage (girls $M = 2.9$, boys $M = 1.8$, $t(105) = 2.07$, $p < .05$), there were no significant sex differences for any of the other measures (MDI: $t(105) = 1.38$, n.s.; PARCA:PR: $t(105) = 0.49$, n.s.; PARCA:PA: $t(105) = 1.16$, n.s.; MCDI:UKSF vocabulary: $t(105) = 1.32$, n.s.). Data from both sexes were therefore combined for all further analyses.

Table 1. Means (*M*) and standard deviations (*SD*) for the MDI, the PARCA parent-report (PR) and parent-administered (PA) components, and MCDI:UKSF scores before and after adjusting for age differences

	Score			Age-residualized score[a]	
	M	SD	r with age	M	SD
MDI	97.3	12.8	.32*	.00	12.1
PARCA parent-report	17.1	3.5	.52*	.00	3.0
PARCA parent-administered	16.1	7.3	.60*	.00	5.8
MCDI:UKSF vocabulary	56.0	26.2	.65*	.00	19.9
MCDI:UKSF sentence complexity	2.5	3.6	.64*	.00	2.8

*$p < .001$.
[a]Age-residualized scores were created by regressing the obtained score on age and using the residuals as an age-adjusted score.
Note. $N = 107$.

Hypothesis 1: Parental assessments of non-verbal cognitive ability predict BSID-II MDI scores

As indicated by the correlations presented above the diagonal in Table 2, both the PARCA parent-report scores and parent-administered scores significantly predicted performance on the MDI even though the MDI includes language as well as non-language tasks. Although parent-report scores were moderately correlated with parent-administered scores, multiple regression analyses revealed that each component

contributed uniquely to the prediction of cognitive development on the MDI (PARCA:PR: $B = 1.60, t(104) = 4.36, p < .0001$; PARCA:PA: $B = 0.54, t(104) = 2.91, p < .005$). This is further illustrated by the pattern of partial correlations presented below the diagonal in Table 2. After partialling out the effects of the parent-administered component, the parent-report measure continued to be significantly correlated with the MDI. Similarly, the significant relation between the MDI and the parent-administered component remained after the effects of the parent-report measure were removed. Moreover, the multiple correlation from the regression of the MDI on the two components of the PARCA was .55 ($p < .0001$). Thus, the combination of the parent-report and parent-administered components together are a better predictor of cognitive ability on the MDI than either component separately. Based on this, we created a total PARCA score by summing the scores from the two components. The correlation between the PARCA total score and the MDI was .52 ($p < .0001$).

Table 2. Zero-order and partial correlations between the PARCA parent-report (PR) and parent-administered (PA) components and the MDI

		PARCA	
	MDI	PR	PA
MDI	—	.49***	.41***
PARCA PR	.39***	—	.40***
PARCA PA	.27**	.24*	—

***$p < .001$; **$p < .01$; *$p < .05$.

Note. Zero-order correlations are in bold above the diagonal. Partial correlations, with the effects of the third variable removed, are below the diagonal.

Hypothesis 2: Adding parent-based measures of language improves the prediction of BSID-II MDI scores

Table 3 presents the intercorrelations between the MCDI:UKSF vocabulary and sentence complexity measures and the MDI and PARCA parent-report and parent-administered scores. Both the MCDI:UKSF vocabulary and sentence complexity were significantly correlated with Bayley MDI scores. However, only the MCDI:UKSF vocabulary scores were related to scores on the PARCA. Multiple regression analyses, including the two parent-based language measures in addition to the parent-report and parent-administered components of the PARCA, revealed that each makes a statistically significant unique contribution to the prediction of cognitive development on the MDI (MCDI:UKSF vocabulary: $B = 0.15, t(102) = 2.65, p < .01$; MCDI:UKSF sentence complexity: $B = 1.09, t(102) = 3.19, p < .01$; PARCA:PR: $B = 1.18, t(102) = 3.23, p < .01$; PARCA:PA: $B = 0.36, t(102) = 2.06, p < .05$). Again, this is further illustrated by the pattern of significant partial correlations between the parent measures and the MDI (see Table 3). That is, each parent measure continued to predict the MDI after the effects of the

other three parent measures were removed. More central to our hypothesis, the multiple correlation between the four parent measures and the MDI was .66 ($p < .0001$). Moreover, the improvement in prediction of the MDI when parent-based language measures were included as predictors in the multiple regression equation was significant (change in $R^2 = .131$, $F(4,100) = 5.77$, $p < .01$), thus confirming our hypothesis.

Because some investigators may prefer to use parent report alone, an additional multiple regression analysis was performed using the parent-report component of the PARCA and the vocabulary and sentence complexity scores from the MCDI:UKSF. These three parent-report measures yielded a multiple correlation with Bayley MDI of .64 ($p < .0001$).

Table 3. Zero-order and partial correlations between the MCDI:UKSF vocabulary and sentence complexity and the PARCA parent-report (PR) and parent-administered (PA) components, and Bayley MDI.

	MCDI:UKSF vocabulary	MCDI:UKSF Sentence complexity	Partial correlation with MDI
MDI	.51***	.38***	—
PARCA PR	.48***	.09	.30**
PARCA PA	.35***	.18	.20*
MCDI:UKSF vocabulary	—	.30**	.25**
MCDI:UKSF sentence complexity	—	—	.30**

***$p < .001$; **$p < .01$; *$p < .05$.

Note. $N = 107$. Zero-order correlations are in bold. The partial correlations are the correlations with the MDI after removing the effects of the other three parent measures.

Hypothesis 3: Parents can discriminate their children's language and non-language abilities

The third major research question for this study was the extent to which parents could discriminate language and non-language abilities. Given the substantial relation to be expected between skills in the two domains, this question is best answered by examining the pattern of correlations of the parent-based measures with Bayley language and non-language scores. For this purpose, the PARCA total score was used as a non-language measure. The MCDI:UKSF vocabulary scale was chosen as the language measure, because vocabulary is widely viewed as the prototypical verbal ability measure (Ammons & Ammons, 1958; Dunn & Dunn, 1981; Sattler, 1988).

The two subscales derived from the Bayley, language and non-language, were correlated .62 ($p < .001$), whereas MCDI:UKSF vocabulary and PARCA total score were correlated .46 ($p < .001$). The fact that the correlation between the domains based on parent measures was lower than the correlation based on tester-based measures might simply reflect lower reliability and/or validity for the former. However, an examination of

the correlations reported in Table 4 suggests that parents are, in fact, responding differentially to their children's ability in these two domains. Of particular interest are the partial correlations, which assess the unique contribution to the prediction made by each parent measure. That is, the correlation between the parent report of vocabulary and the Bayley language and non-language subscales was also examined with the parent-report component of the PARCA partialled out. Partialling the other parent measure removes the effect of general cognitive variance, as well as domain-specific variance, thus providing a more conservative estimate of the predictive power of each parent measure. The two within-domain correlations (.48, .39) are substantially greater than the two cross-domain correlations (.22, .25). These correlations were compared using a modification of Hotelling's T_1 test suggested by Steiger (1980). The MCDI:UKSF vocabulary measure is more highly correlated with Bayley language than with Bayley non-language ($T_2 = 2.55, p < .01$, one-tailed), whereas the PARCA is more highly correlated with Bayley non-language than with Bayley language ($T_2 = 1.80, p < .05$, one-tailed).

Table 4. Zero-order and partial correlations between parent-administered and tester-administered measures of verbal and non-verbal abilities.

	Bayley scale	
	Language	Non-language
MCDI:UKSF Vocabulary	.58***(.48***)	.42***(.25*)
PARCA Total	.43***(.22*)	.51***(.39***)

***$p < .001$; **$p < .01$; *$p < .05$.

Note. $N = 107$. Zero-order correlations are in bold. Partial correlations, with the effects of the other parent measure removed, are in parentheses.

Discussion

The present results indicate that parent reports and parent-administered tasks can be used to provide valid estimates of non-verbal cognitive abilities in 2-year-old children. The parent report and parent-administered components on the PARCA were equally good at predicting performance on the MDI. Moreover, each makes a unique contribution to the prediction of cognitive development as assessed on the MDI. Including both components increases the validity of the measure.

The finding of a multiple correlation of .55 between the parent assessments of non-verbal cognitive ability on the PARCA and the MDI, administered approximately two weeks later, needs to be put into context of other results at this general age level for proper interpretation. For example, the correlation between the MDI of the BSID-II and its predecessor, the BSID (Bayley, 1969), is only .62 (Nellis & Gridley, 1994). This moderate correlation is particularly noteworthy when one considers that the BSID and the BSID-II share a common 'nature and purpose' (Bayley, 1993, p. 1). Moreover, a majority (76 per cent) of the mental scale items on the BSID were retained on the BSID-II (Bayley, 1993).

Thus, the finding that the PARCA correlates with the BSID-II almost as well as the BSID-II correlates with its mother scale is very encouraging.

We are also encouraged by the fact that the correlation between the PARCA and the MDI is similar to correlations between the MDI and other standardized tester-administered measures of non-verbal ability. For example, the performance scales of the Wechsler Preschool and Primary Scale of Intelligence–Revised (Wechsler, 1989) and the McCarthy Scale of Children's Abilities (McCarthy, 1972) and the non-verbal scale of the Differential Ability Scales (Elliott, 1990) correlated .63, .69 and .30, respectively, with the MDI (Bayley, 1993). Within the context of these findings, the ability of parents' responses on the PARCA to predict performance on the MDI is impressive.

As expected, the inclusion of a parental measure of language development in addition to the non-verbal PARCA measure significantly increased the prediction of the MDI. The multiple correlation of .66 between the combined PARCA/MCDI:UKSF battery and the MDI is also similar to the correlation between the MDI and full-scale scores on other standardized tester-administered measures of cognitive ability. For example, the Wechsler Preschool and Primary Scale of Intelligence–Revised (Wechsler, 1989), the McCarthy Scale of Children's Abilities (McCarthy, 1972) and the Differential Ability Scales (Elliott, 1990) have been found to correlate .73, .79 and .49, respectively, with the MDI (Bayley, 1993). The present study clearly demonstrates the concurrent validity of parent-based measures of cognitive abilities. It remains for future research, however, to determine the relative ability of the MDI and parent-based measures to predict cognitive abilities in later childhood.

In addition to providing information that was substantially correlated with a global developmental measure (MDI), parents also demonstrated some ability to discriminate their children's abilities in language and non-language domains. The discrimination was reflected in the modest correlation between language and non-language measures, together with a pattern of significantly higher within-domain predictions than cross-domain correlations (Table 4). It should be acknowledged that the key difficulty in investigating this question is the fact that the 'real' correlation between these domains—the strength of 'g' at this age remains unknown at present due to limitations of both conceptualization and measurement. In the absence of this information, the present data cannot be used to evaluate how fully parents discriminate these skills, they demonstrate only that some, statistically significant, discrimination is being made. It is also of interest that the correlation between language and non-language measures was in fact smaller (.46) for the parent measures than for the tester-administered measures (.62). One interpretation of this difference is that it reflects lower reliability and/or validity for the parent measures. However, the coefficient alpha values for parent report and parent-administered components of the PARCA are .74 and .83. These are somewhat lower than the corresponding value for the full MDI at 24 months, .92; but it is likely the alpha for the two Bayley subscales, which are shorter, are not substantially different from those for the PARCA. Furthermore, both the PARCA and the MCDI:UKSF show validity correlations comparable to those for structured tests. We suggest that it is equally likely that the correlation based on performance on the subscales of the Bayley is inflated because of the shared task requirements of performance on a structured, adult-directed measure. Because parents can aggregate across a much wider range of situations and contexts to formulate their estimates, they may in fact be *more* valid in their ability to discriminate

domains. The nature of the measures may also contribute to superior discrimination of language and non-language abilities by parents; whereas the Bayley language and non-language scales were essentially formed *post hoc*, the PARCA and MCDI:UKSF were specifically designed for assessment of these two domains.

It should be acknowledged that the validation of the parental assessment of cognitive abilities in the present study represents validity under somewhat optimal conditions. That is, the parents in our sample were skewed in favor of higher socio-economic status and were aware that the purpose of the study was to evaluate the measure and not their children. The generalizability of the results to other contexts will require additional research. (It should be noted that these parents were assessing two or even three children, which may have decreased validity to some extent. In the absence of any kind of feedback, it is unlikely that practice effects would improve validity over repeated administration.) Parental assessments, although valid, may nevertheless be prone to bias. For example, parental intelligence or education may influence the assessments. The modest amount of evidence available on parental education and validity of parent report (Diamond & Squires, 1993; Dale, 1996) suggests that parent report is valid over a wide social class range, though it may be reduced for parents with less than a sixth-form education. Similarly, the accuracy of parental reports might be distorted in certain clinical contexts when parents may have some desire to prove that their child does not have a problem or, in certain educational contexts, parents might be motivated to exaggerate their children's abilities. This, however, did not appear to be the case for parents of children with Down's syndrome (Miller, Sedey & Miolo, 1995) or with specific language impairment (Thal & Tobias, 1994); the validity correlations in each of these samples were as high as in studies with normally developing children. Finally, there is a lack of information regarding the role of potential cultural differences in affecting parent report. There are well-documented differences in views of children and childrearing, views of disabilities and their causes, and views of medicine and healing that might affect either parental observation or reporting of their children's abilities (Hanson, Lynch & Wayman, 1990). We would caution that the use of parent-report measures, such as the PARCA, as clinical instruments are inappropriate until the issues of potential parental bias are adequately addressed. Clearly, we were not able to explore these issues in our non-clinical sample, but the demonstrated validity of the PARCA in the present study suggests that parent reports of non-verbal cognitive ability will be useful as a research tool for screening cognitive development in young children.

The Bayley scales have been considered to be the best available measure of infant development (Sattler, 1988) and have been used as a 'gold standard', against which researchers have evaluated other measures of cognitive development (Harris, 1994). However, although administered by trained testers, such tests are conducted by strangers during a brief testing session and, therefore, may not provide a full picture of the child's abilities. Moreover, such measures are expensive in terms of cost and labour. Parents are an untapped resource with regard to their knowledge about their children's cognitive abilities and disabilities. The success of the PARCA and the MCDI suggests that parental assessments can be profitably employed to provide a valid and reliable screening measure of cognitive development in early childhood. Furthermore, the present results demonstrate that, given appropriately specific questions, parents can go beyond a global assessment to providing information about specific domains of development, such as

language and non-language skills. Based on the success of the PARCA for assessing the non-verbal cognitive abilities of 2-year-olds, we are currently developing extensions of the PARCA for use with 3- and 4-year-old children.

Acknowledgments

This research was supported by the UK Medical Research Council and by grants HD-27694 and HD-18426 from the US National Institute of Child Health and Human Development. The authors would like to thank Dorothy Bishop, Mark Greenberg and Nancy Robinson for their valuable suggestions during the development of the PARCA and the preparation of this manuscript. Thank you also to Elizabeth Bryan, Carol Sweetland and the Multiple Births Foundation of London for their assistance in obtaining subjects. A copy of the PARCA can be obtained at cost by contacting Bonny Oliver.

References

Achenbach, T.M. & Edelbrock, C. (1983). *Manual for Child Behavior Checklist and Revised Child Behavior Profile*. Burlington, VT: Achenbach.

Ammons, R.B. & Ammons, C.H. (1958). *Quick Test*. Missoula, MT: Psychological Test Specialists.

Bayley, N. (1969). *Bayley Scales of Infant Development*. New York, NY: The Psychological Corporation.

Bayley, N. (1993). *Bayley Scales of Infant Development, 2nd ed.*. San Antonio, TX: Psychological Corporation.

Bricker, D. & Squires, J. (1989). The effectiveness of parental screening of at-risk infants: The Infant Monitoring Questionnaires. *Topics in Early Childhood Special Education*, 9, 67–85.

Bricker, D., Squires, J. & Mounts, L. (1995). *Ages and Stages Questionnaires: A Parent-Completed, Child-Monitoring System*. Baltimore, MD: Brooks.

Brody, N. (1992). *Intelligence*, 2nd ed. New York: Academic Press.

Carroll, J.B. (1993). *Human Cognitive Abilities*. New York: Cambridge University Press.

Dale, P.S. (1996). Parent report assessment of language and communication. In K.N. Cole, P.S. Dale & D. Thal (Eds), *Assessment of Communication and Language*. Baltimore, MD: Brookes.

Dale, P.S., Bates, E., Reznick, J.S. & Morisset, C. (1989). The validity of a parent report instrument of child language at 20 months. *Journal of Child Language*, 16, 239–250.

Diamond, K.E. & Squires, J. (1993). The role of parental report in the screening and assessment of young children. *Journal of Early Intervention*, 17, 107–115.

Dinnebeil, L.A. & Rule, S. (1994). Congruence between parents' and professionals' judgments about the development of young children with disabilities: A review of the literature. *Topics in Early Childhood Special Education*, 14, 1–25.

Dunn, L.M. & Dunn, L.M. (1981). *Peabody Picture Vocabulary Test–Revised*. Circle Pines, MN: American Guidance Service.

Elliott, C.D. (1990). *Differential Ability Scales*. San Antonio, TX: Psychological Corporation.

Fenson, L., Dale, P.S., Reznick, S.J., Bates, E., Thal, D.J. & Pethick, S.J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, 59 (Serial No. 242).

Fenson, L., Pethick, S. & Cox, J.L. (July, 1994). The MacArthur Communicative Development Inventories: Short Form Versions. Unpublished manuscript, San Diego State University.

Goldsmith, H.H. (1996). Studying temperament via construction of the Toddler Behavior Assessment Questionnaire. *Child Development*, 67, 218–235.

Gradel, K., Thompson, M. & Sheehan, R. (1981). Parental and professional agreement in early childhood assessment. *Topics in Early Childhood Special Education*, 1, 31–39.

Hanson, M.J., Lynch, E.W. & Wayman, K.I. (1990). Honoring the cultural diversity of families when gathering data. *Topics in Early Childhood Education*, 10, 112–131.

Harris, S.R. (1994). Parents' and caregivers' perceptions of their children's development. *Developmental Medicine and Child Neurology*, 36, 918–923.

Hogan, A.E., Scott, K.G. & Bauer, C.R. (1992). The Adaptive Social Behavior Inventory (ASBI): A new

assessment of social competence in high-risk three-year-olds. *Journal of Psychoeducational Assessment*, **10**, 230–239.

Ireton, H. & Thwing, E. (1974). *The Minnesota Child Development Inventory*. Minneapolis MN: Behavior Science Systems.

Lewis, M. (1983). *Origins of Intelligence: Infancy and Early Childhood*. New York: Plenum.

McCarthy, D. (1972). *McCarthy Scales of Children's Abilities*. New York: The Psychological Corporation.

Miller, S.A. (1986). Parents' beliefs about their children's cognitive abilities. *Developmental Psychology*, **22**, 276–284.

Miller, S.A., Manhal, M. & Mee, L.L. (1991). Parental beliefs, parental accuracy, and children's cognitive performance: A search for causal relations. *Developmental Psychology*, **27**, 267–276.

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

Nellis, L. & Gridley, B.E. (1994). Reviews of the Bayley Scales of Infant Development, 2nd ed. *Journal of School Psychology*, **32**, 201–209.

Office of Population Census and Surveys (1991). *Standard Occupational Classification*. Vol. 3: *Social Classifications and Coding Methodology*. London: Her Majesty's Stationery Office.

Reznick, S.J., Robinson, J. & Corley, R. (1997). A longitudinal twin study of intelligence in the second year. *Monographs of the Society of Research in Child Development*, (Serial No. 249, vol. 2, No. 1).

Reuter, J. & Bickett, L. (1985). *The Kent Infant Development Scale Manual*, 2nd. ed. Kent, OH: Kent Developmental Metrics.

Sattler, J.M. (1988). *Assessment of Children*, 3rd ed. San Diego, CA: J.M. Sattler.

Sattler, J.M., Feldman, J. & Bohanan, A.L. (1985). Parental estimates of children's receptive vocabulary. *Psychology in the Schools*, **22**, 303–307.

Saudino, K.J. & Plomin, R. (1997). Cognitive and temperamental mediators of genetic influences on the home environment during infancy. *Merrill-Palmer Quarterly*, **43**, 1–23.

Sheehan, R. (1988). Involvement of parents in early childhood assessment. In T.D. Wachs & R. Sheehan (Eds), *Assessment of Young Developmentally Disabled Children*. New York, NY: Plenum.

Steiger, J.H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, **87**, 245–251.

Thal, D. & Tobias, S. (1994). Relationships between language and gesture in normally developing and late-talking toddlers. *Journal of Speech and Hearing Research*, **37**, 157–170.

Waters, E. & Deane, K.E. (1985). Defining and assessing individual differences in infant attachment relationships: Q-methodology and the organization of behavior. In I. Bretherton & E. Waters (Eds), *Growing Points of Attachment Theory and Research*, pp. 41–65. *Monographs of the Society of Research in Child Development*, **50**, (Serial No. 209).

Wechsler, D. (1989). *Wechsler Preschool and Primary Scales of Intelligence–Revised*. San Antonio, TX: Psychological Corporation.

Received 13 January 1997: revised version received 22 October 1997

Appendix: BSID-II items comprising the language scale[1]

3	101	131	~~177~~
10	106	134	
21	107	136	
22	108	140	
31	109	141	
33	110	142	
61	111	144	
63	113	148	
68	114	~~152~~	
70	117	153	
71	~~118~~	~~154~~	
76	121	156	
78	122	159	
81	124	~~164~~	
94	126	~~166~~	
99	127	167	
100	129	~~172~~	

[1] Items printed in strikeout are classified as both language and cognition in Bayley (1993) and were placed in non-language for the present study.