



PUBLISHED STUDY:

ACADEMIC PROCESS

Until 2008 SuperReading had shown amazing results, but had not been tested by independent, professional researchers. 84% of the dyslexic students finished with higher scores than non-dyslexic university professors. Subsequent courses saw all graduates scoring higher than non-dyslexic students without SuperReading. The Italian study also showed the same results, both before and after translation into Italian.

ACADEMIC PRECISION

Dr. Cooper applied rigorous statistical analysis to our testing methodology. His control group of professors actually scored lower over time. It was expected they could have improved due to testing familiarity. They did not. They received the same instruction as the learning group on their first day before any tools were taught. This ruled out any increases due to chance or knowing what to expect.

LSBU PAPER by Dr. Ross Cooper

Draft Paper submitted to the Journal of Inclusive Practice in FE & HE Vol 1 Number 2 (Spring 2009)

ABSTRACT

The aim of this study was to gauge whether the impact of a reading course for degree level adult dyslexic readers (n=15) was sufficiently robust to justify more extensive research and experimentation. While recognising the limitations of this pilot research and the methodological difficulties of measuring 'comprehension' gains, the 'reading effectiveness' of the group appeared to double in ten weeks. A t-test provided a statistical significance of $p < 0.002$. There was also a statistically significant negative correlation between pre-course TOWRE nonword test scores and the percentage improvement in reading effectiveness. This is surprising and worthy of further investigation in itself, since we might normally predict that those with the most phonological difficulties are likely to make the least progress, not the most. All the participants were enthusiastic about the positive impact of the course on their reading and report a range of affects such as increased stability of print, pleasure and speed of reading. We can conclude that the apparent effect, and the nature of the correlation between the effect and difficulty reading nonwords, is sufficient to justify further research and experimentation.

INTRODUCTION

Background

This research trial arose in a specific context. Ron Cole approached LLU+ after teaching his 'Super Reading' course for fifteen years with the observation that dyslexic readers appeared to make the most progress. The intention was to begin to evaluate this observation and to try to understand the experience of dyslexic readers

on his course. I was particularly interested in his unusual approach to teaching reading improvement, because it was based on an eye exercise. The specific purpose of the trial was to gauge whether there was a measurable impact on dyslexic readers that would justify further investigation, investment and collaboration. This led to a set of research questions: 1. How can we measure improvements in comprehension as well as speed? 2. To what extent might a visual approach to reading overcome phonological difficulties? 3. How might readers with visual processing and tracking difficulties experience a visual approach to reading? 4. To what extent are existing tools to measure reading inappropriate? 5. Might the focus on what is easy to measure have misled researchers away from what is important about the nature of reading? Of all these questions, the most methodologically difficult is how to measure improvements in comprehension when we know that a great many factors are involved (Ellis, 1993), including, interest, prior knowledge, level of tiredness, or stress, vocabulary, difficulty of the text, motivation, language development, working memory, decoding skill, speed of reading, reading styles, reading strategies, culture, background and knowledge of schema & genre, colour of paper and text fonts. Since all the participants were undergraduates, or postgraduates, the texts used to measure comprehension were at an advanced level and there was no differentiation for any difficulties with reading or visual stress. Since four of the participants were below 16th percentile for all four standardised measures of reading prior to the course (WRAT4 single word and comprehension and TOWRE single words) and only two were above the 16th percentile for all four measures, this seemed a very challenging assumption.

Predictions

We made the following predictions: 1. Reading effectiveness would double if the participants practiced 'eye-hops' for half an hour a day. 2. The WRAT single word reading and TOWRE nonword reading scores are likely to remain static over the same time period. 3. WRAT comprehension scores are likely to rise, but as these are untimed sentence level cloze tests, the rise may be minimal. 4. The time taken to do reading tests is likely to fall. 5. TOWRE sight recognition scores may improve due to increased speed of visual recognition. These predictions are predicated on the contention that existing standardised tests are poor measures of real reading (Hansen et al, 1998); that this trial is likely to highlight the inadequacies of the assessment tools as much as the impact of the course. I had hypothesised that those with poor reading skills (four of whom were also bilingual learners) would be unlikely to make as much progress as those with more advanced reading skills (and the advantage of English being their first language). This view was not shared by Ron Cole.

METHODOLOGY

Subjects

The course began with 20 participants. For the purposes of this project, we defined those who were 'compensating' for their dyslexia by pre-course standardised scores on the WRAT lying within an 'average range' (even for the range of scores representing a 95% confidence interval). Mean pre-course WRAT scores Reading Comprehension Participating 108 109 Non-Participating 84 84 All Participants 96.5 98.6 Twelve of the participants fell into the 'compensating' category (although eight of them achieved scores on the TOWRE below the 16th percentile). Eight participants

can be categorised as the 'non-compensating' group. Four of the 'non-compensating' group were also bilingual.

Selection of subjects

London South Bank University Centre for Learning Support & Development emailed all dyslexic students on their database, letting them know that a free reading course was available as part of pilot research. The timing of the course, in the lead up to the summer exams, was not ideal. All interested participants with a diagnosis of dyslexia who were available at the specified times were accepted onto the course. Sixteen students were enrolled onto the course through this means. Four dyslexic staff at LLU+ were also invited onto the course.

Drop out

Four of the students dropped out of the course after the first session. Only one of these drop outs responded to requests to discuss the reasons. There were three: one felt very uncomfortable being 'tested' in a group. She had felt overwhelmed by the first session and uncertain about what 'instructions' she might have missed. She was already overloaded by course work and did not find any time to practice the 'eye-hops' between the first and second session. However, in a post-course interview, she expressed the view that she felt that the techniques she had learned, had she had time to practice them, were 'probably beneficial'. One of the invited dyslexic professionals (an assistive technology tutor) dropped out on the birth of his daughter. He also expressed the view that the course was 'useful'.

The 'Super Reading' course.

The course was taught entirely by Ron Cole over six three-hour sessions. The sessions were held once every two weeks. Participants were taught a range of skills and practices including how to practice 'eye-hops', how previewing and reviewing reading was beneficial, the importance of using their finger to track text and a memory technique. The sessions were intended to be motivational and enjoyable which may have produced a 'Hawthorne' effect (Spratt, 1952). Comprehension was always prioritised over speed. The instruction, 'read this as fast as you can while fully understanding it' was therefore an instruction often repeated. Participants were asked to agree to practice the eye-hop exercises for a minimum of half an hour a day. In the post course interviews, it became clear that very few participants managed this. We averaged around 15 minutes a day. Within each session, participants tested their reading with prepared texts and comprehension questions. 'Reading Effectiveness (RE)' was calculated by multiplying the words per minute by the percentage of correct answers given to the questions. The methodological implications are discussed below.

TESTS USED ON COURSE

The testing process during the course was as follows: 1. Participants were asked to read the test texts as quickly as they could while fully comprehending them. 2. At an agreed moment, test texts were turned face side up, but not yet looked at. 3. At a further agreed moment, participants began to read their text as a large digital clock began timing on the smart board. 4. As soon as they had finished reading, participants turned over their texts and recorded the time taken to read it. 5. They then turned over the questions and answered them as fully as they could, before turning the questions back over. 6. Once everyone had completed this, at an agreed

moment, the process started again, the texts were reviewed, a second time taken was recorded and a second comprehension score recorded.⁷ Participants were then helped to calculate their words per minute and reading effectiveness for 'first' and 'review' reading. All test texts were exactly 400 words long. They included large numbers of numerical and other details that were often included in the questions. During the process, Ron Cole watched carefully for anyone forgetting to check the time, so that timing errors could be reduced. From session two, participants were invited to preview the text for up to the first 30 seconds of reading time during the first read through. This time is included in all calculations of words per minute. For the purposes of the research, all calculations of reading effectiveness were checked. All test texts were randomised during the length of the course so that intrinsic difficulties of particular texts, or the questions, could not play a role in the apparent development of reading effectiveness progress. There was no differentiation of texts for readers of different 'ability'.

Pre & post tests

All participants were given a range of reading tests before and after the course. Standardised tests were chosen that could be administered twice to check on 'progress': WRAT4 Reading & Comprehension, TOWRE Sight and Nonwords. These tests are not without limitations and methodological difficulties. All have been standardised on USA populations which makes it difficult to interpret the results meaningfully. The TOWRE has only been standardised up to the age of 25 and the average age of the participants on the course was 41. This means that the scores must be treated with caution, although the primary purpose of using these tests was to look at comparative results rather than absolute results. Another methodological problem is that these tests are not good tests of reading, particularly the single word tests, since reading words in combination is very different from single word reading (Ellis, 1993, Tadlock & Stone, 2005). The time taken to administer the WRAT4 was recorded because we had predicted that the time taken would change from pre to post course. It was explained to participants that the WRAT4 was 'not a timed test, but I am going to time it to gather more information'. Since the TOWRE is timed, it was hypothesised that the TOWRE sight word scores would rise to reflect the additional speed. Since reading in context provides a range of semantic and syntactical cues to support word recognition, the increased understanding predicted when reading was not necessarily expected to improve single word recognition. The WRAT comprehension test is clearly intended as a reading comprehension test. However, it has a number of flaws. Comprehension is limited to sentence level, rather than discourse. More importantly, it presents 'word finding' problems (Wolf & Bowers, 2000) that often overshadow comprehension. Most of the participants reported that the main difficulty was finding the right word to fit the space. For the four bilingual learners and one of the non-bilingual learners, finding grammatically acceptable words was also reported as a major problem. Using a similar test twice can be methodologically problematic for two distinct reasons. The first is that the testee has a better understanding of the nature of the test and has practiced whatever skill is required. The second is more relevant to children than adults, since we can expect a child to have made progress in their reading skills without any additional intervention in the intervening time. This temporal effect can also apply to bilingual learners, although in this case, all 4 bilingual learners had been learning English for a minimum of 7 years, so a 10-week period is unlikely to account for any

change. The WRAT4 manual claims that test re- test scores can be expected to rise by 2 standardised points.

Action research

An important aspect of the research methodology was to explore the subjective experience of the participants on the course, including my own as a dyslexic reader. This was supported by discussing the experience of the course and tests with participants, including two dyslexic colleagues among the participants. It was expected that this would help provide a range of insights that would promote a better understanding and interpretation of the experience and of the test scores. This runs the risk of influencing my interpretation of tests, but this risk was considered small in an exploratory trial intended to understand the experience of learners as much as measure their progress. Care also had to be taken that no tests were used with which any participant was familiar. Since the WRAT4 was a relatively new test, none of the participants were familiar with the content except me, having begun to use WRAT4 (and TOWRE) with learners. My own test scores on these tests were excluded from the data. None of the other participants had any experience of the TOWRE. One other participant was familiar using WRAT3 with learners. Some of the participants thought that they might have used the WRAT3 as part of their own assessment.

RESULTS Reading Effectiveness

Reading effectiveness, as measured, increased dramatically over the 10 weeks. All participants benefited, from a 22% to a 408% increase. On average, RE increased by 110%. It could be hypothesised that comprehension practice alone could improve the RE scores. However, we would not then expect that those with the lowest test scores prior to the course would gain the most. It is interesting to compare those who were 'compensating' with those that were 'not'. Comparisons remain tentative, because the group sizes are small ($n=8+7=15$). It should therefore be stressed that this comparison is for descriptive purposes, since the differences do not achieve statistical significance. Nevertheless, in this trial the 'non-compensating' group made more progress in reading effectiveness (expressed as a percentage) than the 'compensating' group (140% compared to 80%). Interestingly, in the first session, reading speeds changed very little for both groups between the first reading of the test text and the review reading:

FIRST SESSION: wpm (first read) Comprehension wpm on review Comprehension
Compensating 215 51% 215 76% non-compensating 108 41% 110 66% All

Participants 165 46% 110 66% Whereas the speed changed dramatically during the test in the final session:

LAST SESSION: wpm (first read) Comprehension wpm on review Comprehension
Compensating 228 79% 580 94% non-compensating 179 64% 241 87% All

Participants 205 61% 91% 91% We can also see that comprehension scores rose significantly at both stages. By the end of the course, the 'non-compensating' groups' reading speeds and comprehension both exceeded the scores achieved by the compensating group at the beginning of the course. For each of the differences between pre- and post-test average scores reported here for all participants, the statistical significance of the difference was tested using a paired t-test. Despite having a small sample, statistical significance was achieved for the increased 'comprehension' ($p<0.02$ at first read through, $p<0.001$ at review), and for the increased speed of review ($p<0.001$). At the end of the course, the 'non-

compensating' group are reading a mean of an additional 71 words per minute at the first read stage and able to answer over half as many questions again. Overall the group is reading at a mean of an additional 40 words per minute and answering a mean of 15% more questions correctly. At the review stage, participants are reading a mean of an additional 256 words per minute (doubling their reading speed) and answering an additional 20% of the comprehension questions correctly. Reading effectiveness scores can be calculated for both the first read through and the review reading stages of the test, however, for the purpose of comparison, a 'combined RE' score was calculated. This is because the slower the reading speed at the first read through, the more we can expect to have been understood (or memorised) and the faster the second read through becomes (and vice versa). In other words, the RE scores from the first read through and the review are not independent variables. Combining them therefore provides a better measure of progress. This was done by adding both reading times together, calculating a 'combined wpm', and multiplying by the final comprehension percentage.

Session	Compensating	Non-compensating	All Participants
Session 1	80	36	59
Session 6	153%	86%	118%

The improvement from session 1 to session 6 is highly significant ($p < 0.002$). By the end of the course, the 'non-compensating' group have exceeded the original combined RE score of the 'compensating' group. There is also a statistically significant negative correlation between TOWRE nonword scores and the percentage progress made (-0.767 , $p < 0.01$), meaning the lower the TOWRE nonword scores, the greater the percentage gains in 'reading effectiveness'. There is less correlation with TOWRE sight word scores (-0.310 , which is not statistically significant). Although we expected a correlation between the reported hours of practice and progress, there was very little correlation. However, this was difficult to gauge. Once the participants began to use their 'pattern reading' skills with ordinary text, their real practice times became very difficult to report.

Pre & Post Standardised Test Scores WRAT4 Single reading. As predicted, the standardised scores changed very little.

Pre-course test	Post course re-test
Compensating 107.7	116.4
Non-compensating 83.6	78.3
All Participants 96.5	98.6

Overall, the 'compensating' group achieved their higher mean test result (+0.58 SD) in 82% of the time of the pre-course test. The non-compensating group achieved their lower mean test result (-0.35 SD) in just 33% of the time of the pre-course test. Dealing with such small numbers can be misleading. The combined results of all participants show a mean rise of 2.1 in standardised score, which is consistent with test/re-test expectations.

WRAT Reading Comprehension

These scores remained stable over the 10 weeks.

Pre-course test	Post course re-test
Compensating 109.1	106.1
Non-compensating 84	84.4
All Participants 96.6	95.3

Overall, the mean standardised score changed from 96.6 to 95.3, while participants took 80% of the time for the re-test than was taken for the pre-course test.

TOWRE tests

As the TOWRE subtests are sensitive to reading speed, we expected the sight word scores to increase, but not the nonword scores.

Pre-course test	Post-course test
Compensating 90.2	97.2
Non-compensating 69.5	75.5
All Participants 79.8	86.3

($p < 0.002$) TOWRE nonword Pre-course test Post course re-

test Compensating 9.9 95.9 Non-compensating 71 73.7 All Participants 81.6 85.5
($p < 0.04$)

DISCUSSION

Reading Effectiveness

Since the calculation of 'reading effectiveness' is dependent on both speed and the percentage of correct answers given to the questions, 'reading effectiveness' inevitably includes arbitrary elements. How might a reader have answered a different set of questions? How might their comprehension be affected by their interest in the subject matter, their prior knowledge, their vocabulary? These are difficult questions to address and are best handled by a larger sample than available in this trial. The scale of the apparent gains, their statistical significance and the subjective experience of increased reading speeds with comprehension, are, however, difficult to ignore. It was also surprising to validate Ron Cole's assertion that readers with the lowest reading scores on all measures at the beginning of the course, made better than average progress. For example, the four bilingual readers improved their reading effectiveness by 122%. As already argued, reading text involves much more than phonological decoding. The correlation of reading effectiveness (RE) improvement with difficulties reading the TOWRE nonwords is particularly interesting and appears to support the view that readers with phonological decoding difficulties will make better progress by building on their strengths rather than trying to remediate their weaknesses (Butterworth, 2002). This interpretation of the findings would benefit from further investigation.

Pre & post test results

WRAT4 single word reading Individual test/re-test scores, in the main, stayed within the 95% confidence interval range between pre and post course tests. Two of the 'compensating' group achieved test scores higher than the 95% confidence interval on the post course test (123 to 131, and 104 to 116). Both of these individuals maintained that they experienced better print stability following the 'eye-hop' exercises. One of the 'non-compensating' group achieved a score below the 95% confidence interval (87 to 76). But this score was achieved in 20% of the time taken for the pre-course test. The comparison between pre and post course test scores is interesting for two reasons: As predicted, the scores changed very little. This appears to confirm that untimed single word reading test scores provide little meaningful correspondence with real reading skills, (Tadlock & Stone 2005). The second is that similar scores were achieved in only 40% of the time taken in the pre-course test. This appears to indicate that real improvement in tracking and visual recognition has resulted from the intervention. It could be argued that a saving of 60% time is a valuable gain in itself. Indeed, the 'non-compensating' group achieved their scores in 33% of the time.

WRAT4 Reading Comprehension.

Since this is intended as a test of comprehension, we might have expected these test scores to rise. Consequently, this result appears to undermine the claim of the course to improve comprehension. However, the WRAT4 test scores are affected by both word retrieval difficulties and grammatical expression. Participants often expressed that they understood what they were reading, but could not think of the

right word to fit in the gap. One of the 'compensating' group achieved a retest score below the 95% confidence interval (128 to 117). However, this was achieved in 41% of the time taken for the first test. One of 'non-compensating' group scored above the 95% confidence interval on the re- test (68 to 78). This was achieved in 73% of the time of the pre-course test. This test is also a test of comprehension at single sentence level. This means that the context is restricted unlike a page of text which provides extended cues for expectations and meaning. Overall, the test scores were therefore relatively stable, despite increased speed. The mean time taken to achieve similar test scores was 80% of that taken on the first test. The reduced time in which the scores were achieved has a statistical significance of $p < 0.05$. **TOWRE sight words.** It is important to remember that almost all the participants were above the ceiling age for the TOWRE standardisation. The scores cannot therefore be used as more than a comparative indicator of change for individuals over time. However, as we had predicted, scores on the TOWRE sight words increased. This appeared to be for two reasons. More words were read in the 45 seconds. Less visual misrecognition occurred. This appears to support the subjective views expressed that the 'eye-hop' exercises had improved print stability. One participant had noticed that they now found music easier to read; that they could take in larger passages of music at once and were able to recognise note positions more easily and accurately. Increasing the TOWRE sight recognition scores by almost half a standard deviation in just ten weeks can be considered to be a significant improvement.

TOWRE nonwords.

We had predicted that nonword reading would not improve, since there has been no phonics of any kind as part of this course. Indeed, readers were gradually encouraged to abandon sub-vocalisation until reading for meaning required no phonetic attack or repair. It was therefore a little surprising to discover that these results rose by 0.26 of a standard deviation. The explanation for this probably resides in the improved tracking and stability of print. Some researchers argue that improvements in reading enhances phonological awareness (Morais et al, 1987). I would suggest that this is unlikely in this context- just 10 weeks and the reduction of sub-vocalisation encouraged. However, before nonwords can be decoded, they must first be visually tracked accurately. Consequently, any performance on a nonword test must include some measure of visual processing of text. I would suggest that the increased scores are most likely to be explained by improved speed and accuracy of tracking and stability of print, rather than any improvement in phonological awareness.

Evaluation and Limitations

This research project was designed to identify whether there was an effect that needed further investigation. Results need to be treated with caution because there was no control group and the sample is relatively small ($n=15$). In addition, the TOWRE is only standardised up to the age of 25. The trial was successful in confirming that there is a sizeable effect that needs further investigation. With this small sample, the impact appears consistent and dramatic. The apparent 'reading effectiveness' of the participants has doubled in 10 weeks and all the participants report dramatic improvement in both their speed of reading and the stability of print where this was a prior difficulty. One of the most exciting indications is that those with the most reading difficulty made the most progress (measured as a percentage gain), despite no differentiation of reading material or tests. Even more interesting is

the strong negative correlation between RE progress and pre-course TOWRE nonword scores ($p < 0.01$). This would appear to indicate that the reading course enabled reading effectiveness to improve most dramatically for readers with phonological difficulties without addressing phonological difficulties at all. It was stated at the beginning that a significant methodological problem is that we do not have fit-for-purpose tools for measuring reading comprehension. Indeed, reading comprehension is such a complex process that good tools are very difficult to design. In this methodological vacuum, we used a range of available tools and designed our own measures of 'comprehension'. The project has provided further evidence to challenge the appropriateness of existing single word tests to measure reading skills. They may predict reading difficulty, but they do not necessarily provide clear indications of how to improve reading skills (Torgesen, et al, 2001). However, it must be acknowledged that our own measures of reading 'comprehension' were flawed. We had recognised that using multiple choice questions to measure comprehension can lead to false positives due to factors such as the ability to eliminate unlikely answers, taking risks and sheer chance. We attempted to avoid these by asking highly specific questions that could not be known without detailed reading of the texts, and that were very demanding of the reader. The problem with these is that they also tested detailed short-term memory. This demand slowed the participants reading, because we had to dwell on details which most readers would normally 'look up' if they needed them. The experience of being 'slowed down to memorise detail' was a common one. I would therefore suggest that the RE increases are artificially low as a measure of the benefit of the course. For example, my own reading speed with 'good' comprehension has risen from around 250wpm to 850wpm. This makes reading texts or marking assignments much faster. This speed is similar to the 'review speed' on my last test (857wpm). Discussions with participants and referring to the loosely measured speeds with which they read novels towards the end of the course, appears to confirm that this is more representative of our new reading with comprehension speeds. The mean final review speed of the group was 580wpm (but ranged between 100wpm to 1500wpm). The mean review reading speed for the non-compensated dyslexic group at the end of the course was 241wpm (this is 26wpm faster than the review speed of the 'compensating' group at the start of the course, with 15% more questions answered).

Perceptions

The action research element of the research project provides additional evidence that the reading course was beneficial, since everyone interviewed after the course confirmed that they had experienced direct benefits from the 'eye-hop' exercises and intended to carry on with them. There were 3 participants who experienced particular tracking difficulties at the beginning of the course. Two of these were colleagues who remained unconvinced by the course for the first 4 sessions. They relied heavily on phonetic decoding and sub-vocalisation. Both considered themselves good readers prior to the course. They all found that the use of the finger during 'eye-hops' was distracting. They found the gradual move from sub-vocalisation to visual reading was a difficult process for them. In contrast, others on the course described the process very positively. For example one participant said, 'When I started doing the 'eye-hops' it felt like changing from walking (where I said every word) to beginning to run, where I only said the words when my 'feet' came down. They always seemed to be the important words, I don't know how that happened. As I got faster and faster I ended up 'flying' where my feet didn't touch the ground at all!' In contrast, the three

'sub-vocalisers' resisted the experience and made little progress from session to session. However, on the 4th or 5th session they suddenly found that they were able to comprehend text without full sub-vocalisation and they then made dramatic progress. All of them changed their opinion of the course and expressed the intention to continue with the visual approach, 'since I feel that I've only just started to get the benefit. 'Until this 'break through' I had begun to believe that the course suited those dyslexic readers who had phonological difficulties by building on visual strengths, rather than those who had visual processing difficulties. But this sudden breakthrough appears to indicate that it is merely a matter of time; that skilled reading is essentially a visual process and requires visual tools. A possible hypothesis that progress on the course was simply depressed by visual processing difficulties giving rise to an artefact of a negative correlation with TOWRE nonword scores is inconsistent with the evidence. Progress also correlates negatively with the TOWRE sight word scores (meaning that the lower the sight word scores, the greater the progress), but the correlation (-0.308) is weak compared with the negative correlation with the TOWRE nonword scores (-0.767), and indeed weaker than the negative correlation with the TOWRE combined scores (-0.545). Participants described the experience of increased print stability and improved reading. One described reading a whole book for the first time. Many of how their pleasure in reading has increased. Another finding music easier to read (see above). Another described how he had noticed that, from being a slower reader than his girlfriend, he was now faster and having to wait for her to finish shared reading. I also find that I am taking much less time to assess dissertations. I read three times more books on my summer holidays than I ever have before.

CONCLUSIONS

Evaluation The trial provides very good evidence of a dramatic effect that has improved the reading effectiveness and pleasure of all the participants. It remains to be seen precisely what causes the effect. There were a number of factors involved. The teacher's charisma and ability to engage and motivate the participants is one factor, although it is difficult to imagine that simply motivating the participants could have such a dramatic impact when reading difficulties have been a lifelong and intransigent difficulty for many of the participants. Nevertheless, it will be important to discover whether the effect is transferable; that the effect is a product of strategy rather than charismatic teaching. The most obvious critical explanation for the effect is flaws in the measuring methodology. This would argue that the effect was caused by variable comprehension test validity and participants learning how to do the tests more effectively, rather than the test results measuring any real change in skill. There is some evidence to support this view. Learners learned how to preview more effectively and began to read more strategically, particularly once they realised how detailed the 'comprehension' question were. However, there is also considerable evidence to the contrary, including: 1. The reading tests were randomised. 2. Learning good test strategies alone would be difficult to account for the gains. Let us take one example. Just one of the participants realised that she found reading far more efficient if she knew what the questions were first. She therefore changed strategies to read through quickly the first time, find out what the questions were, and then take more care to read through the 'review' knowing what she was looking for. On the surface this looks like good evidence that strategy can account for much of her improvement. However, the time taken to review the last test (when she achieved 90% 'comprehension') was just 80 seconds. This compares with over 5 minutes to

achieve 90% comprehension in the first test. In addition, although she only skim read the text in 48 seconds during the last test, she achieved a 40% comprehension, compared with almost 6 minutes in the first test when she scored a 50% comprehension. Learning to preview and ask questions of the text are generally considered good reading-for-meaning skills. So the strategies that might account for some of the improvement are part and parcel of good transferable reading strategies. Therefore, rather than be discounted as alternative explanations for reading improvement, they could be considered a legitimate part of the improved skills being evidenced. In addition to this, improvements in the RE scores are also reflected in the improved TOWRE scores and the increased speed with which WRAT4 scores were achieved. Teaching preview skills is an important metacognitive strategy. What the course was very effective in demonstrating is that readers succeeded in answering more questions in less time when they used the first 30 seconds of reading time to preview the text than when they did not. Although I teach the technique, I did not use it myself if I thought that time was of the essence. I have now learned that failing to do so is a false economy. While many dyslexic readers can appear to overcome their reading difficulties, the progress made during this course in 10 weeks is, in my experience, unprecedented. This may be partly because very little research has been undertaken to evaluate reading comprehension, recognising that it is methodologically problematic. Yet improving reading effectiveness must lie at the heart of any reading intervention.

NEXT STEPS

Research funds are now needed to extend the pilot project. This trial has provided very good evidence of an effect, we now need to establish with more certainty precisely what has created it and to what extent it is transferable. This can only be done with further trials involving a larger sample and control group. The participants on the course seem in little doubt that it is the 'eye-hop' exercises that made the difference, but there were other factors at play on the course. The next phase of the research would benefit from reducing the memorisation necessary to achieve 'comprehension' scores. We can expect that the course would be particularly effective for any dyslexic learners progressing to higher level courses that put more pressure on reading skills. This tends to occur quite suddenly as learners progress to A-levels, but in particular when they progress to university. We are very interested in trialling the intervention with students just prior to progressing to university and can foresee a strong argument for the DSA paying for the intervention, since it could be very effective in preparing students for university. Indeed all the students expressed the view that they wish they had been able to take this course before they started their university courses rather than during them (and particularly not during their preparations for exams). In order to develop the framework for further research, we are planning to be trained at LLU+ to teach the Super Reading course. This would give us the capacity needed for the more extensive research and allow us to evaluate the transferability of the course.

References

- Bell, T. (2001). Extensive reading: Speed and comprehension. The Reading Matrix, 1(1). <http://www.readingmatrix.com/articles/bell/index.html> Butterworth, B, (2002) Lost for Words, in Tim Radford (Ed.) Frontiers 01: Science and Technology, Atlantic Books
- Ellis, A. (1993)

Reading Writing and Dyslexia, a cognitive analysis. Psychology Press Ltd. Hansen, J. Johnston, P., Murphy S. & Shannon P. (1998)

Fragile Evidence: A Critique of Reading Assessment, Lawrence Erlbaum Associates. Hill, J.K. (1981) Effective reading in a foreign language, English Language teaching Journal, 35 270-281.

Sprott, W.J.H (1952) Social Psychology, London, Methuen, Tadlock, D. with Stone, R (2005)

Read Right! Coaching your child to reading excellence, McGraw-Hill. Torgesen, J.K., Alexander, A.W., Wagner, R.K., Rashotte, C.A., Voeller, K.K.S. & Conway, T. (2001) Intensive remedial instruction for children with severe reading disabilities. Journal of Learning Disabilities, 34(1), 33-58

Wolf, M. & Bowers, P.G. (2000). Naming-speed processes and developmental reading disabilities: an introduction to the Special issue on the double-deficit hypothesis, Journal of Learning Disabilities, 33[4], 322-324

1. Evaluation of a 'Super Reading' Course with Dyslexic Adults - Dr Ross Cooper

UPDATED RESULTS & STATISTICAL ANALYSIS 1 May 2009

I have now got all the data for the group of 11 that just did the reading tests with no course. From the first to the last test, their reading comprehension and first read wpm all dropped slightly by -2 to -5 %. The review read speed increased slightly by +12%. These are all well within the fluctuations you might expect by chance. The mean difference in what I call the combined RE scores is -7% (This varied between -49% to +27%). So practicing the tests made no difference to scores. None of these results for this group have any statistical significance (which means that the maths says they happened by chance). If we compare this with everyone else who has done the course here:

1. Increase in first read wpm = +23%
2. Increase in first read comprehension = +26%
3. Increase in review wpm = +161%
4. Increase in review comprehension = +18%
5. Increase in first read RE = +53%
6. Increase in review RE = +204%
7. Increase in combined RE = +85% (individually this varied from +9% to +408%. Since the individual test variations in combined RE scores can be up to about +/-50%, it is fairly safe to say that the extreme individual results are probably partly explained by up to that variation).

Statistical Significance

The statistical significance is calculated by how many chances out a hundred the results could happen by sheer chance. Less than 5 times out of a 100 ($p < 0.05$) is the threshold for deciding it is statistically significant. All of these results are much more significant than that.

1. 2 out of 100 [Increase in first read wpm = +23%]
2. Less than 7 out of 10,000 [Increase in first read comprehension = +26%]
3. 1 out of a million [Increase in review wpm = +161%]
4. 4 out of 10,000 [Increase in review comprehension = +18%]
5. 2 out of 10,000 [Increase in first read RE = +53%]
6. 2 out of a million [Increase in review RE = +204%]
7. Less than 1 out of 10 million [Increase in combined RE]

Note to Ron Cole:

We can safely say that the statistical analysis is in your favour. In addition, all the mean post-course wpm and comprehension scores of the dyslexic group are higher than the pre-course scores of the non-dyslexics (including in the non-dyslexic group the volunteers who included some pretty good readers). This means that this

dyslexic group, on average, is now reading slightly faster with slightly better comprehension than the average reader. Pretty amazing.- Dr. Ross Cooper

