

## **WORDSWORTH LITERACY PROGRAMME - STATISTICAL ANALYSIS (2009)**

### **FOREWORD**

In education, the current methods for teaching literacy (reading and spelling) in mainstream and remedial classes are not suitable for everyone. A high percentage of students in primary and secondary main stream schools are not reaching their true potential. Governments continue to allocate funds for remedial services using the same products and methods that have not worked in the first place; this does not address the root cause.

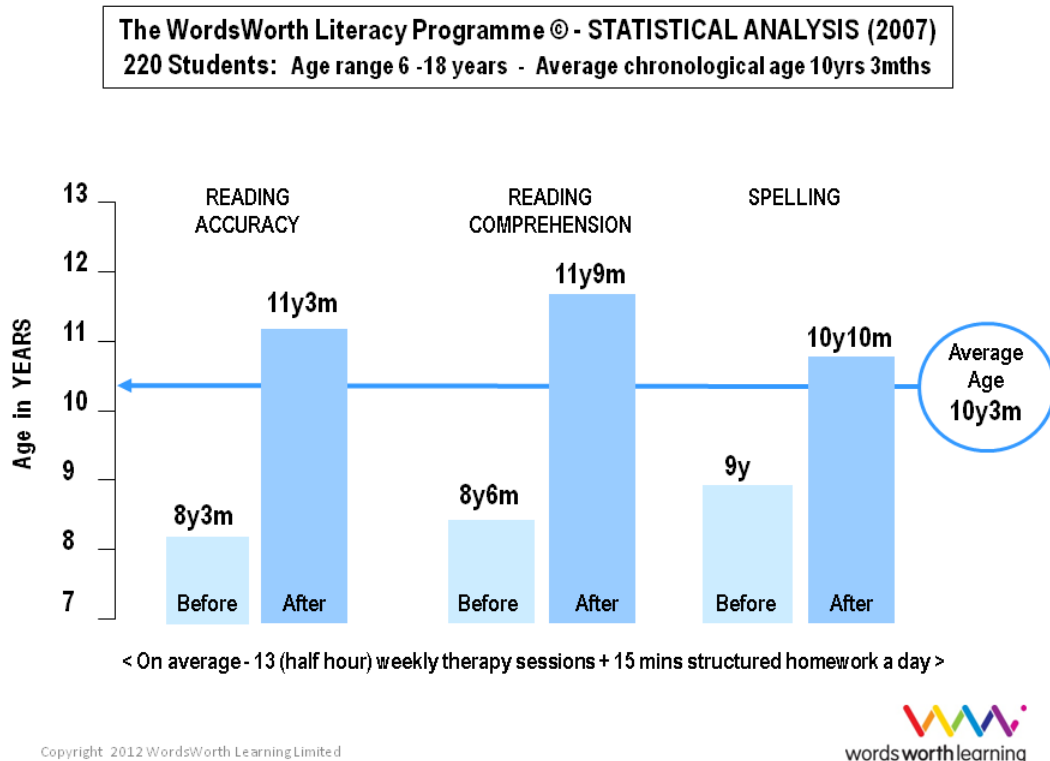
Clients with undiagnosed or diagnosed and untreated “specific learning difficulties” not only have to contend with persistent academic failure but tend also to suffer from concomitant difficulties to varying degrees. Symptomatology ranges from low self-esteem, depression, psychosomatic complaints (generally on school days), to more serious sequelae such as school refusal, conduct disorders, delinquency and more alarmingly suicidal ideation. It is often the presence of some of those difficulties that alert parents and/or professionals to the possibility of other underlying difficulties.

From experience working in Child Psychiatry, one can postulate that if more people who have written language disorders were identified and treated at an early stage or when they were still amenable to intervention, there would be less distress arising from the aforementioned symptomatology, and probably less petty crime and vandalism in our

society. We live in a literate world and to be excluded from that must lead to dissatisfaction and adverse reactions, both personally and socially.

## THE STATISTICAL ANALYSIS

The following statistical analysis to determine the efficacy of the WordsWorth Literacy Programme (WWLP) was professionally conducted by Dr Michael O'Leary, (Senior Lecturer Educational Assessments, St Patrick's College, Drumcondra, Dublin, Ireland).



In all, a total of 217 children took the NARA II (Neale Analysis of Reading Ability) reading accuracy and comprehension tests and 220 took the Vernon Spelling Test before the WWLP intervention. Following the intervention, most of these children were assessed again. However, due to attrition or the fact that students who had performed at the maximum level on the NARA II tests were not included again, the total numbers taking

the assessments at the post intervention stage were 190, 169, and 217 for the reading accuracy, reading comprehension and spelling tests respectively. The data pertaining to the performance of these individuals are presented in Table 1.

**TABLE 1: Pre and post intervention reading and spelling test outcomes**

		Mean Score (months)	SD	Difference Score (months)	SD
NARA II Reading Accuracy (n=190)	<i>Pre</i>	99.4	25.3	+35.9	15.3
	<i>Post</i>	135.3	28.9		
NARA II Reading Comprehension (n = 169)	<i>Pre</i>	101.8	25.2	+39.6	17.5
	<i>Post</i>	141.4	26.9		
Vernon Spelling (n= 217)	<i>Pre</i>	107.9	23.5	+22.1	12.9
	<i>Post</i>	130.0	28.7		

This table provides the average pre and post-test scores (months) for reading accuracy, reading comprehension and spelling as well as a difference score (post test minus pre-test) in each case. It can be seen that, for all three measures, the post-test score was higher. In addition the difference scores ranged from 22 to almost 40 months. Three paired sample t-tests were carried out to determine if the pre and post-test scores were statistically significantly different. In each case, the null hypothesis was rejected:  $t(189) = 32.3, p < .001$  for the reading accuracy score;  $t(168) = 29.4, p < .001$  for reading comprehension and  $t(216) = 25.1, p < .001$  for spelling. From the resultant t-test data, an eta squared statistic was calculated for each test. This statistic allows one to measure the size of a difference between two scores and, in that way, provides complementary data to the outcome of a statistical test. For reading accuracy, reading comprehension and spelling the eta squared values were .85, .84 and .75 respectively. According to Cohen (1988), these values can be interpreted to mean that there was a large effect, with a substantial difference between the pre and post-test scores in each case.

While the differences were large on average, it should also be pointed out that most children involved in the study also made improvements, with some children doing remarkably well. For example, analyses conducted on individual scores revealed that 26 children’s reading accuracy scores improved by 55 months or more (compared to an average improvement of 35 months). One child improved her score by 79 months. Over 35 children improved their reading comprehension score by 55 months or more (compared to the average of 39.6 months) Again one child improved by 95 months. This improvement for individuals is also evident when we examine the difference scores for groups of children at various percentile points. Table 2 provides the difference scores for groups at five percentile points on all three tests. It will be noted that even at the lower end of the scale (e.g. the 10<sup>th</sup> and 25<sup>th</sup> percentiles); gains were made following the intervention.

**TABLE 2: Scores at five points in the difference score distribution**

Percentile	Reading Accuracy (n=190) (months)	Reading Comprehension (n=169) (months)	Spelling (n=217) (months)
10	16	16	6
25	24	27	12
50	33	38	20
75	47	52	31
90	57	64	42

While most of the 220 children participating in the intervention were classed as non-dyslexic, presenting with a ‘specific learning disability’, a total of 24 had been diagnosed as being classically ‘dyslexic’ (i.e. with an identified neurological component to their presenting difficulty). The pre and post-test data pertaining to these children are presented in Table 3.

**TABLE 3: Pre and post intervention reading and spelling test outcomes for participants diagnosed with dyslexia**

		Mean Score (months)	SD	Difference Score (months)	SD
NARA II Reading Accuracy (n = 24)	<i>Pre</i>	89.8	24.9	+31.5	12.4
	<i>Post</i>	121.3	29.1		
NARA II Reading Comprehension (n = 24)	<i>Pre</i>	88.1	24.5	+44.5	19.5
	<i>Post</i>	132.6	28.3		
Vernon Spelling (n = 24)	<i>Pre</i>	94.5	17.0	+19.5	7.5
	<i>Post</i>	114.0	17.5		

While the pre and post-test scores in table three are lower than their counterparts in Table 1, it can be seen that children with dyslexia made substantial gains on the three tests. Indeed, paired sample t-tests indicated that for **all three post-test scores were statistically significantly higher than their pre-test counterparts:**

$t(23) = 12.4, p < .001$ ;  $t(21) = 10.7, p < .001$ ;  $t(23) = 12.7, p < .001$  respectively for reading accuracy, reading comprehension and spelling. Again effects sizes were large at .8 or above.

Further analyses were carried out on the data to determine if there were any gender differences in terms of how boys and girls benefited from the interventions. Relevant data are included in Table 4.

**TABLE 4: Mean differences between post and pre-test scores by gender**

		Difference Mean (SD) (months)	T value(df) (months)	p
NARA II Reading Accuracy	Boys (n =128)	37.4 (15.0)	28.3 (127)	<.001
	Girls (n = 62)	32.7 (15.6)	16.5 (61)	<.001
NARA II Reading Comprehension.	Boys (n =113)	42.1 (18.7)	23.9 (112)	<.001
	Girls (n = 56)	34.6 (13.5)	19.0 (55)	<.001
Vernon Spelling	Boys (n =147)	23.2 (13.3)	21.1 (146)	<.001
	Girls (n = 70)	19.8 (11.7)	14.1 (69)	<.001

In all cases it can be seen that improvements were made by both boys and girls. Again, paired sample t-tests indicated that the difference between each pre and post-test score was statistically significantly different within the genders. It is also apparent from the data in table three that boys have higher improvement scores. The extent to which these scores were statistically significantly was assessed using a one-way analysis of covariance (ANCOVA). In each case, the covariate used was the pre-test score. In other words, the differences in the improvement scores for each gender were evaluated while taking into account and adjusting for the relevant pre-test scores for the students. The outcomes of the three ANCOVAs performed are presented in Table 5.

**TABLE 5: Comparison of male and female difference scores adjusting for pre-test scores**

	Adjusted Mean Difference (Standard errors) (months)		F	p	Partial eta squared
	Male	Female			
NARA II Reading Accuracy	37.4 (1.3)	32.7 (1.9)	4.0	.05	.02
NARA II Reading Comprehension	42.0(1.6)	34.8 (2.2)	7.1	.009	.04
Vernon Spelling	23.3 (1.1)	19.7 (1.5)	3.9	.05	.02

In all three cases it can be seen that the boys' improvement scores on average were statistically significantly higher than the scores for girls. The reason for this is not clear. However, it should be borne in mind that the effect sizes as suggested by the partial eta squared statistics are small (Cohen 1988) and, as a consequence, the differences may be thought of as not being important in practical terms.

A final set of analyses was carried out on the data to examine the extent to which the number of sessions attended impacted on the difference scores of the participants. The children attended between 3 and 21 sessions depending on their needs (i.e. the complexity of their presenting difficulties, whether or not they had reading accuracy, reading comprehension, spelling, associated language deficits or a combination of the aforementioned). The actual number of sessions attended by each participant were organised in five bands from low to high. These categories and the relevant mean scores associated with them are presented in Table 6.

**TABLE 6: Pre and post intervention reading and spelling test outcomes by number of sessions attended**

Sessions	n	Pre-Test Mean (SD) (months)	Post-Test Mean (SD) (months)	Difference Mean (SD) (months)
NARA II Reading Accuracy				
3-6	12	132.2 (18.7)	162.9 (9.9)	30.7 (15.8)
7-10	51	117.8 (21.5)	153.3 (22.9)	35.5 (15.2)
11-12	61	98.9 (19.1)	136.2 (25.4)	37.3 (15.7)
13-14	52	82.6 (15.4)	118.9 (23.6)	36.3 (14.1)
15-21	14	68.8 (13.5)	102.1 (24.5)	33.3 (18.5)
NARA II Reading Comprehension				
3-6	8	125.8 (20.0)	150.0 (21.1)	24.2 (13.0)
7-10	38	116.6 (23.2)	154.2 (23.6)	37.6 (18.2)
11-12	58	105.9 (20.7)	149.5 (21.0)	43.6 (16.9)
13-14	51	89.9 (21.2)	128.9 (26.3)	39.0 (17.7)
15-21	14	73.9 (19.7)	113.6 (25.3)	39.7 (15.0)

Vernon Spelling				
3-6	28	140.8 (16.9)	166.0 (18.8)	25.2 (11.1)
7-10	63	118.1 (16.8)	143.4 (23.5)	25.3 (15.4)
11-12	61	104.2 (17.2)	123.5 (21.8)	19.3 (11.2)
13-14	51	88.9 (13.2)	109.1 (18.9)	20.2 (10.9)
15-21	14	81.6 (11.6)	102.3 (19.9)	20.7 (15.0)

What is clear from the data is that there is an inverse relationship between the number of sessions attended and pre-test scores for all three measures i.e. the more sessions attended the lower the pre-test score. This, of course, was to be expected. However, it can also be seen that for each category of session attendance and for each test, the improvement scores are large (similar, in fact, to those discussed with respect to Table 1 above). Moreover, it is notable that when the mean improvement scores across categories were compared after adjusting for pre-test scores in an ANCOVA, the differences were found to be not statistically significant with respect to the NARA II Reading Accuracy and Vernon Spelling but statistically significantly different with respect to the NARA II Comprehension measure ( $F(4) = 2.8, p < .05$ ). However, the effect size in the latter case, as indicated by the partial eta square statistic, was very small (.06).

Reference: Cohen, J. (1988). *Statistical power and analysis for the behavioural sciences*. Hillsdale, NJ: Erlbaum.

## SUMMARY

The WordsWorth Literacy Programme is highly effective and successful and has been clinically tested over the last fifteen years by a wide range of professionals i.e. Speech & Language Therapists, Teachers, Remedial Teachers. It is now used in working practice both in the education services and in private practices throughout Ireland.