Cost-Effective Prediction of Reading Difficulties

Steve M. Heath
John H. Hogben
The University of Western Australia,
Crawley, Australia

This study addressed 2 questions: (a) Can preschoolers who will fail at reading be more efficiently identified by targeting those at highest risk for reading problems? and (b) will auditory temporal processing (ATP) improve the accuracy of identification derived from phonological processing and oral language ability? A sample of 227 preschoolers was screened for Performance IQ and was tested on phonological awareness (PA). The upper and lower quartiles of the PA distribution were selected as being at lowest and highest risk, respectively, for reading failure. Children with good and poor PA were tested on ATP, phonological short-term memory, rapid automatized naming, oral language, receptive vocabulary, and 2 measures of listening comprehension. Reading outcomes were measured at the end of Year 2. Only 1 child in the good-PA group became a poor reader by the end of Year 2, confirming that being in the top quartile for PA predicts positive reading outcomes. Discriminant analysis using the authors' test battery within the poor-PA group identified poor readers with sensitivity of .91 and specificity of .84, but ATP did not improve classification accuracy afforded by phonological and oral language. A brief screening procedure was formulated using only PA, phonological short-term memory, and demographic variables, with which 80% of children with poor PA who are at risk of reading problems can be identified. Further refinements of this screening procedure would increase accuracy of identification at the cost of only a small increment in required testing time.

KEY WORDS: auditory temporal processing, reading difficulties, dyslexia, prediction, specific language impairment, phonological awareness, phonological short-term memory, phonological retrieval

Between 3% and 10% of school-age children fail to learn to read efficiently, in spite of normal intelligence and adequate educational and socioeconomic opportunity (Shaywitz, 1998; Snowling, 2000). Such surprising reading failure is sometimes called *dyslexia* or *specific reading disability*. Children with severe reading difficulties are denied an essential skill that allows their classmates access to increased standards of education and technology. The academic and later vocational success of these children is likely to be considerably reduced if they do not receive extra support both at home and at school (Maughan, 1995; Stanovich, 1986). Reading failure has also been found to impact negatively on children's social and emotional development (Edwards, 1993). In fact, Prior, Sanson, Smart, and Oberklaid (1995) contended that without effective early intervention, reading failure results in a wide range of negative outcomes for the individual and society. However, these authors also showed that successful reading development in early primary school was a protective factor in children with serious behavioral problems at school entry.

It therefore seems of critical importance to develop a means of accurately identifying children at risk of reading difficulties as early as possible in their development. An extensive list of factors that might predict reading outcomes has been explored in the past several decades (for reviews, see Adams, 1990; Horn & Packard, 1985). The usual approach has been to test groups of prereaders on batteries of developmental and neuropsychological tests, then measure reading between 1 and 3 years later. Either a multiple regression or classification approach is then used to derive an appropriate function for prospective identification of children at risk of reading difficulties. However, at present we know of no reliable and cost-efficient predictive combination that could be widely used in schools for early identification of children at risk. As a result, most children are currently still not diagnosed until they are in third grade (Lyon, Alexander, & Yaffe, 1997).

One very promising line of research has examined alphabetic knowledge of letter names and sounds and phonological processing, which involves the use of information about the speech sound structure of language (Catts, 1989). Letter name knowledge was identified early as a strong predictor of reading achievement (Bond & Dykstra, 1967; Chall, 1967). Since that time, a number of studies (e.g., Mann, 1984; Vellutino & Scanlon, 1987) have confirmed the strength of letter name and sound knowledge as predictors of reading achievement even as far as the seventh grade. The aspect of phonological processing that has been most fully explored as a predictor of reading achievement is phonological awareness (PA), which refers to explicit awareness of and access to the sounds of spoken language (Wagner & Torgesen, 1987). PA has been shown in a range of studies to account for as much or more of the variance in reading outcomes than any other predictor (e.g., Felton & Brown, 1991; Mann, 1984; Share, Jorm, MacLean, & Matthews, 1984). Two other aspects of phonological processing—phonological short-term memory and retrieval of phonological information from long-term memory—have also been found to be strong predictors of reading achievement (e.g., Badian, McAnulty, Duffy, & Als, 1990; Blachman, 1984; Torgesen, Wagner, & Rashotte, 1994).

In light of these data, the majority of recent predictive studies have focused on phonological processing and alphabetic knowledge, with consistent findings that these factors measured before the start of reading education reliably predict outcomes in early primary school (e.g., Majsterek & Ellenwood, 1995; Muter, Hulme, Snowling, & Taylor, 1998; Swank & Catts, 1994) and even over extended periods of follow-up (e.g, MacDonald & Cornwall, 1995; Stuart & Masterton, 1992). In the main, these factors have been explored by the multiple regression approach to prediction. With this approach, the amounts of variance in reading achievement accounted for so far have typically varied between 40% and 60% (e.g., Muter, 1996; Share et al., 1984).

Somewhat better results have been obtained using the classification approach, which is also more suited to early identification of individuals at risk. In classification studies, the critical results are (a) the sensitivity of the function, which is measured by the proportion of children initially identified as at risk who become poor readers (i.e., valid positives), and (b) the specificity of the classification strategy, which is assessed by the proportion of the total number of children classified as not at risk who do not become poor readers (i.e., valid negatives). Rates of sensitivity and specificity have ranged between .85 and .90 in studies classifying children with significant reading deficits (e.g., Badian, 1994; Felton & Brown, 1991).

Such rates seem very promising because of the urgency and importance of identifying children at risk. However, we were concerned about a number of issues in the existing research. First, alphabetic knowledge taps print-related knowledge. This knowledge would clearly be subject to the availability of reading materials in the child's home environment and also to variations in preschool teaching practice and resources. It seemed important to predict reading outcomes using purely prereading variables. Second, two additional potential predictors, oral language ability and auditory temporal processing (ATP), had emerged from the literature preceding the present study. Although both indices appeared likely to improve the accuracy of prediction, neither factor seemed to have been fully explored in the classification research. Third, although it would be generally acknowledged that children with good PA are unlikely to have poor reading, just how unlikely this was had not been established. Neither was it known to what extent good PA may be considered a predictor of positive outcomes, nor had strength in oral language been examined in this way. Fourth, we know that a high proportion of children with poor PA will develop reading difficulties, but we still do not know how to determine which individual children will become poor readers (Bishop, 1991). Finally, in spite of the considerable knowledge now available about the precursors of reading failure, educators still were lacking a cost-effective method for identifying children with poor PA who will develop reading difficulties.

For some years Tallal and colleagues (e.g., Tallal, Miller, Jenkins, & Merzenich, 1997) have advanced the proposition that ATP might be causally related in some children to reading difficulties and specific language impairment (SLI). It is proposed that poor ATP could result in phonological processing deficits that would cause language and literacy problems. Children experiencing difficulties with their reading, their oral language development, or both have been shown to have ATP deficits (Farmer & Klein, 1993; Heath, Hogben, & Clark, 1999; Reed, 1989; Stark, Tallal, & McCauley, 1988; Tallal, 1980; Tallal & Piercy, 1973; Wright et al., 1997). Furthermore, specialized training in ATP of nonverbal and verbal stimuli has been reported to produce amelioration of reading and oral language deficits (Merzenich et al., 1996; Tallal et al., 1996). These findings suggested ATP should be explored as a predictor of reading.

To date, only two groups have investigated ATP measures of any sort. Hurford and colleagues (e.g., Hurford et al., 1993; Hurford, Schauf, Bunce, Blaich, & Moore, 1994) and Share et al. (1984), using rapid discrimination of synthesized speech and the Tallal Repetition Test, respectively, obtained moderate correlation coefficients (.25 to .39) with reading. Neither of these groups used a comprehensive measure of oral language in combination with ATP. Indeed, despite the growing awareness that oral language deficits are potentially related to reading difficulties, aspects of oral language apart from phonological processing have not been included in recent predictive batteries. This is surprising because studies that have investigated the predictive power of syntactic or semantic abilities have generally found these measures to be related to reading outcomes (e.g., Bishop & Adams, 1990; Scarborough, 1990, 1991; Stothard, Snowling, Bishop, Chipchase, & Kaplan, 1998). In addition, retrospective examination of a large number of reading-disabled children and children with SLI, who were tested in our own laboratory, revealed substantial comorbidity between the two disorders. We further demonstrated that at least 50% of the reading-disabled children had a standard score on a clinical oral language measure that was at least 1 *SD* below the population mean for their age (McArthur, Hogben, Edwards, Heath, & Mengler, 2000).

This project aimed to address the concerns outlined above. We sought to maximize efficiency of identification of children at risk by examining predictors in groups of preschool children whom the literature suggests are likely to be at highest and lowest risk for reading difficulties. We acknowledged the demonstrated power of PA as a major predictor by using this factor to select our highest and lowest risk groups (i.e., children with poor and good PA, respectively). We then attempted to further refine prediction after PA is taken into account. We piloted this approach in a study by Heath and Hogben (2000) and then refined our methodology in the study described here.

Our aims in this study were to (a) discover whether ATP and oral language ability, in combination with the two other components of phonological processing identified in the literature apart from PA (i.e., phonological short-term memory and phonological retrieval from long-term memory), could be used to improve at-risk predictions without needing to include alphabetic knowledge; (b) confirm the suggestion from the literature that PA and the other phonological processing factors, along with oral language, might be considered predictors of positive reading outcomes; and (c) use these predictors as the basis for a screening procedure in children at highest risk before the commencement of formal literacy education.

Method

Design

The study was conducted in two phases. In Phase 1, preschoolers were screened during the second half of the school year in Terms 3 and 4 and divided into two groups on the basis of their PA ability. Children were then tested on ATP, oral language, phonological retrieval, and phonological short-term memory. In Phase 2, reading outcomes were measured 2 years later at the end of Year 2.

Participants

Participants were selected according to the exclusionary criteria generally used for selection of dyslexics (adapted from Critchley, 1970, and described in detail by Heath & Hogben, 2000). In sum, children were excluded in the presence of any factor likely to cause reading failure, apart from poor PA. They were also required not to be functional readers at the commencement of the study. Because of the concerns expressed by Heath et al. (1999) about the possible interaction of Performance IQ (PIQ) with ATP, PIQ was rigorously controlled by Heath and Hogben (2000), who required a minimum score of 100. However, we decided that much greater external validity could be achieved in this study by including children with a PIQ of 85 or above.

A total of 262 children from 25 schools were screened as described by Heath and Hogben (2000).[1] Of these, 30 children did not fit the guidelines and 5 children moved to schools not included in the study, leaving 227 children to be tested for PA. Among these children, 114 were identified as having either good or poor PA (60 in the good-PA group, including 29 males and 31 females, and 54 in the poor-PA group, including 30 males and 24 females), and they were assessed on the Performance scale of the Wechsler Preschool and Primary Scale of

[1]For this study, participants were required to have bilateral pure-tone conduction thresholds of 25 dB HL rather than 20 dB HL. This less-stringent hearing criterion was introduced because children's hearing was assessed in situ at the children's schools. Because this study drew small numbers of children from a large number of schools, reliable low-noise conditions could not be negotiated in each school, with the result that ambient noise levels were considerably in excess of those in the pilot study.

Intelligence–Revised (WPPSI-R; Wechsler, 1989). Six of the children in the poor-PA group had a PIQ lower than 85 and were excluded from the study. Two children in the good-PA group moved away from metropolitan Perth and could not be followed up in Year 2. Children who were absent when a particular measure was given were routinely followed up on that measure. In rare instances where this was not possible (N = 3 in any one set of scores), the missing value has been replaced by the mean for the group to which the child belonged.

The two groups differed significantly in both age at the end of preschool and the parents' level of education, but the actual differences were small. For age at the end of preschool, the mean for the good-PA group was 67.3 months (SD = 3.2), and for the poor-PA group it was 65.6 months (SD = 3.5), $t(106) = 2.58$, p = .011, two-tailed. For parents' mean years of education, the mean for the good-PA group was 13.3 years (SD = 2.0), and for the poor-PA group it was 12.1 years (SD = 2.0), $t(98)$[2] = 2.86, p = .005, two-tailed. The groups did not differ significantly in PIQ: good-PA mean = 110.8 (SD = 12.9); poor-PA mean = 106.7 (SD = 11.7).

Psychometric Instruments

Phonological Awareness

All 227 children were mass tested with the Test of Phonological Awareness (TOPA; Torgesen & Bryant, 1994). This test presents two categorization tasks in pictorial form. Children are required either to select from three images the one that starts with the same sound as the target or to identify the image that starts with a different sound from three other pictures. Children at or above the 75th percentile were included in the good-PA group. The poor-PA status of children in the bottom quartile of the TOPA distribution then was cross-validated by administration of the Blending and Deletion tasks described by Swank and Catts (1994). These tasks represented both simple and compound PA tasks as recommended for best prediction by Yopp (1988). If the child's mean performance on these two tasks was equivalent to a percentile rank of 25 or less, then the child was included in the poor-PA group.

Oral Language Ability

Oral language abilities were measured using the six core subtests of the Clinical Evaluation of Language Fundamentals–Revised (CELF-R; Semel, Wiig, & Secord, 1987). This battery examines receptive and expressive language, in tasks tapping phonological short-term memory and semantic, syntactic, and morphological abilities. Listening comprehension was assessed with three measures: (a) the Peabody Picture Vocabulary Test–Revised (PPVT-R; Dunn & Dunn, 1981), for which the child has to select from four pictures the image that best represents a target word; (b) the Listening to Paragraphs supplementary subtest from the CELF-R, which requires children to listen to short paragraphs then answer questions about what they have heard; and (c) the Sentence Structure subtest from the CELF-R, on which children have to select from four pictures the image that best represents a target sentence (this subtest is part of the core battery but was analyzed separately for this purpose).

Phonological Retrieval

Phonological retrieval from long-term memory was measured with the Rapid Automatized Naming (RAN) subtest from the draft form of the Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgesen, & Rashotte, 1999). On this test, a child is presented with a chart, on which simple line drawings of six everyday objects are repeated six times. The child is required to name each object as fast as possible. Standard scores were generated from the summary data supplied with the draft of the RAN subtest.

Phonological Short-Term Memory

Phonological short-term memory for novel, semantically meaningful material was assessed with separate analysis of the Recalling Sentences subtest from the CELF-R battery. In this subtest, children have to repeat sentences of increasing length until a ceiling level is attained.

Reading Measures

Word recognition was tested with the Word Identification subtest from the Woodcock Reading Mastery Tests–Revised (WRMT-R; Woodcock, 1987) Form H, and phonological decoding or nonword reading was assessed with the Word Attack subtest from the WRMT-R Form H. Each of these subtests consists of a graded list of words. On the Word Identification subtest, children have to read aloud a mixture of regular and irregular real words, and on Word Attack, they have to pronounce nonsense words.

Reading accuracy and comprehension of meaningful text were measured with the Neale Analysis of Reading Ability–Revised (NARA-R; Neale, 1988) Form 1. The NARA-R comprises six graded passages, which children have to read aloud until a ceiling level is reached. Reading accuracy ages are based on the number of errors made, and reading comprehension is based on questions pertaining to each passage.

[2]Degrees of freedom reflect the failure of eight parents to respond to the parent questionnaire from which parents' mean years of education was calculated, after reasonable follow-up (1 parent from the good-PA group and 7 from the poor-PA group).

For all tests, reading achievement was expressed by comparison to the level expected for age. For children behind this level, performance was expressed as a lag of reading age behind chronological age. Therefore, for children in advance of the level expected for their chronological age, performance is shown as a negative lag.

ATP

We measured ATP with an adaptive version of the Tallal Repetition Test (TRT) which was piloted by Heath and Hogben (2000). This task examines ATP of rapid sequences by presenting a child with two nonverbal complex sounds of high and low pitch. The child has to identify the tones and specify the order in which they occurred. The child's accuracy in identifying and sequencing the tones is then examined as a function of the interstimulus interval (ISI).

Stimuli and Apparatus

The basic stimuli and apparatus for the perception of rapid sequences task were retained as described by Heath and Hogben (2000). The two sounds, produced using the FM chip of a Sound Blaster 2.0, each comprised four equal-amplitude frequency components. Three frequencies (491.5, 736.1, and 1471.0 Hz) were common in both sounds, and to these were added a frequency of 97.7 Hz for the low tone or 309.8 Hz for the high tone. The sounds closely approximated the description given by Tallal (1976). The amplitude of each sound at stimulus onset increased from zero to a steady maximum over a 10-ms interval and then decreased to zero over a similar period 85 ms after onset. The two tones were paired at random in each of four possible combinations (high–high, low–low, high–low, and low–high), and the child was trained to identify the sounds and their order of presentation.

The tones were generated by computer. Stimuli in the practice phases were presented through loudspeakers and in the test phases stimuli were presented binaurally through headphones at a comfortable listening level (mean intensity across both ears = 76.1 dB SPL). The task was administered on a computer, which was "dressed" in a life-sized clown costume. Children used a button box to register their responses and received nonverbal feedback about the correctness of their responses on the computer screen.

Procedure for Administration of the Task

The task consisted of three practice phases followed by two testing phases. In the first two practice phases, children learned to recognize the two separate tones. In the first phase, the low tone was presented as a truck noise, and the children pressed a button labeled with a truck picture to register their response. During Phase 2, the high tone was paired with a train picture on the button box in a similar way. Children were trained to identify the separate sounds until they could correctly identify 20 consecutive sounds. In the final practice phase, children learned to identify and sequence both tones together with a fixed ISI of 500 ms. The task was presented as a race between the truck and the train. Children had to press the correct buttons to indicate what sounds they heard and which was coming first. The response was demonstrated on the first four trials by the experimenter, and children then completed eight training trials with feedback.

The first testing phase was a criterion phase during which the interval between the tones remained fixed at 500 ms. Participants were required to reach a criterion level of 18 out of 24 correct responses before proceeding further. If children could not reach criterion initially, then they repeated the three practice phases before attempting the criterion phase again. If, after this additional practice, children still got less than 16 correct on the criterion phase, testing was discontinued. If they scored 16 or 17 correct on the criterion phase, they were given up to three more attempts to reach criterion. If these children were still unable to do so, testing was discontinued.

In the second test phase, the task was identical to that for the criterion phase, except that the ISI was varied systematically by using a Parameter Estimation by Sequential Testing (PEST; Taylor & Creelman, 1967) procedure. The PEST algorithm adaptively changes both step direction and size to place trials at the most efficient location on the stimulus axis and converge on a threshold level in a minimum number of trials. Threshold was then calculated by taking the mean of ISIs tested on trials occurring after the fourth reversal of step direction. A ceiling of 500 ms was imposed on the ISI because all participants attempting the PEST procedure had demonstrated at least 75% accuracy at this level in the criterion phase.

Children performed blocks of 100 trials and received stickers and small prizes for their participation. To control for possible effects of exposure to the task and practice effects, the first 100 trials were treated as practice trials. All children completed a minimum of 200 trials. Children were required to produce a stable PEST function on the second block of 100 trials (i.e., a relatively smooth function with no clear trend in either direction for at least the last 30 to 40 trials) for their threshold estimates to be included in the analyses. If the function from this second block was still unacceptable, then each child completed another block of 100 trials and the estimate from the third block was used. Data from unsatisfactory blocks of trials were not included in the analyses. Two children were still unable to perform satisfactorily after 300 trials, but no further testing was undertaken on the task with these participants.

Results

Description of the Data and Comparisons Between the Groups on the Predictor Variables

ATP

Threshold distributions for the perception of rapid sequences on the TRT are presented in Figure 1, which shows that these ATP results were positively skewed. There was a significant between-group mean difference on the task, but very large within-group variability, which created considerable overlap across the range of values for the two groups (see Table 1). The great majority of good-PA children (i.e., children who might be predicted to have strong reading) had threshold values under 200 ms, and indeed it appeared that a number may even have been at floor on this task. The two good-PA children in this group with thresholds above 400 ms were clearly aberrant performers. By contrast, 55.5% of the poor-PA group were poor performers (either unable to reach criterion or had a threshold well outside the 0–200 ms range of the majority of children). Thus, in these data, the poor-PA children were less competent at ATP and less amenable to benefit from practice.

Oral Language and Phonological Processing

Table 1 shows that there were very significant between-group differences on the oral language variables. In the poor-PA group, 54.2% of children had CELF-R Total Language scores of more than one standard deviation below the population mean, compared with only 6.8% of good-PA children. Similar but less divergent distributions were observed for the three listening comprehension measures. Table 1 also shows that there were highly significant differences between the groups in phonological short-term memory. There was a very clear separation of the groups, as almost half the good-PA children (i.e., 47.4%) scored at least one standard deviation above the mean, but only 1 child in the poor-PA group (out of 48) performed this well. There was more overlap

Table 1. Group mean values for both groups on prereading and reading outcome measures.

Variable	Good-PA group (N = 60) M	Good-PA group (N = 60) SD	Poor-PA group (N = 48) M	Poor-PA group (N = 48) SD	t-test result
TOPA (% correct)	88.3	6.5	31.4	9.2	$t(106) = 37.73$, $p < .001$
CTOPP RAN	96.7	9.3	88.9	11.9	$t(106) = 3.82$, $p < .001$
CELF-R Recalling Sentences	110.6	9.6	92.5	10.9	$t(106) = 9.19$, $p < .001$
ATP threshold (ms)[a]	42.3	50.6	169.6	169.5	$t(73) = 4.74$, $p < .001$
CELF-R Expressive	101.6	11.0	82.9	12.4	$t(106) = 8.29$, $p < .001$
CELF-R Receptive	102.5	14.2	87.7	9.2	$t(106) = 6.28$, $p < .001$
CELF-R Total Language	102.2	11.8	84.1	8.4	$t(106) = 8.97$, $p < .001$
PPVT-R	114.7	11.4	101.2	10.1	$t(106) = 6.40$, $p < .001$
CELF-R Listening to Paragraphs	101.3	11.5	92.7	11.3	$t(106) = 3.86$, $p < .001$
CELF-R Sentence Structure	102.6	14.2	93.4	13.9	$t(106) = 3.39$, $p < .001$
WRMT-R WI lag (mo.)	−24.4	15.9	−7.6	12.9	$t(106) = 5.92$, $p < .001$
WRMT-R WA lag (mo.)	−37.9	37.0	−6.0	23.6	$t(106) = 5.18$, $p < .001$
NARA-R Acc lag (mo.)	−21.2	14.1	−2.9	14.8	$t(106) = 6.78$, $p < .001$
NARA-R Com lag (mo.)	−13.9	11.6	−0.8	10.6	$t(106) = 6.06$, $p < .001$

Note. Unless otherwise stated, standard scores are shown (M = 100, SD = 15). PA = phonological awareness; TOPA = Test of Phonological Awareness; CTOPP = Comprehensive Test of Phonological Processing (draft form); RAN = Rapid Automatized Naming; CELF-R = Clinical Evaluation of Language Fundamentals–Revised; ATP = auditory temporal processing; PPVT-R = Peabody Picture Vocabulary Test–Revised; WRMT-R = Woodcock Reading Mastery Tests–Revised (WI = Word Identification subtest; WA = Word Attack subtest); NARA-R = Neale Analysis of Reading Ability–Revised (Acc = Accuracy scale; Com = Comprehension Scale); mo. = months.

[a]ATP data are not included for 11 children; 9 children were not tested due to absences from school and equipment failure at one school; values for 2 other children in the good-PA group (Z = 4.1, Z = 3.9) were deemed to be atypical of group performance and are therefore not included in this analysis.

Figure 1. Modified back-to-back stem and leaf plot (Tukey, 1977) showing individual thresholds (in milliseconds) for perception of rapid sequences in both phonological awareness (PA) groups. The full distribution of threshold values for good-PA children is shown to the left; the distribution for poor-PA children is to the right. Entries in the shaded central frame or "stem" represent 20 ms bins; numbers on the "leaves" to the sides show actual threshold values for the two groups. (Could not reach criterion = unable to perform 18 correct trials out of 24 with interstimulus interval fixed at 500 ms.)

Good-PA (N = 54)												Poor-PA (N = 45)				
3	1	1	1	0	0	0	0	0	0	0	0–19	2	2	5	7	
19	18	17	17	11	8	5	5	5	5	3		11	19			
						39	28	28	25	20	20–39	22	39			
					57	56	50	44	41	40	40–59	52	55			
										63	60–79	60	69			
						94	92	87	87	86	80–99	89	90	94		
										108	100–119	103	104			
							132	130	124	123	120–139	130	132			
											140–159	153				
											160–179					
											180–199					
											200–219					
										231	220–239	229				
											240–259					
											260–279	269				
											280–299					
											300–319	309	311			
											320–339					
											340–359	345				
											360–379					
											380–399					
											400–419					
											420–439					
									456	440	440–459					
											460–479	460	469	472		
											480–499	491	494			
						*	*	*	*	*	Could not	*	*	*	*	*
									*	*	reach	*	*	*	*	*
											criterion	*	*	*	*	*

between the groups for phonological retrieval (CTOPP RAN) even though the between-group mean difference was significant.

Identification of Poor Readers at the End of Year 2. Year 2 reading outcomes presented in Table 1 show that by this time there were very highly significant differences between the groups on all reading variables. A mean decoding lag was calculated for each child based on WMRT-R Word Identification and Word Attack subtests and the NARA-R Accuracy Scale. Children whose mean decoding skill was 6 months or more behind the expected level for their age by the end of Year 2 were considered to be poor readers. Children with a mean decoding lag of less than 6 months were taken to be functional readers. One good-PA child and 11 poor-PA children were identified in the poor-reader group for the classification analyses. This poor-reader group had a mean lag of 11.6 months (SD = 4.1 months). Of these, 7 children (all from the poor-PA group) already showed pronounced reading difficulty (group mean decoding lag = 14.5 months; SD = 2.7 months). All 7 children were at least 12 months behind age level on the NARA-R Accuracy scale, and at least 6 months behind age level on the WMRT-R Word Attack subtest (mean Accuracy = 16.4 months, SD = 3.5 months; mean Word Attack = 13.0 months, SD = 4.1 months).

Prediction of Reading Achievement Using the Prereading Variables

Regression Analyses

Within each of the PA groups, simultaneous regression analyses were conducted with all predictors combined with PIQ for each of the four reading outcome measures taken at the end of Year 2 (i.e., WRMT-R Word Identification and Word Attack and NARA-R Accuracy and Comprehension). As discussed above, the TRT threshold distributions for the perception of rapid sequences task were not normal and the variable entered into the analyses was ATP performance category (good or poor). Similar trends were observed between the predictors and all of the reading measures, so only the regression of WRMT-R Word Identification is detailed here.

Simple within-group correlation coefficients are reported in Table 2 for PIQ and for each predictor with Word Identification. In good-PA children, all three phonological processing measures (TOPA, CTOPP RAN, and CELF-R Recalling Sentences) and two of the oral language measures (CELF-R Total Language and CELF-R Sentence Structure) were significantly related to WRMT-R Word Identification. However, for these children PIQ was not significantly related to Word Identification. In the poor-PA group, CELF-R Recalling Sentences and CELF-R Total Language were also significantly correlated with reading, but in this group, PA (measured by the TOPA) accounted for no further variance in addition to that captured in the group selection process. Neither was CTOPP RAN strongly related to reading in this group, although PIQ was very strongly and significantly related to reading outcomes in the poor-PA children. In spite of the significant between-group differences in ATP, the relation between performance on the perception of rapid sequences task and reading was nonsignificant in both groups.

Semipartial correlations are also presented in Table 2 to show the unique contribution of each factor to reading. The semipartial correlations make it clear that of

Table 2. Results of simultaneous regression analyses: Simple and semipartial correlation coefficients for WRMT-R Word Identification and the predictor variables within each PA group.

	Poor-PA group (N = 48)		Good-PA group (N = 60)	
	Simple	Semipartial	Simple	Semipartial
TOPA	.02	.04	−.30*	−.14
CTOPP RAN	−.03	.00	−.38**	−.23
CELF-R Recalling Sentences	−.35*	−.03	−.29*	.00
ATP performance category	.08	−.02	.20	.13
CELF-R Total Language	−.30*	−.20	−.42**	−.13
PPVT-R	−.26	−.18	−.13	.07
CELF-R Sentence Structure	.18	.28	−.26*	.02
CELF-R Listening to Paragraphs	.16	.24	−.18	−.07
WPPSI-R PIQ	−.43**	−.35*	−.18	.00

Note. As WRMT-R reading measures were expressed as a lag behind chronological age, signs of correlations with this variable must be interpreted accordingly. WPPSI-R PIQ = Wechsler Preschool and Primary Scale of Intelligence–Revised Performance IQ

*p < .05. **p < .01.

the variance captured in the simple bivariate relations between the predictors and Word Identification, a sizeable proportion is shared. There is considerable overlap between these factors, of which the common variance between oral language (CELF-R Total Language score) and PA (TOPA) is most noteworthy. CELF-R Total Language showed a strong simple relation with word recognition in both groups, but once the influence of the other variables was removed, the unique contribution of oral language ability fell away in the good-PA group. For the poor-PA children, CELF-R Total Language did continue to explain more unique variance than in the good-PA group. Furthermore, in this group, all three other oral language variables (Sentence Structure, Listening to Paragraphs, and PPVT-R) also contributed unique variance to reading. This did not occur in the good-PA group, in which these factors added little after the influence of CELF-R Total Language had been excluded. In the poor-PA group, PIQ accounted uniquely for more than 12% of the variation in reading, but the influence of PIQ in good-PA children appeared to be considerably less. ATP seemed to tap little variance that was not shared with a number of the other measures and made a very small unique contribution to reading for only the good-PA group. This overlap between the variables of our battery meant that, using the multiple regression and correlation approach, we were able to explain only 28% (adjusted R^2) of variation in reading in the children of most interest (i.e., those in poor-PA group) and 12% of the variation in the good-PA group.

Prospective Classification of Poor Readers

As mentioned above, only 1 out of the 58 good-PA children who were available for testing at the end of Year 2 was in the poor-reader group. Therefore, formal discriminant analysis was not required for the good-PA children. Without the need for any other variable, prospective classification of children in the top quartile of the PA distribution as not at risk for becoming poor readers would yield a specificity of .98. Simple classification of children as at risk on the basis of PA in the bottom quartile of the distribution is very much less accurate. In our data, this would mean a sensitivity of only .23 with many false positives. However, the regression analyses suggested that the use of PIQ and the oral language variables, which contributed unique variance to reading outcomes, would significantly improve classification accuracy in this group.

We show in Table 3 that a function derived from CELF-R Total Language score, PIQ, Listening to Paragraphs, Sentence Structure, and PPVT-R identified prospective poor readers with a sensitivity of .91 and specificity of .84. Inclusion of actual TOPA percentage correct in this function did not improve accuracy of classification, probably because there is so much shared variance between these factors. As might be expected given the low contribution of unique variance by ATP in the regression equations, adding this factor to the classification functions did not improve these levels of accuracy. To check for multivariate outliers, squared Mahalanobis distances for each case were examined and none was found to reach significance at the .01 level (Tabachnick & Fidell, 1989). Box's M tests for homogeneity of covariance matrices of canonical discriminant functions revealed that these also did not differ significantly from one another.

In line with our goal of developing a screening tool that would target the highest risk children, we attempted to reduce the battery substantially without a significant cost to prediction accuracy. We sought to index the principal components of the classification function in perhaps one or two measures. In particular, it

Table 3. Prospective classification of the poor-PA children into reading groups (poor readers and functional readers) Year 2.

Predictors	Prediction (beginning Grade 1)	Reading Group (Year 2) Poor reader	Functional reader	N	Correct identification rate
CELF-R Total Language +	At risk			48	
Listening to Paragraphs +	Proportion	.91	.16		
Sentence Structure +	Not at risk				
PPVT-R + WPPSI-R PIQ	Proportion	.09	.84		
					.85
Recalling Sentences +	At risk			48	
TOPA % correct	Proportion	.82	.27		
+ age + sex + parents'	Not at risk				
mean years of education	Proportion	.18	.73		
					.75

seemed necessary to exclude the PIQ measure and reduce the time taken to administer the full CELF-R battery. We examined all the CELF-R subtests to discover whether we could obtain an index of oral language functioning using only one subtest. Recalling Sentences was the subtest most highly correlated with the full CELF-R battery in our data. We then combined CELF-R Recalling Sentences with TOPA and three other variables easily accessible to teachers (i.e., child's age and sex, and parents' mean years of education). Classification of the poor-PA children using this combination produced a sensitivity of .82, with a specificity of .73 (see Table 3).

Discussion

In this study we set out to discover, firstly, whether ATP, oral language ability, and the three components of phonological processing identified in the literature (i.e., PA, and phonological short-term memory and retrieval) could be used to improve at-risk predictions without the necessity of using alphabetic knowledge. We used a regression approach to examine the relations between predictor variables within the two groups of children before using discriminant analysis to classify them into good and poor readers. Secondly, we sought to confirm the suggestion from the literature that oral language and these phonological processing variables might be considered predictors of positive outcomes. Finally, we attempted to develop a screening procedure from these predictors for the children at highest risk of reading problems.

Using ATP, Oral Language Ability, and Phonological Processing to Improve At-Risk Predictions

Regression Approach

The simple correlations in this study between reading, ATP, phonological processing, and oral language suggested relations between the variables very similar to those that existed in Heath and Hogben's (2000) sample, even though reading outcomes were measured at the end of Year 3 in the earlier investigation. With the regressions restricted within each group in both this study and Heath and Hogben's sample, interpretation of the correlations is not straightforward. Correlation coefficients may be somewhat attenuated due to the restricted range in our groups. However, when the unique variance contributed by each factor was examined in the present data, it became clear that much of the variance in reading achievement accounted for by these variables was shared and that ATP did not make a meaningful contribution to the variance in reading outcomes.

The simple relations between TOPA and Word Identification were consistent with the strength of relationships between PA and reading outcomes reported in the literature (e.g., Majsterek & Ellenwood, 1995). However, TOPA was related to reading in only the good-PA group. Once a child was below the 25th percentile in PA, relative level did not appear to matter. This result may to some extent be due to our within-group approach because it is likely that most of the variance in PA would not be within either group, but between the two groups. Therefore, a considerably higher coefficient might be expected across the whole sample. Surprisingly, CTOPP RAN, reflecting efficiency of retrieval of phonological information from long-term memory, appeared to have a negligible influence on reading outcomes in poor-PA children after that of the other phonological variables had been excluded. In contrast, this factor was making a marked contribution to reading in good-PA children, where RAN accounted for more than 10% of the variance. In this case the differing strength of relationships within the two groups did not appear to be related only to our within-group analyses, because the two RAN distributions overlapped almost totally and the variance within the groups appeared remarkably similar. In addition, when we correlated RAN and reading outcomes across the whole sample, we in fact obtained slightly lower r values than those within the good-PA group.

The strength of the simple relations between oral language and reading achievement in both groups was convergent with findings in the literature but was greater than had been expected for word recognition. Bishop and Adams (1990) concluded in reference to the effects of oral language deficits on reading comprehension that "syntactic and semantic ability are responsible for the major part of variation in reading ability" (p. 1045), but this was found to be true here for word recognition as well. The influence of oral language on reading outcomes appeared to be greater in the children with poor PA in preschool than in those with good PA. PIQ also seemed to be more important to reading achievement in children with poor PA, but this apparent difference in strength of relationship within the two groups could be due to the fact that selecting for high or low PA does, at least to some extent, partial out PA. This could mean that the correlation within the good-PA group appears stronger.

Using the regression approach within the poor-PA group, we did not achieve an improvement in previous levels of prediction with only phonological and language variables and ATP. Nevertheless, our analyses underlined the fact that different factors seem to predict reading outcomes in children with poor PA compared to those who begin school with good PA. Our data suggest that if children have good PA to begin with, then for these children neither PIQ nor oral language ability makes a real difference to reading outcomes. Nor does actual PA level seem particularly important, and only phonological retrieval, the efficiency with which the child can access phonological information from long-term memory, seems to have some influence. By comparison, children with poor PA initially are affected by all aspects of oral language processing, as well as PIQ—for them, the whole range of their language experience appears to make a material difference to reading.

Classification by Discriminant Analysis

By focusing only on the children likely to be at highest risk, we obtained levels of classification accuracy that compared favorably with the other classification work sampling across the whole range (e.g., Felton & Brown, 1991; Satz & Friel, 1974, 1978). Using only PIQ and the oral language variables (CELF-R Total Language score, Listening to Paragraphs, and Sentence Structure, and PPVT-R), we identified children with poor reading at the end of Year 2 with a sensitivity of .91 and specificity of .84. Our results are particularly encouraging in the context of those of Hurford and colleagues (Hurford et al., 1993, 1994). These authors used discriminant analysis with a predictive set including ATP, PA, Receptive Vocabulary (PPVT-R), and children's reading scores at commencement of education to classify children into groups of nondisabled readers, dyslexics, and poor readers with general cognitive deficits. It is very significant, therefore, that with this set of variables, and using a more direct and economical assessment strategy, an equivalent level of prediction was achieved without the need for prereading scores.

Tabachnik and Fidell (1989) cautioned against "overfitting" (i.e., when a function will make accurate predictions in the original sample but fails to generalize to any other samples) if the number of cases in the smallest group being classified does not markedly exceed the number of predictors. Our functions contained five predictors, when based on only the psychometric variables, and eight predictors when we added demographic factors. With only 11 cases in the poor-reader group, overfitting could still be a possibility. This seems less likely, however, than in the data of Hurford et al. (1994), who had only 10 cases in their smallest group and 10 predictors in their function.

The sensitivity and specificity obtained here without the use of alphabetic knowledge as a predictor were considerably lower than those achieved by Badian et al. (1990), who used regression analysis to derive predictive mean scores for Letter Sounds, RAN of numbers, and Finger Localization as cutoffs for classification of dyslexics and normal readers in a somewhat larger sample (N = 163). However, in the Badian study, the dyslexics were considerably more severe than the poor readers here (2 *SD*s below the mean), which considerably increases the chances of correct identification. Badian also followed children up until the end of Grade 4, which allowed for the severity of the dyslexics' reading problems to become fully apparent. In this study, only 1 child had a reading problem approaching that level of severity by the end of Year 2. In a later study by Badian (1994), which more closely resembled the present study, results were reported that were more consistent with those found here. A total of 118 children were followed up until the end of Grade 1, and a sensitivity of .87 and specificity of .91 was achieved by classifying children into good and poor readers on the basis of a battery of 11 phonological processing, alphabetic knowledge, and visuospatial tasks. This classification also appears to be in danger of overfitting, because the smallest group contained only 15 children.

We found, therefore, that using these five psychometric predictors in poor-PA children would provide very accurate classification in view of what has been achieved so far in the wider literature. However, even this present battery would still be expensive in time and educational resources and would require specialist personnel to administer it.

Can Oral Language, PA, Phonological Short-Term Memory, and Phonological Retrieval Predict Positive Reading Outcomes?

Our data provide a resounding demonstration that children in the top quartile of the PA distribution in preschool are most unlikely to develop significant reading difficulties, even though other factors such as school and home environments will, of course, also be critical influences. Only 1 child out of the 58 children followed up at the end of Year 2 had a mean decoding score more than 6 months behind the level expected for her age. We also observed that more than half of the children in the good-PA group had scores on or above the mean for the CELF-R oral language battery, compared with a small fraction of the poor-PA group. This suggests that children with average or above average oral language are also far less likely to become poor readers than those with weaker oral language. A similar separation between the groups was observed in phonological short-term memory, suggesting that children with above average short-term memory are unlikely to encounter difficulty with reading. The groups overlapped much more on phonological retrieval, measured here by RAN of objects, so this factor does not appear to be so clearly linked to successful reading.

Development of a Screening Procedure for Children at Highest Risk of Reading Problems

Our data suggest that screening children in the bottom quartile of the PA distribution with PA and CELF-R Recalling Sentences combined with age, sex, and parents' educational levels would prospectively identify poor readers with a sensitivity of more than .8 and false positives at the rate of only .7. This two-tiered screening method would require only (a) access to information readily available within the school system to teachers, (b) 20 min per preschool child to administer the TOPA, and (c) an additional 10 min per child for those in the bottom PA quartile to administer the Recalling Sentences subtest. Both tests could be administered using existing resources within the school system (e.g., by appropriately in-serviced teaching aides in a quiet area within a preschool center during Terms 2 and 3 of the school year).

Decisions about whether to use this screening procedure or a function derived from our full battery would need to be made in the context of existing and attainable resources within the particular education system attempting early identification. With the lower sensitivity of the screening procedure, a small proportion of the children at risk would potentially be overlooked. The reduced specificity of this approach could mean a cost of falsely labeling and intervening with children who might potentially self-correct. The implications of false-positive classification were discussed by Satz and Friel (1974, 1978), but such problems could be minimized by presenting the preventative program as enrichment rather than remediation. In any case, with our procedure, all the false positives were from among the poor-PA children, which is the very group that is widely acknowledged to be at highest risk for literacy achievement. It is likely that these children would only benefit from an enhancement program.

General Discussion

Limitations of the Present Design

These results need cross-validation in a much larger sample that would be representative of the whole preschool population. Because the sample for this investigation was obtained by recruitment of parents willing to have their children participate, with response rates of .3 to .4, it seems possible that this sample was biased toward inclusion of children in families where literacy development was highly valued and supported. This could have produced environmental conditions that ameliorated potential reading deficits in children in this sample and therefore distorted relations that would be observed in an unbiased sample. In the future, it would be important to secure school participation such that whole classes could be included in the sample.

The present design was based on wide acknowledgment in the reading research community that children in the lowest quartile are at highest risk of reading failure (e.g., Torgesen & Bryant, 1994). We have deliberately developed a strategy for maximizing accuracy of identification of children likely to develop reading problems from this group. This approach does not address children in the middle quartiles of the PA distribution, who may become poor readers. It will be essential to continue attempts to predict reading outcomes for these children.

It must be remembered that because we have examined predictors within subgroups, rather than in a whole population, the relationships observed here may not be applicable to a more representative preschool sample. The present results suggest that there may be particular problems with the likely association between reading, oral language, and PA in the middle quartiles of the PA distribution.

Implications of These Findings for the Prediction of Reading Outcomes in General

The Failure of ATP as an Independent Predictor

The ATP task in its present form was a poor predictor. Performance category for perception of rapid sequences accounted for similar levels of variance in reading to that explained by threshold on this task in the pilot study. The principal problem is that the task once again failed to tap variance in reading independent of PA and oral language. It may be that different children have deficits in different aspects of auditory temporal processing, so that it is unrealistic to expect to index these deficits with any one task. It is possible that a battery of ATP tasks may need to be administered along the lines of those developed by Ahissar, Protopapas, Reid, and Merzenich (2000), Cacace, McFarland, Ouimet, Schrieber, and Marro (2000), or Waber et al. (2001). An ATP profile for each individual might optimize sensitivity and specificity, which could yield additional predictive power.

Assessment of PA and Oral Language

Data from Fletcher and Leitao (1998) suggested that some children who are very weak in phonological processing in preschool are still processing speech at the whole word rather than the syllable or phoneme level. These observations would need to be taken into account when selecting a PA measure, because the measures used in this investigation may have failed to discriminate optimally at the lower end of the distribution. Adams (1990) stressed that conscious analytic or meta-knowledge of phonemes, rather than merely a working knowledge, is the element of PA that is significant for easy acquisition of grapheme–phoneme correspondences. Therefore, a battery including a phoneme deletion task rather than only categorization would seem preferable in future studies.

In this study, Recalling Sentences, examined as a measure of phonological short-term memory, performed almost as well as the full CELF-R battery in classification of poor readers. It has been shown that the repetition of nonwords, another well-documented approach to measuring phonological short-term memory, is also specifically related to several language abilities during early childhood education (Gathercole, Willis, Baddeley, & Emslie, 1994). It may be that nonword repetition would capture more of the variance accounted for by the full CELF-R battery than the Recalling Sentences subtest does. This seems likely because it has been shown to be a behavioral marker for inherited language impairment in SLI children, even after early language difficulties have resolved (Bishop, North, & Donlan, 1996). Gallagher, Frith, and Snowling (2000) also reported that children at familial risk of dyslexia had significantly more difficulty with the repetition of nonwords than normally developing children.

Other Variables That May Improve Prediction Further

There seems to be clear consensus that a family history of dyslexia is of considerable significance (e.g., Snowling, 1996). It is surprising, therefore, that family history of dyslexia does not feature strongly in the prediction literature as an actual predictor. Of the other predictors that have featured strongly in the recent literature, only alphabetic knowledge (i.e., letter name and sound knowledge) is almost universally credentialed as a strong predictor of reading achievement. Muter et al. (1998) found that the product of letter name knowledge and PA contributed additional variance to reading over and above the contribution of each separately. However, Adams (1990) emphasized that both accuracy and speed should be measured in alphabetic knowledge because it is the overall familiarity with the use of letters that best predicts reading achievement. These variables were purposely omitted here, but they do appear to increase the accuracy of classification. Both family history and letter knowledge data could be very cost-effective because they could readily be collected using time and resources that are already available within existing education systems.

Recent findings by Elbro, Borstrom, and Petersen (1998) suggest that distinctness of mental representations of phonological information might also be a powerful predictor of reading. Elbro et al. found that distinctness measured in kindergarten children of dyslexic and normal-reading parents made a significant contribution to reading independent from that of PA, phonological retrieval, and letter naming. They concluded that distinctness might directly influence acquisition of letter–sound correspondences and so the development of phonological decoding independent of PA. However, it still needs to be established that distinctness would tap variance in reading that was not shared with the phonological processing variables explored here and/or with oral language. Also, a suitable standardized measure for this factor has yet to become available.

Nevertheless, the results from the present investigation could be used immediately by schools with resources to provide prevention programs. The classification strategy described here would identify a large proportion of children at risk of reading failure or who were likely to struggle with reading. A positive family history for dyslexia and poor speed and accuracy in alphabetic knowledge would serve as additional indicators of risk.

Acknowledgments

This research was supported by the Financial Markets Foundation for Children and by the Australian Research Council.

References

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.

Ahissar, M., Protopapas, A., Reid, M., & Merzenich, M. M. (2000). Auditory processing parallels reading disabilities in adults. *Proceedings of the National Academy of Sciences*, *97*, 6832–6837.

Badian, N. A. (1994). Preschool prediction: Orthographic and phonological skills and reading. *Annals of Dyslexia*, *44*, 3–25.

Badian, N. A., McAnulty, G. B., Duffy, F. H., & Als, H. (1990). Prediction of dyslexia in kindergarten boys. *Annals of Dyslexia*, *40*, 152–169.

Bishop, D. V. M. (1991). Developmental reading disabilities: The role of phonological processing has been overemphasized. *Mind and Language*, *6*, 97–101.

Bishop, D. V. M., & Adams, C. (1990). A prospective study of the relationship between specific language impairment, phonological disorders and reading retardation. *Journal of Child Psychology and Psychiatry*, *31*, 1027–1050.

Bishop, D. V. M., North, T., & Donlan, C. (1996). Nonword repetition as a behavioral marker for inherited language impairment: Evidence from a twin study. *Journal of Child Psychology and Psychiatry*, *37*, 391–403.

Blachman, B. A. (1984). Relationship of rapid naming ability and language analysis skills to kindergarten and first grade reading achievement. *Journal of Educational Psychology*, *76*, 610–622.

Bond, G. L., & Dykstra, R. (1967). The co-operative research program in first-grade reading instruction. *Reading Research Quarterly*, *2*, 5–42.

Cacace, A. T., McFarland, D. J., Ouimet, J. R., Schrieber, E. J., & Marro, P. (2000). Temporal processing deficits in remediation-resistant reading-impaired children. *Audiology & Neuro-Otology*, *5*, 83–97.

Catts, H. W. (1989). Phonological processing deficits and reading disabilities. In A. Kamhi & H. Catts (Eds.), *Reading disabilities: A developmental language perspective* (pp. 101–132). Boston: College-Hill Press.

Chall, J. S. (1967). *Learning to read: The great debate*. New York: McGraw-Hill.

Critchley, M. (1970). *The dyslexic child*. London: Heinemann Medical.

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

Edwards, J. (1993). The emotional effects of dyslexia. In S. F. Wright & R. Groner (Eds.), *Facets of dyslexia and its remediation* (pp. 425–433). Oxford, U.K.: Elsevier Science.

Elbro, C., Borstrom, I., & Petersen, D. K. (1998). Predicting dyslexia from kindergarten: The importance of distinctness of phonological representations of lexical items. *Reading Research Quarterly*, *33*, 36–60.

Farmer, M. E., & Klein, R. (1993). Auditory and visual temporal processing in dyslexic and normal readers. In P. Tallal, A. M. Galaburda, R. R. Llinas, & G. von Euler (Eds.), *Temporal information processing in the nervous system: Special reference to dyslexia and dysphasia* (pp. 339–341). New York: The New York Academy of Sciences.

Felton, R. H., & Brown, I. S. (1991). Neuropsychological prediction of reading disabilties. In J. E. Obrzut & G. W. Hynd (Eds.), *Neuropsychological foundations of learning disabilites: A handbook of issues, methods and practice* (pp. 387–410). New York: Academic Press.

Fletcher, J., & Leitao, S. (1998, May). A phonological awareness program for preschool children. In K. Hird (Ed.), *Expanding horizons: Speech Pathology Australia National Conference Proceedings, Fremantle, Western Australia* (pp. 199–206). Perth, Australia: Speech Pathology Australia.

Gallagher, A., Frith, U., & Snowling, M. J. (2000). Precursors of literacy delay among children at genetic risk of dyslexia. *Journal of Child Psychology and Psychiatry*, *41*, 202–213.

Gathercole, S. E., Willis, C. S., Baddeley, A. D., & Emslie, H. (1994). The children's test of nonword repetition: A test of phonological working memory. *Memory*, *2*, 103–127.

Heath, S. M., & Hogben, J. H. (2000). Auditory temporal processing, phonological awareness and oral language ability in pre-readers: Can we identify children at risk for reading disability more accurately? *Australian Educational and Developmental Psychologist, 17,* 32–62.

Heath, S. M., Hogben, J. H., & Clark, C. D. (1999). Auditory temporal processing in disabled readers with and without oral language delay. *Journal of Child Psychology and Psychiatry*, *40*, 637–647.

Horn, W. F., & Packard, T. (1985). Early identification of learning problems: A meta-analysis. *Journal of Educational Psychology*, *77*, 597–607.

Hurford, D. P., Darrow, L. J., Edwards, T. L., Howerton, C. J., Mote, C. R., Schauf, J. D., & Coffey, P. (1993). An examination of phonemic processing abilities in children during their first-grade year. *Journal of Learning Disabilities*, *26*, 167–177.

Hurford, D. P., Schauf, J. D., Bunce, L., Blaich, T., & Moore, K. (1994). Early identification of children at risk for reading disabilities. *Journal of Learning Disabilities*, *27*, 371–382.

Lyon, G. R., Alexander, D., & Yaffe, S. (1997). Progress and promise in research in learning disabilities. *Learning Disabilities*, *8*, 1–6.

McArthur, G. M., Hogben, J. H., Edwards, V. T., Heath, S. M., & Mengler, E. D. (2000). On the "specifics" of specific reading disability and specific language impairment. *Journal of Child Psychology and Psychiatry, 41,* 869–874.

MacDonald, W. G., & Cornwall, A. (1995). The relationship between phonological awareness and reading and spelling achievement eleven years later. *Journal of Learning Disabilities*, *28*, 523–527.

Majsterek, D. J., & Ellenwood, A. E. (1995). Phonological awareness and beginning reading: Evaluation of a school-based screening procedure. *Journal of Learning Disabilities*, *28*, 449–456.

Mann, V. A. (1984). Longitudinal prediction and prevention of early reading difficulty. *Annals of Dyslexia*, *34*, 115–136.

Maughan, B. (1995). Long-term outcomes of developmental reading problems. *Journal of Child Psychology and Psychiatry*, *36*, 357–371.

Merzenich, M. M., Jenkins, W. M., Johnston, P., Schreiner, C., Miller, S. L., & Tallal, P. (1996). Temporal processing deficits of language-learning impaired children ameliorated by training. *Science*, *271*, 77–81.

Muter, V. (1996). Predicting children's reading and spelling difficulties. In M. Snowling & J. Stackhouse (Eds.), *Dyslexia, speech & language: A practitioner's handbook* (pp. 31–45). London: Whurr.

Muter, V., Hulme, C., Snowling, M., & Taylor, S. (1998). Segmentation, not ryhming, predicts early progress in learning to read. *Journal of Experimental Child Psychology*, *71*, 3–27.

Neale, M. D. (1988). *Neale Analysis of Reading Ability–Revised: Manual*. Hawthorn, Australia: Australian Council for Educational Research.

Prior, M., Sanson, A., Smart, D., & Oberklaid, F. (1995). Reading disability in an Australian community sample. *Australian Journal of Psychology*, *47*, 32–37.

Reed, M. A. (1989). Speech perception and the discrimination of brief auditory cues in reading disabled children. *Journal of Experimental Child Psychology*, *48*, 270–292.

Satz, P., & Friel, J. (1974). Some predictive antecedents of specific reading disability: A preliminary two-year follow-up. *Journal of Learning Disabilities*, *7*, 48–55.

Satz, P., & Friel, J. (1978). Predictive value of an abbreviated screening battery. *Journal of Learning Disabilities*, *11*, 20–24.

Scarborough, H. S. (1990). Very early language deficits in dyslexic children. *Child Development*, *61*, 1728–1743.

Scarborough, H. S. (1991). Early syntactic development of dyslexic children. *Annals of Dyslexia*, *41*, 207–220.

Semel, E., Wiig, E. H., & Secord, W. (1987). *Clinical Evaluation of Language Fundamentals–Revised: Manual*. San Antonio, TX: The Psychological Corporation/Harcourt Brace Janovich.

Share, D. L., Jorm, A. F., MacLean, R., & Matthews, R. (1984). Sources of individual differences in reading acquisition. *Journal of Educational Psychology*, *76*, 1309–1324.

Shaywitz, S. E. (1998). Current concepts: Dyslexia. *New England Journal of Medicine*, *338*, 307–312.

Snowling, M. (1996). Dyslexia: A hundred years on: A verbal not a visual disorder which responds to early intervention. *British Medical Journal*, *313*, 1096–1097.

Snowling, M. (2000). *Dyslexia* (2nd ed.). Oxford, U.K.: Blackwell.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*, 360–407.

Stark, R. E., Tallal, P., & McCauley, R. J. (1988). *Language, speech and reading disorders in children: Neuropsychological studies*. Boston: College-Hill Press.

Stothard, S. E., Snowling, M. J., Bishop, D. V. M., Chipchase, B. B., & Kaplan, C. A. (1998). Language-impaired preschoolers: A follow-up into adolescence. *Journal of Speech, Language, and Hearing Research*, *41*, 407–418.

Stuart, M., & Masterson, J. (1992). Patterns of reading and spelling in 10-year-old children related to prereading phonological abilities. *Journal of Experimental Psychology*, *54*, 168–187.

Swank, L., & Catts, H. (1994). Phonological awareness and written word decoding. *Language, Speech, and Hearing Services in Schools*, *25*, 9–14.

Tabachnick, B. G., & Fidell, L. S. (1989). *Using multivariate statistics* (2nd ed.). Northridge, CA: Harper Collins.

Tallal, P. (1976). Rapid auditory processing in normal and disordered language development. *Journal of Speech and Hearing Research*, *3*, 561–571.

Tallal, P. (1980). Auditory temporal perception, phonics, and reading disabilities in children. *Brain and Language*, *9*, 182–198.

Tallal, P., Miller, S. L., Bedi, G., Byma, G., Wang, X., Nagarajan, S. S., et al. (1996). Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science*, *271*, 81–84.

Tallal, P., Miller, S. L., Jenkins, W. M., & Merzenich, M. M. (1997). The role of temporal processing in developmental language-based learning disorders: Research and clinical implications. In B. A. Blachman (Ed.), *Foundations of reading acquisition and dyslexia* (pp. 49–65). Mahwah, NJ: Erlbaum.

Tallal, P., & Piercy, M. (1973). Developmental aphasia: Impaired rate of non-verbal processing as a function of sensory modality. *Neuropsychologia*, *11*, 389–398.

Taylor, M. M., & Creelman, C. D. (1967). PEST: Efficient estimates on probability functions. *Journal of the Acoustical Society of America*, *41*, 782–787.

Torgesen, J. K., & Bryant, B. R. (1994). *Test of Phonological Awareness: Manual*. Austin, TX: Pro-Ed.

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1994). Longitudinal studies of phonological processing and reading. *Journal of Learning Disabilities*, *27*, 276–286.

Tukey, J. W. (1977). *Exploratory data analysis*. London: Addison-Wesley.

Vellutino, F. R., & Scanlon, D. M. (1987). Phonological coding, phonological awareness, and reading ability: Evidence from a longitudinal and experimental study. *Merrill-Palmer Quarterly*, *33*, 321–363.

Waber, D. P., Weiler, M. D., Wolff, P. H., Bellinger, D., Marcus, D. J., Ariel, R., et al. (2001). Processing of rapid auditory stimuli in school-age children referred for evaluation of learning disorders. *Child Development*, *72*, 37–49.

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin*, *101*, 192–212.

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *Comprehensive Test of Phonological Processing: Manual*. Austin, TX: Pro-Ed.

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

Woodcock, R. W. (1987). *Woodcock Reading Mastery Tests–Revised: Manual*. Circle Pines, MN: American Guidance Service.

Wright, B. A., Lombardino, L. J., King, W. M., Puranik, C. S., Leonard, C. M., & Merzenich, M. M. (1997). Deficits in auditory temporal and spectral resolution in language impaired children. *Nature*, *387*, 176–177.

Yopp, H. (1988). The validity and reliability of phonemic awareness tests. *Reading Research Quarterly*, *23*, 159–177.

Received January 13, 2003

Revision received June 26, 2003

Accepted January 23, 2004

DOI: 10.1044/1092-4388(2004/057)

Contact author: Steve Heath, MPsych, PhD, School of Psychology, The University of Western Australia, 35 Stirling Hwy, Crawley 6009, Australia.
E-mail: steve@psy.uwa.edu.au