An Evaluation of a Visual Biofeedback Intervention in Dyslexic Adults

Elizabeth Liddle*, Georgina Jackson and Stephen Jackson

School of Psychology, University of Nottingham, University Park, Nottingham NG7 2RD, UK

A prototype of a biofeedback system designed to treat dyslexia by improving heart-rate variability was evaluated in a single blind study of dyslexic adults. Treatment consisted of four 15 minute exposures to a visual display synchronized with either the participant's own cardiac cycle (intervention condition), or of a synthesized cardiac cycle (placebo condition). Repeated measures were made of picture naming speed, single word reading speed and accuracy, copying speed, heart-rate variability and performance on a lateralized visual temporal order judgement task. Small but significant improvements were found in reading and naming speed in the treatment group relative to the placebo group. No significant improvements were found in unspeeded reading measures. Results from heart-rate measures indicated that treatment had effected a shift in the ratio between parameters reflecting the influence of the sympathetic and parasympathetic autonomic nervous systems (ANS), respectively, in favour of the parasympathetic. In the temporal order judgement task, participants who received treatment showed a reduced level of overall improvement relative to that seen in those who received placebo, coupled with evidence of a shift in visual attention from left to right hemifield in their pattern of performance. The results are interpreted as indicating that the treatment induces a shift in autonomic balance in favour of the parasympathetic ANS, and that this shift is also reflected in increased efficiency of left cerebral hemisphere circuits implicated in the perceptual-motor processes required for naming and reading fluency. Conversely, it is also reflected in lower spatial awareness of peripheral visual stimuli, particularly those presented to left hemifield.

Keywords: dyslexia; biofeedback; heart-rate variability; autonomic balance; fluency

*Correspondence to: Elizabeth Liddle, School of Psychology, University of Nottingham, University Park, Nottingham NG7 2RD, U.K. Tel.: +0115-846-6363; e-mail: lpxfebl@psychology.nottingham.ac.uk; Contract/grant sponsor: Brightstar Learning Limited (UK)

Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/dys.287

INTRODUCTION

This paper reports the results of a pilot evaluation of a treatment for dyslexia developed by Brightstar Learning Limited (UK) on bio-feedback principles, in which visual stimuli were synchronized with on-line measurements of heart-rate. The study was conducted on a prototype version of the intervention. The system is now commercially available.

The designers of the intervention intended that by synchronizing visual stimuli with cardiac phase, heart-rate variability (HRV) would be increased. Heart-rate is governed by both the sympathetic and the parasympathetic autonomic nervous systems (ANS). Increased sympathetic ANS activity tends to increase heart-rate, while increased parasympathetic ANS activity tends to reduce it. The designers anticipated that visual stimuli of a particular type presented during a particular phase of cardiac cycle would result in intermittent lengthening of the beat-to-beat interval, thus 'simulating' the parasympathetic ANS influence on HRV, and increasing overall variability. They anticipated that this would improve the responsiveness of the ANS to attentional demands, facilitating reading fluency.

The sympathetic ANS is sometimes characterized as governing the 'fight or flight' response, in contrast to the 'rest and digest' response governed by the parasympathetic ANS (Dodd & Role, 1991). Sympathetic ANS activity accelerates heart-rate in preparation for energetic action. It is implicated in alerting and vigilance, control of which may be lateralized to frontal and parietal regions of the right hemisphere (Posner & Petersen, 1990; Wittling, Block, Schweiger, & Genzel, 1998; Yoon, Morillo, Cechetto, & Hachinski, 1997). In contrast, parasympathetic ANS activity decelerates heartrate; heart-rate deceleration is in turn associated with the anticipatory period preceding response to an expected stimulus, and appears to be associated with the ability to ignore irrelevant stimuli (van der Molen, 2000). There is some evidence that parasympathetic control of heart-rate may be lateralized to the left hemisphere (Oppenheimer, Kedem, & Martin, 1996; Wittling, Block, Genzel, & Schweiger, 1998). The parasympathetic ANS is implicated in visual accommodation, meeting the needs of near vision, including rapid changes in near focus (Chen, Schmid, & Brown, 2003), and pupil constriction, enhancing acuity. Sympathetic stimulation results in pupil dilation, increasing visual sensitivity. In short, the sympathetic ANS tends to promote sustained vigilance, while the parasympathetic ANS tends to promote focussed concentration. The optimal balance between the sympathetic and parasympathetic branches of the ANS is therefore likely to be different for different kinds of attentional task.

Our aim was to evaluate the effects of the bio-feedback intervention in a group of reading-impaired adults by comparing its effects with those of a placebo version. As the intervention was designed as a treatment for dyslexia, we sought firstly to determine whether the intervention would increase fluency and/or accuracy on literacy tasks. Secondly, as the intervention was designed to influence autonomic control, we sought to determine whether it would have any observable effect on heart-rate variability that persisted after treatment. Fluctuations in heart-rate due to sympathetic ANS influence tend to be lower in frequency than fluctuations due to parasympathetic ANS influence (Hyndman, Kitney, & Sayers, 1971; Stein & Kleiger, 1999; Toichi, Sugiura, Murai, & Sengoku, 1997; Vanravenswaaijarts, Kollee, Hopman, Stoelinga, & Vangeijn,

1993). Sympathetic and parasympathetic influence on heart-rate can therefore be quantified by subjecting a recording of its fluctuations obtained by electrocardiogram (ECG) to spectral analysis (Cerutti, Bianchi, & Mainardi, 2001), in order to determine the relative power of heart-rate fluctuations in the low frequency (LF) band that corresponds to sympathetic ANS influence as compared with those in the high-frequency band (HF) that reflect parasympathetic ANS activity. We therefore sought to determine whether the intervention would alter the ratio between the LF (sympathetic) and HF (parasympathetic) components, and in particular, whether it would induce an increase the HF (parasympathetic) component relative to the LF (sympathetic) component.

Thirdly, the evidence cited above suggests that both sympathetic and parasympathetic ANS are implicated in visual attention, the former being implicated in sustained vigilance and the latter in focussed concentration. It also suggests that control of the two ANS branches may be lateralized to right and left hemispheres, respectively. We therefore postulated that if the intervention altered the ratio between sympathetic and parasympathetic ANS activity, there might be detectable changes in performance on a task designed to tap lateralized visual attention to peripheral stimuli.

METHOD

The Biofeedback System

The prototype of the biofeedback system took the form of a visual display screened via a data projector on to a wall in a darkened room. It consisted of visual stimuli moving linearly against a violet background. The stimuli flashed intermittently and unpredictably, were varied in size, speed and acceleration, and emerged in turn and at random from the four quadrants of the screen. A chest-worn sensor belt relayed on-line heart-rate information by radio transmission to the computer running the intervention. This enabled the stimuli to be synchronized appropriately with cardiac phase.

The designers intended the display of these stimuli to be perceived primarily peripherally while the participants' overt attention was concentrated on a central visuo-motor task. A black rectangle therefore occupied the centre of the screen throughout the intervention, on to which the task was screened intermittently. This task consisted of a white-bordered sinusoidally moving track, within which the participant was asked to guide a red square cursor by adjusting pressure on a hand-held grip-force sensor. This central task was screened 8 times during the treatment, for 1 minute at a time. Each block of the task was followed by a 1 minute rest period during which the black rectangle remained on screen. Participants were asked to remain facing the screen with eyes open during the rest periods.

We asked the designers of the intervention to devise a placebo version of the system that would be as similar to the treatment system as possible. In the placebo version, the stimuli were not systematically synchronized with any phase of the participant's own cardiac cycle. Instead, they were synchronized with a computer-generated synthetic heart beat with a degree of variability selected randomly for each participant from a range within the normal range of

the population. In addition, the moving stimuli took a form that the designers considered less likely to effect the desired changes than the treatment stimuli: they were screened against a grey ground, were larger, faster and fewer, and the quadrants from which they emerged were varied less frequently. The central visuo-motor task was identical to that screened during the intervention.

Participants

Thirty-eight participants, aged between 16 and 60, and either self-assessed or previously assessed as dyslexic, were recruited by placing advertisements in local newspapers. They were offered a monetary reward as an incentive, and informed that participation would involve attendance on five separate occasions. Participants were asked about known brain or heart irregularities, epilepsy, migraine, severe vision impairment and known likely physical cause of dyslexia. One participant reported a perinatal cerebellar lesion, and was excluded *a priori* from group analysis. One subject reported suffering periodic migraine, but was accepted into the trial provisionally, and, as she did not report any associated migraine attack, completed the trial. All those recruited were accepted into the study, and were randomly assigned to either intervention or to the placebo group, via a matching procedure that ensured that sex and handedness were equally represented. The study was single blind, in that while the experimenters knew who was in which group, the participants did not; moreover the participants were not told that the experimenters in fact knew who was in which group. All participants attended all sessions, and were informed at the completion of the study as to whether they had received treatment or placebo. Ethics approval was obtained from the Ethics Committee of the School of Psychology, University of Nottingham. All subjects gave signed consent to their participation in the study.

Procedure

Schedule: Each participant attended five sessions. At the first, an initial assessment and pre-test measures were made. The remaining four sessions were treatment sessions. At these, a 15 minute recording of resting heart-rate (ECG) was followed by 15 minutes' exposure to either the intervention or the placebo. At least 2 days and not more than 5 days elapsed between each treatment session, and as far as possible, each participant's treatment sessions were scheduled to occur at the same time of day. Each participant's final treatment session concluded with a post-test of measures administered at pre-test, in order to assess any changes in performance attributable to *four* sessions of treatment.

HRV measures were made at the beginning of each treatment session in order to maintain consistency, and to avoid noise due to diurnal fluctuations. The first (baseline) recording was thus made immediately before the first treatment. Each subsequent recording allowed for the detection of any changes that had occurred in the 2–5 days since the previous treatment. The final recording therefore allowed assessment of any changes detectable 2–5 days following *three* sessions of treatment.

Assessment measures: All participants were assessed for dyslexia using the dyslexia adult screening test (DAST), developed by Fawcett and Nicholson (1998). Items include measures of single word reading speed, copying speed, spelling, nonsense passage reading, phonemic segmentation, backward digit span, postural stability, verbal and semantic fluency and non-verbal reasoning. The DAST generates an 'At Risk Quotient' (ARQ), where a score of 0.7 is taken to be a mild indication of dyslexia, and a score of 1.0 or above is taken to be a strong indication of dyslexia.

Handedness was assessed using the Edinburgh Handedness Inventory (Oldfield, 1971). In cases of mixed handedness, footedness was taken to be the indicator of lateral dominance for the purposes of group matching.

Pre-test and post-test measures: Pre- and post-test measures consisted firstly of scores from three items from the DAST screening battery chosen to tap fluency, namely, Rapid Naming, One Minute Reading and One Minute Writing. Secondly, at both pre- and post-test, participants performed an untimed reading task consisting of single regular, irregular and non-words. This was scored for accuracy only. Lastly, participants performed a lateralized visual temporal order judgement task.

Procedures for pre- and post-test measures

DAST measures: In the Rapid Naming test, participants are given a card on which is printed a series of pictures of common items. There are eight rows of five pictures, and the pictures in the bottom four rows of the card are identical to those in the top four. Participants are asked to name the pictures from left to right, starting at the top row. The tester first names all the items in the top four rows, and participants are given the opportunity to practice on these 20 items, to ensure that all items are recognized and named correctly. For the test itself, the time taken to name all the items on the card is reported in seconds. No penalty is given for giving an item a different name to that used by the tester, as long as it is appropriate. A time penalty of 5 seconds is added to the total time for any other error, including omissions. The same stimulus card was used for both the initial assessment and post-test.

In the One Minute Reading test, the participants are asked to read as many words as they can in 1 minute from a card on which are printed 120 words, arranged in four of columns of 30. The words are graded in difficulty. Participants are given a short practice card first, and told to say 'pass' rather than waste time on a word they think they cannot read. The score for the test is the number of words read correctly in 1 minute. Two alternative word cards, marked Forms A and B, are provided with the test materials. At initial assessment, all participants were given Form A, and at post-test, Form B.

For the One Minute Writing task, subjects are given a card on which is printed a short passage of text. They are asked to copy the text on a sheet of paper, as quickly and neatly as they can. They are given a short practice, and told to include punctuation as printed. They are also told that that although speed is important, points will be deducted if words are incorrectly spelt or illegible. The score is the number of words copied correctly in 1 minute, with points deducted if words are illegible, and for punctuation errors. In the initial assessment, the

passage provided in the DAST was used. For the post test, a similar passage of the same length was devised.

Test of regular, irregular and non-words: The stimuli used at assessment were those devised by Castles and Coltheart (1993). For the post-test, an alternate list was devised. Real words were matched for frequency range with those of the Castles and Coltheart lists, and non-words were devised by exchanging onsets, rimes and syllables of items in the original list to generate alternate non-words of matched difficulty. Both sets of stimuli consisted of 30 regular, 30 irregular and 30 non-words. All three types of word were presented together in a fixed random order, printed three to an A4 page, and presented in a loose-leaf binder, in 18-point font. Participants were told it was a test of accuracy, not of speed and were encouraged to take their time. Scores for each word-type consisted of the number of words out of the total of 30 read correctly.

Temporal order judgement task: This was designed to tap visual spatial/temporal processing, and to detect any lateralized attentional bias. Participants were asked to report the temporal order of the onsets of two visual stimuli presented sequentially to right and left hemifield on the screen of a laptop computer. Following the presentation of a central black fixation cross, two horizontally aligned squares measuring 2 cm × 2 cm appeared on the screen, either side of fixation and 8 cm apart. The left square could appear shortly before, shortly after or simultaneously with, the right square. The first square appeared 1000 ms following fixation, and both squares remained on screen for up to 2000 ms (Figure 1). The Stimulus Onset Asynchrony (SOA: the time interval between the onset of each square) varied in 15 ms steps from 0 to 105 ms, with an additional baseline SOA of 300 ms. Each participant sat approximately 60 cm from the screen, and was asked to judge whether the left or right square appeared first. Participants indicated their responses by means of a left key-press with the left hand, or a right-key press with the right hand. Response keys were adjacent to each other, and aligned centrally with the screen. The dependent measure was the proportion of correct responses at each SOA.

HRV recordings: All four treatment sessions began with a resting electrocardiogram (ECG), recorded for 15 minutes via three disposable adhesive electrode patches applied to the skin of the forearms. Participants were asked not to smoke, or drink tea, coffee or alcohol for 1 hour preceding the session. During the ECG recording, participants sat in an easy chair, with eyes open and legs uncrossed, listening to light music of their choice. They were asked to remain awake, and not to meditate. Data acquired were the inter-beat intervals (IBIs), measured between successive R wave peaks in milliseconds.

Procedures for treatment: Following the recording of resting HRV, participants were then given either the intervention or placebo version of the treatment. All participants, including those receiving the placebo, wore a sensor belt round the chest next to the skin. An aerial was positioned near the subject to receive on-line heart-rate information from the sensor and to relay it to the computer administering the intervention. Skin contact with the belt was maximized by the use of a weak solution of an electrolyte (potassium chloride) plus a wetting agent.

The treatment stimuli were then projected on to a screen. The use of the grip-force sensor for the central visuo-motor task was demonstrated, and the

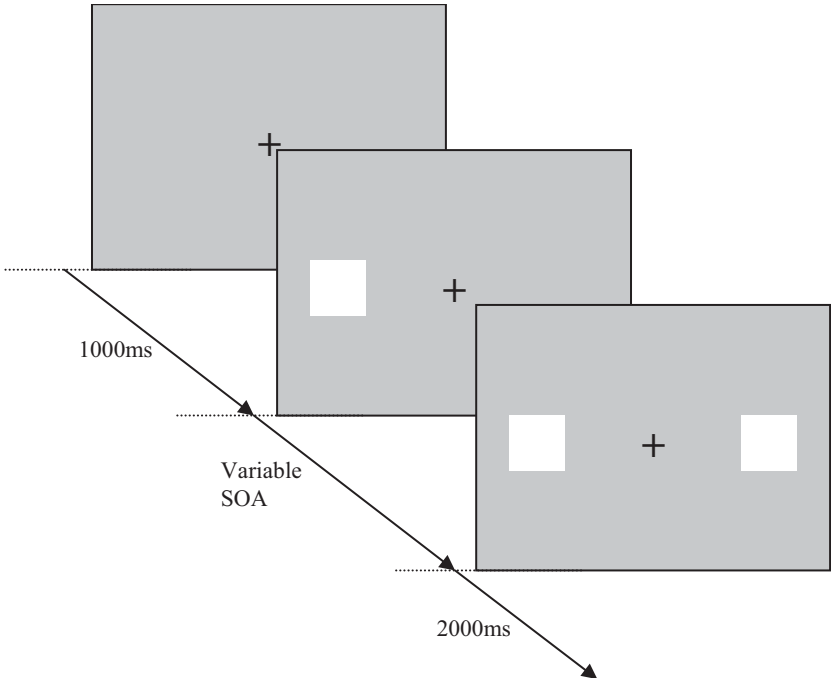

Figure 1. Temporal order judgement task stimuli. A fixation cross was followed after 1000 ms by a white square either to left or right of fixation. A second square appeared in the contralateral hemifield after an SOA that varied randomly in 15 ms steps from 0 to 105 ms, with an additional baseline SOA of 300 ms. The two squares then remained on screen for a further 2000 ms. The participant was asked to report whether the left or right item had appeared first, by means of a key press. Participants were asked to guess if they did not know, and were told that speed of response was unimportant.

participants were told that their task was to keep the red cursor 'on the road' within the sinusoidal track. Participants were asked to remain facing the screen with eyes open during the rest periods. Treatment was terminated after the final block of the task.

Analysis and Results

On examination of the heart-rate variability measurements recorded during the study, three participants were found to have abnormally arrhythmic heart-rates, and were deleted from the study. In addition, a placebo group participant who felt unwell at post-test and performed particularly poorly was deleted as a conservative measure. Together with the participant deleted *a priori*, these deletions left 33 participants in the study, 15 in the placebo group and 18 in the intervention group.

On the baseline assessment, the DAST 'ARQ' scores ranged from 0.4 to 2.4 with a mean of 1.3 and a standard deviation of 0.5. Five subjects scored below the 'at risk' threshold of 0.7; 7 subjects scored above 0.6 but below 1.0; the remaining 21 subjects obtained ARQ scores of greater than 1. All participants reported

functional difficulties in reading, spelling and/or writing and of those with an ARQ of less than 0.7, all produced scores at or below the 11th percentile (Fawcett & Nicolson, 1998) on at least one direct indicator of literacy impairment (Rapid Naming, One Minute Reading, Phonemic Segmentation, Two Minute Spelling, or One Minute Writing). There were no significant initial differences between the groups in either mean or variance of ARQ score, nor of subscores on any items administered in the baseline assessment. There were also no significant differences between groups in either mean or variance of age. Distributions of scores on all measures except regular word reading were approximately normal. Regular word reading scores displayed a truncated range, with a large number of subjects reaching ceiling on this test. This measure was therefore excluded from the analyses. Baseline scores are given in Table 1.

Literacy measures

Main effects: Two multivariate repeated-measures analyses of variance ANOVA were performed on the pre- and post-test literacy measures, one on the three repeated timed DAST measures (Rapid Naming, One Minute Reading and One Minute Writing), and one on the irregular word and non-word reading accuracy from the repeated untimed tests. Experimental group (intervention or placebo) was entered as a between-subjects variable ('group'). On the timed DAST measures, the intervention group showed significantly greater improvement than the placebo group ($F = 3.222; \mathrm{df} = 3, 30; p < 0.05$); on the untimed measures there was no significant difference between groups ($F < 1$).

These five tests were then subjected to individual univariate repeated-measures ANOVAs, the repeated scores of each test in turn being entered as the dependent measure, and experimental group as a between-subjects variable. For all three timed tests, the entire sample of participants showed significant mean improvements at post-test. (One Minute Reading [$F(1, 31) = 40.132, p < 0.001$); Rapid Naming [$F(1, 31) = 35.369, p < 0.001$); One Minute Writing: [$F(1, 31) = 70.180$, $p < 0.001$). There were no significant changes in untimed non-word or irregular word reading scores. However, for two of the timed tests, there was there was also a significant effect of 'group', indicating that the intervention group showed significantly greater improvement than the placebo group in One Minute Reading [$F(1, 31) = 4.827, p < 0.05$], and also in rapid naming [$F(1, 31) = 7.222, p < 0.05$)]. In One Minute Reading, the mean increase in number of words read correctly at post-test by the placebo group was 6.47 (S.D.=9.25) words, whereas the mean increase for the intervention group, was 13.33 (S.D.=8.67) words. The mean percentage increase in Rapid Naming speed for the placebo group was 7%, with a mean improvement of 1.87 (S.D.=3.34) seconds, and 17% for the intervention group, with a mean improvement of 4.94 (S.D.=3.23) seconds. These results are shown graphically in Figure 2.

To test whether baseline scores on either rapid naming or One Minute Reading might have accounted for change in scores at post test, a further ANCOVA was carried out on each of these measures, in which baseline score was entered as a covariate. The group effect in each case remained significant after covarying for baseline score. There was no significant interaction between baseline and treatment group for either measure. Rapid Naming score at baseline was a significant additional predictor of change at post test, with those scoring poorly

Table 1. Baseline assessment scores for all participants included in analysis, ranked by severity as indicated by ARQ on the DAST

DAST sub tests												Castles and Coltheart words		
ARQ	Rapid Naming	One Minute Reading	Postural Stability	Phonemic Segmentation	Two Minute Spelling	Backward Digit Span	Nonsense Passage	Non-verbal Reasoning	One Minute Writing	Verbal Fluency	Semantic Fluency	Regular Words	Non-words	Irregular Words
Placebo group														
0.4	−2	1	−2	0	0	1	0	1	0	0	0	30	29	25
0.8	−1	0	−2	0	−2	0	−1	1	−3	1	1	30	25	27
0.8	−1	−2	−3	−1	0	0	−1	1	0	0	0	30	29	25
0.9	0	−2	*	−2	−3	0	−2	0	0	1	0	30	20	25
0.9	0	−1	0	−2	−1	13	0	0	−3	0	1	30	25	27
0.9	0	1	−3	−3	−2	−2	0	0	0	1	0	30	28	27
1	−2	−2	−1	−2	−3	0	−2	0	−3	0	0	30	24	26
1	0	−2	−3	−3	0	0	−1	1	−1	−1	1	30	25	29
1.2	1	−1	0	−3	−3	−3	−3	1	0	0	0	30	29	27
1.5	−1	−3	−2	−2	−2	−3	−2	0	−1	0	1	29	25	21
1.5	0	−1	−2	−3	−1	−2	−3	−3	−1	1	1	30	21	22
1.5	−1	−3	−3	−2	−2	−2	−3	1	−1	0	0	28	19	20
1.9	−2	−3	−3	−3	−3	−2	−3	−1	−1	0	0	27	14	18
2.2	−3	−3	*	−1	−3	−3	−3	−3	−3	0	0	28	12	26
2.4	−3	−3	−1	−3	−3	−3	−3	−2	−3	12	0	15	4	16
Intervention group														
0.5	0	−2	0	0	−1	0	−2	0	0	1	0	29	29	28
0.5	−1	0	0	−2	−1	1	−1	1	0	1	0	30	28	24
0.6	0	0	−1	0	−2	−2	0	0	−2	1	1	30	30	24
0.7	0	−1	−2	−2	0	−1	0	0	0	0	−2	29	25	21
0.8	−1	0	−3	0	−1	−1	0	1	−2	−1	1	28	30	29
0.9	−3	−2	−3	**	−2	0	**	1	0	1	1	30	30	30
1	−1	−2	*	−3	−2	−1	−1	1	0	0	0	30	29	20
1	0	−2	0	−3	−3	0	−3	1	0	1	0	30	22	26
1.1	−3	−3	−2	−2	0	0	0	0	−1	0	−1	30	29	29
1.4	−1	−2	*	−2	−2	0	−3	−2	0	−2	0	29	21	25
1.4	−1	−2	0	−3	−3	−2	−1	0	−3	0	0	30	25	22
1.4	0	−2	−3	−3	−3	0	−2	0	−2	1	1	30	24	21
1.6	0	−3	−2	−3	−3	−1	−3	−1	−2	1	1	26	22	19
1.8	−3	−3	*	−2	−3	−3	−3	0	−3	1	0	28	18	14
1.9	−1	−3	−3	−3	−3	−1	−3	1	−3	−1	0	24	17	21
1.9	−3	−1	−3	−3	−3	−2	−3	0	−3	1	1	28	15	26
1.9	−1	−3	−3	−3	−3	−3	−2	0	−3	0	0	29	14	22
2	0	−3	−3	−3	−3	−3	−3	−2	−2	0	1	27	10	25

Normed scores on the DAST are given: '−3' = < 5th percentile; '−2'=5th −11th percentile; '−1'=12th–22nd percentile; '0'=23rd–77th percentile and '1'=>78th percentile. Castles & Coltheart word list scores represent number of items out of 30 read correctly.

*Not administered owing to physical fragility.

**Not administered owing to test item familiarity.

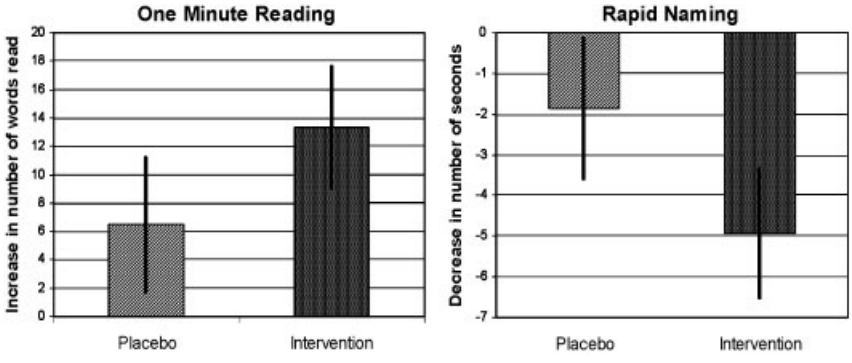

Figure 2. Bar charts showing improvements in 1 minute reading and rapid naming. Error bars represent 95% confidence intervals. For 1 minute reading, the *Y*-axis represents the increase in number of words read correctly in 1 minute. For rapid naming the *Y*-axis represents the change in the number of seconds taken to name all the items on the card. For both these measures the intervention group showed significantly more improvement than the placebo group.

at baseline tending to show greater improvement at post-test [$F(1,30) = 14.244, p < 0.01$], regardless of group membership.

HRV recordings: The heart-rate variability data for each subject at each session was subjected to spectral analysis in order to calculate the power of constituent frequencies (expressed as the area under the curve) in the LF band (0.03–0.15 Hz) assumed to reflect sympathetic ANS influence, and in the HF band (0.15–0.5 Hz) assumed to reflect parasympathetic ANS influence (Cerutti *et al.*, 2001). The LF power was then divided by the HF power in order to arrive at an index of autonomic balance (LF/HF ratio). These ratios showed a skewed distribution in the sample. Natural logarithms of the ratios produced a normalized distribution and were used in statistical tests.

At the baseline session, there was no significant difference between groups in the mean values of the log-transformed LF/HF ratios. In order to determine the time course for any effects of treatment, three repeated-measures ANOVA were performed in which the transformed LF/HF ratios at each session were entered as a within-subjects factor ('treatment session'), and treatment group entered as a between-subjects factor ('group'). In the first ANOVA, only LF/HF ratios from the first two sessions were entered; in the second, ratios from the first three sessions were entered, and in the last ANOVA, ratios from all four sessions were entered. In all three of these ANOVAs multivariate tests showed a significant group effect. In the first ANOVA, representing change after only one session of treatment, a significant 'treatment session' by 'group' interaction [$F(1,31) = 5.229, p < 0.05$] indicated that the LF/HF fell significantly in the intervention group relative to the change in the placebo group which showed a slight rise. In the ANOVA representing changes over two sessions of treatment, the group by session interaction [$F(2,30) = 5.019, p < 0.05$] indicated a continuation of this trend, as did that in the last ANOVA, representing changes over three sessions of treatment [$F(3,29) = 3.353, p < 0.05$].

To ensure that there was no confound from baseline ratios, we then calculated difference values between ratios obtained at the first session and those obtained at sessions 2, 3 and 4, respectively. An ANCOVA was then performed with these

three difference scores as within-subjects factor, and 'group', 'baseline ratio' and a 'group' by 'baseline' interaction term as between-subjects variables. This analysis indicated that higher baseline ratios tended to predict lower subsequent values and *vice versa*, indicating a significant tendency to regress to the mean [$F(1,29) = 4.479, p < 0.05$]. However, the intervention effect remained significant in this analysis; indeed, after thus accounting for variance attributable to regression to the mean, the size of the effect attributable to the intervention increased. The analysis showed that the mean reduction in LF/HF ratio from baseline in the three subsequent sessions was significantly greater in the intervention group than in the placebo [$F(1,29) = 7.822, p < 0.01$]. There was no significant interaction between treatment group and baseline ratio, and no significant differences between any of the last three sessions in the degree of change from baseline observed.

These analyses thus indicated a significant drop in the LF/HF ratios in the intervention group following treatment, relative to changes in the placebo group. However, the mean LF/HF ratio in the placebo group showed an initial rise. The size of the intervention effect may therefore have been inflated by random upward fluctuations in the placebo group. We were therefore sought to determine whether changes observed in either group were also significant relative to their own baseline. Paired sample *t*-tests were conducted separately on the two experimental groups, in which LF/HF ratios at baseline were compared with values at each subsequent session. These tests indicated that at no session did the placebo group show any significant change from the mean recorded for that group at baseline. However, after three treatments, the LF/HF ratios in the intervention group had dropped significantly below the baseline mean for that group [$t(17) = 3.235, p < 0.01$] (Figure 3).

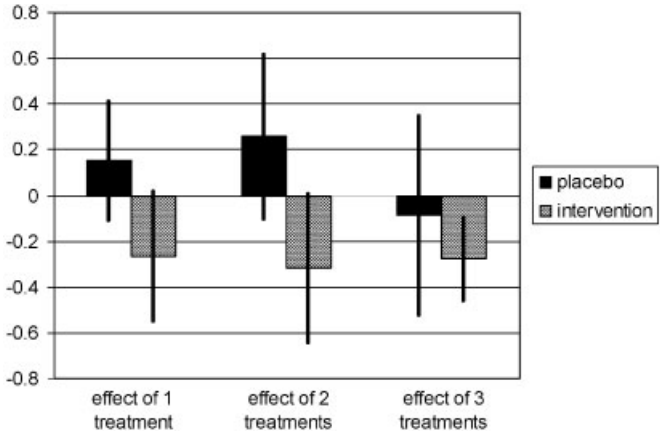

Figure 3. Changes in LF/HF ratio from baseline across sessions for the two groups. The *Y*-axis represents the magnitude of the change from baseline in the log transformed LF/HF ratios. The three successive sessions are plotted along the *X*-axis. The error bars represent 95% confidence intervals; where the error bars do not cross zero, the change is significantly greater than zero. The LF/HF ratio in participants who received the intervention showed a trend to decrease from baseline after one treatment, culminating in a significant net downward change ($p < 0.01$) after three treatments. For the placebo group, there was no significant change from baseline at any of these three sessions. Note that these measures were made 2–5 days after treatment.

Lateralized temporal order judgement task

A repeated measures ANOVA was performed on the TOJ task scores collected at pre-test, with treatment group as a between-subjects factor ('group'). There were two TOJ score factors, namely 'SOA' (eight levels) and 'hemifield order' (two levels: whether the stimuli pairs were presented in left-first or right-first order). The only significant effect was a main effect of SOA, indicating, as would be expected, that accuracy was better at longer SOAs than at short [$F(7,25) = 64.555, p < 0.001$] across the whole sample. There were no significant hemifield effects, and, most importantly, no significant differences between groups (either main effects or interactions between group and within-subject factors) at pre-test.

A repeated-measures ANCOVA was then conducted on the TOJ task scores obtained at both pre-test and post-test. This included a third within-subjects TOJ score factor, namely that of 'testing session' (two levels: pre-test and post-test), in addition to the two factors ('SOA' and 'hemifield order') used in the ANOVA reported above. As we were only interested longitudinal effects, only 'testing session' and interactions with 'testing session' ('testing session' by 'SOA'; 'testing session' by 'hemifield order'; testing session' by 'SOA' by 'hemifield order') were specified as within-subjects variables in the model.

As regards between-subjects effects, there was substantial variance in performance on the task at pre-test, with some participants performing at or near ceiling, particularly at long SOAs. Participants who performed well on the task at pre-test would therefore have had less room for further improvement at post-test than participants who performed poorly. Any learning and/or treatment effect sizes were therefore likely to be inversely proportional to baseline score, and it was therefore important to include a measure of pre-test performance scores as a between-subjects factor in the repeated-measures analysis. Mean scores at pre-test for all SOAs greater than 0 were therefore calculated for each subject, and entered as a covariate ('pre-test score') into the model. Treatment group ('group') was entered as a between-subjects factor. 'Group', 'pre-test score' and the 'group' by 'pre-test score' interaction were specified as between-subjects variables in the model.

Multivariate tests indicated that, when considered across the whole sample, there was a significant learning effect, with mean performance across the sample being significantly [$F(1,29) = 6.525, p < 0.05$] more accurate at post test than at pre-test. As expected, this effect showed a significant interaction with mean pre-test score, with those performing more poorly at pre-test showing greater improvement at post-test than those who initially performed well [$F(1,29) = 5.978, p < 0.05$]. Similarly, a two-way 'testing session' by 'SOA' interaction [$F(7,23) = 3.260, p < 0.05$] indicated that the greatest improvement was to be found at the shortest SOAs (at which the task was most difficult), and a significant three-way 'testing session' by 'pre-test score' by 'SOA' interaction [$F(7,23) = 2.757$, $p < 0.05$ indicated that at longer, easier SOAs the improvement was most marked in those with the lowest baseline scores.

However, there was also an effect of treatment on the degree of change in performance observed at post-test. A significant 'group' by 'testing session' interaction [$F(1,29) = 4.667, p < 0.05$] indicated that the treatment group showed significantly less overall improvement in accuracy than the placebo group at post-test. Moreover, a significant four-way interaction between 'group', 'testing session', 'SOA' and 'hemifield order' [$F(7,23) = 3.400, p < 0.05$] (Figure 4)

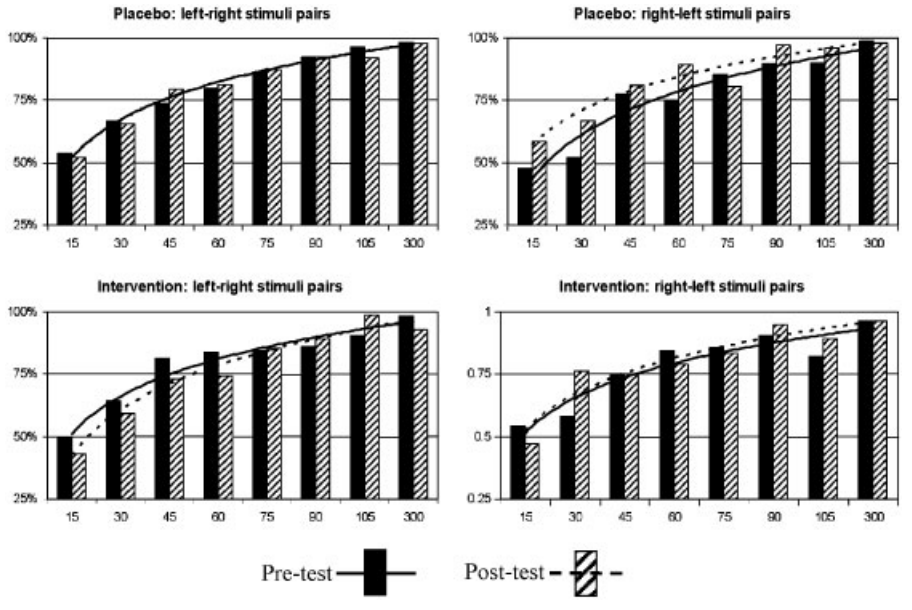

Figure 4. Graph showing the four-way interaction between 'group', 'testing session', 'SOA' and 'hemifield order'. Values are plotted for mean values of the covariate, 'pre-test score'. The categorical axis (*X*-axis) represents SOAs in milliseconds, and the proportion of correct responses at each SOA are plotted on the *Y*-axis. Placebo group data is plotted in the two upper graphs; intervention group data in the lower two. Graphs on the left represent stimuli pairs in which the left preceded the right; graphs on the right represent stimuli pairs in with the right preceded the left. The trendlines are fitted logarithmic functions. The solid trendline indicates pre-test performance, and the dotted trendline indicates performance at post test. In the placebo group, for left–right stimuli, these trendlines overlap , indicating no overall change in performance, whereas for right–left stimuli, both groups shows an overall increase in accuracy across all SOAs. In the intervention group, performance on left-right stimuli shows a fall in accuracy at shorter SOAs. The patterns of change in the placebo group suggest increased sensitivity for right hemifield stimuli at post-test, resulting in an overall increase in accuracy. The intervention group show little overall change in accuracy, but an increased tendency to respond 'right', for both right–left and left–right stimuli pairs.

indicated that whereas both groups tended to become more accurate for right-left stimuli pairs, the intervention group actually tended to become less accurate for left-right pairs, especially at short SOAs. Lastly, there was a significant five-way interaction [$F(7, 23) = 3.019, p < .05$) between 'group', 'session', 'SOA', 'hemifield order' and 'pre-test score'. This indicated that the direction and magnitude of the effects indicated by the four-way interaction was modulated by pre-test score, in such a way that improvements in performance were greater at SOAs where subjects were further from ceiling and decrements in performance were greater at SOAs where subjects were further from floor.

Discussion

The results of this study indicate firstly, that the system tested produced significant improvements in fluency on literacy tasks, as measured by timed tests

of naming, single word reading and copying, with improvements in naming speed and single word reading being individually significant. The results do not indicate any significant overall improvement in accuracy in reading either exception or non-words on untimed tests.

Secondly, the results indicate that the system induced a significant decrease in the LF/HF ratio relative to baseline, indicating a shift in autonomic balance from the sympathetic to the parasympathetic ANS. The differences between the groups were significant after only one treatment session, measured after a period of 2–5 days, at which point the LF/HF ratio in the intervention group had begun to fall significantly relative to that in the placebo group. After two further sessions of the intervention, the reduction in LF/HF ratio in the intervention group had dropped significantly below baseline, whilst LF/HF ratios in the placebo group remained insignificantly different from baseline. These results support the conclusion that the intervention leads to a significant reduction in the LF/HF ratio, after as little as one treatment session, and that these effects are observable at least 2–5 days after a treatment session.

Thirdly, the results indicate that the system induced changes in the performance of a lateralized visual-spatial task, particularly at levels of difficulty where there was scope for change. Whereas the placebo group showed greater overall improvement in accuracy than the intervention group, the latter failed to show any learning effect, and merely manifested an increased tendency to respond 'right-first' rather than 'left-first' to both types of stimuli pair, with no overall increase in accuracy.

It is of note that the positive effects attributable to the intervention were confined to measures of fluency, rather than of accuracy. Although errors are penalized in both the Rapid Naming and the One Minute Reading tasks, both these measures are weighted in favour of speed rather than accuracy, and it was on these measures that a significant effect of the intervention was found. In contrast, when participants undertook reading tasks in which they were explicitly asked to aim for accuracy rather than speed, they showed no effect of treatment. The TOJ task was also untimed, and assessed only for accuracy; on this task, net improvements on this task were confined to the placebo group.

This raises the question as to what the mode of action is likely to have been, and, in particular, whether the improvements in One Minute Reading were likely to have been due to improved lexical or sub-lexical reading processes. A number of points can perhaps be made. Firstly, the way the untimed reading task was administered is likely to have maximized the chance that participants were performing at or near the ceiling of their sight-word knowledge at both pre- and post-test. As regards the exception words, it is therefore likely that any improvement in word recognition would have been too small to have been detected in a study of this size. The failure to find a positive effect of the intervention on this untimed task of irregular word reading does not therefore rule out the possibility that the intervention works by improving word-level reading processes. However, as regards the non-words, the lack of improvement in response to treatment is possibly informative. If the mode of action of the intervention was to improve sub-lexical decoding skills, one would have expected an increase in accuracy in non-word reading, even on an untimed test. Our findings therefore are more consistent with an explanation couched in terms of increased speed of lexical retrieval, rather than improved implementation of

sublexical decoding strategies; this would also be consistent with the finding of a beneficial effect of the intervention on the rapid naming task, a task also likely to tap speed of lexical retrieval.

Regarding the findings from the HRV measures and the TOJ task, Robertson and colleagues (Robertson, Mattingley, Rorden, & Driver, 1998) found that in patients with right parietal lesions who showed both 'neglect' of left field stimuli and lowered states of arousal, attention to left hemifield could be temporarily restored by presenting the patients with alerting stimuli. As well as providing evidence for a right hemisphere locus for control of Posner's postulated adrenergic modulation of arousal (Posner & Petersen, 1990), this finding suggests that vigilance in general, and for left hemifield stimuli in particular, may be enhanced in under-aroused patients by a shift in autonomic balance towards a sympathetically dominated 'fight or flight' response. Our converse finding of reductions, or reduced gains, in accuracy in detecting peripheral stimuli, particularly in left hemifield, in a group whose autonomic balance had shifted significantly in the direction of the parasympathetic ANS is therefore consistent with Robertson's evidence. If, so the question remains as to why improvement in reading and naming fluency, should also have been associated with such a shift.

One possibility is that that the mode of action of the intervention may have been to effect a general shift in the balance of activation from right-hemisphere circuits involved in spatial attention and arousal to left hemisphere circuits involved in parasympathetic ANS control, resulting in a shift of attentional resources to the hemisphere specialized for language, and in which imaging studies have found disrupted connectivity in dyslexia (Paulesu *et al*., 1996; Pugh *et al*., 2001, 2000). In particular, Paulesu and colleagues found underactivation in the left insula in dyslexic participants as compared with controls. There is evidence that left insular cortex may play a specific role in parasympathetic cardiac effects (Oppenheimer *et al*., 1996). It is therefore possible that the system tested in this study has a direct effect on a cortical area implicated in both reading disability and in parasympathetic cardiac regulation. On the other hand, it is of note that both the items on which the treatment group improved significantly more than the placebo group involved rapid lexical retrieval, whether from picture stimuli (Rapid Naming) or from single word stimuli (One Minute Reading). Fast performance on a lexical retrieval task requires rapid response selection from a range of competing but less relevant alternatives. If heart-rate deceleration is associated with the ability to inhibit responses to irrelevant stimuli (van der Molen, 2000), then it may be that a system that increases the dominance of the ANS branch implicated in heart-rate deceleration (the parasympathetic ANS) may enhance participants' performance on tasks requiring rapid response selection, including lexical retrieval tasks, while reducing performance in a task requiring vigilance for spatially unpredictable peripheral stimuli, namely the TOJ task.

From a clinical perspective, although this study provides no evidence that the positive effects on reading and naming speed apparently induced by the system are more than transient, it is reassuring that although the post-test measures were made immediately following a treatment session, and may therefore have reflected a transiently induced state, the final HRV measurement was made 2–5 days following the third treatment session, and showed significant changes from baseline. If the cognitive effects observed are a consequence of the autonomic

shift, it may not be too optimistic to hope that they too may endure for at least a few days following treatment. If so, the fluency gains may be of clinical value. As reading is a complex skill in which success breeds success, and lack of fluency deprives the dyslexic reader of the reading experience that in the non-dyslexic developing reader rapidly leads to development of further fluency, it is possible that a system that even temporarily optimizes the attentional state required for speedy word recognition may substantially increase the gradient of any learning curve resulting from either appropriate tuition or from reading itself.

CONCLUSION

As the study was single blind, all conclusions must be made with caution. Furthermore, we do not have evidence that the system either does, or does not, correct the underlying deficits in dyslexia. However, the study provides suggestive evidence that the system tested is successful not only at inducing increased parasympathetic modulation of heart-rate variability, relative to sympathetic modulation, but also greater fluency in naming and reading aloud.

ACKNOWLEDGEMENTS

The authors wish to thank Anna Plodowski for assistance with data collection, and Professor Peter F. Liddle for helpful discussions regarding interpretation of the data.

We also acknowledge the assistance of Brightstar Learning Limited (UK) for funding and equipment for the study.

References

Castles, A., & Coltheart, M. (1993). Varieties of developmental dyslexia. *Cognition*, *47*(2), 149–180.

Cerutti, S., Bianchi, A. M., & Mainardi, L. T. (2001). Advanced spectral methods for detecting dynamic behaviour. *Autonomic Neuroscience-Basic and Clinical, 90*(1–2), 3–12.

Chen, J. C., Schmid, K. L., & Brown, B. (2003). The autonomic control of accommodation and implications for human myopia development: A review. *Ophthalmic and Physiological Optics, 23*(5), 401–422.

Dodd, J., & Role, L. W. (1991). The autonomic nervous system. In T. M. Jessell (Ed.), *Principles of Neural Science* (3 ed.) (pp. 761–775). New York: Elsevier Science Publishing Co., Inc.

Fawcett, A. J., & Nicolson, R. I. (1998). *The dyslexia adult screening test (DAST)*. London: The Psychological Corporation.

Hyndman, B., Kitney, R., & Sayers, B. (1971). Spontaneous rhythms in physiological control systems. *Nature, 233*, 339–341.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*, 97–113.

Oppenheimer, S. M., Kedem, G., & Martin, W. M. (1996). Left-insular cortex lesions perturb cardiac autonomic tone in humans. *Clinical Autonomic Research, 6*(3), 131–140.

Paulesu, E., Frith, U., Snowling, M., Gallagher, A., Morton, J., & Frackowiak, R. S. J., *et al*. (1996). Is developmental dyslexia a disconnection syndrome? Evidence from PET scanning. *Brain, 119*, 143–157.

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience, 13*, 25–42.

Pugh, K. R., Mencl, W. E., Jenner, A. R., Katz, L., Frost, S. J., & Lee, J. R., *et al*. (2001). Neurobiological studies of reading and reading disability. *Journal of Communication Disorders, 34*(6), 479–492.

Pugh, K. R., Mencl, W. E., Shaywitz, B. A., Shaywitz, S. E., Fulbright, R. K., & Constable, R. T., *et al*. (2000). The angular gyrus in developmental dyslexia: Task-specific differences in functional connectivity within posterior cortex. *Psychological Science, 11*(1), 51–56.

Robertson, I. H., Mattingley, J. B., Rorden, C., & Driver, J. (1998). Phasic alerting of neglect patients overcomes their spatial deficit in visual awareness. *Nature, 395*(6698), 169–172.

Stein, P. K., & Kleiger, R. E. (1999). Insights from the study of heart rate variability. *Annual Review of Medicine, 50*, 249–261.

Toichi, M., Sugiura, T., Murai, T., & Sengoku, A. (1997). A new method of assessing cardiac autonomic function and its comparison with spectral analysis and coefficient of variation of R-R interval. *Journal of the Autonomic Nervous System, 62*(1–2), 79–84.

Van der Molen, M. W. (2000). Developmental changes in inhibitory processing: evidence from psychophysiological measures. *Biological Psychology, 54*(1–3), 207–239.

Vanravenswaaijarts, C. M. A., Kollee, L. A. A., Hopman, J. C. W., Stoelinga, G. B. A., & Vangeijn, H. P. (1993). Heart-rate-variability. *Annals of Internal Medicine, 118*(6), 436–447.

Wittling, W., Block, A., Genzel, S., & Schweiger, E. (1998). Hemisphere asymmetry in parasympathetic control of the heart. *Neuropsychologia, 36*(5), 461–468.

Wittling, W., Block, A., Schweiger, E., & Genzel, S. (1998). Hemisphere asymmetry in sympathetic control of the human myocardium. *Brain and Cognition, 38*(1), 17–35.

Yoon, B. W., Morillo, C. A., Cechetto, D. F., & Hachinski, V. (1997). Cerebral hemispheric lateralization in cardiac autonomic control. *Archives of Neurology, 54*(6), 741–744.

