## The reliability and validity of tasks measuring perception of rapid sequences in children with dyslexia

## Steve M. Heath and John H. Hogben

The University of Western Australia

Background: Claims that children with reading and oral language deficits have impaired perception of sequential sounds are usually based on psychophysical measures of auditory temporal processing (ATP) designed to characterise group performance. If we are to use these measures (e.g., the Tallal, 1980, Repetition Test) as the basis for intervention in language and literacy deficits, we need to demonstrate that they can effectively quantify individual differences. Therefore, questions of standardisation, reliability and construct validity can no longer be ignored. Method: We explored these issues in three studies: (i) 52 Dyslexics and Good Readers aged 8 to 11 years performed a task requiring perception of rapid sequences (PRS) based on the Tallal Repetition Test; (ii) a subgroup of the initial sample was retested on the task three to four months later, and after extended practice; (iii) a further subgroup then completed a rate of auditory processing task using a backward recognition masking paradigm. **Results:** With a standardised methodology, we were able to replicate previous results with the PRS task, and demonstrate moderate reliability of measurement across time and practice. However, there were large effects of exposure and practice, and the task did not seem useful for identifying absolute and continuing deficits in given individuals. Conclusions: Our results call into question the use of this type of task as an individual measure of ATP. Neither is it certain that it is capturing what is currently understood as ATP. Keywords: Auditory temporal processing, auditory backward recognition masking, dyslexia, language impairment. Abbreviations: ATP: auditory temporal processing; PRS: perception of rapid sequences; SLI: specific language impaired/impairment; RAP: rate of auditory processing; LLI: language learning impaired; PIQ: performance IQ; ID: intensity discrimination.

Numerous explanations have been proposed for *developmental dyslexia*. The failure of between 3 and 5% of school-age children to learn to read successfully in spite of adequate intelligence and learning opportunities (Shaywitz, 1998; Snowling, 1998) is mostly attributed to a deficit or developmental delay in one or more aspects of cognitive or perceptual functioning.

One candidate for such a deficit has been auditory temporal processing (ATP) which refers to the effective temporal resolution of brief or rapidly-presented auditory stimuli. Tallal and colleagues have advanced the temporal processing hypothesis which proposes that impaired ATP is causally related either to specific language impairment (SLI) or to dyslexia in some children via the mediating effect of phonological processing deficits (e.g., Tallal, 2000; Tallal, Miller, Jenkins, & Merzenich, 1997). It has been found that individuals who show deficits in ATP are likely to have difficulties with their reading, their oral language development, or both (e.g., Farmer & Klein, 1993; Heath, Hogben, & Clark, 1999; McCroskey & Kidder, 1980; Reed, 1989; Tallal, 1980; Tallal & Piercy, 1973a,b; Wright et al., 1997).

A significant contribution to this research has been made by a task designed by Tallal & Piercy (1973a, b) to measure perception of rapid sequences (PRS). A typical PRS task involves children being

asked to specify the order in which two tones (e.g., of high or low pitch) occurred. Their accuracy is then examined as a function of the interval between the two tones. A number of separate investigations using this type of task across a range of tone durations and frequencies have found significant differences between groups of disabled and normal readers in children and adults (e.g., Ahissar, Protopapas, Reid, & Merzenich, 2000; Farmer & Klein, 1993; Heath et al., 1999; Ludlow, Cudahy, Bassich, & Brown, 1983; Protopapas, Ahissar, & Merzenich, 1997; Reed, 1989; Watson & Miller, 1993). These data have been supported by the demonstration of ATP deficits both in individuals with dyslexia and those with SLI, using a range of other auditory and visual tasks (e.g., Lovegrove, Martin, & Slaghuis, 1986; McCroskey & Kidder, 1980; Stein & Talcott, 1999; Tallal & Piercy, 1974, 1975; Wright et al., 1997).

On the strength of such findings, Tallal and colleagues (Merzenich et al., 1996; Tallal et al., 1996) hypothesised that training in rapid temporal resolution might be therapeutic in SLI and dyslexic children, to whom they have applied the umbrella term *language learning impaired* (LLI). When they trained these children in processing rapidly presented non-verbal and synthesised speech stimuli using a computer task based on the original Repetition Test (i.e., a PRS task), this group reported significant

© Association for Child Psychology and Psychiatry, 2004.

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

improvement in existing reading and oral language deficits. The two studies concerned confound training on nonverbal and synthesised speech stimuli, which means it is not clear that training in PRS was the critical factor responsible for these improvements. Nevertheless, this group is now presenting intervention based on such training for individual LLI children worldwide. This attempt to ameliorate deficits in particular individuals seemed to us to signal the need to shift from an exclusively psychophysical approach to testing for ATP to a psychometric perspective, which seeks to measure a given dimension in an individual with some documented consistency of outcomes.

We were concerned that the prevailing approach in this research continued to rely on between-group comparisons of various measures, to make inferences about individuals. We believed that two urgent questions remained unanswered in relation to the use of ATP as a measure of individual characteristics. Firstly, how stable is ATP across time and practice when measured by a standardised procedure? Although significant practice effects have been observed in the general psychophysical literature (e.g., Hogben & Di Lollo, 1984; Warren, 1974) and in ATP research (McFarland, Cacace, & Setzen, 1998), these seemed rarely to be discussed in relation to ATP. Further, in the work with ATP there had been a marked absence of standardised stimuli and procedure.

Secondly, to what extent is ATP a valid construct, demonstrable across a range of different measures? The problem of construct validity had first been flagged by Studdert-Kennedy and Mody (1995), but appeared to have been largely ignored since that time. Heath et al. (1999) reported concerns about a possible confound in their measurement of temporal resolution due to frequency discrimination ability, which might be thought to chiefly involve the processing of simple spectral stimulus components. We reasoned that if the PRS task does reflect fundamental aspects of temporal processing, then it should still correlate well with a task for which participants could not possibly enhance their performance by utilising greater powers of frequency discrimination.

We selected McArthur's (1998) Rate of Auditory Processing (RAP) task which effectively discriminated SLI children with deficits in reading accuracy from language-normal control children. McArthur's task is a backward recognition masking task which requires recognition of a target tone in a group of varying intensity tones followed by a noise masker. We thought this task might be a relatively pure measure of ATP because it would draw only on temporal and not spectral components of auditory processing. In addition, it controlled for individual differences in intensity discrimination. We also favoured this task because it used the same adaptive procedure as the PRS task. McFarland et al. (1998) demonstrated that the reliability of threshold estimates was strongly influenced by the type of psychophysical method used, even within a given stimulus dimension (e.g., frequency or intensity). Further, the task used a similar computer format with analogous response demands, which would be expected to reduce the potential error variance from nonperceptual task performance correlates.

## **Experiment 1**

In this experiment, we standardised Heath et al.'s (1999) methodology and attempted to replicate their findings in a larger sample of children from a wider selection of schools. We predicted that in this more representative sample a significant proportion of children with dyslexia would again show poor perception of rapid auditory sequences.

## Method

Participants. Fifty-two children aged from 8 years 6 months to 11 years 9 months were recruited from eight primary schools in metropolitan Perth, Western Australia. Children were screened using (1) the Neale Analysis of Reading Ability - Revised (Neale, 1988; NARA-R), where the child reads graded passages aloud until a ceiling level is reached; (2) the Performance Scale of the Wechsler Intelligence Scale for Children - 3rd Edition (Wechsler, 1991; WISC-III); and (3) screening also included a questionnaire completed by parents, which was used to rule out children with comorbid diagnoses such as SLI, and any alternative factors which might account for language or reading disability. The questionnaire also provided a measure of Parents' Mean Years of Education, which was used as an index of the children's home language environment. Children's hearing was assessed by standard audiometric testing using a Maico MA39 manual audiometer. All participants were required to have bilateral pure-tone conduction thresholds of 20 dB HL or less across 125, 250, 500, 1,000 and 2,000 Hz (AS, 2586, 1983).

Groups of Dyslexic children and Good Readers were equated for Performance IQ (PIQ), Parents' Mean Years of Education, and chronological age (see Table 1). Dual control of both chronological and reading age is generally acknowledged in the literature as optimal, because the reading age match controls for effects of differences in reading achievement, and in linguistic skills which may stem from the process of learning to read (Goswami & Bryant, 1989). However, a reading age matched control group was not included in this study, because of developmental trends in ATP previously observed in the literature (Elliott & Hammer, 1993; Tallal, 1976). To create a reading age control group for this study, it would have been necessary to compare disabled readers, whose ages ranged from 10 years 0 months to 8 years 6 months, with Good Readers aged no more than 8 years 6 months, and possibly as young as seven vears. Therefore, a sizeable proportion of the control children would have been well below the age at which

Variable	Good Readers (N = 22) Mean (SD)	Dyslexics (N = 30) Mean (SD)	Statistical test result
Chronological Age (months)	118.9 (11.3)	120.8 (11.7)	t(50) = .60, p = .554
NARA-R Accuracy Lag (months)	-12.6(4.7)	27.2 (7.9)	t(50) = 21.16, p = .000
NARA-R Comprehension Lag (months)	-16.5 (12.3)	21.6 (8.8)	t(50) = 13.04, p = .000
WISC-III Performance IQ	105.0 (9.5)	102.7 (8.5)	t(50) = .95, p = .349
Parents' Mean Years of Education	11.9 (1.8)	11.8 (1.7)	t(48) = .17, p = .868
PRS Threshold (ms)	36.1 (33.5)	92.4 (101.6)	t(50) = 2.50, p = .008

Table 1 Group mean values for	Good Readers and Dyslexics on a	screening variables and PRS in Experiment 1
*	0	

*Note 1*: NARA-R = Neale Analysis of Reading Ability – Revised; WISC-III Wechsler Intelligence Scale for Children – 3rd Edition. *Note 2*: Lag of reading age behind chronological age is reported, therefore performance in advance of chronological age shows as a negative lag.

*Note 3*: Two-tailed levels of significance are shown for the screening variables, because for the equating of groups, no difference was predicted between the groups in either direction.

children's PRS performance is thought to asymptote to adult levels, introducing a potential confound due to age effects.

The Good Reader group consisted of 22 children (12 girls and 10 boys) and the Dyslexic group comprised 30 children (13 girls and 17 boys). Criteria for inclusion in the Dyslexic group, which were adapted from Critchley (1970), were as follows: (i) a NARA-R reading accuracy level of at least 18 months below that expected for chronological age (this criterion was retained from Heath et al. (1999) to ensure a comparable sample of dyslexic children to that used in the earlier study); (ii) at least average intelligence as measured by the Performance Scale of the WISC-III (85 or above); (iii) no history of significant behavioural, emotional or psychiatric problems; (iv) no evidence of sensorineural hearing loss and normal or corrected to normal visual acuity; (v) absence of socio-economic disadvantage; (vi) history of regular school attendance; (vii) English as a 'first language'; (viii) no stimulant medication for Attention Deficit/Hyperactivity Disorder; and (ix) no diagnosis of SLI. Good Readers were selected according to the same criteria as the Dyslexic group except that these children had NARA-R accuracy between 6 and 24 months above the level expected for chronological age. Strong average readers were chosen by Heath et al. (1999) for their comparison group to ensure that there would be minimal overlap between the reading distributions and it was thought sensible to sample from the same population in the present study.

*PRS task.* The child was presented with two complex sounds generated by using the FM chip on a Creative Labs Sound Blaster 16 driven by a computer. These sounds were designed to match those described by Tallal (1980) as closely as possible. Each sound was 95 ms long (including 10 ms rise/falls) and comprised four equal-amplitude sinusoidal components. The 'High' sound had its lowest component at 309.8 Hz, while the 'Low' sound had its lowest component at 97.7 Hz. The three higher frequency components were identical in the two sounds at 491.5, 736.1 and 1471.0 Hz.

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 presented by a computer, which also displayed visual feedback about responses on the monitor. Children were trained to respond by pressing designated keys on the computer. In the practice phases of the task, the sounds were presented to children through stereo loudspeakers at a comfortable listening level. For the testing phases, the sounds were presented binaurally through earphones at a comfortable listening level (mean intensity across both ears = 76.1 dB SPL.).

The Heath et al. (1999) version of the task was presented without modification except for three improvements in the psychophysical methodology which were as follows: (i) The Tallal (1980) Repetition Test format, with 24 trials at fixed ISIs from 8 to 305 ms, was no longer included as it had not differentiated between the groups in the earlier study. (ii) Extreme PRS thresholds (above one second) had proven a significant problem in the Heath et al. (1999) study, resulting in loss of valuable data. In this study, we required a criterion level of performance (75% correct) at a fixed ISI of 500 ms before testing with a variable ISI began, and this was taken to indicate a threshold of 500 ms or less. Therefore a ceiling ISI of 500 ms could be introduced to the PRS task to prevent participants' thresholds from becoming extreme. (iii) Finally, the earlier protocol of Heath et al. (1999) for PRS measurement was altered to provide feedback to participants throughout all phases of the task so that participants would achieve the criterion level of performance under exactly the same conditions as those prevailing for estimation of threshold.

Procedure. The task consisted of three practice followed by two testing phases. In the first two practice phases children learned to recognise the two separate tones. In the final phase children learned to identify and sequence both tones together with a fixed ISI of 500 ms. 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 before proceeding further. If children could not reach criterion, they repeated the three practice phases before attempting the Criterion Phase again. If they were still unable to reach criterion, but had more than 16 correct, they were again presented with the three practice phases, before making a further attempt at the Criterion Phase. Participants who were still unable to meet the criterion were given an additional testing session within 10 days of the first session and permitted two further attempts to reach criterion. If they were still unable to do so, testing was discontinued on this task.

In the second test phase the PRS task was presented using the adaptive procedure developed by Heath et al. (1999) from the Rapid Perception Subtest of the Tallal Repetition Test (Tallal & Piercy, 1973b) but with the above modifications. The ISI was varied according to the child's response accuracy over 100 trials, using the PEST procedure (Taylor & Creelman, 1967) to determine the ISI at which children identified the two stimuli and their order with an accuracy of 75%. Threshold was calculated by taking the mean of ISIs tested on trials occurring after the fourth reversal of step.

*General procedure.* Children were first tested with the NARA-R, and those with appropriate reading lags then completed the Performance Scale of the WISC-III. Most participants completed two testing sessions of one hour's duration. In the first session, participants were practised on the PRS task then completed an initial block of 100 trials. In some children, up to two further blocks of 100 trials were required for them to produce a stable PRS function (i.e., a relatively smooth function with no clear trend in either direction for at least the last 30 to 40 trials of the PEST procedure).

## Results and discussion

*Group validation.* Descriptive statistics for chronological age, NARA-R Reading Accuracy and Comprehension, PIQ and Parents' Mean Education level are presented in Table 1. Independent sample *t*-tests confirmed that there was no significant difference between the groups in chronological age, PIQ or Parents' Mean Education level. However, as expected, significant differences were found between the groups for both NARA-R Accuracy and Comprehension.

Between-group comparisons in PRS. Table 1 also shows that there were highly significant differences between the groups in PRS. The present results can be directly compared with those of Heath et al. (1999) by combining their two groups of disabled readers. The differences in PRS thresholds were consistent with those reported by Heath et al., although the effect size was somewhat smaller than in the previous study (*Cohen's* d = .83, cf. d = 1.06).

*PRS distribution patterns.* Figure 1 shows that the distributions in this study closely resembled those reported by Heath et al. in the following ways: (i) there was very large variability within the groups; (ii) there was marked positive skew in both groups; (iii) the distributions overlapped considerably; (iv) most Dyslexic children performed within the same range as the Good Readers; and (v) a subset of the Dyslexic group (20%) showed extremely high thresholds, well outside the range of all the other participants. As



**Figure 1** Modified back-to-back stem and leaf diagram (Tukey, 1977) for auditory temporal order judgement (PRS) thresholds (ms) in both groups for Experiment 1. Good Readers are shown to the left and Dyslexics to the right. (Note: Entries in the shaded central frame or 'stem' represent 20 ms bins; numbers on the 'leaves' to the sides show actual PRS thresholds for the two groups.)

previously, the Good Readers' thresholds were significantly less variable than those of children in the Dyslexic group (F[29,21] = 10.894, p = .002).

The characteristics of the distributions reported here appear consistent with results on PRS tasks in the literature as far as they are reported. Overlap across a large proportion of the clinical and normative groups, but noticeably larger variance in Dyslexic than control children, was reported by Tallal (1980), and is seen in the data of Ludlow et al. (1983), Watson (1992), and Protopapas et al. (1997). Most recently, this effect again seems evident in Ahissar et al.'s (2000) results, where this characteristic pattern seems more pronounced in this task compared with frequency discrimination tasks with and without backward masking which also produced differences between reader groups. Thus, it does appear that the present patterns may be typical of performance on this type of task in dyslexic and normal reading children.

#### Experiment 2

The second experiment examined reliability of individual measurements across time and practice. Few published studies appeared to have made repeated ATP measures over time or trained children or adults on these tasks under controlled conditions, but two particular studies informed our predictions.

Bishop, Carlyon, Deeks, and Bishop (1999) reported significant correlations in 9- to 15-yearolds between performance on the Tallal Repetition Test and backward masked tone detection thresholds, and pitch discrimination thresholds measured two years earlier. When Hurford, Schauf, Bunce, Blaich, and Moore (1994) used a syllable discrimination task to measure ATP four times as children increased in age from 6 to 8, group patterns in disabled and nondisabled readers appeared stable over time.

We therefore predicted that significant relationships would be found between PRS thresholds measured in 8- to 12-year-olds on two occasions three to four months apart, and also when measured before and after intensive practice. We predicted that all children would improve with practice, but it was not clear from the literature whether dyslexic children would show differential improvement to normal readers.

## Method

*Design.* Dyslexic and Good Reader groups drawn from the participant pool for Experiment 1 were matched on PRS thresholds measured in Experiment 1 (Time 1). These groups were then measured on the same task between three and four months later (Time 2). One week after Time 2, thresholds were measured on four more occasions (Times 3 to 6) within a two-week period.

*Participants*. Children with PRS thresholds of less than 25 ms in Experiment 1 were removed from the pool, as their potential range for improvement was already very small. Four Dyslexic children out of the six with elevated thresholds were still available, and these children were allocated to a separate subgroup (Poor Initial PRS subgroup) to be examined independently of the Dyslexic and Good Reader groups.

Groups of Dyslexic children (n = 10; females = 4, males = 6) and Good Readers (n = 10; females = 4, males = 6) were formed from the remaining children by matching Dyslexic participants and Good Readers as closely as possible on their initial PRS threshold at Time 1. This was done in order to fix the baseline for measurement. Care was taken to ensure that the essential characteristics of the groups used in Experiment 1 were not lost. Independent sample *t*-tests confirmed that the new Good Reader and Dyslexic groups were again equated for PIQ, chronological age and Parents' Mean Years of Education, but differed significantly on Reading Accuracy (t[18] = 11.47, p = .000).

Descriptive statistics for the Poor Initial PRS subgroup revealed that these children did not differ markedly from the main Dyslexic group for this experiment on any characteristic except PRS threshold. Results for this subgroup are presented in contrast to the two principal groups in this experiment, but no statistical analyses have been conducted because the size of this subgroup was too small to be meaningful.

*Procedure.* No alteration was made to the basic PRS procedure described in Experiment 1, except that this time, before doing the task at Time 3, participants were told their Time 2 threshold. Children were informed that they would perform four more blocks of 100 trials each, and they could win a prize worth up to \$2.50 on each new block if they improved on their previous threshold. If their threshold had previously decreased to below 10 ms, children were required only to equal that level. At Times 2 and 3, full instructions and training were given as described for Experiment 1. However, for all following testing sessions, children completed only the Criterion Phase of the task as a 'warm-up' before proceeding to the PEST procedure.

## Results and discussion

Effects of repeated exposure to the PRS task. At Time 2, both the Good Reader and Dyslexic groups had improved significantly on their threshold measured at Time 1 (see Table 2). The mean for each group fell to around half what it had been originally, and six Good Readers and five children from the Dyslexic group reduced their threshold to half their Time 1 estimate. As will be obvious from Figure 2, there was no significant difference between the groups in amount of improvement. The mean threshold for the Dyslexic subgroup with High Initial PRS was reduced by more than one-third from Time 1 to Time 2, with three of these children improving dramatically, and only one child adding 130 ms to her original threshold. These results demonstrate a very significant effect of simple exposure to the task, rather than practice per se.

*Reliability of PRS over time.* Table 3 shows that even though there was such a strong effect of exposure, children who had a low threshold at Time 1 tended to perform well again, and the reverse. Across the Good Reader and Dyslexic groups, and the High Initial PRS subgroup, performance at Time 2 was highly significantly correlated with Time 1 (r = .67). The correlation appeared to be largely carried by the four Dyslexic children with High Initial PRS, but even

**Table 2** Group mean differences in PRS threshold in Experiment 2 at Times 1, 2 and 6, and group mean improvement in threshold from Time 1 to Time 2, and Time 2 to Time 6

Measure	Good Readers $(N = 10)$ Mean (SD)	Dyslexics $(N = 10)$ Mean (SD)	Dyslexics with High Initial PRS (N = 4) Mean (SD)	Statistical test result (Good Readers vs. Dyslexics only)
PRS threshold at Time 1 (ms)	59.8 (35.3)	70.8 (30.4)	279.7 (83.4)	t(18) = .75, p = .233
PRS threshold at Time 2 (ms)	25.0 (18.5)	38.5 (31.8)	175.8 (177.4)	t(18) = 1.6, p = .130
Improvement: Time 1–Time 2	34.9 (40.1)	32.4 (32.5)	103.9 (166.5)	t(18) = .15, p = .879
PRS threshold at Time 6 (ms)	1.5 (3.7)	19.4 (38.1)	123.9 (243.5)	t(18) = 1.47, p = .076
Improvement: Time 2–Time 6	23.5 (16.4)	19.1 (35.8)	51.9 (74.1)	t(18) = .35, p = .733



**Figure 2** Group mean performance of Good Readers and Dyslexics in PRS from Time 1 to Time 6 in Experiment 2 (Error Bars =  $1 \times SEM$ )

**Table 3** Pearson product-moment correlation matrix for repeated measures of PRS and difference scores; at Times 1, 2 and 6 across the whole sample (N = 24) for Experiment 2

		1	2	3	4	5
1	PRS threshold at Time 1	_				
2	PRS threshold at Time 2	.67**	-			
3	PRS threshold at Time 6	.48**	.93**	-		
4	Improvement Time 1–Time 2	.46**	36*	51**	_	
5	Improvement Time 2–Time 6	.26	16	52**	.52**	-

*Note*: \* *p* < .05; \*\* *p* < .01.

when these data were excluded, the Time 1–Time 2 coefficient for the other Dyslexic children remained moderately high (r = .45).

However, it was also clear in these data that the PRS measure was not consistent across time in indicating an absolute individual deficit in ATP. In Experiment 1, a notional 'normal' range of threshold measured without practice appeared to extend from 0 to 140 ms. Three of the High Initial PRS children (i.e., they were outside this range at Time 1) dropped to thresholds within this range at Time 2.

Performance after extended practice. Figure 2 shows that after training, the means for both groups at Time 6 had dropped to below 20 ms, and it appears that with more trials, both groups would reach asymptote at around 0 ms. There were neither significant differences between the groups at Time 6, nor differential effects of practice from Time 2 to Time 6 between the groups. As already noted, by far the steepest gradient in both the group functions occurs between Times 1 and 2. This pattern is consistent with the findings of two recent studies of learning in ATP. McFarland et al. (1998) found that in normal adults, a considerably greater proportion of learning occurred over the first several sessions of

repeated measures of auditory temporal order judgment (in both frequency and intensity), with more subtle improvement occurring during the following sessions. Bishop et al. (1999) found that in SLIs and language-normal children most improvement in performance occurred between Times 1 and 2, with little learning taking place after that.

The substantial practice effects observed here were convergent with the literature discussed above. They also provide strong evidence for Tallal and colleagues' (Merzenich et al., 1996; Tallal et al., 1996) contentions that they were actually modifying ATP thresholds in their training programme, which similarly provided substantial motivation to progressively improve performance.

The only children who showed a differential response to practice by failing to improve significantly were four Dyslexic children; three of these were from the Dyslexic group and one belonged to the High Initial PRS subgroup. However, it seems possible that because many children reached asymptote near zero, there were significant floor effects in these data, leaving the better Good Readers and Dyslexic children no further room for improvement. This may have masked a potential differential effect of practice which might have shown up in a task for which threshold range did not approach 0 ms after practice.

Stability of PRS after extended practice. Data in Table 3 suggest that the reliability of PRS threshold is very high even after extensive practice. Time 2 versus Time 6 threshold estimates were highly significantly related to each other, but the scatter plot for the whole sample indicated that the very high coefficient here rested solely on the outlying estimate of one participant. When the estimates of the High Initial PRS group were removed from the data, the correlation for Time 2 versus Time 6 decreased to r = .50 (p = .006). This seems a more valid quantification of the relationship, but even so, the very restricted range in the Good Readers' performance would undoubtedly have attenuated this correlation. Indeed, it seems that some Dyslexic children also hit a ceiling of performance, and that if there had been more discrimination available in this measure, these estimates of reliability could have been substantially higher.

#### **Experiment 3**

Experiment 3 examined the construct validity of the PRS task by comparing PRS and RAP performance. We believed that the construct validity of the PRS task would be supported. Only a few studies appeared at that time to have correlated performance across any ATP tasks within dyslexic and normal reading children: Farmer and Klein (1993) found that click fusion and PRS were not significantly related (r = -.16), although, as might be expected,

their two-tone PRS and four-tone sequencing tasks were more strongly correlated (r = .38). Watson and Miller (1993) reported that even though they did not get equivalent between-group separations on all tasks, all their temporal measures were significantly correlated with each other (coefficients ranged from .3 to .5). Therefore, on the basis of this and the work by Bishop et al. (1999) mentioned above, we hypothesised that performance on the tasks would be at least moderately related.

#### Method

*Design.* RAP thresholds were measured in a subset of participants from Experiment 1. These thresholds were then compared with participants' earlier PRS performance.

Participants. Thirteen Good Readers (seven female, six male) and 22 Dyslexic children (9 female, 13 male) were selected from the original Experiment 1 pool to represent the full range of PRS threshold values. Groups were arranged to ensure that as far as possible they retained the essential characteristics of the original sample. As in Experiment 1, Good Readers did not differ significantly from Dyslexic children on chronological age (Mean, Good Readers: 119.0 months, SD = 10.1 months; Mean Dyslexics: 125.4 months, SD = 11.8 months), WISC-III PIQ (Mean, Good Readers: 103.4, SD = 9.9; Mean Dyslexics: 101.3, SD = 9.1), or Parents' Mean Years of Education (Mean, Good Readers, 11.8 years, SD = 1.9 years; Mean Dyslexics: 11.5 years, SD = 1.6years). However, Good Readers were significantly better than Dyslexics in NARA-R Reading Accuracy (Mean Reading Good Readers: -12.3 months, Lag, SD = 6.2 months; Mean Reading Lag, Dyslexics: SD = 8.2 months; t[33] = 14.96, 27.0 months, p = .001).

*RAP task.* The basic stimuli for the task consisted of two groups of tones, separated by a one-second interval. Both groups contained a standard, comparison and masking tone. The standard and comparison tones were 25 ms 1000 Hz tones (ramped on and off for 5 ms) separated by a 500 ms interval. The 200 ms masker consisted of the first three harmonics of a sawtooth wave with a fundamental frequency of 1 kHz (0 ms rise/fall times).

The task was closely based on McArthur (1998) and conducted in two phases. In the first phase, individual thresholds for discrimination of intensity were measured (Intensity Discrimination, ID Phase). These individual differences were then built into the testing phase (RAP phase) where the child had to discriminate, before the onset of a masking tone, between two customised pairs, one of which contained a standard and comparison tone of identical intensity, and one of which contained the standard and a softer comparison tone (RAP phase). The child's two-alternative forced-choice task was to say whether the first or second group contained the soft tone.

We modified McArthur's original task (as described by McArthur & Hogben, 2001) by inserting the mask to the ID phase, as well as the RAP phase, because McArthur's results, and our own piloting with the task, suggested that children's intensity thresholds should be measured under the exact conditions that would later prevail in the testing phase. In the ID phase, the mask followed the comparison after 500 ms, but in the RAP phase, it followed the comparison tone after a variable interval.

All tones were presented at 63 dB SPL except for the softer comparison tone. In the ID phase, the 'soft' comparison tone was varied in intensity, and a PEST procedure was used to estimate the child's intensity discrimination threshold. In the RAP phase, the 'soft' comparison tone was set at 3 dB below the child's difference threshold and the interval between the comparison tone and the mask was varied using a PEST procedure to determine the child's rate of auditory processing. We also introduced a ceiling ISI of 500 ms in both phases for the same reasons as in the PRS task in Experiment 1. Children usually completed the ID task together with the first block of trials for the RAP task in one session, followed by a further one or two sessions within 10 days of completing the ID task.

#### Results and discussion

*Between-group comparisons.* Compared to PRS, the RAP task discriminated far less well between Good Readers and Dyslexic children. Figure 3 shows that the distributions overlap for most of the range, and as a result the group mean RAP thresholds did not differ significantly (Good Readers: 129.6 ms [83.7]; Dyslexics: 166.7 ms [93.1]). These results

Good Readers (N = 13)			Dyslexics $(N = 22)$					
			0 - 19					
	52	40	20 - 39	40				
	52	49 66	40 - 39 60 - 79	72				
95	87	87	80 - 99	88	89	(92)	93	94
,,,	107	100	100 - 119	104	116	()2)	,,,	1
	107	134	120 - 139	101	110			
		140	140 - 159					
			160 - 179	163	(176)			
		185	180 - 199	187	188	193	(197)	
			200 - 219				. /	
		233	220 - 239	221	229			
			240 - 259	249				
			260 - 279					
			280 - 299	(292)				
			300 - 319					
			320 - 339					
		348	340 - 359					
			360 - 379					
			380 - 399					
			400 - 419					
			420 - 439					
			440 – 459	441				
			Could not	*	(*)			
			Reach					
			Criterion					
			Criterion					

**Figure 3** Modified back-to-back stem and leaf diagram for Rate of Auditory Processing (RAP) thresholds (ms) in both groups for Experiment 3. (Note: Dyslexics with High Initial PRS (>140 ms) in Experiment 1 are shown by brackets enclosing their threshold value.)

were surprising because McArthur's (1998) findings suggested that our larger sample had the power to detect any between-group effect. Also, the improvements we made should have tightened up the error variance in this data set. Finally, it was also clear that there was no ceiling effect in this task. Therefore, there seems no possibility that greater separation between the groups could have been achieved if the Good Readers had had further room for improvement, which was a possible inference from the results of Experiments 1 and 2.

*Relationship between RAP and PRS.* Three participants (two Dyslexic children and one Good Reader) had elevated RAP thresholds (i.e., well outside those of the remainder of the two groups). As can be seen in Figure 3, only one of these belonged to the subgroup with elevated initial PRS thresholds, but both other children had original PRS thresholds lower than their respective group means.

Performance on the two measures was only weakly related across the whole sample (r = .18, p > .05). However, the coefficient may have been significantly attenuated by the extremely poor RAP performance of two children with low PRS thresholds. When the data from these two children were excluded from this analysis, the two tasks were moderately and significantly related in the remaining children (r = .38, p = .019). A coefficient of almost that value was also obtained for the Dyslexic group alone (r = .33), though it was nonsignificant, with a weaker relationship in the Good Reader group (r = .27). Therefore, it seems that these values may be more representative of the relationship between performance on the two tasks in most children.

## **General discussion**

This investigation showed that by using a standardised procedure for PRS measurement, very similar results to those in the Heath et al. (1999) study could be obtained even in a wider school population of children aged on average some 16 months older than the original participants. Once again, PRS deficits were found in only a small subset of Dyslexic children, and Good Readers did not appear to show these deficits. This finding is consistent with results from other studies indicating that this type of ATP deficit affects only a proportion of the dyslexic group (e.g., Reed, 1989; Tallal, 1980). The size of this proportion was somewhat less than the .45 reported by Tallal (1980), but this could be due to the fact that although our stimuli closely matched the description given by Tallal (1980), it now seems possible that this description was inadequate. As a result, our tones may be inherently less discriminable from one another than those used in the original Tallal work.

## Is PRS a stable, measurable construct?

The standardised PRS procedure did appear to deliver moderate reliability of measurement with repeated exposure to the task and after systematic practice. Taken together with the strong correlations reported by Bishop et al. (1999), the present results suggest that ATP is a stable construct, which can be captured on different occasions by tasks such as this. However, our data demonstrate clearly that the great majority of Dyslexic children showed equivalent effects of learning to those observed in Good Readers, at least on this task. So, if this task is tapping a stable characteristic of temporal processing, how could it be so strongly modified by exposure and training? And indeed, what exactly is being measured in those individuals who exhibit apparent deficits on an ATP task in studies which give small amounts of practice in the absence of feedback (e.g., Tallal, 1980)?

These data also strongly suggest that poor performance on this task does not indicate an absolute deficit in PRS. This is demonstrated by the fact that three out of four of the High Initial PRS children improved dramatically with exposure and/ or practice. Further, although four other Dyslexics did not appear to be progressively learning, these may in fact be the children who really should be considered to have ATP deficits. It is important to note that only one of these children would have been identified by the methodology which we had previously used with this task. We had assumed that it was possible to obtain a valid measure of PRS by taking a threshold estimate as soon as a child achieved a stable function. For many children this would mean that an estimate could be made after as little as only 100 trials. It now appears that children may need to be practised until they reach asymptote before we can measure their PRS.

Our data suggest that to quantify an individual's ATP by taking the first credible threshold estimate, which is the approach characteristic of many ATP investigations as well as our own (e.g., Amitay, Ahissar, & Nelken, 2002; Protopapas et al., 1997; Rosen & Manganari, 2001; Wright et al., 1997), seems at the very least to be inadequate, but may in fact be dramatically misleading in terms of identifying individuals with true ATP deficits. Indeed Ahissar (2001) points out the great individual differences in rate of perceptual learning. Therefore, we suggest routine preliminary administration of a substantial block of trials under exact test conditions, before commencing testing proper, in order to prevent children being identified with apparent ATP deficits, which would be likely to disappear with greater exposure to the task.

## The problem of practice effects

Our results reinforce Bishop et al.'s (1999) warning that threshold estimates on ATP tasks are mean-

ingless unless the level of exposure to the task is known. It seems that exposure to the adaptive procedure itself may need to be controlled for. Mengler (1996) noticed that the PEST functions of SLI and language-normal children were generally more stable after a preliminary block of 100 trials to accustom children to that phase of the task. We would recommend that a routine practice block of 100 trials be introduced to better control the location on individual learning curves at which children's thresholds are measured.

The substantial practice effects observed were convergent with both the wider psychophysical and the ATP literature (e.g., Ahissar & Hochstein, 1993; Tomblin & Quinn, 1983). In fact, the present results appear more dramatic than those reported by Tallal and colleagues (Merzenich et al., 1996; Tallal et al., 1996), even though the children here received far less intensive practice. Such significant practice effects underline the difficulty of comparing ATP results across studies which may have given participants different levels of exposure to the task and practice, but often do not give sufficient details about either (e.g., Farmer & Klein, 1993; Tallal, 1980). Even in a recent study by Ahissar et al. (2000), where individuals were given a range of psychoacoustic tasks examining spectral and temporal resolution ability, no information is given about the level of exposure and practice on tasks before threshold measurement. Hence, it remains very difficult to compare these results to others in the field.

Finally, we must remember that some of the problems of interpreting the data on practice effects for this experiment may be inherent in the parameters of this particular task. Because so many of the participants are within close range of the floor at initial testing on this task, one cannot know for certain that there are no differential effects of practice. A similar effect may be present in some of the data reported by Ahissar et al. (2000). It is possible that an artificial overlap is created between their groups by floor effects comparable to those seen here. It seems essential to use tasks where the range of thresholds is not constrained at the optimal performance end. An example of such a task would be McFarland et al.'s (1998) temporal order judgement for intensity.

# Implications of these results for the construct validity of ATP tasks

Although the results of the first two experiments suggested that the task may be tapping a stable dimension of some sort, the data from Experiment 3 did not permit a clear inference that this task is primarily tapping some underlying ATP dimension. The chances of observing a strong relationship between two ATP tasks, if there were one, ought to have been optimal in this study. There were some very strong controls in place here, with two tasks being presented on the same computer, with similar formatting and response modes, and the same adaptive method of threshold estimation. However, in spite of indications of consistency across tasks in the middle range of performers, the aberrations at the ends of the PRS and RAP distributions pose significant problems for the construct validity of these tasks.

If both tasks do capture some essential temporal dimension such as has been postulated to underlie all the ATP work, and if Heath et al. (1999) were correct in their suggestion that children with temporal processing deficits were able to use their frequency discrimination abilities to process spectral components of the stimulus, then they might be good or moderately good on the PRS task but very bad on the RAP task. This is what we found, but it seems that the converse would not apply. If the two tasks are tapping some essential temporal dimension then children should not be able to be very poor at PRS but good at RAP, and we found two children who were.

There is another possible interpretation of our data. Recently Ahissar et al. (2000) conceptualised ATP tasks as having primarily spectral or temporal demands. They found that adults with a history of childhood reading disorder differed significantly from normal readers on pure tone frequency discrimination with and without backward masking, and on a tone sequencing task similar to our PRS task, but did not show deficits on purely temporal tasks such as backward detection masking. This group failed to report correlations between these two groups of tasks, and at least one other group has reported differences between dyslexic and normal readers in backward detection masking (e.g., Rosen & Manganari, 2001).

However, it is possible that a dissociation between spectral and temporal tasks could explain what we found. If these dimensions are in fact separate, then children could excel at one and be abysmal at the other. This also does not preclude the possibility of some sort of interaction between these two dimensions as originally suggested by Heath et al. Indeed, Ahissar et al. (2000) noted that some dyslexic individuals seemed to have sufficient ability to adequately perform simple spectral tasks (e.g., unmasked pure tone frequency discrimination), but that once temporal constraints were introduced impairment became evident in these individuals (e.g., in the backward masking or two tone sequencing task).

This further suggests that dyslexic individuals may be more susceptible to deficits in performance when their auditory systems are put under unusual pressure. When Mengler (1996) introduced a procedure to control for individual differences in frequency discrimination in the PRS task itself, she failed to find a difference between SLI and languagenormal children. Thus, in two studies of different auditory dimensions, when individual differences on that dimension were controlled, an effect was not found. It may be that what has been captured in the work with ATP is auditory discrimination ability, i.e., differences in ability to deal in fine gradations of the stimulus dimension in which ATP is being measured.

A clue to this possibility seems to have been provided by Tallal's (1980) finding that dyslexic children were significantly less accurate at same/different discriminations of the 'High–Low' tone pairs when these were presented at shorter ISIs. This finding is echoed in Protopapas et al.'s (1997) observation of clear differences between good and poor adult readers asked to make same–different discriminations between pairs of high and low frequency tones with backward masking.

Cacace, McFarland, Ouimet, Schrieber, and Marro (2000) examined temporal order discrimination in both auditory and visual tasks measuring frequency and intensity. They concluded that children with reading impairments showed deficits in discrimination which were not specific to brief or rapidly presented stimuli, and no deficits in simple detection tasks. When Waber et al. (2001) presented learning impaired children with a two-tone discrimination task, they found significant differences between the groups in discrimination ability which were unrelated to rate of stimulus presentation. More recently, Amitay, Ben-Yehudan, Banai, and Ahissar (2002) have reported deficits in reading disabled adults in both visual and auditory tasks requiring fine discrimination of frequency, which were present without temporal constraints.

Thus, there is evidence from several sources that auditory discrimination may be critically involved in tasks which have typically been thought of as ATP measures. This would be consistent with the present results. As the RAP task is specifically designed to equate the pressure placed on intensity discrimination ability across children, it is possible that the vulnerable child's auditory system is simply not under sufficient pressure for deficits in auditory processing to manifest in this task, whereas the more demanding PRS task will do so. Even so, we are still left with having to explain why children with good discrimination do well on the PRS task but perform poorly on the RAP task, by postulating the involvement of some other additional factor (e.g., temporal processing) in the RAP task. Hence, the present data again underline the construct validity problems within ATP research.

## Conclusions

The work described here has led us to seriously question both the use of ATP tasks as measures of individual difference and the very notion of auditory *temporal* processing deficits in children with dyslexia

and SLI. Our data emphasise the construct validity problems in this field and point to an urgent need for a systematic examination of this issue, rather than opportunistic exploration of likely-looking tasks in groups of control and clinical participants. Consistent with conclusions by Ramus (2003) in his review of recent work in this field, we believe that to continue with the primary emphasis on the temporal components of auditory processing is no longer tenable.

We believe that a conceptual shift is required and propose that viewing the deficits typically reported in this body of literature as dysfunctional perception of finer gradations in nonverbal auditory stimuli would more accurately reflect the data. This view would be intuitively easier to integrate with the most recent propositions of leading researchers such as Snowling (2000) that the core deficit in dyslexia lies in inability to form fine-grained mental representations of phonological or speech sound information. It would also still be consistent with the general notion that auditory deficits might give rise to phonological and therefore language and literacy problems.

However, the problem remains that only a percentage of dyslexic children exhibit these auditory deficits. Ramus (2003) estimated this proportion across a range of studies to be 39%. As we have pointed out, there is likely to be enormous error variance in this aggregated sample because the contributory studies will have used a huge range of tasks with widely varied methodologies. Most importantly, different studies are likely to have measured individuals at different places on their learning curves. Nevertheless, we would suggest that the fraction of children with true auditory deficits, which persist even after extended practice, may be somewhere around one-third. Hence, at this stage it does not appear that auditory deficits, whatever their nature, can possibly provide a full causal explanation for the phonological deficits present in many dyslexic individuals.

In view of this fact, our data have very significant implications for the Fast ForWord auditory training programme proposed by Tallal and Merzenich and colleagues (Scientific Learning Corporation, 2003) as an intervention for children with reading difficulties. So far, neither the Tallal group nor anyone else seems to have demonstrated that training in ATP is the necessary ingredient in the putative success of their programme (Ramus, 2003). The present results suggest that if the critical element in their programme is - as Merzenich et al. (1996) claim training in nonverbal auditory processing, two out of three children with dyslexia may require far fewer learning trials than the number offered in their programme to achieve and sustain performance within normal range. This finding would be consistent with longitudinal evaluation of this programme by Hook, Macaruso, and Jones (2001), which suggested that it

is no more effective in improving reading skills than a more traditional reading intervention, even after a two-year follow-up. Hook et al. concluded that 'the intensive amount of time needed to complete the Fast ForWord activities and the expense involved did not seem warranted' (p. 92). However, our data also suggest that approximately one in three dyslexic children will not improve so quickly. These may in fact be the children who require programmes of the extended length proposed by Merzenich et al. (1996).

Ramus (2003) also suggests that recent methodology in this field has greatly improved and we agree, but contend there is still some way to go. Ramus appears to take 'at face value' many of the results he reviews, but when individual studies are closely examined, it is clear that the issues we have raised continue to impede the successful demonstration of reliability and construct validity in this area (e.g., Amitay et al., 2002). Therefore, it is still not possible to know which actual dyslexic *individuals* have what deficits in which area of auditory perception.

Understanding the auditory processing difficulties of some individuals with dyslexia and SLI as deficits in discrimination could provide us with a more viable conceptual framework within which we can move forward. Such a framework could stimulate a more coherent range of better-focused tasks with which to specify the perceptual dysfunction underlying language and literacy deficits. However, if we continue to ignore the issues raised here we are unlikely to know more from a new range of tasks than we do with those presently available.

## Acknowledgements

This research was supported by the Financial Markets Foundation for Children, the Rischbieth Memorial Trust, administered by the Meerilinga Young Children's Foundation, Inc., and the Australian Research Council.

We wish to thank Dr Stuart Rosen for his generous and detailed input to this paper.

## **Correspondence to**

Steve Heath, Department of Psychology, The University of Western Australia, 35 Stirling Hwy, Crawley WA 6009, Australia; Email: steve@psy.uwa.edu.au

## References

- Ahissar, M. (2001). Perceptual training: A tool for both modifying the brain and exploring it. *Proceedings of the National Academy of Sciences*, 98, 11842–11843.
- Ahissar, M., & Hochstein, S. (1993). Attentional control of early perceptual learning. *Proceedings of the National Academy of Sciences*, 90, 5718–5722.

- Ahissar, M., Protopapas, A., Reid, M., & Merzenich, M.M. (2000). Auditory processing parallels reading abilities in adults. *Proceedings of the National Academy of Sciences*, 97, 6832–6837.
- Amitay, S., Ahissar, M., & Nelken, I. (2002). Auditory processing deficits in reading disabled adults. *Journal of the Association for Research in Otolaryngology*, 03, 302–320.
- Amitay, S., Ben-Yehudah, G., Banai, K., & Ahissar, M. (2002). Disabled readers suffer from visual and auditory impairments but not from a specific magnocellular deficit. *Brain*, 125, 2272–2285.
- Bishop, D.V.M., Carlyon, R.P., Deeks, J.M., & Bishop, S.J. (1999). Auditory temporal processing impairment: Neither necessary nor sufficient for causing language impairment in children. *Journal of Speech*, *Language and Hearing Research*, 42, 1295–1310.
- 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 and Neuro-Otology*, *5*, 83–97.
- Cohen, J. (1988). Statistical power analysis for the behavioural sciences. Hillsdale, NJ: Erlbaum Associates.
- Critchley, M. (1970). *The dyslexic child*. London: William Heinemann Medical Books, Ltd.
- Elliott, L.L., & Hammer, M.A. (1993). Fine-grained auditory discrimination: Factor structures. *Journal* of Speech and Hearing Research, 36, 396–409.
- 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.
- Goswami, U., & Bryant, P. (1989). The interpretation of studies using the reading level design. *Journal of Reading Behaviour*, 221, 413–424.
- 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.
- Hogben, J.H., & Di Lollo, V. (1984). Practice reduces suppression in metacontrast and apparent motion. *Perception and Psychophysics*, 35, 441–445.
- Hook, P.E., Macaruso, P., & Jones, S. (2001). Efficacy for Fast ForWord training on facilitating acquisition of reading skills by children with reading difficulties – a longitudinal study. *Annals of Dyslexia*, *51*, 75–96.
- 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.
- Lovegrove, W., Martin, F., & Slaghuis, W. (1986). A theoretical and experimental case for a visual deficit in specific reading disability. *Cognitive Neuropsychology*, *3*, 225–267.
- Ludlow, C.L., Cudahy, E.M., Bassich, C., & Brown, G.L. (1983). Auditory processing skills of hyperactive, language-impaired, and reading disabled boys. In E.Z. Lasky, & J. Katz (Eds.), Central auditory processing disorders: Problems of speech, language and

*learning* (pp. 163–184). Baltimore: University Park Press.

- McArthur, G.M. (1998). *Rate of auditory processing in specific reading disabled and specific language impaired children*. Unpublished doctoral dissertation, The University of Western Australia.
- McArthur, G.M., & Hogben, J.H. (2001). Auditory backward recognition masking in children with a specific language impairment and children with a specific reading disability. *Journal of the Acoustical Society of America*, *109*, 1092–1100.
- McCroskey, R.L., & Kidder, H.C. (1980). Auditory fusion among learning disabled, reading disabled, and normal children. *Journal of Learning Disabilities*, *13*, 69–76.
- McFarland, D.J., Cacace, A.T., & Setzen, G. (1998). Temporal-order discrimination for selected auditory and visual stimulus dimensions. *Journal of Speech, Language and Hearing Research, 41,* 300–314.
- Mengler, E. (1996). Frequency discrimination and auditory temporal order judgement in language disabled children with and without concomitant reading problems. Unpublished honours dissertation, The University of Western Australia.
- 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.
- Neale, M.D. (1988). *Neale analysis of reading ability revised: Manual.* Hawthorn: Australian Council for Educational Research.
- Protopapas, A., Ahissar, M., & Merzenich, M.M. (1997). Auditory processing deficits in adults with a history of reading difficulties. Society for Neuroscience Annual Meeting, New Orleans, October 25–30.
- Ramus, F. (2003). Developmental dyslexia: Specific phonological deficit or general sensorimotor dysfunction? Current Opinion in Neurobiology, 13, 212–218.
- 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.
- Rosen, S., & Manganari, E. (2001). Is there a relationship between speech and nonspeech auditory processing in children with dyslexia? *Journal of Speech, Language, and Hearing Research,* 44, 720–736.
- Scientific Learning Corporation. (2003). http://www.scientificlearning.com.
- Shaywitz, S.E. (1998). Current concepts: Dyslexia. New England Journal of Medicine, 338, 307–312.
- Snowling, M. (1998). Reading development and its difficulties. *Educational and Child Psychology*, 15, 44–58.
- Snowling, M. (2000). *Dyslexia*. Oxford: Blackwell Publishers.
- Standards Association of Australia. (1983). Australian standards on audiometers (number 2586). Sydney: Author.
- Stein, J.F., & Talcott, J. (1999). Impaired neuronal timing in developmental dyslexia – the magnocellular hypothesis. *Dyslexia*, 5, 59–77.
- Studdert-Kennedy, M., & Mody, M. (1995). Auditory temporal perception deficits in the reading-impaired:

A critical review of the evidence. *Psychonomic Bulletin and Review*, *2*, 508–514.

- Tallal, P. (1976). Rapid auditory processing in normal and disordered language development. *Journal of Speech and Hearing*, *3*, 561–571.
- Tallal, P. (1980). Auditory temporal perception, phonics, and reading disabilities in children. *Brain and Language*, 9, 182–198.
- Tallal, P. (2000). Experimental studies of language learning impairments: From research to remediation. In D.V.M. Bishop, & L.B. Leonard (Eds.), *Speech and language impairments in children: Causes, characteristics, intervention and outcome* (pp. 131–155). Hove, East Sussex: Psychology Press Ltd.
- Tallal, P., Miller, S.L., Bedi, G., Byma, G., Wang, X., Nagarajan, S.S., Schreiner, C., Jenkins, W.M., & Merzenich, M.M. (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: Lawrence Erlbaum Associates.
- Tallal, P., & Piercy, M. (1973a). Deficits of non-verbal auditory perception in children with developmental aphasia. *Nature*, *241*, 468–469.
- Tallal, P., & Piercy, M. (1973b). Developmental aphasia: Impaired rate of non-verbal processing as a function of sensory modality. *Neuropsychologia*, *11*, 389–398.
- Tallal, P., & Piercy, M. (1974). Developmental aphasia: Rate of auditory processing and selective impairment of consonant perception. *Neuropsychologia*, *12*, 83–93.
- Tallal, P., & Piercy, M. (1975). Developmental aphasia: The perception of brief vowels and extended stop consonants. *Neuropsychologia*, 13, 69–74.
- Taylor, M.M., & Creelman, C.D. (1967). PEST: Efficient estimates on probability functions. *Journal of the Acoustical Society of America*, 41, 782–787.
- Tomblin, J.B., & Quinn, M.A. (1983). The contribution of perceptual learning to performance on the repetition task. *Journal of Speech and Hearing Research*, *26*, 369–372.
- Tukey, J.W. (1977). *Exploratory data analysis*. London: Addison-Wesley.
- Waber, D.P., Weiler, M.D., Wolff, P.H., Bellinger, D., Marcus, D.J., Ariel, R., Forbes, P., & Wypij, D. (2001).
  Processing of rapid auditory stimuli in school-age children referred for evaluation of learning disorders. *Child Development*, 72, 37–49.
- Warren, R.M. (1974). Auditory temporal discrimination by trained listeners. *Cognitive Psychology*, 6, 237– 256.
- Watson, B.U. (1992). Auditory temporal acuity in normally achieving and learning disabled college students. *Journal of Speech and Hearing Research*, 35, 148–156.
- Watson, B.U., & Miller, T.K. (1993). Auditory perception, phonological processing, and reading ability/ disability. *Journal of Speech and Hearing Research*, *36*, 850–863.

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

Manuscript accepted 22 October 2003