The Effect of Elementary Grade Retention on Subsequent School Achievement and Ability*

Marilyn Westbury
concordia college

This study compared the achievement and ability of a group of retained elementary students with that of a group of continually promoted elementary school students. The two samples were matched on the basis of gender, school entry reading achievement, grade, and regular school program. Short-term and long-term comparisons between the groups were made using Grade 1 Edmonton Public Schools' Reading Comprehension Test scores, Grade 3 and Grade 6 Canadian Cognitive Ability Test (CCAT) scores, and Grade 6 subject matter achievement tests. The author concluded that grade retention was ineffective for improving achievement and ability.

Cet article traite d'une étude qui comparait le rendement et les aptitudes d'un groupe de doubleurs au primaire avec celles d'un groupe d'élèves continuellement promus d'une classe à l'autre, également au primaire. Les deux échantillons ont été appariés en tenant compte du sexe, de l'aptitude à la lecture à l'entrée à l'école et de la classe; tous les élèves étaient inscrits dans le programme scolaire ordinaire. Des comparaisons à court terme et à long terme ont été effectuées entre les deux groupes à l'aide du *Grade 1 Edmonton Public School's Comprehension Test, du Grade 3 and Grade 6 Cognitive Ability Test (CCAT)* et de divers tests de rendement ayant aux matières de la 6e année. L'auteure conclut que le redoublement n'améliore ni le rendement de l'élève ni ses aptitudes.

Grade retention, which schoolchildren often term flunking or failing, is an old and persistent practice in North American schools. Retention means that a child who has spent a full school year in a particular grade must repeat the entire grade just completed. In the 1990s, grade retention remains widespread, contrary to public perception that very few students fail a grade. Considering how prevalent the practice is, it is surprising that school boards and government agencies do not consistently collect promotion and retention rates. The Canadian Education Association's (1989) survey report on various retention and promotion practices employed by Canadian school boards offers no hard facts on the frequency of the practice. Other research suggests retention rates have been climbing as teachers come under increasing pressure to ensure that students have mastered the curriculum (Johnston, Markle, & Mims, 1985). Indeed, crusaders

*The research discussed in this article was funded by Alberta Advisory Committee for Educational Studies. I also thank Dale Armstrong of Edmonton Public Schools for his assistance.

for "educational excellence" and higher educational standards often refer to one obvious redress for slack standards — nonpromotion of students who have not mastered the grade level.

In the current political atmosphere, it is imperative to examine whether retention produces the desired educational excellence. One assumption behind grade retention is that holding a student in the same grade for an additional year will improve the student's achievement or grades. The research in the area, however, does not support this assumption. Retained children appear to score, at best, no better than comparison groups of continually promoted children. C. Thomas Holmes (Holmes & Matthews, 1984) updates his earlier meta-analysis on grade level retention effects by including 19 new controlled studies done in the interim (Holmes, 1989). Sixty-three studies are integrated using the sophisticated meta-analysis approach to average actual differences between experimental and control groups across studies. Holmes' findings are consistent with his earlier study and with other reviews of the research (Jackson, 1975; Rose, Medway, Cantrell, & Marus, 1983), which find largely negative effects for retention. When compared to socially promoted pupils, retained students show both poorer academic results and inferior personal adjustment. Only 14% of studies on grade retention show positive gains for the retained over the promoted group. However, certain research conditions appear to increase the likelihood of demonstrating beneficial effects of retention. The positive studies often: (a) confined the follow-up period to one year; (b) used only academic outcomes; (c) compared grade peers and not age peers; (d) gave intensive remediation to the retained students without an equivalent remediation to the promoted students; and (e) gathered achievement measures from competency-based tests (Shepherd & Smith, 1989, p. 28). In other words, it was difficult to state conclusively that the retention was the single, probable cause of the gain in achievement.

Rose et al. (1983) are critical of the general failure of the research to look at the long-term effects of retention. In one exception, Abidin, Golladay, and Howerton (1971) compare sixth-graders who were retained with a random sample of regularly promoted children who scored below the 25th percentile on the Metropolitan Readiness Test in the first grade. Although the ability of the retained group was significantly higher at the beginning of the first grade, in the fourth through sixth grades the retained group scored lower than the promoted group on the same test. More longitudinal research spanning the total school experience of children would fill a gap in the research. Some studies suggest that occasionally there are short-term gains for students repeating a grade but that the gains diminish in the long run. Moreover, grade repeaters are more likely to drop out of school than continually promoted students (Lloyd, 1978). These long-term repercussions should be taken into consideration when retention decisions are being considered in the early grades. Of course, all of the studies will only be valuable if they become the basis on which subsequent educational decisions are made.

The prime objective of the research I report in this article is to ascertain whether lower achieving students who were retained once in elementary school perform as well on subsequent achievement and ability tests as a matched group of students who were continually promoted. The study addresses the question: What are the short-term and long-term effects of grade retention on student achievement and ability?

LIMITATIONS

It is essential to understand the limitations and constraints under which field research on schoolchildren is conducted. First, experimental research on school populations is considered unethical. That is, one cannot arbitrarily or randomly hold children in the same grade for a second year, or perform any extraordinary experimental procedure which has the potential to negatively affect a child. Second, cohort studies in which a group of matched children are selected and followed for a number of years to determine the relative progress of each have one major drawback. Modern student mobility is so high that the majority of the sample students may not be located after a four- or five-year interval. Third, there are unique problems in conducting an historical study such as the one I report here. The only reliable source of data on the child's educational history is the student record card. However, not all variables are recorded on every student's card. School principals and teachers can be more or less attentive to record keeping. Moreover, students may have been absent from school for one or more of the recorded tests. One must expect at least one variable to be missing on most of the students' record cards. Despite these limitations, much valuable information may be gained from school-based research on pupil populations.

METHOD

My study took place in Edmonton Public Schools, a large metropolitan, western Canadian school district with an enrollment of approximately 79,000 students. The sample students were selected in the following manner. A school district computer list identified all students (N=329) in a regular program who wrote Grade 6 achievement and ability tests in June 1989, but were in their seventh year of schooling. A random sample (n=135), or 41% of the total population, was selected from this group. The sample was geographically representative of one-time grade repeaters. A control group (n=103) of continually promoted students who also wrote Grade 6 achievement and ability tests in June 1989 was selected from the same schools distributed throughout the district. The two groups were matched on the basis of school entry-level achievement test scores, gender, grade level, and regular, as opposed to special, school program. The first important control, entry-level achievement test scores, reduced the final subgroup

sizes. The samples had students who had written different elementary school entry tests although they fell within the same range of scores. It was decided to include only those students who wrote the *same* test, Edmonton Public Schools Grade 1 Reading Comprehension Test. Early reading comprehension scores are often considered a good predictor of how students will perform in elementary school. This control reduced the total number to 125 retained and 84 continually promoted pupils who wrote the same pre-retention test.

Gender is a second important control variable in matching retained and promoted students. Most studies, this one included, indicate that males are far more likely than females to repeat an elementary school grade, with the gender difference persisting when achievement is controlled. This gender bias was described by Shepard and Smith (1989, pp. 180, 227) as the result of teachers' "invalid" beliefs about children's physiological readiness for schooling which led to decisions to retain more of the slower maturing males. Table 1 indicates that 60% of retained students were males. The control group was matched on the basis of the same 60:40 male to female ratio.

Attendance in a regular school program and only one grade retention are also controls. Excluded from this study are all those extraordinary students retained more than once in elementary school and/or students in such special programs as English-as-a-Second-Language (ESL), bilingual, immersion, opportunity, handicapped, impaired, disabled, or behaviour-disordered. Finally, 85% of the retainees repeated grades 1, 2, or 3, the early primary grades.

To summarize, the methodology is designed to isolate two groups of students who are as alike as possible on all the recorded variables with one exception—one group repeated one grade in elementary school and the other did not. Second, the subjects are regular-program, English-speaking schoolchildren without any apparent mental, physical, or social handicaps. Both of these design parameters are important to increase the probability that differences or similarities between the two groups are attributable to failing or passing a grade.

TABLE 1

Retained and Promoted Students by Gender

	Retained			*Promoted*		
	Male	*Female*	*Totals*	*Male*	*Female*	*Totals*
Number	75	50	125	51	33	84
Percent	60	40	100	60.7	39.3	100

Data were collected from student record cards in various district schools. The two group means on Edmonton Public Schools Grade 1 Reading Comprehension Test were subsequently compared. The two group means on Canadian Cognitive Ability Text (CCAT) Scores in grades 3 and 6 were also compared. Similarly, Grade 6 school district Science, Mathematics, Social Studies, and Language Arts raw achievement scores of retained students were compared with those of continually promoted students.

RESULTS

Table 2 shows the entry-level scores of both groups on the same Grade 1 preretention test. In total, 209 students, 125 retained students and 84 continually promoted pupils, had Edmonton Public Schools Grade 1 Reading Comprehension Test percentiles on their student record cards. The average percentile of 30.0 for the retained group is very similar to the control group average percentile of 30.6; there is no significant difference in reading comprehension scores between the two groups at school entry. This finding is expected, given that the retained and promoted students were matched on the basis of the school entry-level achievement test scores.

The results of the standardized test measures of ability are shown in Table 3. The data demonstrate that the retained group's verbal and quantitative scores were significantly below those in the promoted group on the Canadian Cognitive Ability Tests administered in Grade 3. The verbal ability of the promoted group averaged more than seven points higher than that of the retained group, despite the fact that average reading comprehension scores for both groups were in the 30th percentile at entry level. It is interesting to note that there was no significant difference in nonverbal CCAT scores between the Grade 3 children who had

TABLE 2

Mean Grade 1 Reading Comprehension Percentiles for Retained and Promoted Students

	Retained		*Promoted*		
	N[a]	$\overline{X}$	*N*[a]	$\overline{X}$	*Significance*
Grade 1 Edmonton Public Schools Reading Comprehension Test	125	30.0	84	30.6	.831*

*No significance.

TABLE 3

Mean Ability Scores for Retained and Promoted Students

	Retained		*Promoted*		
	N[a]	$\overline{X}$	*N*[a]	$\overline{X}$	*Signficance*
Grade 3 Canadian Cognitive Ability Test (CCAT)					
Verbal	109	93.8	75	101.1	.000*
Quantitative	110	92.1	76	97.5	.004*
Nonverbal	110	94.7	76	96.7	.319**
Grade 6 Canadian Cognitive Ability Test (CCAT)					
Verbal	112	94.4	82	95.5	.492**
Quantitative	112	96.0	81	97.9	.230**
Nonverbal	112	100.3	81	96.7	.069**

[a]Variation in subject Ns occurs because not all sample students were present for all tests.
*$p<.005$
**no significance

been retained and those who had been continually promoted. The Grade 3 retained and promoted students were fairly equally matched in terms of their spatio-visual ability scores. Also of interest is the fact that, on average, the retained students demonstrate slightly stronger nonverbal skills than verbal skills, whereas the converse is true for continually promoted students.

By Grade 6, there were no significant differences in verbal, quantitative, and nonverbal ability between the promoted and retained students. Mean verbal score of the promoted group dropped significantly between grades 3 and 6 whereas nonverbal scores of the retained group increased.

Table 4 shows the raw scores on Grade 6 subject matter, school district achievement tests in Science, Mathematics, Social Studies, and Language Arts. There were no significant differences between the retained and promoted groups in subject area mastery by the end of Grade 6. Having started their elementary schooling with below-standard performances, the students in both groups were still achieving lower than school district means in all subject areas.

TABLE 4

Mean Achievement Raw Scores for Grade 6 Retained and Promoted Students

Edmonton Public School District Tests	*School District Mean* $\overline{X}$	*Retained*		*Promoted*		*Significance*
		N	$\overline{X}$	*N*	$\overline{X}$	
Science	(32.08)	125	28.2	81	26.2	.057*
Mathematics	(30.04)	121	26.6	78	25.1	.176*
Social Studies	(47.44)	122	43.6	78	42.0	.281*
Language Arts	(54.82)	120	50.5	82	48.7	.212*

* No significance

DISCUSSION

The results of this study are consistent with the meta-analysis of 63 studies on grade level retention effects done by C. Thomas Holmes. Only 14% of those studies show significant positive gains for the retained over the promoted group (Shepard & Smith, 1989, p. 27). In the study described here the retained students show neither positive academic gains nor negative academic losses when compared to the promoted group on ability and subject matter achievement.

Two sets of outcome measures were employed in this study—ability and achievement. Schools are mainly concerned with achievement. Theoretically, more schooling should produce higher achievement in subject areas generally considered to be the business of schooling. On the other hand, the CCAT is a test of general cognitive skill, a standardized measure resembling an I.Q. test. Such tests are believed to measure innate ability, a relatively fixed capacity for learning.[1] Ability measures the potential for learning whereas achievement measures stored knowledge.

One obvious question arising from the ability data is: Why did the Grade 3 CCAT significant differences between the two groups become nonsignificant by Grade 6? No clear explanations emerge from the data. Age would have no bearing (the fact that the retainees were 11 months older) because the CCAT is expressed in standard age scores. Similarly, an extra year of schooling may increase achievement but not a so-called fixed or innate capacity for learning. It may well be, as many psychologists believe, that measures of intelligence or

ability can vary in a given student. One study revealed similar ability discrepancies over time in another group of elementary school students. Students placed in an academic challenge program for gifted students on the basis of early elementary (grades 1 to 3) ability scores were often excluded from the gifted student category when retested in grades 7 to 9 (Stelmaschuk, 1986). In other words, it is plausible that the Grade 3 CCAT verbal scores in this sample are also inflated and that the Grade 6 tests would prove more reliable. It is also worth noting that it was a drop in the verbal scores of the promoted group rather than an increase in the retained group that closed the gap between the two groups, again suggesting inflation and instability in verbal scores in early elementary. Verbal scores tend to be better predictors of school achievement than nonverbal or quantitative scores.

Achievement is the second outcome measure employed in this study. The achievement of the retained students was not significantly greater than that of the continuously promoted students. Nevertheless, there are several reasons why it is logical to *expect* significant gains in achievement of retained over promoted students given the equal achievement of both groups on Grade 1 tests. First, the retained students had the benefit of an additional year of schooling to increase stored knowledge and subsequent achievement in subject matter tests. Second, the retained students were approximately 11 months older when they wrote Grade 6 achievement tests and developmental psychologists would argue that at least some children had reached a higher developmental stage necessary for increased achievement (Smith & Shepard, 1987). Third, this study isolated the "best" retainees by excluding students with diagnosed problems who were in special programs or who repeated a grade more than once in elementary school. Some students who continue to experience difficulty in mastering curriculum would likely become second repeaters. One could hypothesize that the retained students' mean scores on the various outcome variables would drop with the addition of students who had learning problems or who were multiple repeaters. Fourth, the expressed reason for retention is to help children achieve academically, and given the historical persistence of the practice one would expect the desired outcome. Otherwise, educational recycling is a costly and ineffective intervention that should not have persisted.

Finally, this study concentrates only on academic outcomes. Student record cards contain little or no information about the social and emotional impact of grade repetition on students. Therefore, we can only speculate from other studies, which found social and emotional costs for young students who were required to repeat a grade while their peer group moved on. For example, when Byrnes and Yamamoto (1985) interviewed early elementary pupils, the majority described the experience of repeating a grade as sad or upsetting. It seems plausible that the emotional strain attached to repeating a grade would lower self-esteem in repeaters.

CONCLUSION

Although this and similar studies indicate that students either show declines in achievement or perform no better after repeating a grade than continuously promoted students, the conclusion that *no* students improve is, nevertheless, incorrect. Indeed, a small minority of the repeaters did significantly improve achievement scores after retention. The problem is that there is no foolproof way to predict which students are likely to benefit from the practice or if they would have shown the same improvement without retention. In the absence of such indicators, it is wise to establish practices that benefit the majority.

Many educators defend retention on the basis that continually passing students only pushes the achievement problem further down the road so that poor students arrive at high school totally unprepared to do the work. However, this study suggests that grade repetition does not correct the original learning problem. Therefore, failing a student does nothing to improve high school readiness. Educators must seek alternatives to grade repetition that correct learning problems early and hold students through high school graduation. Some recent remediation programs that focus on individual tutoring for lower achievers while keeping students with their grade peers have proven successful (D. Armstrong, personal communication, 17 March 1992). My findings suggest these alternative measures should be explored and tested.

NOTE

[1] There are numerous ongoing debates among psychologists regarding the measurement of I.Q. or ability. Many believe that most tests do not measure the full range of intelligence while others question whether "innate" ability can ever be accurately measured independent of environmental influences.

REFERENCES

Abidin, R. R., Golladay, W. M., & Howerton, A. (1971). Elementary school retention: An unjustifiable, discriminatory and noxious educational policy. *Journal of School Psychology, 9,* 410–414.

Byrnes, D., & Yamamoto, K. (1985). Academic retention: An inside look. *Education, 106,* 208–214.

Canadian Education Association. (1989). *Grade promotion and retention: Practices in Canadian school boards* (CEA Information Note). Toronto, ON: CEA.

Holmes, C. T. (1989). Grade level retention effects: A meta-analysis of research studies. In L. A. Shepard & M. L. Smith (Eds.), *Flunking grades: Research and policies on retention* (pp. 16–33). London, UK: The Falmer Press.

Holmes, C. T., & Matthews, K. M. (1984). The effects of nonpromotion on elementary and junior high school pupils: A meta-analysis. *Review of Educational Research, 54,* 225–236.

Jackson, G. (1975). The research evidence on the effects of grade retention. *Review of Educational Research, 45,* 613–635.

Johnston, J. H., Markle, G. C., & Mims, E. (1985). What research says about grade retention. *Middle School Journal, 17,* 8–10.

Lloyd, D. N. (1978). Prediction of school failure from third-grade data. *Educational and Psychological Measurement, 38,* 1193–1200.

Rose, J. S., Medway, F. J., Cantrell, V. I., & Marus, S. H. (1983). A fresh look at the retention-promotion controversy. *Journal of School Psychology, 21,* 201–211.

Shepard, L. A., & Smith, M. L. (Eds.). (1989). *Flunking grades: Research and policies on retention.* London, UK: The Falmer Press.

Smith, M. L., & Shepard, L. A. (1987). What doesn't work: Explaining policies of retention in the early grades. *Phi Delta Kappan, 69,* 129–132.

Stelmaschuk, M. A. (1986). *Identified and unidentified gifted: A comparative analysis.* Unpublished doctoral dissertation, University of Alberta, Edmonton, Alberta.

Marilyn Westbury is Dean of Continuing Education, Concordia College, 7128 Ada Boulevard, Edmonton, Alberta, T5B 4E4.