



Symbol-string sensitivity and adult performance in lexical decision

Kristen Pammer^a, Ruth Lavis^a, Charity Cooper^a, Peter C. Hansen^b, Piers L. Cornelissen^{a,*}

^a *University of Newcastle, Newcastle-Upon-Tyne, UK*

^b *University Laboratory of Physiology, Oxford, UK*

Accepted 17 January 2005

Abstract

In this study of adult readers, we used a symbol-string task to assess participants' sensitivity to the position of briefly presented, non-alphabetic but letter-like symbols. We found that sensitivity in this task explained a significant proportion of sample variance in visual lexical decision. Based on a number of controls, we show that this relationship cannot be explained by other factors including: chronological age, intelligence, speed of processing and/or concentration, short term memory consolidation, or fixation stability. This approach represents a new way to elucidate how, and to what extent, individual variation in pre-orthographic visual and cognitive processes impinge on reading skills, and the results suggest that limitations set by visuo-spatial processes constrain visual word recognition.

© 2005 Elsevier Inc. All rights reserved.

Keywords: Vision; Magnocellular; Phonological; Reading; Lexical decision; Anagram; Dyslexia; Pre-orthographic

1. Introduction

Amongst reading researchers, there is consensus that abstract letter identity, independent of font type and case, represents the basic perceptual unit underlying visual word recognition (Besner & McCann, 1987; Grainger & Jacobs, 1996). Recent support for this view was provided by Pelli, Farell, and Moore (2003), who showed that the efficiency of visual word recognition (i.e., the ratio of an ideal observer's word identification threshold to that of a human observer) is inversely proportional to word length. This means that, for a page of text, contrast energy (i.e., the product of squared contrast and 'ink' area) in the image must be divided equally amongst the letters, so that every letter is as visible as every other letter. The alternative possibility, which Pelli et al. refuted, was that contrast energy be shared equally at the word level regardless of word-length. This would mean that individual letters within a word become less

visible as words get longer. By comparing these psychophysical results to an ideal observer model, Pelli et al. showed that human performance in visual word recognition never exceeds that attainable by strictly letter- or feature-based models. Thus, individual letter identification provides the gateway to more complex sublexical and lexical processes, such as phonological and morphological processing.

As with many aspects of object recognition, there is a great deal of uncertainty about how letter representations are extracted from input to the visual system. It has commonly been assumed that letters are identified by analyzing pre-orthographic features (e.g., lines, angles, and curves). Consistent with this view, recent theoretical models of visual word recognition have shown how a level of feature extraction can be included prior to the level at which letters are represented (e.g., Whitney, 2001). Empirically, several studies have indicated that letter features play a role in letter recognition and that similar features occurring in different locations interact with each other, where the degree of interaction depends on the spatial distance between the features

* Corresponding author.

E-mail address: p.l.cornelissen@ncl.ac.uk (P.L. Cornelissen).

(Bjork & Murray, 1977; Chastain, 1977; Krumhansl & Thomas, 1976; Strangert & Brännström, 1975). More recently, Pelli, Burns, Farell, and Moore (2004) found that efficiency for letter identification is independent of stimulus duration, overall contrast, and eccentricity, and is only weakly dependent on size, suggesting that letters are identified similarly across this wide range of viewing conditions. Critically, efficiency is inversely proportional to complexity (perimeter squared over “ink” area) and nearly independent of everything else. This, and the fixed ratio of detection and identification thresholds, lead Pelli et al. (2004) to suggest that letter identification is mediated by detection of some 20 visual features.

From a developmental perspective, research has highlighted the need for better understanding of the relationship between visual processing and the extraction of orthographic information from print. A number of studies have shown relationships between low-level dynamic visual sensitivities and reading skills in school age children and adults. For example, Talcott et al. (2000) found that sensitivity to coherent motion is correlated with orthographic skills, and Sperling, Zhong-lin, Manis, and Seidenberg (2003) have shown that sensitivity to a phantom contour illusion is correlated with children’s orthographic choice accuracy as well as their exception word reading (see also Chase, Ashourzadeh, Kelly, Monfette, & Kinsey, 2003; Cornelissen et al., 1998a, Cornelissen, Hansen, Hutton, Evangelinou, & Stein, 1998b). Results like these suggest that natural variability in pre-orthographic visual processing may constrain visual word recognition.

In our own work, we have pursued the question whether individual differences in pre-orthographic visual processing are related to children’s reading accuracy. To do this, we have departed from studies which investigated early vision (e.g., contrast sensitivity and coherent motion detection). Instead, we have moved upstream to investigate processes that are more closely related to object recognition, like feature extraction. For example, we used a ‘symbol-string’ task to assess children’s sensitivity to the position of briefly presented non-alphabetic but letter-like symbols (Pammer, Lavis, Hansen, & Cornelissen, 2004a). We found that sensitivity in the symbols task explains a unique proportion of the variability in children’s contextual reading accuracy, even when short term memory, intellectual and phonological skills are taken into account. Since the information that differentiates the two alternatives in the symbols task is symbol position, it is plausible that the task is measuring some aspect of pre-orthographic position encoding. However, there are a number of alternative explanations. First, individual variation in the symbols task may simply be related to individual differences in either speed of processing and/or participant’s vigilance. This argument has been thoroughly illustrated with respect to contrast sensitivity losses in developmen-

tal dyslexia (Stuart, McAnally, & Castles, 2001). In this case, computer simulations of observers with poor concentration yielded reduced estimates of contrast sensitivity that were a constant proportion of the true sensitivity across spatial frequencies. Therefore, it is entirely plausible that individual differences in concentration levels in our samples of children could account for much of the observed variability across tasks. Second, the symbols task required participants to view briefly presented, masked stimuli, about which they had to make a decision. Therefore, both the reading and symbols tasks are likely to make demands on the visual working memory system. So, it is plausible that the correlations we found between symbols task sensitivity and reading accuracy might in fact be due to differences in the visual working memory system. Third, it is possible that individual differences in eye-movement control, particularly fixation stability, could account for variable symbols task performance (see e.g., Jordan, Patching, & Thomas, 2003; Murakami, 2004). Fourth, the correlation between symbols task performance and reading in children might reflect some general developmental trend in either the cognitive or visual domain (e.g., Braddick, Atkinson, & Wattam-Bell, 2003), rather than anything specifically to do with the interface between vision and written language. Therefore, if we can show similar correlations in adults, who have mature visual systems and tens of thousands of hours reading experience, we can be sure that such individual differences are not spurious, and may reflect genuine pre-orthographic visual constraints on reading (Pammer et al., 2004a).

In the current study, we explore the relationship between symbols task performance and visual word recognition in adults. We wanted to focus on word recognition processes that occur during fixation, and avoid any additional complexity involved with saccadic eye-movements. Therefore we measured performance in a visual lexical decision task and used sensitivity in this task as our outcome variable. We show that individual variation in the symbols task predicts performance in lexical decision. Then, based on a number of control tasks, we show that this relationship cannot be explained by other factors including: speed of processing and/or concentration (dots task, Experiment 1), short term memory consolidation, (attention dwell time, Experiment 2), or fixation instability (eye-movement recording, Experiment 3).

2. Experiment 1

2.1. Methods

2.1.1. Participants

All our methods conformed to the tenets of the Declaration of Helsinki and also had local ethics committee

approval. Fifty undergraduate students from the University of Newcastle consented to take part in the study and were paid for their participation. All participants performed at or above normal on the Wide Range Achievement Test (WRAT) of reading, and wore their normal acuity correction. The experiments were carried out over two 1–2 h sessions. Five students were excluded from the final analysis either because they responded below chance level in the symbols task, or they failed to complete all experimental tasks. Details of the 45 participants remaining are presented in Table 1.

2.1.2. Stimuli and procedures

All visual stimuli were presented on a Windows PC and an Iiyama Vision Master Pro 17 in. CRT monitor (100 Hz vertical refresh). Stimuli were generated and participant responses (key-press or mouse button press depending on the task) recorded by bespoke software that locked the timing of all events to the vertical screen refresh. This was confirmed with a light sensitive resistor connected to a digital oscilloscope. Participants sat 60 cm from the computer monitor in a windowless room illuminated by a single lamp, placed to minimise glare and specular reflections.

2.1.3. Lexical decision task (LDT)

To assess visual word recognition skills, we used a modified lexical decision task (cf. Cornelissen et al., 1998a; Pammer et al., 2004a). Stimuli were five-letter words, anagrams or nonwords, the latter taken from Castles and Coltheart (1993). Since most lexical decision tasks reported in the literature use nonwords, rather than anagrams, we included some nonwords in our task for direct comparison. To do this, we compared d -prime scores calculated from responses to words and anagrams (LDTanagrams) with d -prime scores calculated from responses to words and nonwords (LDTnonwords). In general, d -prime (d') was calculated as: $d' = Z(\text{Noise}) - Z(\text{Signal} + \text{Noise})$ where $Z(\text{Noise}) = Z$ transform $(1 - p_{\text{false alarm}})$ and $Z(\text{Signal} + \text{Noise}) = Z$ transform $(1 - p_{\text{hit}})$ (Gescheider, 1997). Provided that there is good agreement between these two d -prime scores, we argue that it is reasonable to make inferences from our modified lexical decision task to visual word recognition in general.

Anagrams were produced by switching the internal letter position of five-letter words in a counter-balanced fashion: 1/3 of the anagrams contained 2nd and 3rd letter position swaps (e.g., HOUSE to HUOSE), 1/3 contained 3rd and 4th letter position swaps (e.g., HOUSE to HOSUE) and 1/3 contained 2nd and 4th letter position swaps (e.g., HOUSE to HSUOE). The mean Kucera-Francis frequency of the words was 168.2 ($SD = 240.8$, range = 42–1815). Systematic bigram frequency differences between the three classes of anagram were sought by extracting all the position dependent token frequencies of bigrams from the CELEX psycholinguistic database (Centre for Lexical Information, Nijmegen, The Netherlands). We calculated a position sensitive bigram frequency score for each anagram class and then compared scores across the three classes of anagram. A one factor ANOVA of token frequency determined no statistically significant difference between the three bigram groups $F(2, 213) = 0.65$, $p > .05$. This suggests that there was little information, other than letter position, that distinguished the three types of anagram.

During the experiment, participants were presented one randomised block of 168 stimuli (24 nonwords, 72 anagrams, and 72 words) for lexical decision. Letter-strings appeared for 66 ms (i.e., four screen refreshes of 16.67 ms). See Appendix for justification of presentation times. Each trial consisted of a fixation cross lasting 300 ms followed immediately by the letter string stimulus. Letter strings were then masked for 100 ms, and followed by a blank screen that persisted on screen until the participant responded. Participants were required to respond as accurately as possible by pressing the appropriate mouse button to indicate whether they thought the stimulus was a word or not. There were no time constraints. The computer recorded both the participants' mouse button press as well as the time taken to respond. All letter strings appeared as white letters on a dark grey background, in upper-case Times font. Michelson contrast was ~ 0.9 , and each letter subtended a vertical visual angle of 0.6° .

2.1.4. Symbols task

The symbols task was intended to measure the accuracy with which participants could discriminate

Table 1
Participant characteristics for Experiments 1 and 2

	Experiment 1 $N = 45$		Experiment 2 $N = 34$	
	Mean (SD)	Range	Mean (SD)	Range
Age (years, months)	20.4 (20.9 months)	18.4–26.1	22.1 (36.7 months)	18.8–30.11
WAIS-R SIMILARITIES (standard score)	11.6 (1.7)	8.0–16.0	11.4 (1.9)	8.0–15.0
WAIS-R BLOCK DESIGN (standard score)	13.6 (2.9)	8.0–19.0	13.3 (2.5)	8.0–19.0
WAIS-R DIGIT SPAN (standard score)	10.2 (1.9)	6.0–16.0	—	—
Ishihara	—	—	19.7 (1.5)	14–21

the position of briefly presented non-alphabetic but letter-like symbols. Fig. 1 illustrates the stimulus sequence.

Thus a symbol-string appeared briefly on screen, was masked, and then two alternative symbol-strings were displayed one above the other in a 2AFC (spatial) paradigm. Participants were asked to pick which of the two alternatives they had just been shown; one was the same string, and the other was a string with the same symbols, but arranged in a different order. The correct choice appeared at random in either the upper or lower position, with equal probability. The ‘incorrect’ choices were generated in the same way that anagrams were generated in the lexical decision task: 1/3 contained 2nd and 3rd symbol position swaps, 1/3 contained 3rd and 4th symbol position swaps and 1/3 contained 2nd and 4th symbol position swaps. The stimulus set for this task comprised 26 symbols constructed from only vertical or horizontal lines and they were designed to contain a similar number of line elements to actual letters.

During the symbols task, participants carried out a total of 216 trials divided into three separate blocks. Each trial consisted of a fixation cross lasting 300 ms followed by a target string lasting 100 ms. The target was then masked for 100 ms, and was followed by the force choice alternatives that remained visible until the

participant responded by key press. This initiated the next trial. Stimuli had the same visuo-spatial characteristics (such as luminance and size) as the lexical decision task stimuli.

2.1.5. Dots task

In Experiment 1, we used a correlational design to identify potential relationships between pre-orthographic visual processing (symbols task) and visual word recognition (lexical decision task). It is feasible that some participants may be more vigilant than others when carrying out such an experiment. Consequently, superior performance might be attributable to increased vigilance alone, rather than natural variation in visual processing per se (see Stuart et al., 2001). Alternatively, some participants may process the dynamic properties of these stimuli quicker than others. If either of these explanations were true, we would expect to see high correlations between the lexical decision task, the symbols tasks and any other visual task that has similar temporal processing requirements. In order to search for such effects, we included a modified version of the metacontrast masking paradigm of Pammer and Lovegrove (2001), henceforth referred to as the dots task.

Metacontrast is a backward masking technique in which a mask follows a target in time, but the mask does not overlap the target spatially; instead it appears in a location adjacent to where the target was. When plotting accuracy of target perception (y -axis) against the delay between the onset of the target and onset of the mask (SOA), a U-shaped function is typically found; accuracy is greatest when the target and mask are separated by either short or long intervals, and least (the point of maximum masking) at intermediate intervals. Although the SOA at which maximum masking occurs tends to vary with stimulus parameters, as well as individual differences in speed of processing, a value of approximately 50 ms is common (e.g., Hogben & Di Lollo, 1984; Ogmen, Breitmeyer, & Melvin, 2003; Pammer & Lovegrove, 2001).

For present purposes we wanted a visual control task where good performance most likely depended on a combination of fast visual processing and good concentration. The metacontrast paradigm clearly fulfils this requirement (Pammer & Lovegrove, 2001). Since we were not interested in metacontrast masking per se, we did not systematically vary target/mask SOA in order to measure the complete metacontrast masking function. Instead, we identified a standard target/mask SOA of 50 ms for use with all subjects, where, based on previous research, some degree of masking and hence individual variability in response accuracy was expected (Pammer & Lovegrove, 2001). Accordingly, the dots task, illustrated in Fig. 2, required participants to make a decision about the internal configuration of the whole stimulus, in much the same manner as the symbols and

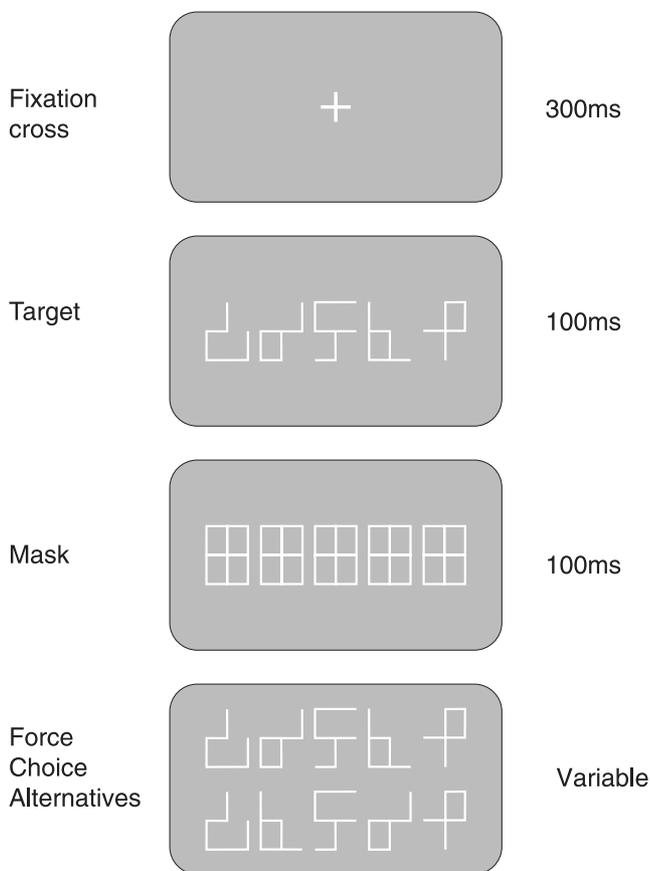


Fig. 1. Schematic to illustrate the stimulus sequence for the symbols task in Experiment 1. See text for details.

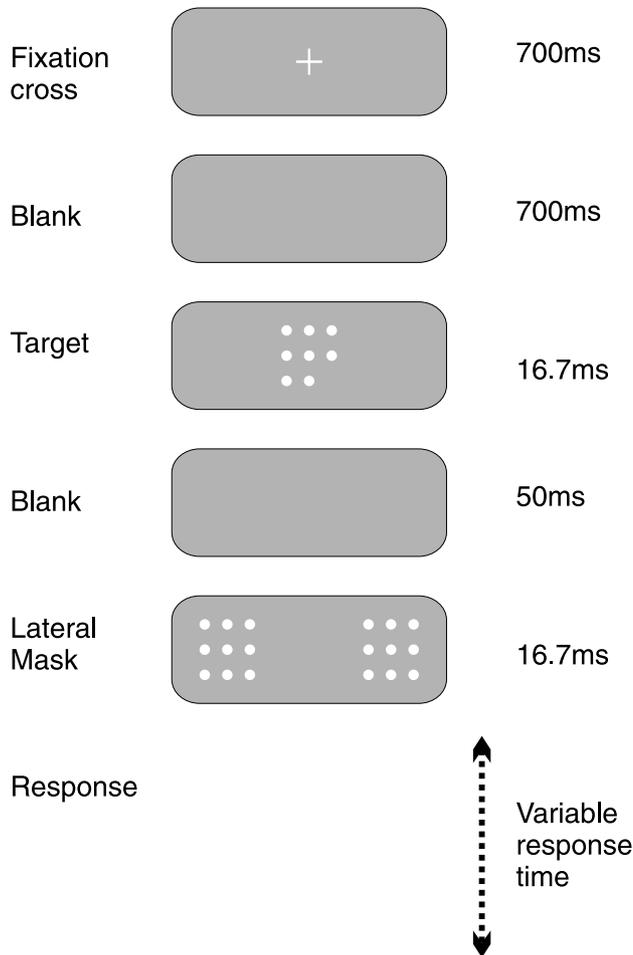


Fig. 2. Schematic to illustrate the stimulus sequence for the dots task in Experiment 1. See text for details.

lexical decision tasks, while at the same time minimising any similarity to word-like stimuli.

Participants were presented with a fixation dot in the centre of the screen followed by a three by three dot matrix for 16.67 ms (i.e., one vertical screen refresh). On each of 144 trials, one dot, randomly determined, was missing from the matrix. This target was followed by a 50 ms delay, then a mask consisting of two whole matrices appeared (for 16.67 ms), one to the left and the other to the right of the first matrix. On each trial, the participant was required to indicate which dot was missing from the target by using the numeric keypad of the computer keyboard. The target/mask dots each subtended a visual angle of 0.07° , and each full nine-dot matrix subtended a visual angle of 0.3° by 0.3° . The mask matrices were separated from the target by 0.2° visual angle on either side of the target. Each display was white on a dark background with a Michelson contrast ~ 0.8 . This task is equivalent to a 9-alternative forced-choice paradigm. Percentage correct responses were converted to d -prime (Gescheider, 1997).

If the between-subject variability in the dots task, the symbols-task and the lexical decision task is simply due

to individual variation in speed of visual processing and/or participant concentration, then d -prime scores for all these tasks should be strongly correlated, because they all tap the same source of variance. However, if real differences in task sensitivity exist which cannot be accounted for by individual variation in vigilance/speed of processing, we expect to see positive correlations between the lexical decision task and symbols task, but not between the dots task and symbols task.

2.1.6. Psychological tests

Our hypothesis is that sensitivity to the symbols task should predict a significant proportion of the variance in lexical decision. If so, this would suggest that pre-orthographic visual processes set limits on the extraction of orthographic information during visual word recognition. However, it is well documented that individual differences in intelligence may also account for inter-subject variability in psychophysical performance (Hirsh & Watson, 1996) as well as lexical decision (Baddeley, Emslie, & Nimmosmith, 1993). For these reasons, all participants were tested on three subtests of the revised Wechsler Adult Intelligence Scale (WAIS-R): (1) *SIMILARITIES*: in which participants are asked to identify the conceptual connection between three items (e.g., “In what way are an ‘orange’ and a ‘banana’ alike?”); (2) *BLOCK DESIGN*: in which participants are asked to reconstruct a red and white geometric pattern using coloured blocks. The pattern is shown to participants on a card and they must produce the correct result as quickly as possible; (3) *DIGIT SPAN*: in which the longest string of numbers that participants can retain in short term memory is identified. These measures allowed us to statistically control for variance associated with such ‘top-down’ cognitive influences.

2.2. Results

2.2.1. Reaction time analysis

In the lexical decision task, participants were asked to respond as accurately as possible without any demand to respond quickly. Nevertheless, strategic differences between individuals could have led to some participants trading speed of response for accuracy. Therefore, to investigate such effects, distributions of reaction times for each participant were individually inspected. To remove the influence of outlying data-points, we applied a relatively conservative criterion by trimming individual distributions for each subject for each condition at their 10th and 90th percentiles (Miller, 1991; Ratcliff, 1993). The average reaction times in the lexical decision task for correct responses to words and anagrams were 637.9 ms (SE 9.9 ms) and 744.3 ms (SE 11.9 ms), respectively ($t = 6.9$, $p < .0001$). Spearman correlations between percentage correct responses and reaction times for words and anagrams were $r = -.01$ ($p > 0.5$) and

$r = -.21$ ($p > .1$), respectively. Thus, encouraging participants to concentrate on accuracy rather than speed of response did not seem to precipitate a speed-accuracy trade-off in this sample of adults.

2.2.2. *d*-Prime analysis: Univariate statistics

Percentage correct and incorrect responses for the lexical decision, symbols and dots tasks were collated and converted to *d*-prime scores. Table 2 shows the overall performance on each of the visual tasks. Table 2 also shows that, on average, participants' *d*-prime scores were systematically higher for the dots task than any other task, while the range for percentage correct in all three tasks is very comparable. Moreover, Shapiro-Wilk tests showed that the distributions for all three tasks did not deviate significantly from normal. The high *d*-prime values for the dots task are attributable to the fact that it is a 9-alternative force choice task, where the guessing level is 0.11, compared to the 2-alternative choice symbols tasks, where guessing is 0.5.

2.2.3. *d*-Prime analysis: Correlations

Pearson correlations between the *d*-prime scores for each of the three blocks of the symbols task all exceeded $r = .6$ ($p < .0001$), suggesting an acceptable degree of task reliability. Table 3 shows the Pearson correlations between the lexical decision task, the symbols task, the dots task and the cognitive measures. Since the symbols task was not correlated with the dots task, this suggests that inter-participant variability in the symbols tasks is unlikely to be explained by differences in concentration/speed of processing alone. Moreover, the absence of a correlation between the dots task and the symbols tasks is not due to a lack of variability in one task relative to another. The ranges of *z*-scores for the two tasks

are comparable: -2.3 to 2.5 for the symbols task; -2.4 to 1.8 for the dots task.

We found a positive correlation between symbol-position and BLOCK DESIGN, which is consistent with the idea that the symbols task taps pre-orthographic, visuo-spatial skills. It should be noted however, that we did not find correlations between the cognitive and lexical decision tasks. We suggest that the very short stimulus exposures in the lexical decision task limited the influence of 'top-down' cognitive factors on performance. If so, this negative result is consistent with the idea that we succeeded in preferentially tapping early processes in visual word recognition. Detailed time-course experiments would be required to test this hypothesis further. Finally, it should be noted that the Pearson correlation between *d*-prime for LDTanagrams and LDTnonwords was $r = .79$ ($p < .0001$), and *d*-prime for LDTnonwords was significantly correlated with the symbols task (Table 3). While this result is reassuring, in that it suggests that our findings may have general relevance to visual word recognition, it should be treated with caution since the number of nonwords that participants were presented was relatively small; a total of 24, as compared to 72 anagrams, and 72 words.

2.2.4. *d*-prime analysis: Multiple regression

Fig. 3 shows scatterplots of *d*-prime for the lexical decision task as a function of *d*-prime for symbols, and *d*-prime for the dots task as a function of *d*-prime for symbols; only the former illustrates a positive association between tasks.

To quantify these effects, we ran multiple regression analyses using PROC REG in SAS (SAS Institute, North Carolina, US). All variables were first converted to *z*-scores for two reasons. First, this is one of a number of convenient ways to centre the data for regression analysis (Altman, 1991). Secondly, it provides a simple way for the magnitudes of the regression coefficients to be compared by inspection. We estimated the variance in lexical decision explained by the symbols task, while simultaneously controlling for effects attributable to the dots task, DIGIT SPAN, SIMILARITIES, BLOCK DESIGN and age. First we fitted the model in full, the results of which are shown in Table 4.

Table 2
Performance on the visual tasks in Experiment 1

	Mean	SE	Range
LDTanagrams (<i>d'</i>)	1.35	0.08	0.31–2.72
LDTnonwords (<i>d'</i>)	1.56	0.11	0.16–3.02
Symbols task (<i>d'</i>)	1.34	0.04	0.58–2.20
Dots task (<i>d'</i>)	2.52	0.11	0.70–3.87

Table 3
Pearson correlations between the visual and cognitive tasks in Experiment 1

	Lexical decision (anagrams)	Lexical decision (nonwords)	Symbols task	Dots task	SIMILARITIES	BLOCK DESIGN
Lexical decision (nonwords)	0.79**	—				
Symbols task	0.39*	0.30*	—			
Dots task	0.35*	0.40*	0.07	—		
SIMILARITIES	–0.16	–0.05	–0.18	0.04	—	
BLOCK DESIGN	0.15	0.08	0.37*	0.08	–0.05	—
DIGIT SPAN	0.06	0.08	0.24	0.24	0.11	0.10

* $p < .05$.

** $p < .005$.

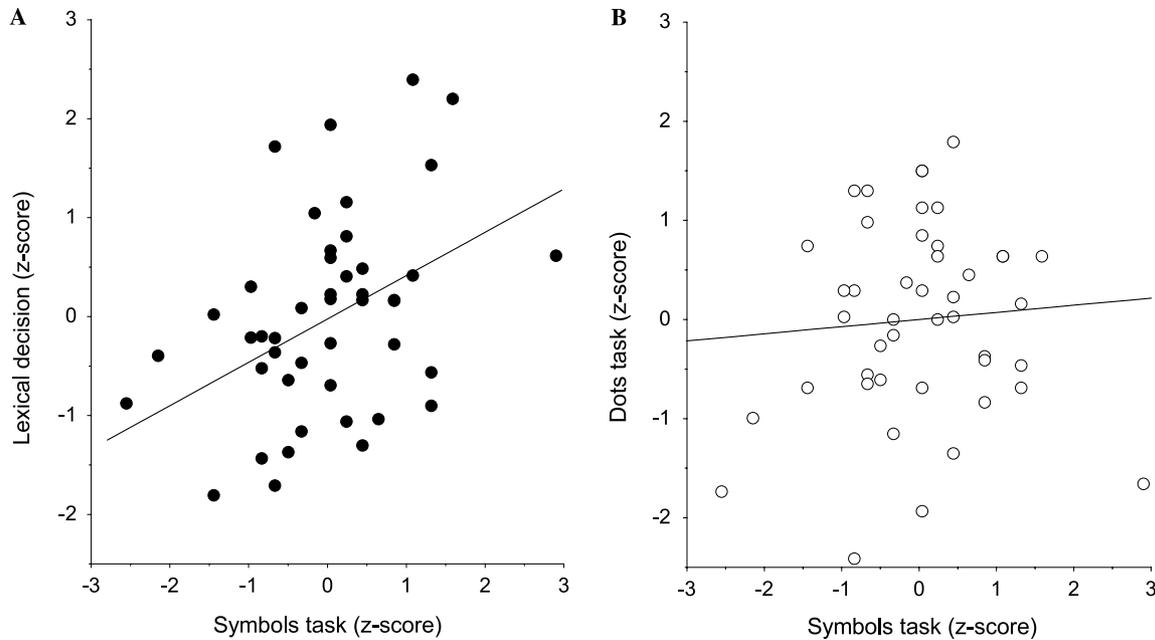


Fig. 3. Plots of: (A) d' -prime for lexical decision as a function of d' -prime for the symbols task, (B) d' -prime for the dots task as a function of d' -prime for the symbols task in Experiment 1. For each dataset, solid lines represent the simple regression of x on y .

Table 4
Output of the full multiple regression models from Experiment 1

Explanatory variable	Regression coefficient	SE	t value	p value
Symbols task	0.34	0.15	2.3	0.02
Dots task	0.38	0.14	2.8	0.008
Age	-0.26	0.14	1.9	0.07
DIGIT SPAN	-0.17	0.15	1.1	0.3
BLOCK DESIGN	0.03	0.14	0.2	0.8
SIMILARITIES	-0.15	0.13	1.1	0.3

Next, we used PROC REG to optimise the model by simultaneously minimising Mallows' C_p statistic, maximising R^2 , and only permitting explanatory variables significant at $p < .05$. According to these criteria, the best fit model explained 25% of the total variance in lexical decision and is:

$$y = 0.4x_1 - 0.32x_2 - 0.015,$$

where $y = d'$ LDT, $x_1 = d'$ symbols ($p = .01$), and $x_2 = d'$ dots ($p < .05$).

2.3. Conclusion

We investigated the relationship between individual performance in the symbols task and lexical decision. In Experiment 1, we found that better performance on the symbols task is associated with greater sensitivity in the lexical decision task even when other cognitive factors are taken into account. Moreover, this association cannot be explained by individual differences in speed of processing/poor concentration alone, because the symbols and dots tasks were not correlated. These findings are

consistent with our previous results with contextual reading in children (Pammer et al., 2004a), and suggests a link between pre-orthographic visual processing ability and visual word recognition. However, an important question remains: what does the symbols task measure?

Since the information that differentiates the two alternatives in the symbols task is symbol position, it is plausible that the task is measuring some aspect of position encoding. To understand the potential relevance of this to visual word recognition, it is important to distinguish between two possibilities: absolute versus relative position encoding schemes. For words, absolute position encoding means that there exist separate representations of each letter in each position in a string, i.e. there are separate units representing "A" in the first position, "A" in the second position, etc. Therefore, for a given input, an optimal match can only occur for the same letter in the same position. Relative-position encoding means that the order of letters is encoded, without specification of their absolute positions, although anchoring of the initial and final letters is often assumed. Evidence for relative-position coding for visually presented words has been obtained from priming studies (Humphreys, Evett, & Quinlan, 1990; Peressotti & Grainger, 1999). Accordingly, the string "GRDN" can prime the target "GARDEN" just as much as "G-RD-N." However, "GDRN" does not prime "GARDEN." Thus, when letter order is preserved, even when absolute position is not, there is priming, but when order is not preserved there is no priming.

The symbols we used were unfamiliar to the participants who took part in the study, so they were very

unlikely to have had access to abstract representations of these stimuli. Since relative-position encoding schemes require abstract representations of ordered sequences, it is unlikely that the symbols task was measuring relative position sensitivity. It is much more likely that, if the symbols task is a measure of position sensitivity at all, then it is a measure of absolute-position sensitivity. When detecting that ABCD differs from ACBD (where A, B, C, and D are any symbols), it is sufficient to notice that the absolute position of features in string-position 2 do not match those in string-position 3.

2.3.1. Within task analysis of the symbols task

An important question is whether there is any evidence to suggest that the representation of the symbol string in memory is in the form of symbol-position subgroups, such as pairs (e.g., AB, BC, CD, etc.) or higher order groups, as opposed to a purely spatial pattern (like a bitmap) where no differentiation exists between individual symbols, their features and their positions. If symbol strings were encoded as undifferentiated patterns, then there should be no differences in sensitivity to the three kinds of position swap (i.e., between positions 2–3, 3–4, and 2–4). However this was not the case;

for the 45 participants mean d -prime scores for these three conditions were: 1.74 (SE 0.11), 1.99 (SE 0.09), and 0.80 (SE 0.07), respectively. A repeated measured ANOVA showed a significant effect of position ($F_{2,134} = 46.2, p < .0001$). Scheffe post hoc comparisons ($\alpha = 0.05$, critical difference = 0.32) confirmed that the non-adjacent, position 2–4 swaps were significantly harder to detect than the two kinds of adjacent position swaps. This suggests one of two possibilities: symbol-strings are indeed encoded as undifferentiated patterns, and it is participants' viewing strategies that resulted in these within task difference. For example, they may have preferred to focus their attention towards one or other end of the symbol-string rather than the middle. Alternatively, the spatial coding system for symbols may operate on a scale smaller than the complete 5-element string and some kind of symbol grouping scheme is one plausible candidate. While the following analysis shows evidence consistent with variable viewing strategies, nevertheless, both possibilities may be still be true, and require further research to elucidate.

Fig. 4A illustrates a target symbol-string and the three kinds of foil that were included as forced choices in the symbols task: position swaps 2–3, 3–4, and 2–4.

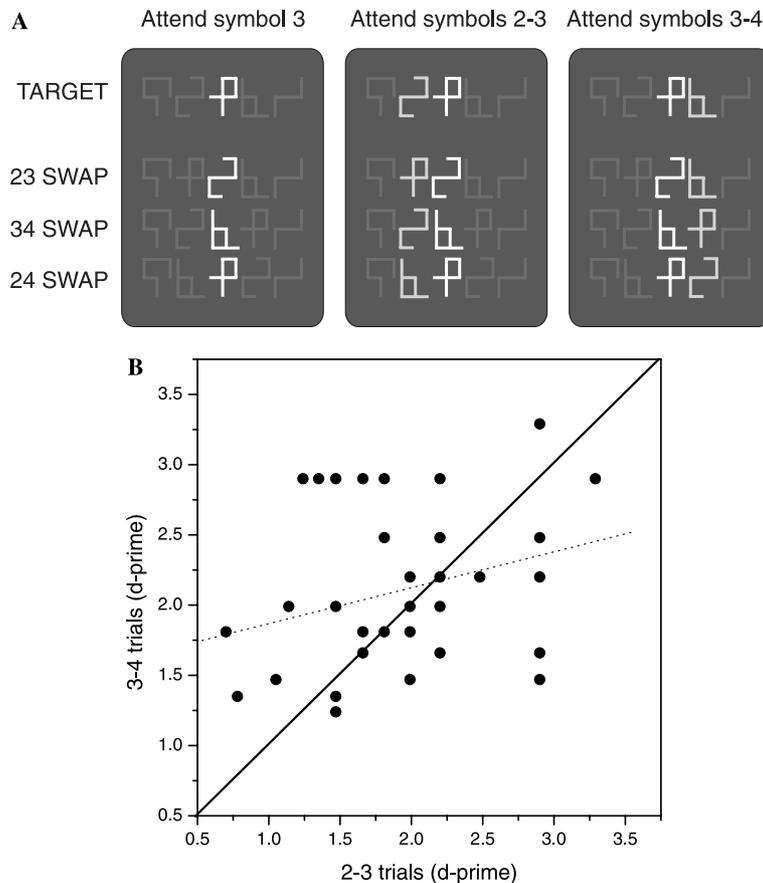


Fig. 4. (A) Schematic to illustrate three possible viewing strategies that participants could have adopted while they carried out the symbols task. See text for details. (B) Plot of d -prime for symbols task trials containing 3–4 foils as a function of d -prime for trials containing 2–3 foils in Experiment 1. The solid line represents a hypothetical regression slope of unity. The dotted line represents the simple regression of x on y .

In addition, Fig. 4A illustrates three different viewing strategies that participants could, in principle, use during the task: (a) Attend to symbol 3 alone; (b) Attend to symbols 2 and 3; (c) Attend to symbols 3 and 4. In Fig. 4A, symbols that are being attended to are indicated by pale grey.

First let us consider the simplest viewing strategy that participants could adopt, ‘Attend symbol 3.’ Since the location of the fixation cross was spatially coincident with the middle symbol of the upcoming target string, participants could choose to attend only to the shape of the central symbol, and ignore all the others. If they had done this, detecting that there was a different symbol in the 2–3 and 3–4 foils would be sufficient to reject them, and solve those trials. However, this strategy would not work for trials containing 2–4 foils, because the central symbol does not change, and participants would have to guess. A breakdown of d -prime scores shows that for 35/45 subjects, sensitivity was indeed lowest for trials containing 2–4 foils (mean = 0.68; $SE = 0.08$). At first glance, this is consistent with an ‘Attend symbol 3’ strategy. [For these 35 subjects, mean d -primes for trials containing 2–3 and 3–4 foils were: 1.9 ($SE 0.11$) and 2.1 ($SE 0.10$), respectively]. But, for this to be a sufficient explanation, we should also expect to see a high correlation between d -prime scores for trials containing 2–3 and 3–4 foils, and this was not the case (Spearman $r = .3$, $p > .05$). Instead, Fig. 4B shows that for this sub-sample of 35 participants, those who successfully rejected 2–3 foils (i.e., higher d -prime scores) tended to be poor at rejecting 3–4 foils (i.e., lower d -prime scores), and vice versa.

Individual differences in viewing strategy would explain this pattern of results. At one extreme, participants mainly attend symbol 3, and partially attend symbol 2 (a leftward bias). At the opposite extreme, participants mainly attend symbol 3, and partially attend symbol 4 (a rightward bias). In the former case (i.e., Attend symbols 2–3 in Fig. 4A) symbol overlap, and therefore potential confusion, exists between the target, 3–4 and 2–4 foils, but not 2–3 foils. In the latter case (i.e., Attend symbols 3–4 in Fig. 4A), symbol overlap exists between the target, 2–3 and 2–4 foils, but not 3–4 foils.

Another important question is whether different viewing strategies in the symbols task are associated with better performance at lexical decision? Spearman correlations between d -prime for 2–3, 3–4, and 2–4 foils in the symbols task and lexical decision were: $r = .31$ ($p < .05$), $r = .27$ ($p = .07$) and $r = .09$ ($p > .1$), respectively. This does suggest that sensitivity to 2–3 swaps in the symbols task is best related to lexical decision performance, but the evidence is not compelling since the difference between the correlation coefficients for 2–3 and 3–4 swaps is just 0.04. Moreover, none of these correlations is larger than that for overall performance in the symbols task and lexical decision (see Table 3). In

a further analysis, we classified whether participants were ‘good’ or ‘poor’ at lexical decision based on a median split. We then carried out a one between groups (i.e., ‘good’ versus ‘poor’ lexical decision), one repeated measures (i.e., 2–3, 3–4, versus 2–4 symbol swaps) ANOVA of symbols d -prime. This failed to show a significant interaction between the location of the symbols swap and lexical decision performance ($F_{2,86} = 0.41$, $p > .5$). Therefore, there is no compelling evidence that specific attention/activation patterns across the symbol strings are associated with good lexical decision performance.

In summary, Experiment 1 suggests that the symbols task is estimating some aspect of pre-orthographic position/shape sensitivity which is relevant to lexical decision. While different individuals appear to use different strategies in the symbols task, nevertheless these strategic differences do not appear to influence the relationship between symbols and lexical decision. Thus, irrespective of participant strategy, the symbols task seems to index between-subject variability in early feature-level processing and this to some extent determines lexical decision performance.

3. Experiment 2

Both the symbols and lexical decision tasks require participants to view briefly presented, masked stimuli, about which they must subsequently make a decision. Therefore, both tasks will make demands on the visual working memory system. The first component of this system, iconic memory, is a purely perceptual component and is thought to have a storage capacity of only three to four items (Sperling, 1960; Vogel, Woodman, & Luck, 2001). The second component of the visual working memory system consolidates transient perceptual representations into durable working memory representations (Vogel & Luck, 2002). It is slow and attention demanding (Jolicoeur & Dell’Acqua, 1998; Potter, 1976). In Experiment 2, we explore whether between-subject variation in the symbols and lexical decision tasks may be attributable to individual variation in the consolidation process of visual working memory. To do this, we sought correlations between *attentional blink* duration and performance in the symbols/lexical decision tasks.

An attentional blink occurs when an observer attempts to detect two targets embedded in a rapid stream of stimuli. Items are displayed at a rate of one every 100 ms or so, each target being masked by the next item in the stream. If the two targets are displayed within a critical temporal interval, identification is nearly perfect for the first but is substantially reduced for the second (Broadbent & Broadbent, 1987; Reeves & Sperling, 1986). The robustness of this second-target deficit (when the first target is being attended) has been confirmed in a

number of psychophysical and electrophysiological conditions (Chun & Potter, 1995).

A number of studies have shown that while the second target is in fact perceived during the attentional blink, nevertheless it cannot be reported accurately (Jolicœur & Dell'Acqua, 2000; Maki, Frigen, & Paulson, 1997; Shapiro, Driver, Ward, & Sorensen, 1997; Vogel, Luck, & Shapiro, 1998). This suggests that impaired second target performance arises at a postperceptual stage. Most explanations of the attentional blink have therefore focused on impairments in working memory in general and the process of working memory consolidation in particular (Chun & Potter, 1995; Giesbrecht & Di Lollo, 1998). These explanations propose that the rapid rate of stimulus presentation in the attentional blink paradigm makes it necessary to protect the perceptual representations of the first and second targets from being overwritten by subsequent stimuli. However, because this consolidation process is relatively slow and attention demanding, the first target may still be undergoing consolidation when the second target is presented. As a result, the second target may be overwritten by subsequent stimuli (Vogel & Luck, 2002).

In Experiment 2, we used the attention dwell time task of Duncan, Ward, and Shapiro (1994) to index short term memory consolidation, by measuring the duration of the attentional blink for each participant. Specifically, we estimated how long after the first target the second target had to be presented in order that participants could detect it correctly 75% of the time (see e.g., Hari, Valta, & Uutela, 1999). We assume that individuals who demonstrate longer attentional blinks are less efficient and/or slower at consolidating visual information into ordered symbolic representations. Therefore, if working memory consolidation plays a role in the lexical decision and symbols tasks, we should find marked correlations between performance in these tasks and attention dwell time.

3.1. Methods

3.1.1. Participants

Fifty-one students at Newcastle University consented to take part in the study, and were paid for their participation. All participants performed at or above normal on the Wide Range Achievement Test (WRAT) of reading, and wore their normal acuity correction. None had anomalous colour vision as defined by the Ishihara Test for Colour Blindness. In Experiment 2, reliable attention dwell time thresholds were impossible to identify in sixteen participants, and they were therefore excluded from the final analysis. A further participant was excluded for performing well below chance (50%) on the symbols task. Descriptive statistics for the remaining 34 participants are reported in Table 1.

3.1.2. Stimuli and procedures

The lexical decision and symbols task were identical to Experiment 1, except that we did not include the non-words in the lexical decision task.

3.1.3. Attention dwell time

During each trial, there was a fixation display ($5^\circ \times 5^\circ$) present at all times, consisting of a central dot and four frames indicating possible stimulus positions; either horizontally to the left and right of fixation, or vertically above and below fixation. Participants initiated each trial by pressing a key. A green digit (2 or 5) was then presented briefly (100 ms) in one of the horizontal locations (left or right) after a random delay of 0–300 ms. This is the first target, T1. A red letter (L or T) then appeared for 100 ms in one of the vertical locations after a further delay of 0–900 ms measured from onset to onset—i.e., Stimulus Onset Asynchrony (SOA). This is the second target, T2. Fig. 5 illustrates the stimulus sequence.

On each trial, the computer selected the letter (L or T) and the digit (2 or 5) at random. Each character was immediately followed by a 250 ms mask to limit visual persistence. At the end of each trial participants identified either the number or the letter (for the single task condition) or both (for the dual task condition). Participants carried out a total of 140 single task and 140 dual task trials. When participants were asked to attend to both characters, they found it harder to identify the second stimulus (T2). Percentage correct responses for single and dual task trials were recorded. For the dual task trials, only those in which T1 had been correctly identified were included. (In our sample, accuracy in identifying T1 never fell below 65% in the dual task condition). Operationally, we defined attentional dwell time as the SOA at which participants correctly identified T2 on 75% of trials in the dual task condition (cf. Hari et al., 1999). We used logistic polynomial regression (see below) to identify this value for each individual.

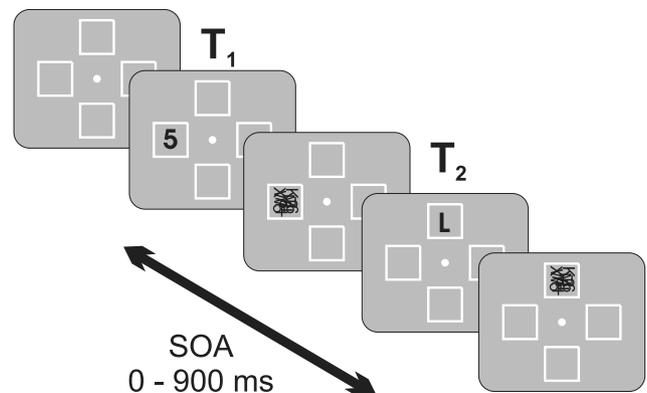


Fig. 5. Schematic to illustrate the stimulus sequence for the attention dwell time task in Experiment 2. See text for details.

3.2. Results

3.2.1. Analysis of attention dwell time

Fig. 6 illustrates the raw data from the attention dwell task. It shows the sample mean probability of a correct response as a function of SOA. Data-points for the single task and dual task are plotted separately. The typical asymmetrical U-shaped function for dual task performance is clear from Fig. 6.

For each individual, we estimated the SOA at which T2 in the dual task was correctly identified on 75% of trials, and used this as an operational definition of an individual's attentional dwell time. To do this, we first inspected the raw data for each individual to check that it conformed approximately to the expected inverted U-shape. In addition, we required that each individual's data fulfil two criteria to ensure reliability: (1) To avoid potential floor and ceiling effects, the dip in the U-shape should occur at a percentage correct somewhere between 50 and 75%; (2) Best performance in the dual task should be 80% or greater, consistent with recovery from interference. On these stringent criteria, the data from 12 participants were excluded.

For the remaining 38 subjects, we used third order polynomial logistic regression to model the dual task data. In the absence of an a priori model, we felt that this was the simplest way to capture the asymmetric shape of these curves. Moreover, we required that the proportion of variance explained (r^2) by each model should exceed 50%, to ensure a reasonable goodness-of-fit for each participant. On this additional criterion, we had to exclude a further four subjects. We used logistic regression because the logit transform (i.e., log odds, or $\log_e(p/(1-p))$) is a convenient way to stabilize the var-

iance of proportionate data (see Altman, 1991). Fig. 7 illustrates the procedure.

The model we fitted to each participant's data was:

$$\log_e(p/1-p) = i + b_1x_1 + b_2x_2 + b_3x_3 + e,$$

where p = proportion T2 correctly named, provided T1 correct,

$$x_1 = \text{SOA}, \quad x_2 = (\text{SOA})^2, \quad x_3 = (\text{SOA})^3.$$

3.2.2. d -Prime analysis: Univariate statistics and correlations

As before, percentage correct and incorrect responses for the lexical decision task and symbols task were converted to d -prime scores. Table 5 demonstrates participants' overall performance on each of the three visual tasks in Experiment 2.

Table 6 shows the Pearson correlations between symbol-position, attention dwell time, SIMILARITIES and BLOCK DESIGN. Interestingly, the lack of correlation between attention dwell time and the symbols task suggests it is very unlikely that the symbols task was limited by working memory consolidation. Moreover, the lack of correlation is not due to a lack of between-subject variability for one task relative to another. The range of z -scores for the two tasks is very comparable: -1.5 to 2.5 for symbols; -1.4 to 2.6 for attention dwell time. We did find a significant negative correlation between attention dwell time and lexical decision; participants with shorter attention dwell times did better at lexical decision.

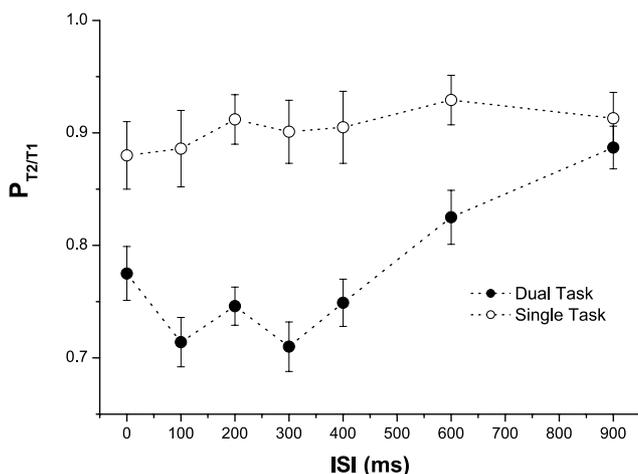


Fig. 6. Plot of the sample mean proportion of correct responses in the attention dwell time task. Single task (open circles) and dual task (closed circles) responses are plotted separately. Error bars represent 1 standard error of the sample mean calculated separately for each SOA.

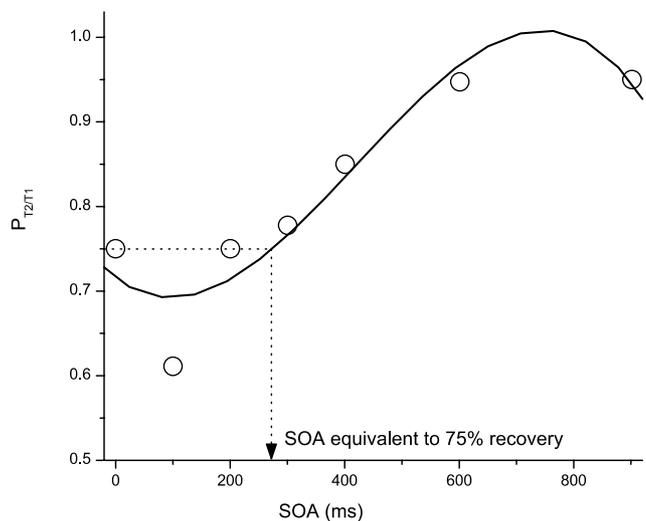


Fig. 7. Procedure used to determine an individual attention dwell time in Experiment 2. The plot shows the individual data points (proportion correct response) as a function of SOA between T1 and T2, together with a third order polynomial fit for one participant. The regression equation was used to locate the SOA (marked by the arrow) at which there was 75% recovery in the dual task condition. See text for further details.

Table 5
Performance on the visual tasks in Experiment 2

	Mean	SE	Range
LDT (d')	0.96	0.08	0.30–2.14
Symbols task (d')	0.74	0.06	0.14–1.73
Attention dwell time (ms)	436.6	32.0	143.4–900.2

We replicated the correlation between the symbols task and BLOCK DESIGN, consistent with both tasks making demands on visuo-spatial skills. But unlike in Experiment 1, we also found a positive correlation between SIMILARITIES and lexical decision.

3.2.3. Multiple regression analysis

Fig. 8 shows a scatterplot of the relationship between d -prime (converted to z -scores) for lexical decision as a function of d -prime for the symbols task (Fig. 8A) and attention dwell time (Fig. 8B), respectively.

We used the same multiple regression procedures as in Experiment 1 to quantify these effects. The full fitted model is shown in Table 7.

The optimised model, in which Mallow's Cp statistic is minimised and R^2 is maximised explained 42% of the total variance in lexical decision and is:

$$y = 0.33x_1 - 0.43x_2 + 0.35x_4,$$

where $y = d'$ LDT, $x_1 = d'$ symbols ($p < .05$), $x_2 =$ attention dwell time ($p < .005$), $x_4 =$ SIMILARITIES ($p < .05$).

Therefore, we replicated the basic result from Experiment 1: individual variation in the symbols task significantly predicts performance in lexical decision. In addition, we found an additional independent effect of attention dwell time.

3.3. Conclusion

In Experiment 2, we asked whether between-subject variation in the symbols and lexical decision tasks may be attributable to individual variation in working memory consolidation. If so, we should have found high correlations between attention dwell times and both the symbols and lexical decision tasks, but we did not. Only attention dwell time was significantly correlated with lexical decision; symbols was not.

One reason for these findings might be that the symbols task is indeed estimating some aspect of pre-orthographic position/shape sensitivity which is relevant to lexical decision, but which is separable from the consolidation process in visual working memory. For example, the symbols task may be indexing early feature-level processing as suggested earlier. Since no prior symbolic representations of these symbols exist—they have no learnt labels—then short term memory processes need not be involved beyond the stage of a transient perceptual representation, and consolidation is not required. If so, we would not expect to see a correlation between attention dwell time and symbols, and we did not. However, in the lexical decision task, not only do features have to be extracted, but also efficient consolidation is required to generate an ordered, symbolic representation of the letter identities and their positions. Since both feature extraction and consolidation are involved, we would expect to see correlations between symbols and lexical decision, as well as between attention dwell time and lexical decision. And we did.

This interpretation implies a sequence of events for lexical decision: first feature extraction (primarily indexed by the symbols task) and then consolidation (primarily indexed by attention dwell time). While our results confirm that between-subject variability at each level is uncorrelated (i.e., no correlation between attention dwell time and symbols), nevertheless there may be interactions between the two. For example, some individuals may be more efficient at feature extraction than others. However, since feature-extraction feeds into the consolidation process, any impact of this variability on lexical decision may in fact depend on the efficiency of the consolidation step. If this is slow and inefficient, lexical decision performance will be reduced whatever the quality of feature extraction. But, if consolidation is fast and effective, then lexical decision performance may then depend much more on the efficacy of feature extraction.

These considerations lead us to an explicit prediction: re-running the multiple regression analysis in Experiment 2 should also reveal a significant interaction term for *symbols task* \times *attention dwell time*. The bottom half of Table 7 shows the output from just such a regression model, and it explains 57% of the variance in lexical

Table 6
Pearson correlations between the visual and cognitive tasks in Experiment 2

	Lexical decision (anagrams)	Symbols task	Attention dwell time	SIMILARITIES	BLOCK DESIGN
Symbols task	0.41*	—			
Attention dwell time	−0.33*	0.13	—		
SIMILARITIES	0.40*	0.37*	0.16	—	
BLOCK DESIGN	0.07	0.35*	0.22	0.29	—

* $p < .05$.

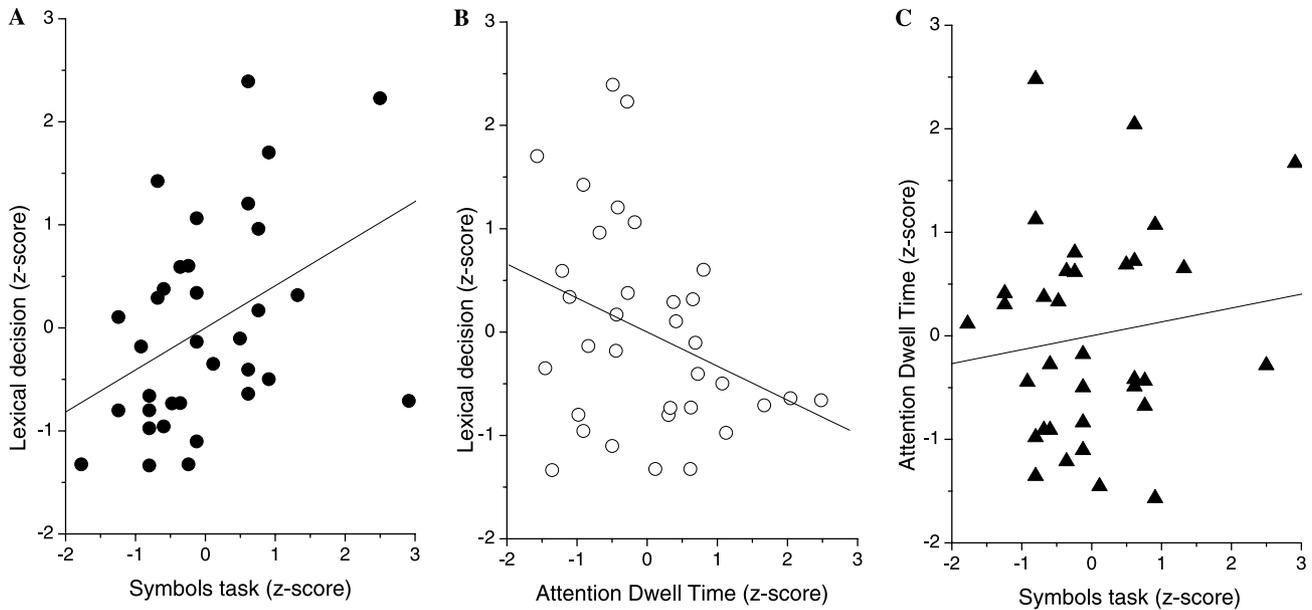


Fig. 8. Plots of: (A) d -prime for lexical decision as a function of d -prime for symbols task, (B) d -prime for lexical decision as a function of attention dwell time, (C) attentional dwell time as a function of d -prime for symbols task in Experiment 2. For each dataset, solid lines represent the simple regression of x on y .

Table 7
Output of the full multiple regression models from Experiment 2

Explanatory variable	Regression coefficient	SE	t value	p value
Symbols task	0.33	0.15	2.1	0.04
Attention dwell time	-0.43	0.14	3.0	0.005
Age	-0.18	0.14	1.3	0.2
BLOCK DESIGN	-0.07	0.15	0.5	0.6
SIMILARITIES	0.39	0.15	2.5	0.02
Symbols task	0.53	0.15	3.3	0.002
Attention dwell time	-0.39	0.13	2.9	0.3
Symbols task \times attention dwell time	-0.37	0.13	2.7	0.01
Age	-0.19	0.13	1.5	0.1
BLOCK DESIGN	-0.07	0.14	0.5	0.6
SIMILARITIES	0.27	0.15	1.9	0.07

decision. As a post hoc control, we repeated the same exercise with the data from Experiment 1, by looking for an interaction between the symbols and dots tasks. As expected, this failed to produce a similar result, suggesting that the *symbol task* \times *attention dwell time* interaction in Experiment 2 is not just a spurious finding. The 3D surface plot in Fig. 9 illustrates the full model shown in the bottom plot of Table 7. According to this model, individuals with slow attention dwell times do indeed demonstrate relatively low sensitivity in lexical decision, irrespective of their sensitivity in the symbols task; consolidation is the rate limiting step. However, as attention dwell times become faster consistent with fast efficient consolidation, then there is an ever increasing advantage for those who are most sensitive to symbols; limitations

are now dictated by the effectiveness of feature extraction.

Finally, as noted in the results section, we rejected the attentional dwell time data in 16 of the 51 original subjects. This happened for several reasons: normally attentional blink tasks, including the attentional dwell time task we used, are run as laboratory experiments by small numbers of highly motivated, highly practiced observers. In addition, participants tend to perform many more trials than in our experiments [e.g., 1600 trials per subject in Kristjansson and Nakayama (2002), compared to our total of 280 trials]. In this study of individual differences, it was simply not feasible to ask many participants to carry out thousands of trials just for one task. Therefore, in order to be reasonably confident

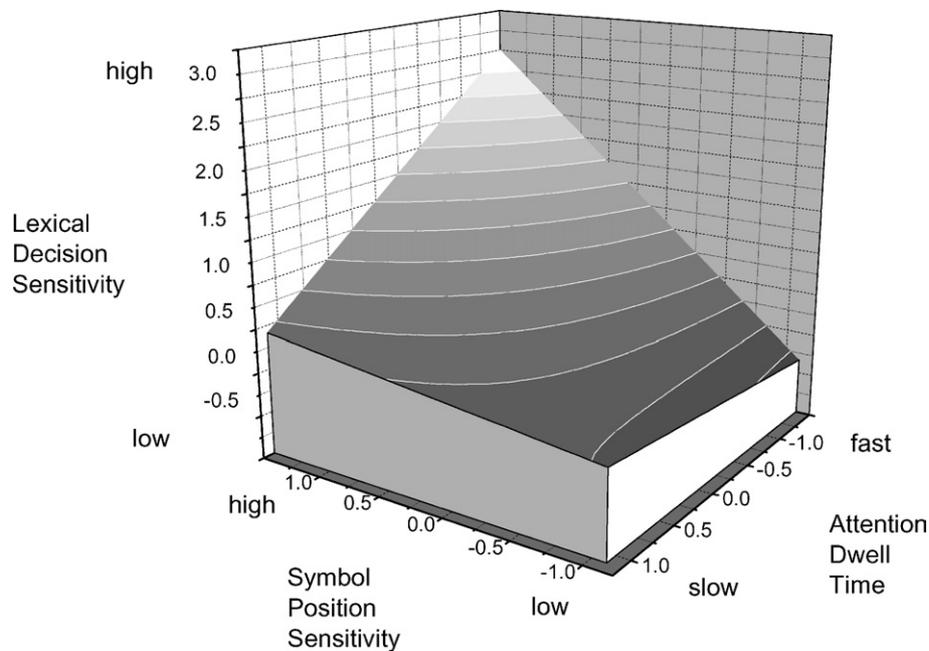


Fig. 9. 3D surface plot to illustrate the best fit multiple regression model for the LDT of Experiment 2, including the interaction term *symbol-position* \times *attentional dwell time*. Symbols task (d') = x-axis, attention dwell time (ms) = y-axis, LDT (d') = z-axis.

that we measured performance in the attention dwell time task as reliably as possible, we had to reach a compromise by running fewer trials, but setting a low threshold for rejecting data.

4. Experiment 3

The central theme of this study is an investigation of inter-participant variability in pre-orthographic position/shape sensitivity and its relation to visual word recognition. In theory, one cause of positional uncertainty could be image movement on the retina caused by variation in fixation movement. Previous researchers have certainly discussed visual performance in relation to small eye movements (McKee & Levi, 1987; Smeets & Brenner, 1994; Tulunay-Keesey & VerHoeve, 1987). Murakami (2004) has shown dramatically how people who are poorer at maintaining fixation are actually poorer at identifying unreferenced motion, and this is because poorer fixation increases internal velocity noise. Here, “unreferenced motion” refers to the situation where a central motion stimulus is presented with nothing in the surround (i.e., the surround was kept at the same luminance as the background). This of course is similar to the symbols and lexical decision tasks in which a central stimulus is surrounded by an extensive, uniform dark background. Although we were not measuring motion per se, our stimuli comprise a rapid sequence of spatio-temporal changes. Therefore, we feel it is justified to determine whether fixation instability might be a confounding factor in these experiments. Hence we mea-

sured participants’ eye-movements while they performed the lexical decision task.

4.1. Methods

4.1.1. Participants

We asked 11 participants from Experiment 1 to return and carry out the lexical decision task while their eye movements were recorded. We could not use the data from one participant because eye-blinks could not be removed satisfactorily from the eye-movement recording. The remaining 10 students were chosen to have amongst the five best, and five worst results on the symbols and lexical decision tasks. As Table 8 shows, apart from WAIS-R BLOCK DESIGN, there was otherwise no difference between the two groups on the cognitive measures.

4.1.2. Apparatus and procedure

We recorded horizontal right eye movements by infrared reflectometry. We used an ASL Model 310 limbus tracker, which has a horizontal recording range of 30°, accurate to 1° horizontally (i.e., the expected difference in degrees of visual angle between true eye position and mean computed eye position during a fixation) and precision 0.25° horizontally (i.e., the amount of jitter in the eye position measurement when the eye is stationary). Analogue signals from the ASL system were sampled at 500 Hz per eye, and recorded onto a PC for off-line analysis. This sample rate is suitable for characterising dynamic movements of the eyes such as saccades and fixation stability.

Table 8

Performance on the cognitive and visual tasks for the participants from Experiment 1 who took part in the eye movement recordings

	Mean (<i>SE</i>)		<i>t</i> value
	5 Best performers	5 Worst performers	
Age (months)	254 (20.5)	241.3 (7.0)	0.7
WAIS-R SIMILARITIES	10.5 (0.64)	11.2 (0.31)	−1.0
WAIS-R BLOCK DESIGN	16.7 (1.44)	12.2 (0.79)	2.8*
WAIS-R DIGIT SPAN	10.0 (1.08)	8.7 (0.76)	1.0
LDT task (<i>d'</i>)	1.98 (0.19)	0.87 (0.14)	4.8**
Symbols task (<i>d'</i>)	1.75 (0.16)	1.08 (0.08)	3.7*
Dots task (<i>d'</i>)	2.31 (1.1)	1.95 (0.37)	0.56

* $p < .05$.** $p < .005$.

Participants were positioned in front of the computer screen at the same distance as for Experiments 1 and 2, and their head supported by a combined chin and cheek rest. The lexical decision task was carried out in exactly the same way as Experiments 1 and 2, except that it was divided into two blocks of 64 trials only. This was because wearing head mounted eye-movement recording equipment is moderately uncomfortable, and infrared illumination of the eyes has a tendency to dry the cornea and conjunctiva. Therefore a compromise was struck between optimising data acquisition, and minimising participant discomfort.

At the start of each block, the eye tracking system was calibrated. Participants were instructed to remain as still as possible throughout the block, which lasted approximately 4 min. Participants were then given a break, after which the ASL system was re-calibrated, in preparation for the second block of 64 trials.

4.2. Results

4.2.1. Analysis of eye-movement data

Inspection of the raw data showed no compelling evidence for participants making saccades in the period around target presentation, so we did not attempt to identify and quantify eye-movements of this kind. To analyse the eye-movement data, an algorithm first identified and discarded epochs contaminated by stereotypical blink artefacts. The remaining blink-free data were divided into individual fixation periods. We defined a fixation period as the 150 ms prior to target presentation plus the 60 ms of target plus the 100 ms of mask. The time points marking these events were identified by TTL pulses generated by the stimulus PC and co-registered with the eye position data by the eye-movement recording PC.

To estimate the stability of each fixation, we calculated the standard deviation (*SD*) of eye position for each fixation period (cf. Cornelissen, Munro, Fowler, & Stein, 1993). This comprised 155 ± 1 samples of eye position per fixation. In this way, we obtained 596 and 677 useable fixations from ‘good’ and ‘poor’ performers,

respectively. Mean fixation stability for each group was 0.072° (*SE* 0.002) and 0.068 (*SE* 0.002), respectively. Fig. 10 shows cumulative probability distributions for the frequency of eye-position *SD* values (0.02 bin size), and reveals that there is negligible difference between the two groups on this measure of eye-movement behaviour.

To quantify these results, we computed a between groups (‘good’ versus ‘poor’ performers in the visual tasks), repeated measures (trial block in the lexical decision task) ANOVA of fixation stability (i.e., *SD* of eye position). The main effect of group was not significant at $p < .05$ ($F_{1,8} = 0.71$, $p > .1$), but the effect of trial block was significant ($F_{1,7} = 19.4$, $p < .01$). Therefore, the results from these eye movement measurements suggest that inter-participant differences in the lexical decision and symbols tasks are unlikely to be attributed to differences in eye movement behaviour. However, the systematic reduction in *SD* of eye position from the first to the second block of trials may suggest some kind of

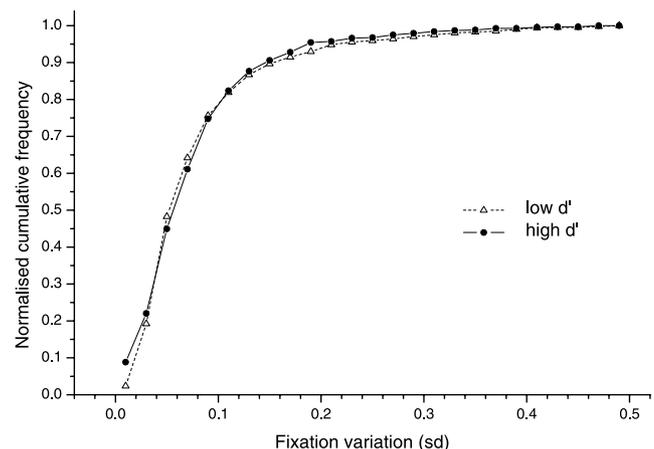


Fig. 10. Cumulative probability distributions for the frequency of eye-position *sd* values (0.02 bin size) calculated from 310 ms fixation periods. Eye movements were recorded during the LDT from the five best (closed circles) and five poorest (open triangles) performers on the lexical decision task and symbols tasks in Experiment 1. See text for details.

adaptation by participants to the equipment. It is possible they were able to ignore the encumbrance of the equipment and focus better on the task. Alternatively, this result may reflect some process like systematic reduction in signal strength, caused by slight eye drying for example.

We acknowledge that the number of participants who took part in Experiment 3 represents only a subset (~25%) of those who took part in Experiment 1. However, the size of the fixation samples ($n = 596$ and $n = 677$) was respectable, and the individuals from whom they were obtained represented the extremes of the relationship between symbols and lexical decision. Therefore, we argue that if fixation instability had been a confounding factor in Experiments 1 and 2, this analysis should have revealed it.

5. General conclusion

The motivation behind this study was to investigate why the symbols task is correlated with contextual reading accuracy in primary school children (Pammer et al., 2004a). First and foremost, we wanted to exclude the possibility that this finding was due to some non-specific developmental effect. Therefore, in the current study we looked for a similar pattern of results in adult readers, who have mature visual systems and tens of thousands of hours reading experience. Secondly, we focused on word recognition processes that occur during fixation, and excluded effects related to saccadic eye-movements by measuring performance in a visual lexical decision task. What we found in two samples of university students is that the symbols task predicts performance in

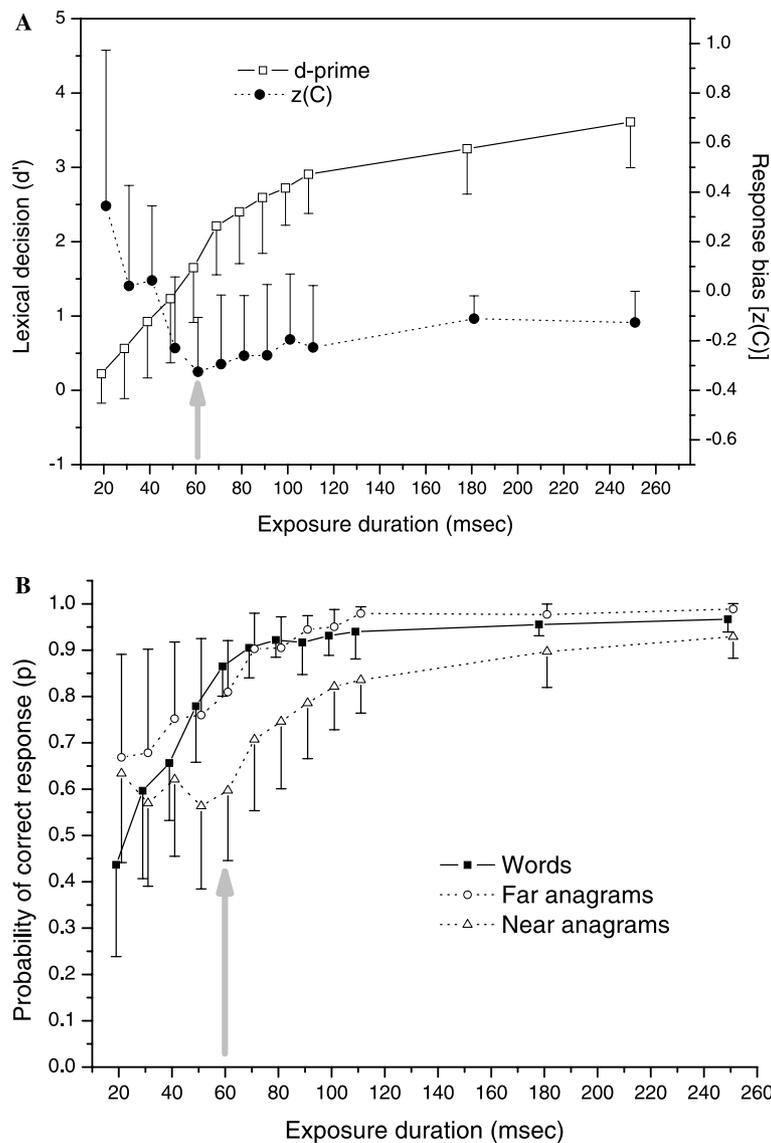


Fig. 11. (A) Plot of d' -prime and response bias as a function of stimulus exposure duration for the pilot study in Appendix. Error bars represent one standard deviation. See text for details. (B) Plot of proportion correct in the lexical decision task as a function of exposure duration. Error bars represent one standard deviation. See text for details.

lexical decision, and that this association cannot be explained by other factors including speed of processing and/or concentration (Experiment 1), short term memory consolidation (Experiment 2), or fixation stability (Experiment 3). These results are therefore consistent with the idea that limitations set by pre-orthographic visuo-spatial processes may constrain visual word recognition and hence reading accuracy.

Traditionally non-words or pseudo-words have been used for comparison with words in lexical decision tasks. However, the use of anagrams in lexical decision tasks is not without precedent (see e.g., Andrews, 1996; Chambers, 1979). We used anagrams here, because we wanted a minimalist intervention which allowed us to disturb the structure of a letter-string so as to break any automatic contact between graphemic, phonological and semantic representations as early as possible in the chain of events that underlie visual word recognition. As discussed in the introduction, we know from behavioural evidence that abstract letter identity represents the basic perceptual unit of visual word recognition (Besner & McCann, 1987; Grainger & Jacobs, 1996; Pelli et al., 2003). Thereafter, in order that letter identities can be mapped onto whole-word representations in memory, evidence suggests that we also compute letter-position (Humphreys et al., 1990; Mason & Katz, 1976; Mason, 1981; Mason, 1982; Peressotti & Grainger, 1999). Therefore, we argue that anagrams of words, in which internal letter positions are swapped, provide us with such a minimalist tool. Only changes in letter position, rather than letter identity, determine whether or not the subject immediately perceives a word. Consistent with this aim, recent magnetoencephalographic (MEG) studies of visual word recognition (Cornelissen, Tarkiainen, Helenius, & Salmelin, 2003; Pammer et al., 2004b; Salmelin, Service, Kiesila, Uutela, & Salonen, 1996; Tarkiainen, Helenius, Hansen, Cornelissen, & Salmelin, 1999) show negligible differences in the early brain responses to words, anagrams, nonwords and random consonant strings (i.e., up to 200–250 ms post-stimulus). Furthermore, the correlation between *d*-prime scores for LDTanagrams and LDTnonwords was high ($r \sim .8$), and the pattern of correlations between LDTnonwords, LDTanagrams and the visual tasks was very similar. In short, it is reasonable to extrapolate from the current results to visual word recognition in general.

The undergraduates who took part in these experiments were naïve participants. While this is ideal for investigating individual variability in visual word recognition, it immediately raises the problem about how well such people perform as experimental subjects; how much of the variability in task performance might be attributed to global factors such as poor concentration or some general notion of speed of processing? Looked at in this simplistic way, both the dots task in Experiment 1 and the attention dwell time task in Experiment

2 can be thought of as appropriate controls, because all visual tasks in both experiments require concentration to perform well, and both the dots task and the attention dwell time task have fast temporal dynamics. Therefore, if there had been some non-specific contamination of the data, related to variable concentration and/or variable speed of processing, we would have found high correlations between all visual tasks, and we did not.

The results presented here demonstrate clear associations between pre-orthographic visual processing and visual word recognition. Further research is required to elucidate what mechanisms underlie these associations, whether they may be causal, and to what extent they apply to contextual reading in adults. At this stage it is also worth considering whether the direction of causality could be the other way round; could extensive practice at learning to read facilitate the ability to rapidly encode any string-like stimulus presented visually? By analogy, there is a debate whether phonological awareness is a cause or a consequence of learning to read, though much evidence supports the idea of reciprocal causation. Learning to read promotes phonological awareness, especially at the sub-syllabic level, as well as vice versa (Goswami, 2003; Perfetti, Beck, Bell, & Hughes, 1987). With respect to the current study, Whitney (2001) and Nazir, Ben-Boutayab, and Decoppet (2004) argue that, in expert readers, the visual system becomes adapted at a low level to process print in special ways. If so, given that the symbols are letter-like, they may trigger the same low-level, bottom-up, print-specific processing as actual letters. Thus the reading specific adaptations in expert readers may help with performance of the symbols task. If so, it is plausible that the symbols task may even be indexing a very early aspect of orthographic processing.

Acknowledgments

The authors thank Donald Boogert for writing the software and Jane Raymond and Burt Rosner for advice with experimental design and data analysis. We thank Joel Talcott for comments on early drafts of this paper, as well as three anonymous referees. The research was supported by the Wellcome Trust and the Research Committee of Newcastle University.

Appendix

We wanted our modified lexical decision task to fulfil two criteria: (1) It should emphasise the role of early processes in visual word recognition by stimuli appearing as briefly as possible, before being masked. (2) It should capture variable performance across the sample

of participants, while avoiding floor or ceiling effects. To identify a stimulus duration for the main experiments that satisfied these criteria, we carried out a time-course experiment with seven adult subjects. Over the course of several days, each subject carried out the complete lexical decision task at 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 180, and 250 ms exposure durations. The order of exposure durations was randomised for each subject. The pilot study was run on a PC system that has a vertical screen refresh of 10 ms, using SuperLab version 2 software for Windows. The plots in Fig. 11A show mean d -prime and mean $z(C)$, which is a measure of response bias, as a function of exposure duration. Fig. 11B shows the mean probabilities of a correct response for words, far anagrams (i.e., 2–4 swaps) and near anagrams (2–3 and 3–4 swaps combined) as a function of exposure duration. For all three plots, error bars represent one standard deviation of the mean. It is clear from these data that an exposure duration of 60msec in the lexical decision task achieves three things, consistent with our requirements. First, it reduces the likelihood of simultaneous ceiling effects for both word and far anagram responses, as well as floor effects for near anagrams (Fig. 11B). Second, it minimises response bias (Fig. 11A). Third, there is clearly respectable variability between individual subjects for d -prime at this time point (Fig. 11A).

References

- Altman, D. G. (1991). *Practical statistics for medical research*. London: Chapman and Hall.
- Andrews, S. (1996). Lexical retrieval and selection processes: Effects of transposed-letter confusability. *Journal of Memory and Language*, 35, 775–800.
- Baddeley, A., Emslie, H., & Nimmosmith, I. (1993). The Spot-the-word test—a robust estimate of verbal intelligence based on lexical decision. *British Journal of Clinical Psychology*, 32, 55–65.
- Besner, D., & McCann, R. (1987). Word frequency and pattern distortion in visual word identification and production: An examination of four classes of models. In M. Coltheart (Ed.), *Attention and Performance XII: The Psychology of Reading*. Hillsdale, NJ: Erlbaum.
- Bjork, M. J., & Murray, J. T. (1977). On the nature of input channels in visual processing. *Psychological Review*, 84, 472–484.
- Braddick, O., Atkinson, J., & Wattam-Bell, J. (2003). Normal and anomalous development of visual motion processing: Motion coherence and 'dorsal-stream vulnerability'. *Neuropsychologia*, 41, 1769–1784.
- Broadbent, D. E., & Broadbent, M. H. P. (1987). From detection to identification: Response to multiple targets in rapid serial visual presentation. *Perception & Psychophysics*, 42, 105–113.
- Castles, A., & Coltheart, M. (1993). Varieties of developmental dyslexia. *Cognition*, 47, 149–180.
- Chambers, S. M. (1979). Letter and order information in lexical access. *Journal of Verbal Learning and Verbal Behaviour*, 18, 225–241.
- Chase, C., Ashourzadeh, A., Kelly, C., Monfette, S., & Kinsey, K. (2003). Can the magnocellular pathway read. Evidence from studies of color. *Vision Research*, 43, 1211–1222.
- Chastain, G. (1977). Feature analysis and the growth of a percept. *Journal of Experimental Psychology*, 3, 291–298.
- Chun, M. M., & Potter, M. C. (1995). A two-stage model for multiple target detection in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 109–127.
- Cornelissen, P. L., Hansen, P. C., Gilchrist, I. D., Cormack, F., Essex, C., & Frankish, C. (1998a). Coherent motion detection and letter position encoding. *Vision Research*, 38, 2181–2191.
- Cornelissen, P. L., Hansen, P. C., Hutton, J. L., Evangelinou, V., & Stein, J. F. (1998b). Magnocellular visual function and children's single word reading. *Vision Research*, 38, 471–482.
- Cornelissen, P. L., Munro, N., Fowler, S., & Stein, J. (1993). The stability of binocular fixation during reading in adults and children. *Developmental Medicine and Child Neurology*, 35, 777–787.
- Cornelissen, P. L., Tarkiainen, A., Helenius, P., & Salmelin, R. (2003). Cortical effects of shifting letter-position in letter-strings of varying length. *Journal of Cognitive Neuroscience*, 15(4).
- Duncan, J., Ward, R., & Shapiro, K. (1994). Direct measurement of attentional dwell time in human vision. *Nature*, 369, 313–315.
- Gescheider, G. A. (1997). *Psychophysics, the fundamentals*. London, New Jersey: Erlbaum.
- Giesbrecht, B. L., & Di Lollo, V. (1998). Beyond the attentional blink: Visual masking by object substitution. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 1454–1466.
- Goswami, U. (2003). Why theories about developmental dyslexia require developmental designs. *Trends in Cognitive Sciences*, 7(12), 534–540.
- Grainger, J., & Jacobs, A. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review*, 103, 518–565.
- Hari, R., Valta, M., & Uutela, K. (1999). Prolonged attentional dwell time in dyslexic adults. *Neuroscience Letters*, 271, 202–204.
- Hirsh, I. J., & Watson, C. S. (1996). Auditory psychophysics and perception. *Annual Review of Psychology*, 47, 461–484.
- Hogben, J. H., & Di Lollo, V. (1984). Practice reduces suppression in metacontrast and apparent motion. *Perception & Psychophysics*, 35, 441–445.
- Humphreys, G. W., Evett, L. J., & Quinlan, P. T. (1990). Orthographic processing in visual word identification. *Cognitive Psychology*, 22, 517–561.
- Jolicœur, P., & Dell'Acqua, R. (1998). The demonstration of short-term consolidation. *Cognitive Psychology*, 36, 138–202.
- Jolicœur, P., & Dell'Acqua, R. (2000). Selective influence of second target exposure duration and Task1 load effects in the attentional blink phenomenon. *Psychonomic Bulletin & Review*, 7, 472–479.
- Jordan, T. R., Patching, G. R., & Thomas, S. M. (2003). Assessing the role of hemispheric specialisation, serial-position processing, and retinal eccentricity in lateralised word recognition. *Cognitive Neuropsychology*, 20(1), 49–71.
- Kristjansson, A., & Nakayama, K. (2002). The attentional blink in space and time. *Vision Research*, 42, 2039–2050.
- Krumhansl, C. L., & Thomas, E. A. C. (1976). Extracting identity and location information from briefly presented letter arrays. *Perception & Psychophysics*, 20, 243–258.
- Maki, W. S., Frigen, K., & Paulson, K. (1997). Associative priming by targets and distractors during rapid serial visual presentation: Does word meaning survive the attentional blink. *Journal of Experimental Psychology: Human Perception and Performance*, 23, 1014–1034.
- Mason, M. (1981). Recognition time for letters and non-letters as a function of retinal location and array length. *Bulletin of the Psychonomic Society*, 18, 75–76.
- Mason, M. (1982). Recognition time for letters and non-letters—effects of serial position, array size, and processing order. *Journal of Experimental Psychology: Human Perception and Performance*, 8, 724–738.

- Mason, M., & Katz, L. (1976). Visual processing of non-linguistic strings: redundancy effects and reading ability. *Journal of Experimental Psychology: General*, *105*, 338–348.
- McKee, S. P., & Levi, D. M. (1987). Dichoptic hyperacuity: The precision of nonius alignment. *Journal of the Optical Society of America A*, *4*, 1104–1108.
- Miller, J. O. (1991). Reaction time analysis with outlier exclusion: Bias varies with sample size. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *43*, 907–912.
- Murakami, I. (2004). Correlations between fixation stability and visual motion sensitivity. *Vision Research*, *44*, 751–761.
- Nazir, T. A., Ben-Boutayab, N., Decoppet, N., et al. (2004). Reading habits, perceptual learning, and recognition of printed words. *Brain and Language*, *88*(3), 294–311.
- Ogmen, H., Breitmeyer, B., & Melvin, R. (2003). The what and where in visual masking. *Vision Research*, *43*, 1337–1350.
- Pammer, K., & Lovegrove, W. (2001). The influence of colour on transient system activity: Implications for dyslexia research. *Perception & Psychophysics*, *63*, 490–500.
- Pammer, K., Lavis, R., Hansen, P., & Cornelissen, P. L. (2004a). Symbol-string sensitivity and children's reading. *Brain and Language*, *89*, 601–610.
- Pammer, K., Hansen, P. C., Kringelbach, M. L., Holliday, I., Barnes, G., Hillebrand, A., Singh, K. D., & Cornelissen, P. L. (2004b). Visual word recognition: the first half second. *Neuroimage*, *22*, 1819–1825.
- Pelli, D. G., Farell, B., & Moore, D. C. (2003). The remarkable inefficiency of word recognition. *Nature*, *423*, 752–756.
- Pelli, D. G., Burns, C. W., Farell, B., & Moore, D. C. (2004). Identifying letters. *Vision Research* (in press).
- Peressotti, F., & Grainger, J. (1999). The role of letter identity and letter position in orthographic priming. *Perception & Psychophysics*, *61*, 691–706.
- Perfetti, C., Beck, I., Bell, L., & Hughes, C. (1987). Phonemic knowledge and learning to read are reciprocal: A longitudinal study of first grade children [Special issue]. *Merrill-Palmer Quarterly*, *33*, 283–320.
- Potter, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory*, *2*, 509–522.
- Ratcliff, R. (1993). Methods for dealing with outliers. *Psychological Review*, *114*, 510–532.
- Reeves, A., & Sperling, G. (1986). Attention gating in short-term visual memory. *Psychological Review*, *93*, 180–206.
- Salmelin, R., Service, E., Kiesila, P., Uutela, K., & Salonen, O. (1996). Impaired visual word processing in dyslexia revealed with magnetoencephalography. *Annals of Neurology*, *40*, 157–162.
- Shapiro, K., Driver, J., Ward, R., & Sorensen, R. E. (1997). Priming from the attentional blink: A failure to extract visual tokens but not visual types. *Psychological Science*, *8*, 95–100.
- Smeets, J. B. J., & Brenner, E. (1994). The difference between the perception of absolute and relative motion: A reaction time study. *Vision Research*, *34*, 191–195.
- Sperling, G. (1960). The informational available in brief visual presentations. *Psychological Monographs*, *74*(11), 1–29.
- Sperling, A. J., Zhong-lin, L. U., Manis, F. R., & Seidenberg, M. S. (2003). Selective magnocellular deficits in dyslexia: A “phantom contour” study. *Neuropsychologia*, *41*, 1422–1429.
- Strangert, B., & Brännström, L. (1975). Spatial interaction effects in letter processing. *Perception & Psychophysics*, *17*, 268–272.
- Stuart, G. W., McAnally, K. I., & Castles, A. (2001). Can contrast sensitivity functions in dyslexia be explained by inattention rather than a magnocellular deficit? *Vision Research*, *41*, 3205–3211.
- Talcott, J., Witton, C., McLean, M., Hansen, P., Rees, A., Green, G., & Stein, J. (2000). Dynamic sensory sensitivity and children's word decoding skills. *Proceedings of the National Academy of Sciences of the United States of America*, *97*, 2952–2957.
- Tarkiainen, A., Helenius, P., Hansen, P. C., Cornelissen, P. L., & Salmelin, R. (1999). Dynamics of letter-string perception in the human occipito-temporal cortex. *Brain*, *122*(11), 2119–2131.
- Tulunay-Keesey, U., & VerHoeve, J. N. (1987). The role of eye movements in motion detection. *Vision Research*, *27*, 747–754.
- Vogel, E. K., Luck, S. J., & Shapiro, K. L. (1998). Electrophysiological evidence for a postperceptual locus of suppression during the attentional blink. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 1656–1674.
- Vogel, E. K., & Luck, S. J. (2002). Delayed working memory consolidation during the attentional blink. *Psychonomic Bulletin & Review*, *9*(4), 739–743.
- Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 92–114.
- Whitney, C. (2001). How the brain encodes the order of letters in a printed word: The SERIOL model and selective literature review. *Psychonomic Bulletin & Review*, *8*, 221–243.