BrainWare SAFARI Pilot at Saint Andrew's Scots School
May-August 2013

Saint Andrew's Scots School in Buenos Aires, Argentina conducted a pilot of BrainWare SAFARI (BWS) with 4th grade students during the 2013 school year. A total of 41 students were selected to be part of the program, with 20 assigned to the treatment group (received BWS as an intervention), and 21 assigned to the no-treatment (did not receive the intervention) matched comparison group. The students in the treatment group started using BWS in May 2013 and completed their work at the end of August 2013. Both the treatment and no-treatment groups were tested with the Diagnostic Achievement Battery – 3rd edition (DAB-3) at the beginning and end of the 14 weeks of the intervention. The test was given to the students in English, rather than their native language.

BWS Usage Information

The students in the treatment group completed a median of 32 ± 2 sessions with a minimum of 27 and a maximum of 34 over 13 ± 2 weeks, yielding an average of 2.5 ± 0.1 sessions per week. During this time they completed 125 ± 19 levels with a minimum of 90 and a maximum of 150. This is an excellent level of fidelity, as the frequency of usage is within the expected 30 to 50 sessions in 10 to 14 weeks. This also represents excellent intensity, which can be determined looking at the number of levels completed when compared to grade-level peers who have used BWS in other studies. Other 4th graders who have used BWS with fidelity completed a median of 111 levels in 36 sessions.

About the DAB-3

The DAB-3 is an individual test of school success designed for ages 6 to 15. It uses 14 brief subtests to identify a child's strengths and weaknesses in five areas: Listening, Speaking, Reading, Writing and Mathematics. The battery is individually administered and takes between 90 and 120 minutes to complete; some of the subtests can be administered in groups. The test results provide a profile of the student, allowing teachers to identify specific areas of strength and weakness. For example, the test might indicate that a child is poor in math reasoning but average in math computation, or below average in listening skills but superior in reading ability. This enables teachers to be more targeted in their instruction and determine what sorts of interventions would be appropriate. The students in this study were native Spanish-speakers, taking the DAB-3 in English.

The reliability and validity of each of the subtests has been established, so a portion of DAB-3 can be administered without hurting reliability. In this study, 8 of the 14 subtests were used, with each skill area represented by at least one subtest, as shown in Table 1. This test was administered in English.

Table 1: Subtests Used in This Study	
Test Area	**Subtests Used**
Listening	Story Comprehension Characteristics
Speaking	Synonyms
Reading	Reading Comprehension
Writing	Writing: Contextual Language Writing: Story Construction
Mathematics	Math Calculation Math Reasoning

The analysis for this report used the percentile rankings on the DAB-3 which provide an indication of how the student is performing compared to other students of the same age on the same test. This means that a student that scores in the 75th percentile, is scoring higher than 75% of others their age on the test and 25% of the students of the same age are scoring higher
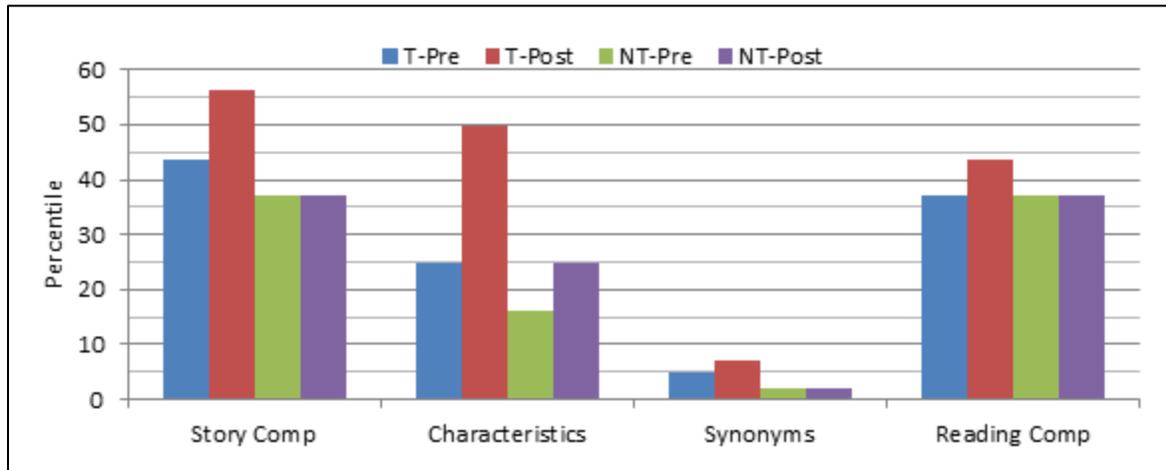
DAB-3 Results

The median percentile for the tests of Listening, Speaking and Reading in the treatment group and the no-treatment are represented in Figure 1.

Figure 1: Listening, Speaking and Reading Results

The Story Comprehension (Listening Area) test requires the student to listen to the examiner read a story and then answer oral questions about the story. The treatment group (T) increased their median percentile score from 43.5 ± 24 percentile points to 56.5 ± 27.6 percentile points on the post-assessment. This is a 13 ± 26 percentile increase for the students. The no-treatment group (NT) had no change in percentile score on this test with a pre- and post- median score of 37 ± 21 and 37 ± 22, respectively. This is an effect size of 0.2 which is statistically significant. In addition, 55% of the students in the treatment group increased their score, compared to 43% for the no-treatment group.

The Characteristics (Listening Area) test requires the student to listen to brief statements being read by the examiner and then respond with whether the statement is true or false. On this test, the treatment group's pre-test median score was 25 ± 26.4 percentile points while their post-test median score increased to 50 ± 29.6 percentile points, a difference is 25 ± 28.4 percentile points. In contrast, the no-treatment group increased 9 ± 30.5 percentile points from a pre-test median score of 16 ± 30 percentile points to post-test median score of 25 ± 30 percentile points. This is an effect size of 0.26 which is statistically significant. There was no significant difference between the groups in the percentage of students increasing their scores or changing their ability level on this test.

The Synonyms (Speaking Area) test requires the student to provide synonyms for words that are read by the examiner. Not surprisingly, given that the students were not taking the test in their native language, this test was a weak area for both the treatment and no-treatment groups, and the use of BWS did not show a significant impact on this particular test, during the time frame of the study. The treatment group increased slightly, while the no-treatment group performed the same between on the pre- and post-assessments.

The Reading Comprehension (Reading Area) test requires that the student read short stories independently and then answer questions presented by the examiner. The students in the treatment group increased their median score 6.5 ± 21.6 percentile points from a pre-test median score of 37 ± 43.5 to a post-test median score of 43.5 ± 21 percentile points. The no-treatment group had no change in median score with a pre-test median score of 37 ± 16.8 and a post-test median score of 37 ± 21.1 percentile points. This is not a statistically significant difference with an effect size of 0.15. However, a clear trend is evident

The median Percentile Scores on the Writing and Mathematics tests for the treatment group and no-treatment groups on pre- and post-assessment are shown in Figure 2.

Figure 2: Writing and Mathematics Results

The Writing: Contextual Language and Writing: Story Construction scores used for this study come from a single test. This test requires the student to write a story in response to three pictures that represent a modified version of a classic fable "The Tortoise and The Hare." The story quality is evaluated according to 14 aspects of contextual language and 11 aspects of story construction, with a score being provided for each element.

There was a dramatic different between the pre-test scores for Writing: Contextual Language for the treatment (median 37 ± 32.1 percentile points) and no-treatment (75 ± 22.7 percentile points) groups. The difference narrowed on post-test because the treatment group's score was unchanged (0 ± 28.7 percentile points), while the no-treatment group's score decreased (-12 ± 26.9 percentile points). This has a statistically significant effect size of 0.2. The treatment group also performed better in terms of the percentage of students staying at the same score, rather than decreasing (Table 3).

On the Writing: Story Construction score, the treatment group improved their median by median 6.5 ± 28.9 percentile points, while the no-treatment group median remained unchanged at 0 ± 29.5 percentile points. The treatment group had a pre-test median score of 50 ± 30.6 and a post-test median of 56.5 ± 26.9 percentile points. The no-treatment group started out 13 percentile points ahead at a pre-test median of 63 ± 30 percentile points. The post-test median for the no-treatment group of 63 ± 29 percentile points allowed the treatment group to be only 6 percentile points behind. There was a notable difference in the percentages changes in scores and ability ranking between the groups. In the treatment group, 40% increased their score where 24% increased their score in the no-treatment group.

Table 3: % Students Writing: Contextual Language			
	Increase	Same	Decrease
T-score	25%	35%	40%
NT -Score	29%	14%	57%

On the Mathematics Reasoning test, the student is presented with mathematical information in the form of pictures for younger students, or oral statements for the older students. The students are required to use the information to solve math problems. It is on this test that the largest difference is noted between the treatment and no-treatment groups. The treatment group's median score increased 19.5 ± 22.3 percentile points, from 37 ± 25.5 on pre-test to 56.5 ± 16.3 percentile points on post-test. The no-treatment group median score remained the same with a 0 ± 21.5 point difference between the pre-test median of 50 ± 20.6 and a post-test median of 50 ± 22.8 percentile points. The effect size of 0.40 is statistically

significant. Additionally, the treatment group had 65% increasing their scores while 38% increased their scores in the treatment group.

The Mathematics Calculation test requires the student to solve 36 written calculation problems. The treatment group's pre-test median score of 25 ± 22.5 percentile points increased 6 ± 23 points to a post-test median score of 31 ± 24.1 percentile points. In comparison, the no-treatment group had a median that was unchanged with a pre-test median of 37 ± 17.4 and post-test median of 37 ± 20 percentile points. While there is an obvious trend on this test, consistent with the other tests administered, the effect size is too small for statistical significance. The treatment group had 45% of the students increasing their score compared with 14% increasing their score in the no-treatment group.

Explanation of Results

The purpose of the pilot was to examine the impact of developing students' cognitive skills on their academic performance levels. BrainWare SAFARI is not a reading, writing or math, intervention; it develops 41 cognitive skills that are necessary for performance of reading, writing and math tasks. The improvements seen in academic performance of the treatment group compared to the no-treatment group would be unlikely to be seen with standard reading, writing and math interventions since those do not address the underlying cognitive processes exercised in BrainWare SAFARI.

Overall Conclusions

Students who received the BrainWare SAFARI intervention as part of the treatment group improved their performance to a greater degree on all areas of the DAB-3 test, compared to the no-treatment group.

- The greatest difference between the results of the two groups was on the Mathematics: Reasoning Test, where the median scores improved for the treatment group and not for the no-treatment group. The difference in performance was significant (a medium effect size of .4). 27% more increased and 33% fewer decreased their scores in the treatment group than the no-treatment group.
- Two other areas demonstrated a significant impact based on the median percentile ranking differences: Characteristics and Writing: Contextual Language. (Effect size of 0.2). On the Characteristics subtest, the treatment group experienced a 25 percentile-point increase, compared to a 9 percentile point increase for the no-treatment group. In Writing: Contextual Language the treatment group stayed the same and the no-treatment group dropped in score, significantly.
- The treatment group's improvements on the two comprehension measures compared to the no-treatment group were also notable. On the Story Comprehension test, the treatment group's median score improved, while the no-treatment group's median score stayed the same. This difference was also significant, although the effect size was small (.19). On this test, 55% compared to 43% improved their scores. Similarly, on the Reading Comprehension test, the treatment group improved their median score, while the no-treatment group's median score remained unchanged.

The overall trend across all areas test showed greater improvement for the treatment group than the no-treatment group. In the areas of Mathematics Reasoning, Characteristics, Writing: Contextual Language and Story Comprehension, the differences were significant, an exciting set of findings given the small sample size. It is interesting to note that the strongest area of improvement and difference between the treatment and no-treatment groups was in Mathematics Reasoning, an area that is arguably less dependent on verbal skills, a consideration for non-native English speakers such as the students in this study.