

Reflex Improves Math Fact Fluency in Elementary School Students

Evidence from a Randomized Control Trial Experiment in Rural and Urban Title I Schools

Study Sample:

- 178 2nd, 3rd, and 4th grade students
- 6 classes from Rosehill Elementary school in Tomball ISD (rural)
- 4 classes from Bruce Elementary school in Houston ISD (large city)
- Sample included 55% non-white, 47% economically disadvantaged, 25% academically at-risk students
- 2016-2017 school year

Research Methodology:

- Classes randomly assigned to treatment group (Reflex usage) or control group (business as usual)
- Control and treatment groups completed a paper and pencil measure of math fact fluency in September 2016 (pretest) and May 2017 (posttest)

Analyses Conducted:

- Repeated samples t-tests to analyze fall-spring fluency growth within Reflex usage students
- Independent samples t-test to analyze difference in fall-spring fluency growth across usage groups
- Partial correlations between usage and assessment growth, controlling for fall achievement

This research brief presents the results of a research study designed to explore the impact of ExploreLearning's Reflex program on math fact fluency for students in Grades 2-4 across two school districts during the 2016-2017 school year. Reflex is an adaptive, online program that helps students develop math fact fluency—the automatic (quick and effortless) recall of basic math facts. It is recommended that students use Reflex at least three days per week, with an individual session taking 15 minutes on average to complete (“earning the green light”).

Methodology

The treatment and control groups were spread across the two Title I schools. Rosehill Elementary in Tomball ISD is located in a rural area with approximately 70% students of colors, 60% Hispanic students, and 57% students on free or reduced-price lunch. Bruce Elementary in Houston ISD is located in a large city with approximately 25% Hispanic students, nearly 100% students of color, and 98% students on free or reduced-price lunch.

The full sample included 178 students in 10 classrooms (two Grade 2, six Grade 3, and two Grade 4), half of which were randomly assigned to the treatment group. The sample included 55% non-white, 47% economically disadvantaged, 25% academically at-risk students. 27 students were removed from the final analyses for missing data (24 students) and/or non-compliance of Reflex usage (7 students).¹ The final sample included 74 Reflex users and 77 control students.

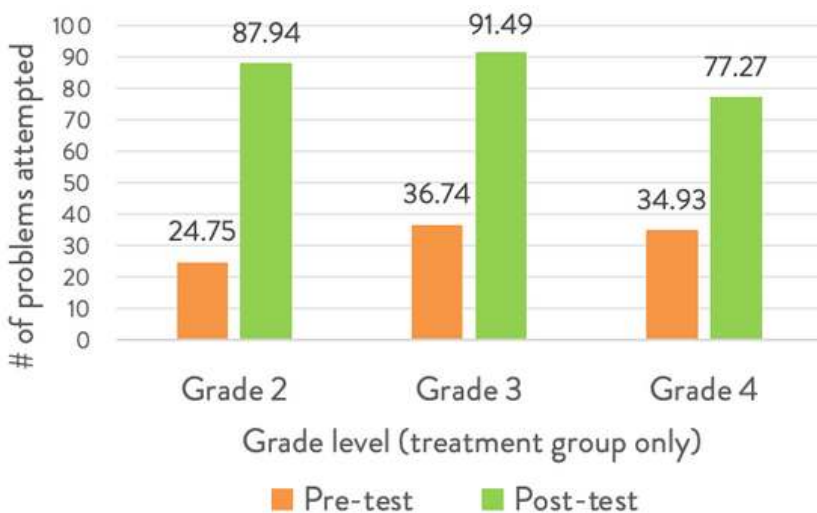
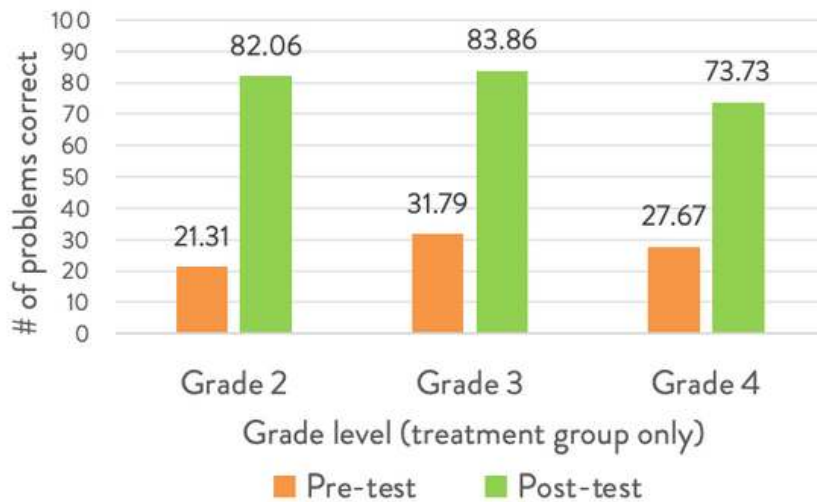
Students in both the treatment and control groups completed a paper and pencil measure of math fact fluency both at pretest (September) and at posttest (May). Students were given 2 minutes to answer as many problems as possible. Student's performance was assessed by three variables: the number of problems answered correctly out of 272 (fluency), the number of problems attempted out of 272 (speed), and the percentage of problems answered correctly out of attempted questions (accuracy).

Results

Outcome 1: Students who used Reflex significantly improved in their math fact fluency.

Students in the treatment group completed an average of 53 days of Reflex, with an average usage across students of 18 minutes per day. Students who used Reflex went from a pretest average of 29 correct problems to a posttest average of 81 correct problems: an improvement of 184%.

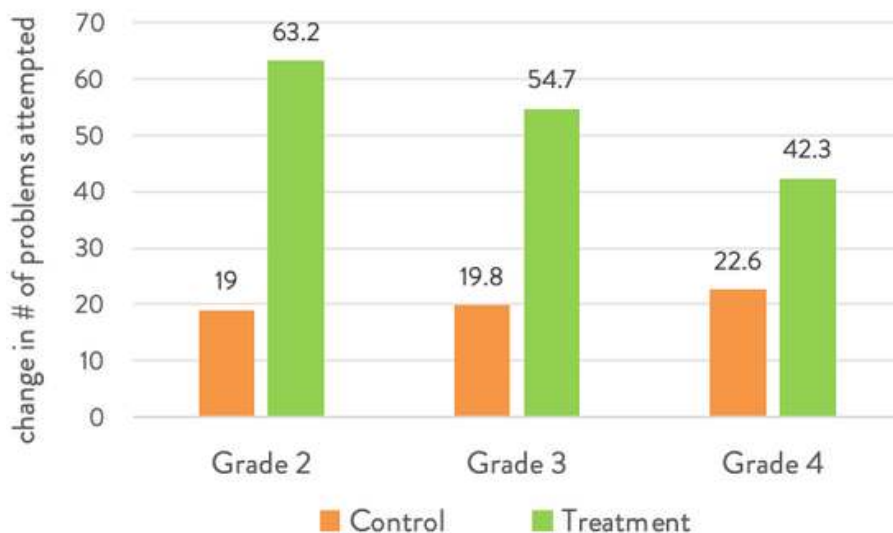
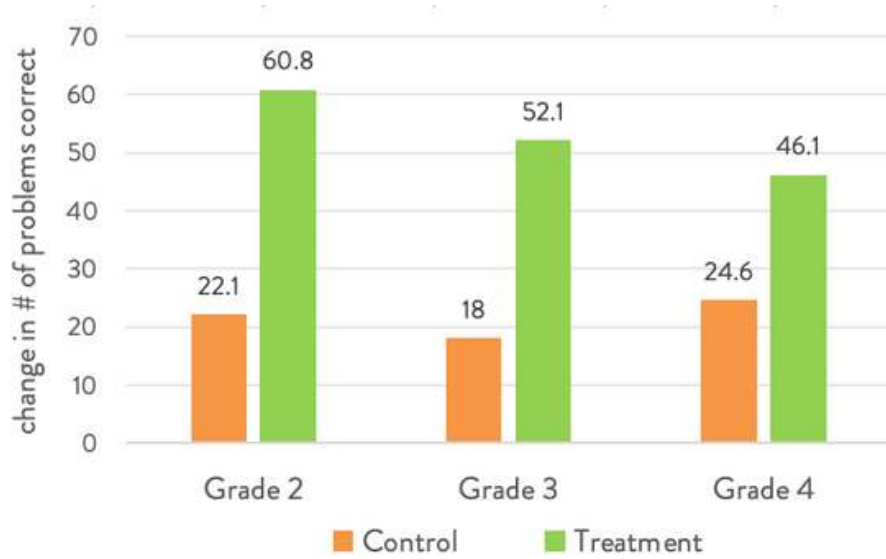
As depicted in the graphs below, all grade levels improved in math fact fluency after using Reflex. Grade 2 students were able to correctly answer, on average, 61 more problems at posttest than at pretest. Grade 3 students were able to correctly answer, on average, 52 more problems at posttest than at pretest. Grade 4 students were able to correctly answer, on average, 46 more problems at posttest than at pretest. For all 3 grades, this change was significant and can be attributed to an increase in both speed and accuracy².



Results (cont'd).

Outcome 2: Students who used Reflex experienced greater improvement in their math fact fluency than students who did not use Reflex

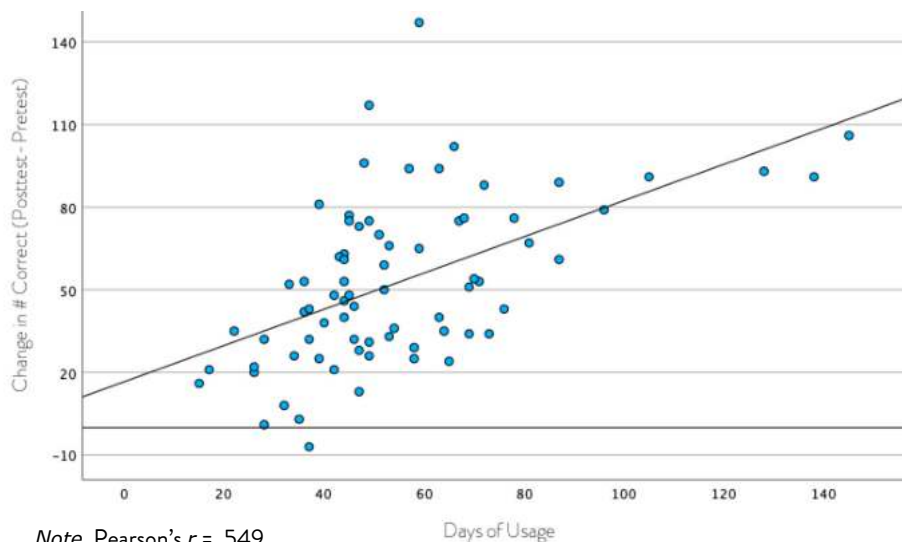
Despite starting out identical at pre-test³, students who used Reflex outperformed control students by 2.6x at posttest. Grade 2 students who used Reflex improved by 285%, while grade 2 students in the control group only improved 113%. Similar results were found for Grade 3 and Grade 4, as show in the graphs below. For all 3 grades, this difference between groups was significant for both change in number correct (fluency) and change in number attempted (speed).⁴



Results (cont'd).

Outcome 3: Students who used Reflex more often experienced greater improvement in their math fact fluency than students who used the program less often

Differences between control and treatment groups can often be caused by confounding variables outside of the educational program being evaluated (e.g. teaching style, individual differences, etc.). To rule out competing hypotheses, correlational analyses were conducted to look for a dosage response. If Reflex was primarily responsible for the change in scores, we would expect to see a significant, positive correlation between days of usage and score increase. When controlling for pretest fluency scores, we found a significant positive relationship between usage and growth with respect to fluency in all three grades.⁵ Treatment group students who used Reflex more often experienced greater improvement in math fact fluency than treatment group students who used the program less often.



Conclusions

Treatment students' math fact fluency significantly increased from pretest to posttest. Additionally, treatment students experienced significantly larger growth in comparison to their peers. Despite starting off the school year with similar math fact fluency scores, students who used Reflex ended the school year with greater math fact fluency and automaticity compared to students who did not use Reflex. Additionally, students who used Reflex more experienced even greater gains in fluency, further demonstrating evidence that the treatment was responsible for the growth observed in these students. Together, this evidence suggests Reflex can be a useful tool for supporting the growth of students in economically disadvantaged schools.

Technical Notes

1. Independent samples t-tests revealed no significant differences in pre-test scores between students who completed both the pretest and posttest and the students who were excluded due to failure to complete the posttest and/or Reflex program non-compliance (all p 's > .05).
2. The data table below presents the changes in outcomes from pretest to posttest for the treatment group only, including significance results from paired samples t-tests.

Outcome	Grade 2			Grade 3			Grade 4		
	Mean diff. (SD)	t	Hedge's g	Mean diff. (SD)	t	Hedge's g	Mean diff. (SD)	t	Hedge's g
Change in # Correct (Fluency)	60.75 (29.32)	8.29***	.87	52.07 (31.23)	10.93***	1.64	46.07 (25.05)	7.12***	1.13
Change in # Attempted (Speed)	63.19 (28.56)	8.85***	.95	54.74 (33.88)	10.60***	1.80	42.33 (28.66)	5.72***	1.35
Change in % Correct (Accuracy)	6.5 (13.68)	1.90*	.47	8.22 (18.99)	2.84**	.51	14.6 (17.52)	3.23**	.51

* $p < .05$, ** $p < .01$, *** $p < .001$

3. Independent samples t-tests were conducted to look at differences in pre-test scores between students in the control condition and the Reflex condition for all three measures (fluency, speed, and accuracy). Within all three grades levels, there were no significant differences between conditions on any of the three variables (all p 's > .05)
4. The data table below presents the differences in posttest and pretest change between the treatment and comparison groups, including significance results from independent samples t-tests.

Outcome	Grade 2			Grade 3			Grade 4		
	Mean diff. (SD)	F	Hedge's g	Mean diff. (SD)	F	Hedge's g	Mean diff. (SD)	F	Hedge's g
Change in # Correct (Fluency)	38.69 (7.82)	24.95***	1.71	34.07 (5.56)	3.88***	1.28	21.50 (8.57)	2.11**	.91
Change in # Attempted (Speed)	44.19 (7.99)	12.82***	1.91	34.98 (6.02)	4.77***	1.24	19.69 (9.78)	1.77*	.73
Change in % Correct (Accuracy)	-6.83 (5.77)	2.38	-4.1	6.11 (3.33)	3.25*	.38	4.37 (6.19)	.64	.26

* $p < .05$, ** $p < .01$, *** $p < .001$

5. A partial correlation was conducted looking at relationships between number of days of Reflex usage and posttest fluency scores (# correct), controlling for pretest fluency scores. The correlation was significant, $r(71) = .549, p < .001$. Looking individually at each grade, all three partial correlations were significant:
 - a. Grade 2: $r(13) = .728, p = .002$
 - b. Grade 3: $r(40) = .430, p = .005$
 - c. Grade 4: $r(12) = .531, p = .05$