

The Xtreme Reading Program

Overview of the Program

The Xtreme Reading Program is a year-long course specifically designed for adolescents in grades 6 through 9 who are reading above the 3.5 grade level and who have deficits in reading. Typically, students selected for the program in each school are two or more years below grade level in reading. The course is taught by a teacher who preferably has some experience teaching reading skills. Class size is 12 to 15 students.

The course is based on the concept of strategic instruction and is founded on the hypothesis that at-risk students can be taught cognitive strategies to improve their learning and compensate for their performance deficits. Indeed, research has shown that low-achieving students and students with high-incidence disabilities do not use such strategies. Fortunately, research with adolescents who do not use cognitive strategies has shown that they can learn and apply complex strategies so that they perform well on a wide variety of academic tasks, including reading textbook chapters, studying for tests, and writing essays (see Schumaker & Deshler, 2006 for a review).

The Xtreme Reading Program is based on this research and focuses on the strategic knowledge and skills that students need to read a wide variety of written materials. Students participating in the program are taught the cognitive and overt processes to use in sequence as they read, along with information about text structure, reading materials, the circumstances under which to use certain strategies, and the metacognitive processes that will guide their strategy use.

The Xtreme Reading Program is founded, in part, on the theory associated with a “cognitive apprenticeship.” An “*apprenticeship*” is a relationship between an expert and a novice in which the expert imparts his/her knowledge and skills to the novice. In a “*cognitive apprenticeship*,” the emphasis is on the expert’s demonstration of cognitive processes as well as overt skills. While the expert initially leads the instruction, the novice gradually assumes more and more responsibility for applying the knowledge, procedural skills, and metacognitive strategies introduced by the expert. In the case of Xtreme Reading Program, the knowledge and skills that students (novice learners) learn are reading strategies. The teacher provides information and demonstrations, and the curriculum materials provide support while the novice learner acquires knowledge and skills related to the mental processes involved in reading.

Further, the instruction in the Xtreme Reading Program is based on a blend of the instructional methodologies that have been found to be most successful with at-risk students: direct instruction and strategic instruction. Successful direct instruction involves breaking a complex skill into parts, teaching each part to mastery, teaching the integration of the parts, providing feedback for correct and incorrect responses, beginning with easy tasks and moving toward more complex tasks, and fading out support as the student becomes more and more skilled. The most successful strategic instruction involves describing and modeling a strategy for a student, providing multiple opportunities for practice, providing detailed feedback after each practice attempt, and requiring mastery. The combination of all these methods specifically found to be successful with at-risk students ensures that students benefit from participation in the Xtreme Reading Program.

The Strategic Curriculum

The curriculum used in the Xtreme Reading Program involves several components. In essence, several reading strategies are taught within a context in which students become part of a learning community where they work toward individual goals while helping other students reach their goals. Students initially are introduced to the course and the

concept of strategic reading, and they learn several skills related to participating in the learning community. They also begin to assess their own strengths and weaknesses, express their hopes and dreams for the future, and set some goals.

Next, they learn some word-level strategies. They learn the Word Mapping Strategy, which enables them to predict the meaning of new words, and they learn the Word Identification Strategy, which enables them to decode the multi-syllable words that they are likely to encounter in secondary course textbooks and other materials.

Once these word-level strategies have been mastered, they learn a group of reading comprehension strategies that can be applied to a variety of reading materials. The Self-Questioning Strategy is used by students to ask questions to themselves and make predictions while reading. The Visual Imagery Strategy is used by students to make movies in their minds of the information being read. The Paraphrasing Strategy is used by students to find the main ideas and details in paragraphs and put that information into their own words as they read. The Inference Strategy is used by students to make inferences about the information they are reading.

While students are learning all the strategies, they are reading novels and other texts together as a group, and they use other specially designed materials to practice using the strategies individually and in pairs. The teacher provides regular and individual feedback to students. Progress monitoring is built into the program so that each student and the teacher know how the student is performing on each strategy daily. Additionally, the students revisit their goals and set new goals for themselves throughout the year.

Results

Results associated with the Xtreme Reading Program have been promising. Thirteen Xtreme Reading teachers recently reported data that they had gathered in their classrooms. These teachers were part of a project in which several federal agencies participated. For the federal report on the project, go to ies.ed.gov/pubsearch/pubsinfo.asp?pubid=NCEE20094036 Since the testing conditions for the students were not optimal during the federal project (e.g., ninth graders were bussed to another school by testers; students who had received reading instruction were tested with students who had not received instruction in a cafeteria setting; teachers with whom they were familiar were not present), some data were also gathered by 13 of the students' Xtreme Reading teachers who administered the Gates McGinite Reading Test at the beginning and end of the school year. A total of 502 students participated in this testing.

The data for three groups of students were analyzed separately.* First, data for the whole set of 502 students were analyzed. Second, because the researchers and school administrators selecting students for the Xtreme Reading Program allowed students who were reading above grade level to participate in the program, the data for the subset of students who were reading below grade level at the beginning of the year were analyzed

* All data sets were analyzed using the General Linear Mixed Model (GLMM) or HLM approach. A three-level model in which test scores (pre and post) were nested within students and students were nested within teachers was employed. SAS PROC MIXED was used to conduct the analyses. Maximum Likelihood estimation procedures were used to determine the fixed effects in the model, and Restricted Maximum Likelihood estimation was used to determine the random effects. The final analysis results are reported for the REML models. Degrees of Freedom were determined using the Kenward-Roger method. The change in raw scores over time (pre and post) was the primary focus of the analysis.

separately. This “below-grade-level group” was comprised of 417 students. Third, because some students earned lower scores on the posttest than on the pretest, and because students cannot “unlearn” how to read, the data for the subset of students who earned posttest scores at the same level as their pretest scores or posttest scores that were above the level of their pretest scores were analyzed. This “test-responder group” was comprised of 348 participants. The results for these three groups of students are reported separately below.

Results for the Whole Group

Table 1 shows each teacher’s results and the number of students that were present in each teacher’s classes during the pretest and posttest. It also shows the mean pretest and posttest raw scores on the Gates McGinitie, the amount of change, the effect size, and the mean grade-equivalent scores. As can be seen in the table, mean raw scores increased between the pretest and posttest for the students of all teachers except Teacher 6. There is, however, some variability in the amount of increase, with the largest increase made by the students of Teacher 10 (an increase of 10.0 raw score points or 2.6 grade levels in reading), and the smallest positive increase (.63 raw score points or .5 grade levels in reading) made by the students of Teacher 8. The mean gain across all the students was 1.2 grade levels in reading comprehension. Because of this variability among the teachers, “teacher” was included as an effect in the model.

The best fitting model included time, teacher, and the interaction between teacher and time. All effects were significant. There was a significant change over time, $F(1,488)=123.9, p < .0001$, indicating that the mean of the posttest scores was significantly different from the mean of the pretest scores. The effect for “Teacher” was significant, $F(12,488)=3.03, p < .0004$, indicating differences in the overall means between the students of the different teachers. Additionally, there was a significant interaction between “Time” and “Teacher,” $F(12, 488) = 6.87, p < .0001$, indicating that the change over time was not the same for all teachers. In other words, some teachers produced greater gains than others. Follow-up tests examined the change over time within teachers using the LSMEANS procedure in PROC MIXED to test for the simple effects. The results from these follow-up procedures indicated that the change from pretest to posttest was significant at the .05 level for all teachers except Teachers 6, 8, 9, and 13. The results of these follow-up analyses are shown in the Table 2 below.

Overall, the correlation between the pretest and posttest scores was .367. Using this correlation between the scores, the effect size for the overall time effect is .41, which, according to Cohen (1988), is somewhat smaller than a medium effect. Since there is a significant interaction effect involving the teachers, the effect sizes for the individual teachers were examined. In fact, the positive effect sizes ranged from .07 (very small, Teacher 8) to 1.05 (a large effect, Teacher 10). These results for the differences in means and effect sizes are also shown in Table 1.

Table 1. Mean pretest and posttest raw scores, change, effect size, and grade-equivalent (GE) scores by teacher--Full sample

Teacher		Pre	Post	Change	Effect Size	Mean Pre GE	Mean Post GE
1	Mean	27.85	30.50	2.65	.28	7.0	8.2
	Std.						
N=40	Deviation	6.735	8.013				
	2	24.45	27.10	2.65	.28	6.3	7.3
N=49	Std.						
	Deviation	7.380	7.959				
3	Mean	28.26	30.62	2.36	.25	7.3	8.2
	Std.						
N=47	Deviation	8.897	9.877				
	4	24.03	27.77	3.74	.39	6.2	7.4
N=39	Std.						
	Deviation	9.247	8.836				
5	Mean	22.66	30.22	7.56	.79	5.9	8.1
	Std.						
N=32	Deviation	10.095	7.374				
	6	26.50	24.50	-2.00	-.21	6.8	6.6
N=28	Std.						
	Deviation	7.406	8.417				
7	Mean	22.98	28.76	5.78	.60	6.0	7.7.
	Std.						
N=58	Deviation	8.090	7.040				
	8	25.50	26.13	0.63	.07	6.5	7.0
N=32	Std.						
	Deviation	8.191	8.323				
9	Mean	24.80	26.73	1.93	.20	6.3	7.2
	Std.						
N=40	Deviation	7.290	7.818				
	10	18.83	28.83	10.00	1.05	5.1	7.7
N=40	Std.						
	Deviation	7.964	8.146				
11	Mean	19.16	28.12	8.96	.94	5.2	7.5
	Std.						
N=25	Deviation	10.242	8.950				
	12	22.25	27.47	5.22	.55	6.1	7.1
N=36	Std.						
	Deviation	8.510	7.796				
13	Mean	29.80	30.94	1.14	.12	7.8	8.3
	Std.						
N=35	Deviation	7.661	7.757				

Total	Mean	24.47	28.39	3.92	.41	6.35	7.55
N=501	Std. Deviation	8.717	8.273				

Table 2. Follow-up Results for Change within Teacher --Full Sample

Teacher	Num DF	Den DF	F Value	Pr > F
1	1	488	4.81	0.029
2	1	488	5.91	0.015
3	1	488	4.49	0.035
4	1	488	9.36	0.002
5	1	488	31.34	<.0001
6	1	488	1.92	0.167
7	1	488	33.13	<.0001
8	1	488	0.21	0.644
9	1	488	2.54	0.112
10	1	488	68.5	<.0001
11	1	488	34.37	<.0001
12	1	488	16.81	<.0001
13	1	488	0.78	0.377

Results for the Below-grade-level Group

The results for the reduced sample of students who scored below grade level at the pretest are similar to those for the full sample. (See Table 3.) The raw-score means increased between the pretest and the posttest for the students of all teachers except Teacher 6. The largest increase was exhibited by students of Teacher 10 (an increase of 11.29 raw score points), and the smallest positive increase (1.87 raw score points) was exhibited by students of Teacher 13. The best model for the data includes the same three fixed effects: Time, Teacher, and the interaction of Time and Teacher. All effects were statistically significant: Time, $F(1,404) = 137.02, p < .0001$; Teacher, $F(12,603) = 6.28, p < .0001$; and the interaction between Time and Teacher, $F(12, 404) = 5.76, p < .0001$. Follow-up tests for simple effects examined the change over time within teachers, again using the LSMEANS procedure in PROC MIXED. These results are in Table 4 below.

The effect size for the overall time effect is .58, which according to Cohen's (1988) guidelines is slightly larger than a medium effect. Since there is a significant interaction effect involving the teachers, the effect sizes for the individual teachers were examined. The positive effect sizes ranged from .23 (small, Teacher 13) to 1.36 (a very large effect, Teacher 10).

Table 3. Mean Pretest and Posttest Raw Scores, Change, Effect Sizes, and Grade-equivalent Scores by Teacher—Sample Below Grade Level

Teacher		Pretest	Posttest	Change	EF	Mean Pre GE	Mean Post GE
1	Mean	25.76	28.82	3.06	.37	6.5	7.7
N=33	Std. Deviation	5.292	7.568				
2	Mean	22.77	26.26	3.49	.42	5.9	7.1
N=43	Std. Deviation	6.148	7.544				
3	Mean	24.03	27.39	3.36	.40	6.2	7.4
N=33	Std. Deviation	6.748	9.124				
4	Mean	21.3	25.52	4.22	.51	5.7	6.9
N=33	Std. Deviation	7.17	7.571				
5	Mean	20.04	28.3	8.26	1.00	5.4	7.6
N=27	Std. Deviation	8.596	6.268				
6	Mean	23.25	21.3	-1.95	-.23	6.0	6.0
N=20	Std. Deviation	5.628	7.226				
7	Mean	21.42	27.58	6.16	.74	5.7	7.4
N=52	Std. Deviation	6.971	6.424				
8	Mean	22.73	24.73	2.00	.24	5.9	6.7
N=26	Std. Deviation	6.258	7.877				
9	Mean	24.03	26.71	2.68	.32	6.2	7.2
N=39	Std. Deviation	6.714	7.526				
10	Mean	16.54	27.83	11.29	1.36	4.6	7.5
N=35	Std. Deviation	5.447	8.024				
11	Mean	17.65	27.13	9.48	1.14	4.8	7.3
N=23	Std. Deviation	9.193	8.636				
12	Mean	19.94	26.16	6.22	.75	5.7	8.1
N=31	Std. Deviation	6.668	7.59				
13	Mean	25.45	27.32	1.87	.23	6.5	7.4
N=22	Std. Deviation	6.17	6.771				
Total	Mean	21.91	26.73	4.82	.58	5.78	7.25
N=417	Std. Deviation	7.125	7.618				

Table 4. Follow-up Results for Change within Teacher -- Sample 2

Teacher	Num DF	Den DF	F Value	Pr > F
1	1	416	5.26	0.022
2	1	416	8.91	0.003
3	1	416	6.36	0.012
4	1	416	9.97	0.002
5	1	416	31.36	<.0001
6	1	416	1.29	0.256
7	1	416	33.53	<.0001
8	1	416	1.77	0.184
9	1	416	4.59	0.033
10	1	416	75.91	<.0001
11	1	416	35.18	<.0001
12	1	416	20.46	<.0001
13	1	416	1.30	0.255

Results for the Test-responder Group

The results for this reduced sample are generally better than for the previous two samples. (See Table 5.) The mean raw scores earned by students of all teachers, including the students of Teacher 6. Once again, there is sufficient variability among the teachers such that the best-fitting model includes effects for Time, Teacher, and the interaction between Teacher and Time. There was a significant change over Time, $F(1,335) = 341.06, p < .0001$. The effect for Teacher was also significant, $F(12,335) = 2.27, p < .009$, indicating differences in the overall means between the teachers. Additionally, there was an interaction between Time and Teacher, $F(12, 335) = 4.29, p < .0001$, indicating that the change over time was not the same for all teachers. Once again, follow-up tests were conducted to explore the simple effects of change over time within teachers. These results are reported in Table 6.

The effect size for the overall time effect is .91, a large effect. The effect sizes for the individual teachers ranged from .45 (slightly smaller than medium) to 1.50, a very large effect. These results are reported in Table 5 below along with the mean raw scores, mean change from pretest to posttest, and mean grade-equivalent scores.

Table 5. Means Pretest and Posttest Raw Scores, Change, Effect Sizes, and Grade-equivalent Scores by Teacher—Sample 3

Teacher		Pretest	Posttest	Change	Effect Size	Mean Pre GE	Mean Post GE
1 N=28	Mean Std. Deviation	25.5 5.634	30.71 6.341	5.21	0.66	6.5	8.2
2 N=37	Mean Std. Deviation	22.14 6.001	27 7.307	4.86	0.62	5.8	7.3
3 N=27	Mean Std. Deviation	23.96 6.694	29.78 7.668	5.82	0.74	5.7	7.9
4 N=28	Mean Std. Deviation	20.79 6.898	27 6.842	6.21	0.79	5.5	7.3
5 N=24	Mean Std. Deviation	19.25 8.232	29.21 5.209	9.96	1.27	5.2	7.9
6 N=10	Mean Std. Deviation	20.9 4.202	25.6 4.195	4.7	0.60	5.6	6.9
7 N=45	Mean Std. Deviation	20.2 6.43	28.18 6.503	7.98	1.02	5.4	7.6
8 N=21	Mean Std. Deviation	22.29 5.658	25.81 7.291	3.52	0.45	5.9	6.9
9 N=30	Mean Std. Deviation	22.93 6.81	28.27 7.638	5.34	0.68	6.0	7.6
10 N=34	Mean Std. Deviation	16.65 5.493	28.38 7.435	11.73	1.50	4.7	7.6
11 N=22	Mean Std. Deviation	17.09 8.997	27.27 8.811	10.18	1.30	4.7	7.4
12 N=25	Mean Std. Deviation	19.04 6.705	28.32 5.786	9.28	1.18	5.5	7.6
13 N=17	Mean Std. Deviation	24.76 6.741	29.29 6.273	4.53	0.58	6.3	7.9
Total N=348	Mean Std.	21.08 7.007	28.18 6.932	7.1	0.91	5.75	7.55

Deviation						
-----------	--	--	--	--	--	--

Table 6. Follow-up Results for Change within Teacher -- Sample 3

Teacher	Num DF	Den DF	F Value	Pr > F
1	1	335	18.03	<.0001
2	1	335	20.74	<.0001
3	1	335	21.62	<.0001
4	1	335	25.60	<.0001
5	1	335	56.36	<.0001
6	1	335	5.23	0.023
7	1	335	67.82	<.0001
8	1	335	6.17	0.013
9	1	335	20.21	<.0001
10	1	335	110.88	<.0001
11	1	335	54.01	<.0001
12	1	335	50.98	<.0001
13	1	335	8.26	0.004

Summary

These results indicate that the Xtreme Reading Program can be successfully used with average (those scoring at grade level) and at-risk adolescent readers to boost their reading achievement. An important finding is that different teachers produce different results with this group of students. In fact, some teachers produce an average gain of close to 3 grade levels (Teacher 10), while others produce more modest gains of 1.3 grade levels, as indicated by the results with the group of students who responded seriously to the test. Such differences are likely to relate to whether the teachers implemented the program with fidelity and conditions surrounding the program (i.e., when the program was started during the school year, the proportion of students reading above grade level, the prior training and status of the teacher, the number of students moving in and out of the classes, interruptions of the course, etc.).