Evaluation of an Exercise-based Treatment for Children with Reading Difficulties

David Reynolds[1,*], Roderick I. Nicolson[2], and Helen Hambly[1]

[1] *School of Education, University of Exeter, Heavitree Road, Exeter EX1 2LU, UK*
[2] *Department of Psychology, University of Sheffield, Sheffield S10 2TP, UK*

An evaluation is reported of an exercise-based approach to remediation of dyslexia-related disorders. Pupils in three years of a Warwickshire junior school were screened for risk of literacy difficulty using the Dyslexia Screening Test (DST). The 35 children scoring 0.4 or over on the DST were divided randomly into two groups matched for age and DST score. One quarter of the participants had an existing diagnosis of dyslexia, dyspraxia or ADHD. Both groups received the same treatment at school but the intervention group used the DDAT exercise programme daily at home. Performance on the DST and specialist cerebellar/vestibular and eye movement tests were assessed initially and after six months. Cerebellar/vestibular signs were substantially alleviated following the exercise treatment whereas there were no significant changes for the control group. Even after allowing for the passage of time, there were significant improvements for the intervention group in postural stability, dexterity, phonological skill, and (one-tailed) for naming fluency and semantic fluency. Reading fluency showed a highly significant improvement for the intervention group, and nonsense passage reading was also improved significantly. Significantly greater improvements for the intervention group than the control group occurred for dexterity, reading, verbal fluency and semantic fluency. Substantial and significant improvements (compared with those in the previous year) also occurred for the exercise group on national standardized tests of reading, writing and comprehension. It is concluded that, in addition to its direct effects on balance, dexterity and eye movement control, the benefits of the DDAT exercise treatment transferred significantly to cognitive skills underlying literacy, to the reading

*Correspondence to: David Reynolds, School of Education, University of Exeter, Heavitree Road, Exeter EX1 2LU, UK.

DYSLEXIA 9: 48–71 (2003)
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/dys.235

process, and to standardized national literacy attainment tests.

INTRODUCTION

Acquisition of literacy has become ever more important in this information age. Failure to become literate at the appropriate time in school can have extremely serious consequences for an individual's development, happiness and employment prospects. Not surprisingly, improvement of literacy is a major government target both for children and for adults. Literacy is a complex skill, requiring fluent interplay of a number of subskills, and it appears that there is no 'short cut' to literacy, with normal acquisition requiring hundreds of hours practice. Despite increasing resources being committed to literacy in many Western countries, 10-20% of children will be delayed in the acquisition of literacy to the extent that specialist support needs to be given to attempt to help them catch up with their peers. There are many potential causes of low literacy, but arguably the most prevalent is developmental dyslexia. There is still considerable debate over diagnostic methods, but a standard criterion is that provided by the World Federation of Neurology (1968)—'a disorder in children who, despite conventional classroom experience, fail to attain the language skills of reading, writing and spelling commensurate with their intellectual abilities'.

There has been extensive research on ways of reducing reading difficulty, with Marie Clay's Reading Recovery scheme in New Zealand representing one of the most systematic analyses and approaches (Clay, 1993), with the work of Slavin (1996) being perhaps the best known educational intervention in this area. Recently the US National Institute for Child Health and Human Development has undertaken a comprehensive evaluation of the relative effectiveness of different teaching methods (NICHD, 2000). In summarising the results of this and his own NICHD research, Torgesen (2001) comes to the sobering conclusions that:

(i) 'Cure' is extremely time consuming and far from perfect, with on average 5 h intensive one-to-one instruction per point increase in standard score for reading.[†]
(ii) Even with the intensive interventions above, the increase was purely in word reading accuracy, with no concomitant increase in fluency.
(iii) 'Prevention' is by far the most cost-effective solution. Early identification and intervention, though still costly in terms of time, can lead to increases both in accuracy and fluency.

There is a strong genetic component to dyslexia (e.g. Gayan & Olson, 1999; Pennington, 1999) and therefore, in principle, it should be possible to diagnose dyslexia before a child starts learning to read. Appropriate support should substantially reduce subsequent reading difficulties. The search for precursor

[†] The Standard Score is determined such that exactly average is 100, with 85 representing one standard deviation below the mean. In round terms, one would expect 25% of the population to have an SS below 90, 10% below 80, and 2% below 70. One might typically wish to raise standard score by at least 10 points from say 85 to 95 by means of intervention.

symptoms has provided one major motivation for much research into the underlying cause. However, a parallel motivation is that, if the underlying cause could be identified, it might be possible in some way to 'treat' the cause, thereby reducing all subsequent difficulties (whether reading-related or not).

One of the major findings in dyslexia research is that children who later are diagnosed as dyslexic have less well-developed phonological awareness at five years, and these phonological difficulties tend to remain entrenched unless explicit teaching is undertaken. Given the importance of phonological awareness in learning to read one of the major hopes of dyslexia researchers was that phonological support would prove sufficient to mitigate the reading difficulties. Unfortunately, phonological intervention alone appears to be insufficient to fully overcome the reading difficulties, and fluency support also is needed—see Torgesen above, also Hatcher, Hulme, and Ellis (1994). These findings are consistent with the 'double deficit' hypothesis for dyslexia (Wolf & Bowers, 1999).

In view of these difficulties, pragmatic UK dyslexia researchers have focused on providing methods of identifying dyslexia in the initial school period, thereby providing the 'stitch in time' benefits deriving from early diagnosis and support. Fully developed early screening instruments include the Dyslexia Early Screening Test and Singleton's Cognitive Operations Profiling System (e.g., Nicolson & Fawcett, 1996; Fawcett, Singleton, & Peer, 1998). Since the DST (the equivalent of the DEST for slightly older children) is one of the major test batteries used in the present study, it may be valuable to give some information about how these batteries may be used for screening and then intervention.

In an evaluation of effectiveness of the DEST (Nicolson, Fawcett, Moss, & Nicolson, 1999), classes in four UK infant schools were screened to identify children most at risk of reading failure (62 in total, mean initial age 6.0 years). The selected children were given an individually adaptive, curriculum-based, support programme with the emphasis on word building and phonics skills in the broad reading context. The programme was administered to children in groups of four for two half-hour sessions per week for 10 weeks. The intervention group improved significantly in mean reading standard score, whereas a matched control group (no intervention) made no overall improvement. The intervention proved cost-effective, with mean 'effect size' comparable to those reported for Reading Recovery, yet with only 10% of the costs. Nonetheless, as one might predict from Torgesen's summary (above), there was a significant minority of children who made negligible progress during the study. Despite the clear progress of the intervention group overall, 25% remained 'problem readers' (with reading still at least 6 months behind). 88% of these problem readers had initial 'at risk' or 'borderline risk' scores on the DEST screening test, compared with only 28% of the 'recovered readers'.

In short, Torgesen's summary represents a reasonable composite picture of the strengths and weaknesses of intervention for literacy failure. It is a rather forbidding picture, confirming that there is a formidable problem in helping all children to become both competent and fluent at reading.

Complementary approaches to Literacy Support

In view of the limited effectiveness of conventional support methods, there has always been an opportunity for alternative/complementary approaches to literacy support. Many of these have been frankly lacking in theoretical basis, pedagogical plausibility or empirical support. Such approaches have cast serious doubt not only on themselves but also on perfectly 'respectable' complementary approaches. A review of complementary approaches is beyond the scope of this article. See Fawcett (2001) for an overview. Here we note briefly approaches based on the established theoretical approaches to dyslexia.

Both the double deficit hypothesis and the phonological deficit hypothesis are couched in terms of the cognitive symptoms of dyslexia, and therefore do not clearly indicate the underlying cause in terms of brain structures. Tallal and her colleagues have developed an extensive remediation system (Fast ForWord) based on her discovery that many children with specific language impairment show significantly impaired ability to sequence auditory tones presented close together, a difficulty that is attributed to the auditory magnocellular system (Tallal, Merzenich, Miller, & Jenkins, 1998). Fast ForWord attempts to overcome this difficulty by presenting speech with artificially slowed consonant transitions, with the idea that the brain can be retrained to become sensitive to the transition. The transitions are then systematically reduced until the speech is normal. Unfortunately, despite promising initial results (Merzenich *et al*., 1996) Fast ForWord has become mired in controversy, with mainstream dyslexia researchers arguing that only a small minority of dyslexic children show magnocellular problems, and that the approach is not effective for the remainder (Snowling, 2000; Hook, Macaruso, & Jones, 2001). Other researchers have suggested that there is an abnormality in the visual magnocellular system, showing up primarily as a difficulty in detecting low contrast slow movement (Eden *et al*., 1996), but we are not aware of any systematic approach to remediation based on visual magnocellular treatment. Stein, who is also a committed proponent of the visual magnocellular hypothesis (Stein & Walsh, 1997) has long argued that dyslexic children have difficulties with binocular vision, and that these can be treated via the occlusion technique (Stein, Richardson, & Fowler, 2000).

We return to these approaches in the final discussion, but turn now to the theoretical background to the exercise regime that forms the basis of the complementary treatment approach investigated in this study.

The Cerebellar Deficit Theory (CDT)

Nicolson and Fawcett looked for the underlying cause in terms of some dysfunction of the learning. Initially (Nicolson & Fawcett, 1990) they argued that dyslexic children had a pervasive difficulty in making skills automatic—so that they occur fluently and with no need for conscious effort. This hypothesis is best seen as a cognitive level hypothesis—similar in level of analysis to the phonological deficit and speed deficit hypotheses. Following an extensive research programme, they claimed that the automatisation deficits were attributable to abnormal function of the cerebellum—the 'hind brain', an evolutionarily ancient system known to be involved in motor skill execution (Eccles, Ito, & Szentagothai, 1967), and more recently identified as having a central role in skill automatisation and language-based skill (Allen, Buxton, Wong, & Courchesne, 1997; Leiner, Leiner, & Dow, 1989; Thach, 1996). There remains controversy over the role of the cerebellum in cognitive skills not involving speech or 'inner speech' (Ackermann, Wildgruber, Daum, & Grodd, 1998; Glickstein, 1993), but there is now overwhelming evidence of the importance of the cerebellum in language (Ackermann & Hertrich, 2000; Fabbro, Moretti, & Bava, 2000; Silveri & Misciagna, 2000), including a recent demonstration of specific cerebellar involvement in reading (Fulbright *et al*., 1999).

Nicolson and Fawcett established a range of evidence, from behavioural data to brain imaging findings to anatomical data, directly consistent with the hypothesis (Nicolson & Fawcett, 1999; Nicolson, Fawcett, & Dean, 2001a). Furthermore, by developing a model of how cerebellar problems at birth would lead to problems in skill (including articulation) fluency and automaticity, they claimed to subsume both components of the double deficit hypothesis within a single, more specific, framework. The theory was initially somewhat controversial, but is now established as one of the major explanatory theories in the area (Frith, 1997). Levinson (1988) had claimed that there might be a vestibular abnormality, and other theorists (e.g. Rudel, 1985) had highlighted cerebellar-type motor problems, but these approaches were largely discounted (Silver, 1987) once the phonological deficit hypothesis became established. Perhaps most distinctive in the CDT framework is that the theory relies on 'mainstream' cognitive neuroscience discoveries relating to the role of the cerebellum both in skill acquisition and in skill execution. The role in learning led to studies (Nicolson & Fawcett, 2000) that suggested that dyslexic children might have particular difficulty in acquiring complex skills—the longer a skill takes to acquire, the more disadvantaged dyslexic children will be.

It should be noted that, in common with other causal hypotheses for dyslexia, the CDT remains somewhat controversial. Proponents of the magnocellular deficit hypothesis suggest that the cerebellum may in fact be an 'innocent bystander' while the true cause of the difficulties is input to the cerebellum from sensory systems (Stein & Walsh, 1997; Zeffiro & Eden, 2001). Other theorists suggest that because the cerebellum is such a large structure, it is necessary to present a more detailed model of the specific regions affected. Still others suggest that the motor and balance difficulties established in many dyslexic children reflect a subtype of dyslexia also with attention deficit (Denckla, Rudel, Chapman, & Krieger, 1985; Wimmer, Mayringer, & Raberger, 1999).

Analysis of these theoretical issues is beyond the scope of this article. A considered discussion is provided elsewhere (Nicolson *et al*., 2001a; Nicolson, Fawcett, & Dean, 2001b). Suffice it to say here that the applied issue under evaluation is whether or not an exercise-based treatment system based on claimed 'retraining the cerebellum' will prove effective in helping children with reading difficulties.

The Cerebellar Treatment Hypothesis (CTH)

While Nicolson and Fawcett limited themselves to implicating the cerebellum as cause and symptom, Dore and Rutherford (2001) took the hypothesis to its logical conclusion. They proposed that, given that the cerebellum remains plastic throughout childhood, it should be possible to retrain the dyslexic cerebellum so that it becomes more normal, scaffolding learning in a much more efficient fashion. This hypothesis required something of a leap of faith, in that it is generally believed that the cerebellum comprises a very large number of independent 'cerebro-cortical microzones' (Ito, 1984), and so it is not clear why training on one sort of task should generalize to unrelated tasks. Nonetheless, given the limitations of current understanding about how these processes and interactions take place, this is by no means an indefensible hope.

Smith, with Dore and Rutherford went on to develop a systematic Balance Remediation Exercise Training programme based on longstanding principles of balance training (e.g., Belgau & Belgau, 1982), but distinctive in that the approach was based not only on the coherent CDT theoretical framework but also in that they used sophisticated methods of testing vestibular and cerebellar function using an electronystagmography (ENG) system for assessing all parameters of eye movement control and a posturography balance system. On the basis of pilot work using this remediation system, they set up the DDAT Clinic in Kenilworth and offered their treatment to paying clients. They also built in systematic procedures for data collection and external evaluation of the clinic in order to test their claims and develop the treatment regime further.

Interestingly enough, for many years there have been suggestions from practitioners that motor skill intervention can help learning disabled children (Kephart, 1971; Farnham-Diggory, 1992). Not surprisingly, research studies into motor deficits show that motor training produces direct improvement in motor tasks (e.g., Knight & Rizzuto, 1993; Cammisa, 1994). However, rigorous scientific research in motor skill training has been limited, with reports mainly anecdotal or based on case studies, and little evidence of skill transfer to academic and social domains. Moreover, properly controlled studies show that improvement is often found in the control group as well (Bluechardt & Shepard, 1995), suggesting that improvements are largely attributable to the Hawthorne effect (although see McPhillips, Hepper, & Mulhern, 2000 on primitive reflexes). Consequently, motor skill training has not generally been well received within the field of educational research. However, to our knowledge there have been no formal controlled evaluations of these techniques.

Of course, if the CTH does indeed turn out to be even partly valid, then this could revolutionise the treatment of literacy in dyslexia. Rather than attempting laboriously to scaffold the inefficient learning processes of the dyslexic child, one would first treat the cause (the cerebellum) and then learning would take place relatively normally—in all areas of skill. The purpose of the study reported here was to provide a rigorous and wide-ranging evaluation of the DDAT treatment intervention.

EVALUATION OF THE DDAT INTERVENTION IN A JUNIOR SCHOOL ENVIRONMENT

Design

The study reported here was designed to investigate the effects of the DDAT exercise regime.[‡] It was based in a Warwickshire junior school and used an

[‡] The study was carried out by a University of Exeter team led by Reynolds. Nicolson was not involved in the running of the study but advised on the initial design and on data analysis issues.

intervention supported by the school that involved the DDAT exercise regime. Out of 269 children in three years, those with the highest (most at risk) scores on the Dyslexia Screening Test (see below for an overview of the test), were identified, and divided randomly into two groups matched for age and DST score. The groups were not formally identified at the school, and were not in any way distinguished during the school day, but the intervention group undertook the DDAT exercise treatment daily at home, with the others forming a control group. The DST test and specialist tests of cerebellar/vestibular function were administered initially and after six months of treatment.

It is important to note the strengths and limitations of this design. A fully controlled design would ensure that the control group undertook an additional activity, equivalent in duration to the DDAT exercises, daily at home. However, it is difficult to envisage just what such an activity would entail given the need for 180 separate parent-administered sessions over the 6 months. It would be particularly problematic on ethical grounds to ask control group parents to supervise activities that were not likely to be of benefit. In our view it was much more likely that such an activity would lead to harmful effects of boredom and alienation, thereby artificially inflating any positive effects found for the DDAT group. Overall, therefore, we considered but rejected this possibility. The strength of this study's design is that it assesses the predicted'value-added' by the DDAT treatment. Both groups have exactly the same support regime within the school. The only difference is that, in addition, the DDAT group have the DDAT exercise support. This design corresponds directly to the likely use of the DDAT treatment (where it is undertaken in parallel with normal school support). Any difference between the DDAT group performance and the control group performance therefore corresponds to the 'value added' by the treatment. From a parental perspective this is the critical issue.

Participants

A request was circulated to children's parents in three years of children at the school requesting participation in an evaluation of a novel method of literacy support, for 35 children identified as potentially at risk of reading difficulty, using a criterion of at risk quotient of at least 0.4 on the DST (see below for an overview of the DST). Two children with DST score less than 0.4 were also given the DDAT treatment, but these are not included in the analyses. The pre-test characteristics of the intervention group and control group were matched as follows: gender: 10 m, 8f vs 9 m, 8f: mean age: 9;4 vs 9;4: age range 7;11 to 10;06 vs 8;00 to 10;05: mean DST score 0.74 vs 0.72; DST range 0.4 to 1.5 vs 0.4 to 1.6. The school administered NFER reading tests annually as part of its normal performance management system. These were not accessed by the research team, but when analysed later indicated a slight imbalance between the groups. The intervention group had a mean reading delay of 10.6 months (range 33 months delay to 6 months ahead) whereas the control group had a reading delay of 4.4 months (range 45 months delay to 22 months ahead).[§] Nine participants had an external diagnosis of dyslexia (4 in the exercise group, 2 in controls),

[§] For equity, the control group were provided with the DDAT exercise treatment soon after the 6 month post-test.

dyspraxia (1 in each group) or ADHD (1 control). Twelve participants (7 in the intervention group, 5 in the control group) were withdrawn from two lessons per week for small group support.

Method

Unlike other interventions, the DDAT approach bases the intervention on repeated administration of an extensive set of cerebellar, vestibular, and literacy/skill tests, thereby allowing a wide range of analyses of the improvements resulting from the intervention.

DDAT exercise therapy has been developed to 'stimulate simultaneously the central nervous system mechanisms found to be immature in learning disabled children on electroneurophysiological assessment'. The DDAT researchers refer to the condition as Cerebellar Developmental Delay arguing that there appears on testing to be a gradual improvement in the symptoms of the disorder as age increases, and the condition seems to improve rapidly with appropriate remediation. By combining groups of exercises in a way that avoids habituation, over stimulation and automatisation, DDAT exercise therapy aims to bring about significant neurological improvements.

Clients are initially assessed using Computerised Dynamic Posturography and Electronystagmography, eleven tests of academic and skill ability, and interview and examination by a qualified medical examiner. DDAT exercise therapy is applied for ten minutes twice daily every day, and clients are reassessed on a regular basis in order to continually monitor their response, and adjust treatment to allow the maximal response.

The DDAT Exercise Treatment

The DDAT exercise treatment is an extensive and adaptive course mapped out over many months. It is a complex programme of integrated sensory stimulation incorporating visuomotor and vestibular therapy which has been uniquely structured in the combinations and weightings of the sensory inputs. It was devised initially through a trial and error regime and then fine-tuned using the neurophysiological tests to assess progress until significant progress was observed and full physiological resolution occurred within an acceptable timeframe. Key elements include use of a balance board; throwing and catching of bean bags (including throwing from hand to hand with careful tracking by eye); practice of dual tasking; and a range of stretching and coordination exercises. The complete sequence is commercially sensitive and has been the subject of considerable, and ongoing, evolution since its inception.

Issues Investigated

The CDT framework includes a number of logically independent claims. Claims 1 and 2 are supported by existing literature, but nonetheless need to be substantiated. Claims 3 and 4 go beyond the existing literature, and would represent significant support for the CTH.

Claim 1: Incidence of Cerebellar/Vestibular signs
A high percentage of children with dyslexia, dyspraxia or attention deficit will show cerebellar or vestibular problems. There will therefore be a high incidence of cerebellar/vestibular abnormalities in the initial testing.

Claim 2. Remediation of Cerebellar/Vestibular signs
The DDAT exercise programme will, if followed for an appropriate period, lead to effective remediation of the cerebellar/vestibular skills revealed by the posturography and ENG tests.

Claim 3. Near-transfer to Fundamental Cognitive Skills
In addition to the specific cerebellar/vestibular learning indicated in claim 2, the training will transfer to cognitive skills such as phonology, working memory and speed of processing that form the fundamental basis of cognitive functioning, and underpin the subsequent development of literacy skills.

Claim 4. Far-transfer to Literacy skills
In addition to the above near-transfer effects, the 'accelerated' learning ability ensuing from treatment of the cerebellum will lead to normal or above normal acquisition of literacy. There would presumably be a delay in this effect, with progress acceleration occurring primarily after the treatment of the cerebellum is complete. Progress should nonetheless occur without the need for further exercise treatment, assuming that normal reading support is available via the child's school.

Dependent Measures

A battery of tests was administered at the school before treatment in line with standard practice at DDAT Centres. These included the DST screening test (Fawcett & Nicolson, 1996) and a range of cerebellar and vestibular tests. In addition the majority of these tests were re-administered after six months (whereas for standard treatment at a DDAT Centre the full set of tests would be re-administered only once treatment was considered complete). Both pre and post-tests were administered by a tester blind to the child's group. In addition, following standard practice at the school, standardised national attainment tests were administered by the school at the end of each school year.

School tests

School tests included the NFER test of reading together with the three national standardised attainment (SATS) tests, namely writing, comprehension and numeracy. The intervention started early in the school year (September 2001) and the school SATS testing took place in July 2001 and July 2002—hence the SATS follow-up was some three months after the other follow-up tests. For most participants the previous year's SATS and reading data (July 2000) were also available. This allowed relative progress pre- and post-intervention to be assessed.

Dynamic Posturography
These tests involved standard use of the Dynamic Posturography equipment (Neurocom International Inc.). A series of tests was applied including the

Sensory Organization Test (assesses the client's ability to make effective use of visual, vestibular, and proprioceptive information and to appropriately suppress disruptive visual and/or proprioceptive information under sensory conflict condition), the Motor Control Test (assesses the client's ability to reflexively recover from unexpected movements of the platform or walls quickly and with appropriate movement patterns) and the Adaptation Test (assesses the ability to modify reflexive motor reactions when the support surface is irregular or unstable). Each of these protocols is administered and analysed objectively and automatically under programmed control of a computer. Each of these tests was recorded as a score out of 100. A low score indicates poor vestibular function. A score of 50-56 may be expected from children between the ages of 9 and 13 years with good postural control (Shimizu, Asai, Takata, & Watanabe, 1994; Rine, Rubish, & Feeney, 1998)..

Electronystagmography

Eye movements were plotted while the client watched a moving target using specialised CHARTR ENG equipment (ICS Medical). A series of tests was applied, including smooth pursuit, saccade latency and accuracy and optokinetic flow. Each of these protocols is administered and analysed objectively and automatically under programmed control of a computer. The client's visual tracking ability, saccade accuracy and saccade latency were all recorded as a score from 1-100. Norms are not currently available.

The Dyslexia Screening Test

The DST comprises 11 sub-tests in five areas (literacy skills, phonological awareness and verbal memory, motor skill and balance, and memory retrieval fluency). The sub-tests are as follows.

(i) Literacy skills

One minute reading. The number of single words (in ascending order of difficulty) that can be read in 1 min. A composite test of single word reading accuracy and fluency. *Nonsense passage reading*—Jabberwock type passage mixing real words and pseudowords. Pseudowords require knowledge of grapheme/phoneme correspondence to be read correctly. Known to be a sensitive index of dyslexia, even when a child can read real words reasonably fluently. Scoring takes into account both speed and accuracy. *Two minute spelling test*—How many words the child can spell correctly in 2 min, with the tester dictating the next word as soon as the child finishes the previous one. A combined test of spelling accuracy and fluency. *One minute writing*—This is designed to assess speed of writing. Dyslexic children are slower (and less neat) writers. Writing speed is currently one of the key issues for examination allowances. Score is primarily the number of words transcribed in 1 min, with adjustments made for errors.

(ii) Phonological awareness and verbal working memory

Phonemic segmentation—This tests the child's ability to play with the constituent sounds in words. Deleting specific phonemes. A standard phonological test. *Backwards digit span*—A string of single digits is presented on tape, and the child has to repeat the string of digits in the reverse order. The tape starts with 2 digits and increases up to 8 (terminating when consecutive failures occur). A standard test of verbal working memory.

(iii) Cerebellar/vestibular tests

Bead threading—How many beads can be threaded in 30 s, a standard test of manual dexterity. *Postural stability*—How much the child wobbles when pushed gently in the back using a pre-calibrated stability tester, a test of cerebellar/vestibular function (balance).

(iv) Memory retrieval fluency

Rapid automatized naming—involves the time taken to speak the names of pictures on a page full of common objects, a test of general linguistic fluency. *Verbal fluency* simply how many words beginning with S the subject can think of in a min. *Semantic fluency* is how many animals the child can think of in a min. The tests assess speed of retrieval from long term memory. Dyslexic children are thought to be impaired on verbal but not semantic memory retrieval.

Norms based on a UK national sample of over 1000 children are published with the test (with separate norms for each year), and allow each 'raw' score on each subtest to be allocated an 'at risk index' as follows: --- (very strong risk) for centiles 1-4, -- (strong risk) for centiles 5 to 11, - (at risk) for centiles 12-22, o (average band) for centiles 23 to 78, + (above average) for centiles 79 to 100. The overall DST 'at risk quotient' is essentially the average of the scores on the individual sub-tests (counting --- as 3, -- as 2, - as 1 and the remainder 0). An ARQ of 0.9 or greater is considered a 'strong risk' indicator, and an ARQ of 0.6 to 0.89 is considered a mild risk indicator. ARQs from 0 to 0.5 are classified as 'not at risk'. For the purposes of this study, unpublished norm data, allowing deciles to be derived for each individual score, were also used, thereby allowing a more uniform and sensitive index of performance than the at risk index to be derived.

Results

(i) Incidence of cerebellar/vestibular signs

The DST test was administered at the start of treatment and again after 6 months, as were the ENG and posturography tests. Other specialist tests were administered only at the initial assessment, in that re-administration normally takes place only when the treatment is complete (as indicated by posturography and ENG). Consequently we start with the initial testing. This allows the incidence of cerebellar/vestibular problems to be assessed (Claim 1).

There was a considerable incidence of signs in the participants in the study, with 24 out of the 37 tested (65%) scoring lower than 50 (lower limit of normal range) on the dynamic posturography balance test. The absence of norms on the ENG tests makes it difficult at present to evaluate the incidence of eye movement abnormalities.

(ii) Changes in cerebellar/vestibular signs

Figure 1 indicates the changes in cerebellar and vestibular signs following the six month treatment for both groups.[¶] As would be expected given the nature of the treatment, the benefits are particularly striking for the vestibular/cerebellar indicators on posturography. Three out of four improvements for the DDAT group

[¶] The power of the ENG analyses was reduced owing to an unfortunate delay on data collection of the 6 month ENG for some members of both groups, leading to missing data on the post-test for the ENG data.

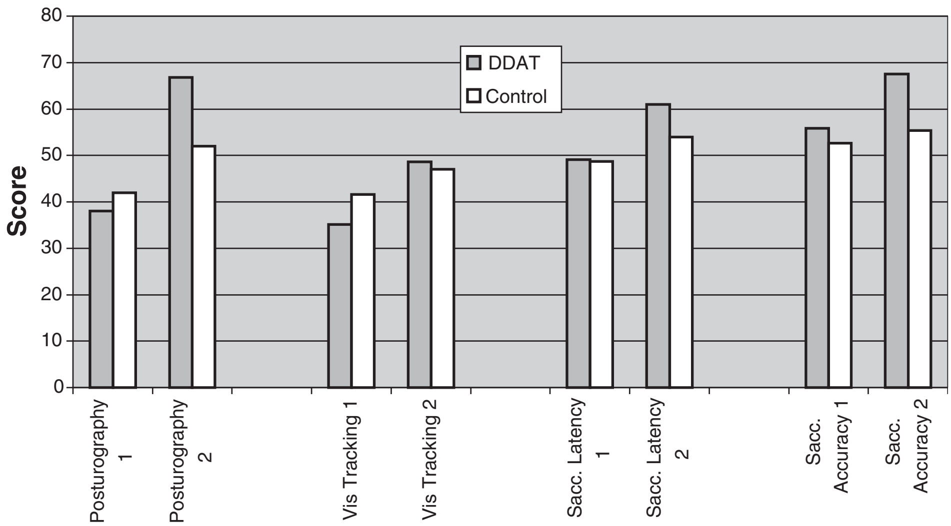

Figure 1. Changes in Cerebellar/Vestibular performance over 6 months. The data represent the initial performance (labelled 1) and the performance after 6 months (labelled 2). The posturography score and the visual tracking score are on a scale from 1 to 100, with 50 being a target 'average' performance. The saccadic latency and accuracy scores are on a percentage scale, but normative data are not currently available.

were significant [$t(17)=5.87$, $p<0.0001$; $t(8)=1.11$, NS; $t(8)=4.83$, $p<0.01$; $t(8)=2.68$, $p<0.05$ for posturography, visual tracking, saccade accuracy and saccade latency resp.]. There were no significant improvements for the control group.

A two factor analysis of variance was undertaken to allow the relative improvements for the DDAT group and the control group to be compared directly. The missing data on post-test seriously weakens the statistical power of the tests. Nonetheless there was a significant interaction [1 tailed], indicating greater improvement for the DDAT group than controls, for all four tests: [$F(1,24)=3.02$, $p<0.10$; $F(1,8)=4.14$, $p<0.10$; $F(1,8)=10.39$, $p<0.05$, $F(1,8)=5.13$, $p<0.10$] for posturography, visual tracking, saccade accuracy and saccade latency respectively.

Changes in DST Scores

In terms of overall DST at risk quotient, there was a noticeable reduction for both groups. The DDAT group decreased from mean 0.74 to 0.39 and the control group decreased from 0.72 to 0.44. Both these changes are highly significant [$t(17)=4.97$, $p<0.0001$; $t(16)=4.09$, $p<0.001$ resp.]. Incidence of 'risk' on the DST also decreased substantially. For the DDAT group, incidence of strong risk (ARQ$\geqslant$0.9) fell from 33 to 11% and incidence of at least mild risk (ARQ$\geqslant$0.6) fell from 56 to 33%. For the control group the corresponding figures are 35 to 24% and 53 to 29%, resp.

For each participant the decile corresponding to the raw score and age was calculated. For ease of interpretation these are presented here as percentile scores. Figure 2 indicates the changes in mean percentile score for the DST sub-tests. Note these explicitly take age into account, and so the changes are less marked than for the raw scores. It is not possible to give a corresponding figure for the Posturography and ENG data since we do not have detailed normative data available.

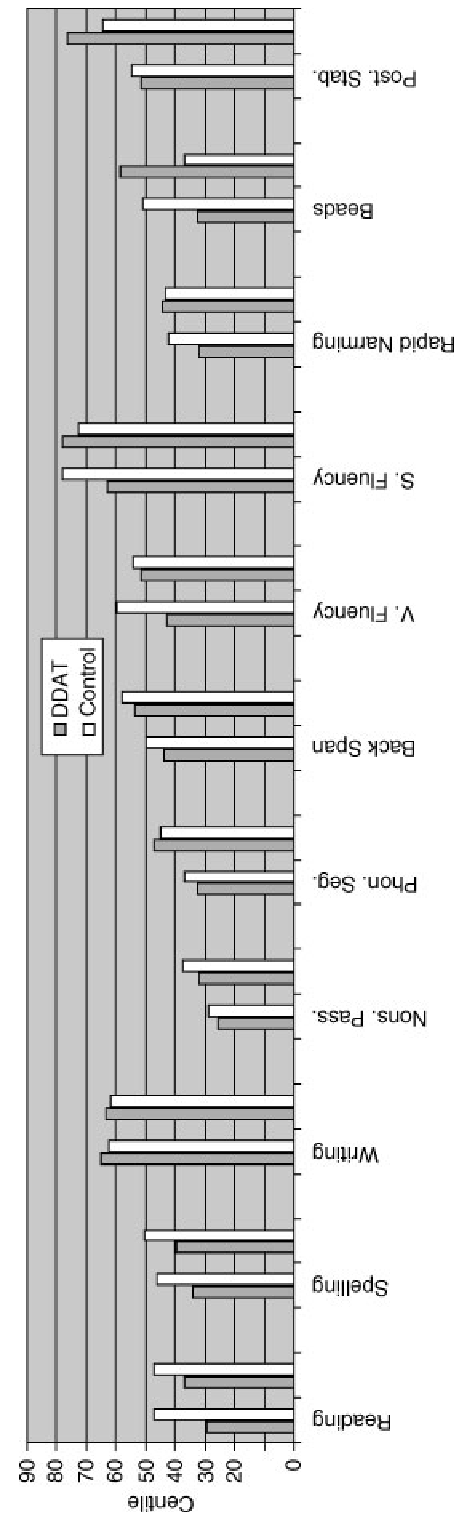

Figure 2. Changes in DST scales. Centiles given are at the start and after 6 months. For example, the first four columns give reading at start (DDAT group then control group), Reading after 6 months (DDAT group then control group).

In order to assess the significance of these changes, two sets of analyses were undertaken. First, paired score *t*-tests were undertaken for each group. We consider the use of a parametric test justifiable here since the underlying decile scale used has equal intervals. However, the data do not satisfy the normal criteria for use of a *t*-test and so a non-parametric Wilcoxon matched pairs signed ranks test was also used. Both sets of tests yielded similar results (see Table 1).

For the DDAT group there were significant improvements ($p < 0.05$ or better) for reading, semantic fluency, phonemic segmentation, bead threading and postural stability, with one-tailed significance ($p < 0.10$) for nonsense passage reading, with near-significance (p=0.12) for rapid naming. For the control group there was a significant improvement for nonsense passage reading and (one-tailed) for phonemic segmentation and backwards span.

Secondly, a series of two factor analyses of variance was undertaken, with the independent variables being group (DDAT and control) and time of test (pre-test vs post-test). These analyses allow us to determine (by means of the interaction between the two factors) whether there are any significant differences between the two groups in the change over the period of the study. Significant interactions were found for reading, bead threading and semantic fluency, though the interaction approached significance (p=0.12) for verbal fluency.

Effect Sizes for changes following treatment

In intervention studies of this type it is considered good practice to convert the improvement to an 'effect size' that gives an index of the improvement relative to the original performance mean and variation of the class (Cohen, 1969; McCartney & Rosenthal, 2000). In this study, we shall be using primarily a differential effect size (that is, the amount of progress over and above that of the control group). For comparability with the extensive data produced by the US National Reading Panel (2000), this effect size is calculated as the difference in the amount of improvement between the groups divided by the averaged standard deviation of the progress of each group. This effect size provides an index of improvement that is independent of the scoring system used and may therefore be used to compare across measures and studies. It is generally held that an effect size of 0.5 or more may be considered 'moderate' and one of 0.8 or more may be considered 'large' (Cohen, 1988) (Table 2).

Standard School Measures

Standardized tests of attainment were administered at the end of the school year. Three SATS tests (writing, comprehension and maths) and one reading test (NFER/Nelson) were administered at the end of each school year. The reading test is reported in terms of reading age (months). The SATS tests are normally graded in terms of levels (digits and letters). Possible grades for level 2 include (in descending order) 2a, 2b, 2c, 2, 2d. To facilitate comparison these have been coded as 2.75, 2.50, 2.25, 2.0 and 1.9 respectively. An equivalent coding was applied for level 3 and 4. The mean scores for the intervention group only[‖] are

[‖] The control group have been omitted from this analysis since they were given the DDAT treatment after the initial 6 months and so their final scores include around 3 months of DDAT treatment.

Table 1. Statistical tests on DST percentiles

	Reading	Spelling	Writing	Nonsense passage	Phon. seg.	Back span	Verbal fluency	Semantic fluency	Rapid naming	Beads	Post. Stab.
DDAT group (post- vs pre-test)											
Paired *t*-test: *t* (sig) df=17	3.20**	1.61	0.29	1.89†	2.55*	1.53	1.38	2.66*	1.73	4.04***	3.59**
Wilcoxon test: *z* (sig) df=17	2.65*	1.40	0.51	1.82†	2.35*	1.43	1.25	2.33*	1.43	2.99**	2.84**
Control group (post- vs pre-test)											
Paired *t*-test: *t* (sig) df=16	0.00	1.05	0.09	3.67**	1.82†	1.64	0.87	1.26	0.18	0.64	1.15
Wilcoxon test: *z* (sig) df=16	0.00	0.90	0.14	2.64**	1.78†	1.83†	0.92	1.20	0.05	1.11	1.36
2 Factor ANOVA											
Interaction: F (sig) df=(1,33)	5.08*	0.02	0.02	0.25	0.56	0.05	2.55	8.2**	1.29	14.4**	1.83

Key: $^{**}p<0.01$; $^{*}p<0.05$; $^{\dagger}p<0.10$.

Table 2. Effect Sizes for the DST centiles and the posturography and ENG scores

DST Percentile (adjusted for age)		Posturography/ENG	
Reading	0.35	Cerebellar/vestibular	1.13
Spelling	0.04	Visual tracking	1.04
Writing	−0.04	Saccadic accuracy	1.30
Nonsense passage	−0.14	Saccadic latency	0.41
Phonemic segmentation	0.24		
Backwards span	0.08		
Verbal fluency	0.46		
Semantic fluency	0.75		
Rapid naming	0.39		
Bead threading	1.26		
Postural stability	0.49		
DST at risk quotient	0.18		

Table 3. Performance of the DDAT group on standardized school tests

Test	July 2000	July 2001	July 2002	Ratio	Effect size	*t*-test
NFER reading (months)	97 (15)	103 (12)	122 (21)	3.30	1.14	2.19 (*)
SATS maths	2.78 (0.46)	3.19 (0.54)	3.74 (0.55)	1.34	0.39	1.03 (NS)
SATS comprehension	2.78 (0.60)	2.94 (0.52)	3.72 (0.52)	4.75	1.24	3.01 (**)
SATS writing	2.53 (0.27)	2.56 (0.44)	2.95 (0.55)	17.05	0.98	2.86 (*)

Key: **$p < 0.01$; *$p < 0.05$.
The DDAT exercise intervention took place September 2001 to February 2002.
Column 5 (ratio) is the ratio of progress in the second year (following DDAT treatment) to that in the first year (pre-DDAT).
Column 6 (effect size) is the effect size of the improvement in the second year to that in the first year. Standard deviations are in parentheses.

given in Table 3. Effect sizes have been calculated (note that here the effect size is calculated in terms of the relative progress post-DDAT to that pre-DDAT).

It is clear that for all four tests there was substantially more improvement in the year including the DDAT exercises than for the previous year. Taking the NFER reading, before DDAT the mean improvement was only 6 months in the year. After treatment the improvement was 19 months—a ratio of 3.30 to 1. The relative effect of the DDAT treatment was smallest for the SATS maths (ratio 1.3 to 1); substantial for SATS comprehension (ratio 4.75 to 1) and extraordinary for SATS writing (17.05 to 1)—though to an extent this reflects the poor improvement in the previous year. The final column indicates that the difference in progress was significant for reading, comprehension and writing, but not for maths. The effect sizes indicate a similar pattern of large effect (effect size 1.0 or more) for reading, writing and comprehension with a relatively modest (but still noticeable) effect size for maths.

DISCUSSION

Overall summary

After 6 months treatment the DDAT group showed the expected physiological changes, with substantial improvements in vestibular function and visual tracking. The group started with a mean vestibular score of 38 and improved to a mean of 67, this is above the vestibular score expected for a group with a mean age of 10;4 years (Shimizu *et al*., 1994; Rine *et al*., 1998). Improvements were also found in visual tracking, with mean score improving from 35 to 49; and in saccadic control, with mean saccade accuracy improving from 56 to 68 and mean saccade latency improving from 49 to 61. By contrast the control group showed only small and non-significant changes in these four measures.

A similar general pattern obtained for the DST measures, with the DDAT group showing significant improvement on the majority (even allowing for age). In this case, though, the control group showed improvements in some scales of the DST (and significant improvement on DST overall score). The following sections attempt to dissect out the theoretical implications of the pattern of changes established.

Theoretical Issues

As noted in the introduction, the study was designed in such a way that four logically independent issues could be investigated. We take them in turn.

(i) *Incidence of cerebellar/vestibular symptoms*.

This school-based study cannot provide definitive data on the true incidence of cerebllar signs. Definitive norms for the posturography and ENG equipment would be needed, as would an explicitly representative set of participants. Nonetheless, some remarks may be made. First, the incidence of the posturography difficulties was 67%, with 74% for the visual tracking. Second, 67% of the group showed significant delay (centiles 1–30) on bead threading, 67% on rapid naming, and 50% on balance. These are high incidences, comparable with those for phonological skill (72%), reading and nonsense passage reading (78 and 72%, resp.), and considerably greater than the incidence for other DST sub-scales except spelling (55%). Since DST score was used in selecting the participants, one would not wish to read much into these data. It is noteworthy that the incidence of balance difficulties found by the objective dynamic posturography was considerably higher than that revealed by the much more limited postural stability test in the DST, indicating that the latter may not be as effective as one would like in identifying balance difficulties.**

(ii) *Effectiveness of DDAT intervention on cerebellar/vestibular symptoms.*

There were clear and significant improvements in cerebellar/vestibular performance and in eye movement performance. It is evident from Figure 1 that

**One specific suggestion is that the DST postural stability measure is only of forward/backward stability, whereas left/right stability as measured by posturography may be a more sensitive index. Interestingly, the original Nicolson and Fawcett (1990) balance study used the classic Romberg balance test in which the child stands with one foot in front of the other, and hence measured lateral stability.

there was considerable progress in all the tests administered at beginning and end, with perhaps the most notable being the reduction of the below-normal-range posturography scores from 75% incidence to 11% incidence for the DDAT group. By contrast, there were no significant improvements for the control group. For posturography the before/after incidence rates were 56 and 60%, respectively. Similar results obtain for the corresponding tests in the DST. The DDAT group improved significantly on both postural stability and bead threading (incidence of risk dropping from 50–0% and 67–22% resp.). The controls showed an improvement (35–12%) for postural stability but not for bead threading (47–53%).

These data suggest strongly that impaired cerebellar, motor and eye movement performance do not recover without a specific intervention, and do recover with the DDAT intervention. The significance of this finding for literacy skill depends upon the role of the improved cerebellar/eye movement skill in transfer to literacy-related skills, as discussed below.

(iii) *'Near transfer' to fundamental cognitive skills*.

It is encouraging, but not particularly surprising that progress is made on balance following the balance training. The first key theoretical issue is the extent to which the training generalises to the fundamental cognitive skills such as working memory (assessed in DST by backwards span), speed of processing (assessed by rapid naming), phonological skill (phonemic segmentation) and general cognitive fluency (verbal and semantic fluency). It is evident from Table 1 that there were significant improvements for semantic fluency, for phonological skill (and near significant for rapid naming—see also Figure 2) but that the improvement on working memory was relatively modest. Incidence of difficulty changed from 15–5% for verbal fluency, from 67–56% for rapid naming, from 72–33% for phonological skill and from 15–10% for working memory. The latter finding suggests that the backwards span measure may be somewhat insensitive. The improvement in phonological skill is very encouraging since this is a sign of dyslexia that is traditionally hardest to eradicate. For the control group, phonological skill also improved significantly unlike rapid naming, working memory, or semantic fluency. The comparative, effect size, analyses reveal solid effect sizes (>0.25) for verbal fluency, semantic fluency and rapid naming, a fair effect size (0.24) for phonological skill and negligible effect size for working memory. In summary, there does appear to be clear near transfer of the DDAT intervention to most of the fundamental cognitive skills underpinning literacy skill, especially those involving fluency.

(iv) *'Far transfer' to literacy*.

The DDAT group improved significantly on reading and on nonsense passage reading, but not on spelling or writing. The control group improved significantly only on nonsense passage reading. There was a significant interaction for reading, indicating that the DDAT group improved significantly more than the controls. In terms of incidence of difficulty, the DDAT group incidence dropped from 78 to 56% for reading, from 61 to 56% for spelling, from 17 to 6% for writing and from 75 to 50% for nonsense passage. For the controls the corresponding figures are 47 to 35%, 53 to 35%, 24 to 18% and 71 to 47%. The only solid effect size was for reading (0.35). It would appear therefore that there is some far transfer to literacy, specifically in the reading. This dissociation between the different components of literacy is a particularly interesting finding, as discussed below.

(v) *National attainment levels*

Finally, it is important to consider the transfer to the national SATS and NFER Reading scores (Table 3). The very much improved progress following the DDAT intervention is particularly encouraging.[††] There were substantial and significant improvements in reading age (reversing the decline caused by only 6 months progress in the previous year by a 19 month progress over the year including the intervention). Substantial and significant relative improvements also occurred for SATS comprehension and SATS writing, though the improvements were much less marked for the mathematics. This slight dissociation is particularly interesting in that it suggests that the improvements consequent on the DDAT exercises transfer primarily to literacy, language-based skills rather than to mathematical skills.

Discussion of the improvement in reading

As noted in the introduction, literacy skills are notoriously difficult to improve significantly, with the largest effects being obtained with the youngest children. The systematic survey undertaken by the National Reading Panel indicates a mean effect size of only 0.27 for the preferred 'systematic phonics instruction' in grades 2–6, and only 0.15 for low achieving readers in those grades (pp. 2–133). The report does not break down the effect size in terms of the cost in hours of the interventions, but most of the interventions lasted at least 6 months. Consequently the effect sizes per hour's intervention of dedicated reading intervention are tiny. It should also be noted that even in cases where reading improved, there was no concomitant improvement in fluency.

In this context, the differential improvement on DST reading within the initial six month period of treatment is striking. The CTH predicts that far transfer to reading will occur after the cerebellar function was significantly enhanced. Presumably therefore one would expect little literacy improvement initially, followed by accelerating improvement after several months. The excellent improvement in SATS literacy results (taken roughly 8-9 months after the start of the DDAT exercise treatment) appear to support this interpretation, though longer-term follow-up assessments are needed to evaluate it fully.

In order to tease out the possible mechanisms behind the improvement in literacy, a Pearson correlational analysis (with pairwise deletions for missing data) was undertaken, comparing the correlations of the reading improvement with all the other 52 available measures (Table 4).

It may be seen that the saccade measures (latency and accuracy) are among the strongest correlates.[‡‡] In addition, good working memory (post-test and progress), good dexterity (post-test and progress) and good semantic fluency progress are the major positive correlates. Intervention group also has a significant effect, as one would expect from the prior analyses. Negative correlations arise for: initial reading, initial semantic fluency, initial nonsense passage reading and follow-up overall DST score.

[††] Indeed, it is particularly encouraging to the school itself, since 'league tables' of performance on SATS measures are now an important national indicator.

[‡‡] The reduced power (attributable to missing post-test data for some control participants) causes the lower significance values.

Table 4. Major correlation with Reading Progress

Saccade latency progress	0.488†
Saccade latency (2)	0.388
DST backwards span (2)	0.370*
Saccade accuracy (2)	0.362
Saccade accuracy progress	0.351
Intervention group	0.350*
DST Semantic fluency progress	0.308†
DST Backwards span progress	0.286†
DST Beads (2)	0.282†
DST Beads progress	0.277
Initial RA deficit	0.233
DST Nonsense passage (1)	–0.231
DST overall (2)	–0.253
DST semantic fluency (1)	–0.339*
DST reading (1)	–0.532***

The correlations shown are between progress on the DST one minute reading test and the other measures taken. All DST test scores have been converted into age-appropriate centile equivalents before analysis. Only those correlations of at least 0.20 are shown. Significance values are shown as *($p<0.05$) and †($p<0.10$). The remainder are not significant.

Taken together, the pattern of correlations suggests that the treatment is particularly effective in increasing reading for children with high DST scores and poor initial reading. It appears that the improvement in saccade accuracy and latency following the DDAT treatment mediate some of this progress, as does increased speed of access to long term memory (semantic fluency).

This speed/eye movement hypothesis is of course speculative, and would need specifically designed studies to tease out these variables. Nonetheless, it is supported by the pattern of literacy improvements shown. In particular, the only differential improvement in literacy skills for the DDAT group was for DST reading, rather than spelling or writing. The DST reading test is a test of single word reading, involving reading as many words as possible - presented one per line in increasing difficulty - in one minute. Consequently, unlike many reading tests, it involves three of the four critical aspects for fluent reading: visual decoding, fluency and eye movements (but not comprehension). The other two tests are timed but do not significantly involve eye movements (or visual decoding).

It may be of particular interest that this hypothesis relates directly to the need to consider not only phonological and literacy skills in reading, but also the less tangible skills involving motor skill (eye movement control) and fluency (speed of cognitive operation). These concerns support observations by Rayner (1998);[§§] Stein *et al*. (2000) and (for fluency) the double deficit theory (Wolf & Bowers, 1999).

CONCLUSIONS

In summary, the study included a sample of children selected as needing further support within a junior school, and involved home administration of the DDAT

[§§] though Rayner, Foorman, Perfetti, Pesetsky, and Seidenberg (2001) surprisingly ignore eye movements in their review relating to dyslexia.

exercise regime. The progress of the DDAT intervention group was compared with that of a matched non-intervention group within the same school. Cerebellar/vestibular signs were substantially and significantly alleviated following the DDAT treatment whereas there were no significant changes for the control group. Even after allowing for the passage of time, there were significant improvements for the DDAT group in postural stability and in bead threading dexterity. There were also significant improvements in fundamental cognitive skills including phonological skill, and (one-tailed) for naming fluency and semantic fluency. Reading fluency showed a highly significant improvement for the DDAT group, and nonsense passage reading was also improved significantly. By contrast, for the control group significant age-adjusted improvement occurred only for nonsense passage reading and (one-tailed) for phonological skill. Significantly greater improvements for the DDAT group than the control group occurred for dexterity, reading, verbal fluency and semantic fluency. Incidence of strong risk on the DST reduced from 33% to 11% for the DDAT group. Performance of the intervention group on national tests of literacy indicated substantial and significant 'acceleration' of progress, allowing them to catch up with their peers. It is particularly notable that the effect sizes of the exercise group were 1.0 and greater for standardised reading, writing and comprehension tests. These scores are considerably better than the mean scores of 0.19 to 0.32 for reading (and 0.12 to 0.32 for comprehension) reported for specific phonics-based reading interventions for children of this age in the National Reading Panel study (2000, p. 2–133) and yet the exercise intervention did not involve reading at all (though of course literacy teaching was undertaken by the school). The relatively modest improvement on the SATS maths test suggests that the literacy improvement was not merely some generalised Hawthorne effect.

It is concluded that, as expected, the DDAT treatment was of direct benefit for balance, dexterity and eye movement control. There was also significant transfer to some of the cognitive skills underlying literacy; including three aspects of fluency—rapid naming, semantic fluency and verbal fluency. Most important, there was transfer to the reading process itself, as indexed by the DST one minute reading test and by the NFER Reading test, together with the SATS tests of writing and comprehension. It should be stressed that this is only a small study, and considerably larger scale research is needed to confirm these preliminary findings and to explore the ways by which the exercise mediates the literacy improvements. Nonetheless, the results do suggest that the exercise treatment was effective, not only in its immediate target of improving cerebellar function but also in the more controversial role of improving cognitive skills and literacy performance.

ACKNOWLEDGEMENTS

We wish to acknowledge the support from the staff and children at Balsall Common Junior School, Balsall Common, for their agreement to participate in these evaluations.

References

Ackermann, H., & Hertrich, I. (2000). The contribution of the cerebellum to speech processing. *Journal of Neurolinguistics*, *13*(2–3), 95–116.

Ackermann, H., Wildgruber, D., Daum, I., & Grodd, W. (1998). Does the cerebellum contribute to cognitive aspects of speech production? A functional magnetic resonance imaging (fMRI) study in humans. *Neuroscience Letters*, *247*, 187–190.

Allen, G., Buxton, R. B., Wong, E. C., & Courchesne, E. (1997). Attentional activation of the cerebellum independent of motor involvement. *Science*, *255*, 1940–1943.

Belgau, F., & Belgau, B. V. (1982). *Learning Breakthrough Program*. Port Angeles: WA: Balametrics.

Bloom, B. S. (1984). The two sigma effect: the search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, *13*, 3–16.

Bluechardt, M. H. W. J., & Shepard, R. J. (1995). Exercise programs in the treatment of children with learning disabilities. *Sports Medicine*, *19*, 55–72.

Cammisa, K. M. (1994). Educational kinesiology with learning disabled children: An efficacy study. *Perceptual and Motor Skills*, *78*, 105–106.

Clay, M. M. (1993). *An Observation Survey of Early Literacy Achievement*. Auckland, NZ: Heinemann.

Cohen, J. (1969). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed)*. New York: Academic Press.

Denckla, M. B., Rudel, R. G., Chapman, C., & Krieger, J. (1985). Motor proficiency in dyslexic children with and without attentional disorders. *Archives of Neurology*, *42*, 228–231.

Dore, W., & Rutherford, R. (2001). *Closing the gap.* Paper presented at the BDA 6th International Conference on Dyslexia, York, UK.

Eccles, J. C., Ito, M., & Szentagothai, J. (1967). *The cerebellum as a neuronal machine*. New York: Springer-Verlag.

Eden, G. F., VanMeter, J. W., Rumsey, J. M., Maisog, J. M., Woods, R. P., & Zeffiro, T. A. (1996). Abnormal processing of visual motion In dyslexia revealed by functional brain imaging. *Nature*, *382*, 66–69.

Fabbro, F., Moretti, R., & Bava, A. (2000). Language impairments in patients with cerebellar lesions. *Journal of Neurolinguistics*, *13*(2–3), 173–188.

Farnham-Diggory, S. (1992). *The Learning-Disabled Child*. Boston MA: Harvard University Press.

Fawcett, A. J., Singleton, C. H., & Peer, L. (1998). Advances in early years screening for dyslexia in the United Kingdom. *Annals of Dyslexia*, *48*, 57–88.

Fawcett, A. J. (2002). *Evaluating therapies excluding traditional reading and phonological based therapies*. London: DfES SEN Publications. Available at http://www.dfes.gov.uk/sen

Frith, U. (1997). Brain, mind and behaviour in dyslexia. In C. Hulme, & M. Snowling (Eds), *Dyslexia: Biology, cognition and intervention*. London: Whurr.

Fulbright, R. K., Jenner, A. R., Mencl, W. E., Pugh, K. R., Shaywitz, B. A., Shaywitz, S. E., Frost, S. J., Skudlarski, P., Constable, R. T., Lacadie, C. M., Marchione, K. E., & Gore, J. C. (1999). The cerebellum's role in reading: A functional MR imaging study. *American Journal of Neuroradiology*, *20*, 1925–1930.

Gayan, J., & Olson, R. K. (1999). Reading disability: Evidence for a genetic etiology. *European Child & Adolescent Psychiatry*, *8*, 52–55.

Glickstein, M. (1993). Motor skills but not cognitive tasks. *Trends in Neuroscience*, *16*, 450–451.

Hatcher, P. J., Hulme, C., & Ellis, A. W. (1994). Ameliorating early reading failure by integrating the teaching of reading and phonological skills: the phonological linkage hypothesis. *Child Development*, *65*, 41–57.

Hook, P. E., Macaruso, P., & Jones, S. (2001). Efficacy of Fast ForWord training on facilitating acquisition of reading skills by children with reading difficulties - A longitudinal study. *Annals of Dyslexia*, *51*, 75–96.

Ito, M. (1984). *The cerebellum and neural control*. New York: Raven Press.

Kephart, N. C. (1971). *The Slow Learner in the Classroom*. Ohio: Bell and Howell.

Knight, D., & Rizzuto, T. (1993). Relations for children in grades 2, 3 and 4 between balance skills and academic achievement. *Perceptual and Motor Skills*, *76*, 1296–1298.

Leiner, H. C., Leiner, A. L., & Dow, R. S. (1989). Reappraising the cerebellum: what does the hindbrain contribute to the forebrain. *Behavioural Neuroscience*, *103*, 998–1008.

Levinson, H. N. (1988). The cerebellar–vestibular basis of learning disabilities in children, adolescents and adults: Hypothesis and study. *Perceptual and Motor Skills*, *67*(3), 983–1006.

McCartney, K., & Rosenthal, R. (2000). Effect size, practical importance, and social policy for children. *Child Development*, *71*(1), 173–180.

McPhillips, M., Hepper, P. G., & Mulhern, G. (2000). Effects of replicating primary reflex movements on specific reading difficulties in children: a randomised, double blind, controlled trial. *The Lancet*, *335*, 537–541.

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

NICHD. (2000). *Report of the National Reading Panel: Teaching children to read*. Washington DC: National Institute for Child Health and Human Development.

Nicolson, R. I., & Fawcett, A. J. (1990). Automaticity: A new framework for dyslexia research? *Cognition*, *35*(2), 159–182.

Nicolson, R. I., & Fawcett, A. J. (1996). *The Dyslexia Early Screening Test*. London: The Psychological Corporation.

Nicolson, R. I., & Fawcett, A. J. (1999). Developmental Dyslexia: The role of the cerebellum. *Dyslexia: An International Journal of Research and Practice*, *5*, 155–177.

Nicolson, R. I., & Fawcett, A. J. (2000). Long-term learning in dyslexic children. *European Journal of Cognitive Psychology*, *12*, 357–393.

Nicolson, R. I., Fawcett, A. J., & Dean, P. (2001a). Developmental dyslexia: the cerebellar deficit hypothesis. *Trends in Neurosciences*, *24*(9), 508–511.

Nicolson, R. I., Fawcett, A. J., & Dean, P. (2001b). Dyslexia, development and the cerebellum - discussion. *Trends in Neurosciences*, *24*(9), 515–516.

Nicolson, R. I., Fawcett, A. J., Moss, H., & Nicolson, M. K. (1999). Early reading intervention can be effective and cost-effective. *British Journal of Educational Psychology*, *69*, 47–62.

Pennington, B. F. (1999). Toward an integrated understanding of dyslexia: Genetic, neurological, and cognitive mechanisms. *Development and Psychopathology*, *11*, 629–654.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*, 372–422.

Rayner, K., Foorman, B. R., Perfetti, C. A., Pesetsky, D., & Seidenberg, M. S. (2001). How psychological science informs the teaching of reading. *Psychological Science, Suppl.*, 31–74.

Rine, R. M., Rubish, K., & Feeney, C. (1998). Measurement of sensory system effectiveness and maturational changes in postural control in young children. *Pediatric Physical Therapy*, *10*, 16–22.

Rudel, R. G. (1985). The definition of dyslexia: Language and motor deficits. In F. H. Duffy & N. Geschwind (Eds), *Dyslexia: A neuroscientific Approach to Clinical Evaluation*. Boston, MA: Little Brown.

Shimizu, K., Asai, M., Takata, S., & Watanabe, Y. (1994). The development of equilibrium function in childhood. In K. Taguchi & M. Igarashi & S. Moru (Eds), *Vestibular and neural front* (pp. 183–186). New York: Elsevier Science BV.

Silver, L. B. (1987). The 'magic cure': A review of the current controverisal approaches for treating learning disabilities. *Journal of Learning Disabilities*, *20*, 498–505.

Silveri, M. C., & Misciagna, S. (2000). Language, memory, and the cerebellum. *Journal of Neurolinguistics*, *13*(2–3), 129–143.

Snowling, M. J. (2000). *Dyslexia*, *2nd edition.* Oxford. Blackwell:

Stein, J., & Walsh, V. (1997). To see but not to read: The magnocellular theory of dyslexia. *Trends in Neurosciences*, *20*, 147–152.

Stein, J. F., Richardson, A. J., & Fowler, M. S. (2000). Monocular occlusion can improve binocular control and reading in dyslexics. *Brain*, *123*, 164–170.

Tallal, P., Merzenich, M. M., Miller, S., & Jenkins, W. (1998). Language learning impairments: integrating basic science, technology, and remediation. *Experimental Brain Research*, *123*, 210–219.

Thach, W. T. (1996). On the specific role of the cerebellum in motor learning and cognition: Clues from PET activation and lesion studies in man. *Behavioral and Brain Sciences*, *19*, 411–431.

Torgesen, J. R. (2001). Theory and practice of intervention. In A. J. Fawcett (Ed.), *Dyslexia: Theory and good practice*. London: Whurr.

Wimmer, H., Mayringer, H., & Raberger, T. (1999). Reading and dual-task balancing: Evidence against the automatization deficit explanation of developmental dyslexia. *Journal of Learning Disabilities*, *32*, 473–478.

Wolf, M., & Bowers, P. G. (1999). The double-deficit hypothesis for the developmental dyslexias. *Journal of Educational Psychology, 91*, 415–438.

World_Federation_of_Neurology. (1968). *Report of research group on dyslexia and world illiteracy*. Dallas: WFN.

Zeffiro, T., & Eden, G. (2001). The cerebellum and dyslexia: perpetrator or innocent bystander? Comment. *Trends in Neurosciences*, *24*(9), 512–513.

