1. Increased Student Self-Esteem as a Non-Targeted Outcome of an Early Literacy Intervention

2. Barbara Hummel-Rossi PhD and Jane Ashdown PhD

3. Department of Applied Psychology, Department of Teaching and Learning

New York University — New York University

4. Barbara Hummel-Rossi
53 Cherry Lane
Huntington, NY 11743

Jane Ashdown
Ansonia Court, Apt. P2L
420 12th St.
Brooklyn, NY 11215

[home addresses as Barbara on sabbatical & Jane's office moving]

5. barbara.hummel.rossi@nyu.edu jane.ashdown@nyu.edu

6.Abstract.

Choosing among educational programs is difficult. Often programs have outcomes that are not measured by educational evaluations. In a cost-effectiveness analysis of an early literacy program the authors discovered that the program might not only be increasing student reading scores, but also positively affecting their self-esteem. In a follow-up study of 61 first graders and their 5 teachers they examined developing self-esteem and reading achievement over the course of a school year in three groups of children with possible literacy-learning risks: "high risk", "at risk", and "no risk". Teacher evaluations indicated substantial increases in developing self-esteem as well as in reading achievement in the "high risk" and "at risk" groups.

Increased Student Self-Esteem as a Non-Targeted Outcome of an Early Literacy Intervention

Barbara Hummel-Rossi and Jane Ashdown

New York University

The decision as to which education or social programs to spend precious dollars on is often a difficult one. Some programs are very effective, but their costs appear to make them prohibitive. At the same time, program evaluations often focus on a limited range of outcome variables. Access to a complete picture of a program's outcomes, including unanticipated outcomes, would provide a stronger basis for decision-making, particularly when financial resources are limited. In investigating the role of cost-effectiveness analysis in educational decision-making (Hummel-Rossi & Ashdown, 2002), we found that secondary program outcomes such as student satisfaction (Levin, 1988; Levin, 1995; Levin & McEwan, 2001), improvements in self-esteem, or good citizenship (Rice, 1997), which are more difficult to measure than achievement, received little attention in cost and outcomes analyses.

Research Objectives

Specifically this research addresses the unanticipated and unknown outcomes of a literacy intervention targeted at low-performing first grade students. Besides improving literacy achievement in the targeted students, the intervention also appeared to improve students' self-esteem and to promote an increased sense of professionalism in the teachers providing the intervention. Identification and appropriate evaluation of these non-targeted outcomes could offer a more complete picture of a program's impact and, thus, provide a stronger basis for decision-making about program selection. In this paper

we focus on possible effects of the intervention on children's self-esteem and employ measures to examine possible changes in the self-esteem of children who received the intervention. Results from previous studies have been mixed, possibly due to weak instrumentation (Rumbaugh & Brown, 2000). Self-esteem and school achievement are positively correlated (Harter, 1988). Harter theorizes that if a child achieves an acceptable level of accomplishment with academic work and if this achievement is important to him/her and is likewise valued by significant others in his/her life (e.g., parents, teachers, peers), then the child will develop positive self-esteem. This positive self-esteem makes the child feel good about him/herself and leads to motivation to do well in school. High levels of achievement help maintain high levels of self-esteem. Marsh (1990) has examined this relation with large samples of high school students in a sophisticated study employing structural equation models.

Given the importance of literacy learning for academic achievement in school, ensuring children's success in this area in the early grades is a high priority for schools. We hypothesized that a secondary effect of this prioritization is to strengthen children's self-esteem. Reading Recovery is a literacy intervention that has proven to be effective in rapidly improving children's literacy achievement, and is the intervention that is examined in this investigation. Reading Recovery is an intensive, daily, one-to-one tutoring program for "literacy at-risk" first grade students. It has a targeted curriculum and testing program and requires considerable staff development.

Methodology

In exploring the full range of outcomes of Reading Recovery, we discovered through focus groups conducted with parents and first grade classroom teachers that

students who participated in this literacy intervention appeared to increase their self-esteem, as well as their literacy skills. In fact, first grade teachers viewed this as the most important outcome of the intervention because increased self-esteem pervaded all aspects of the students' school behavior.

Harter and Pike (1983, 1984) and Harter (1993) theorize that children do not develop self-esteem until they are about eight years old. They contend that children younger than age eight cannot make evaluative judgments of the self and these authors view perceived competence and perceived social acceptance as precursors of self-esteem and self-worth. In this research we have adopted their theory of self-esteem. Harter and Pike also note that young children may not be accurate judges of their competencies and social acceptance and suggest that teacher ratings along the same dimensions might be employed and the incongruities studied.

To explore possible changes in the developing self-esteem in first grade students who participated in Reading Recovery as compared to first graders who did not participate in this intervention, we assessed the developing self-esteem of all first graders whose parents returned permission slips. Measures were made on the students at the beginning of the school year, mid year, and at the end of the school year. Because the student is not necessarily the most accurate appraiser of his/her own self-esteem (Harter & Pike, 1984), we also asked the student's teacher to rate the student's self-esteem at the same time points. Reading achievement measures were obtained at these time points as well.

Participants

The students attended a suburban school serving middle to upper middle SES families who are primarily white. Sixty-one first grade students comprised the first sample wave. There were thirty-four boys and twenty-seven girls; all the children were white except for three students from India. All children were six years old at the time of the fall testing. Eight of the children received Reading Recovery.

Five first grade teachers participated in the study. All are experienced teachers and all are female. The investigators were in their classrooms on many occasions and observed each of them to be highly competent teachers.

We currently are replicating the study in the same school with another 61 first grade children and five first grade teachers, three of whom participated in our first study. We anticipate an additional eight to fifteen of these children will be recipients of Reading Recovery. These data currently are being analyzed.

Instruments

Self-esteem as rated by the children. The developing self-esteem in children was measured using the first grade version of the Pictorial Scale of Perceived Competence and Social Acceptance for Young Children (Harter & Pike, 1983). This scale is individually administered to the child and is comprised of 24 items. Each item presents two pictures for which the child chooses the picture that is most like him or her and then indicates the degree to which he or she is like the picture. Each item receives a score of one, two, three, or four, with one representing a low score on the construct and four representing a high score. There are separate test booklets for boys and girls. In addition, there are separate forms depicting Caucasian boys and girls, Hispanic girls and

boys, African American boys and girls, and Asian boys. The Caucasian forms were used in this study, as they were most appropriate for the sample, although we had the other forms available if deemed appropriate for a particular student. The test contains four subscales that are theorized to represent the precursors to self-esteem: cognitive competence, physical competence, peer acceptance, and maternal acceptance. There is evidence for adequate reliability and convergent predictive validity of the test (Harter & Pike, 1984).

Self-Esteem of the Children as Rated by their Teachers. The teachers participating in our investigation rated the students' cognitive competence and peer acceptance on the Teacher's Rating Scale of Child's Actual Competence and Social Acceptance, Form 1-2, which is found in the manual for the Pictorial Scale of Perceived Competence and Acceptance for Young Children (Harter and Pike, 1983). The items parallel the cognitive competence, peer acceptance, and physical competence scales of the children's test for first and second graders. No evidence of reliability or validity was given; however, the authors cite moderate correlations between children's and teachers' cognitive and physical competency ratings and note that there is convergence of ratings for children in the top and bottom quartiles on the cognitive competence scale (Harter and Pike, 1984). We did not use the physical competence scale found on this measure as the teachers in this school do not participate in the students' physical education classes or supervise them during recess; consequently, they have little opportunity to observe the students' physical competencies.

Reading Achievement. A Text Reading Level measure was used to assess reading achievement (Clay, 1993). This measure consists of a series of graded books organized

according to a scale of text difficulty from level 0 to level 30. Students read texts until the highest text level is determined with at least 90 percent accuracy. The test administrator records reading behaviors during the oral reading task. The texts used represent expected reading performance for the following grade levels: levels 00-02 represent reading readiness, 3-8 are pre primer, 9-12 equal primer, 14-16 are end of first grade, 18-20 represents all of second grade, 22-24 are all of third grade, and 26-30 are achieved in fourth through sixth grades. The first grade teachers or the Reading Recovery teachers administered this measure in this investigation. Text Reading Level is an established measure, used in the field as a basis for instruction and instructional decision-making, including placement in Reading Recovery and special education services.

Reading achievement also was measured by teacher assigned grades. Teachers use a five-point scale to grade students at midyear and at the end of the school year. The teacher selects one of five descriptors of reading performance from (1) Emergent Reader, (2) Beginning Reader, (3) Developing Reader, (4) Capable Reader, to (5) Strong Reader. Each descriptor is accompanied by a full explanation of reading behaviors at that level.

Results

The results reported herein represent the first wave of data, which included 61 children. Although the entire Pictorial Scale of Perceived Competence and Social Acceptance for Young Children (Harter and Pyke, 1983) was administered, only the cognitive competence and peer acceptance scales were used in this analysis as they are the most relevant to student performance in school. The mean values obtained with our total sample of 61 first grade children on the cognitive competence scale and peer

acceptance scales were very similar to those that Harter and Pike (1984) report on a sample of 65 middle class first graders. Our means on the cognitive competence scale were 3.31 (*SD* = .45), 3.46 (*SD* = .41), 3.60 (*SD* = .42), for the first, second, and third testings, respectively, while Harter and Pike report a mean of 3.4 (*SD* = .37). For the peer acceptance scale, our means were 3.13 (*SD* = .55), 3.19 (*SD* = .57), and 3.27 (*SD* = .74), for first, second, and third testings, respectively, and Harter and Pike's mean was 3.1 (*SD* = .55). Thus, we conclude that our first graders look similar on this test to the children on which the test was developed.

For the purposes of data analysis, the 61 children in our study were divided into three subgroups based on their degree of risk of encountering literacy-learning difficulties. Students receiving Reading Recovery were identified as being at "high risk" for literacy-learning difficulties and comprised one group (*n* = 8). Students who were identified at the beginning of the school year as potentially needing Reading Recovery, but who did not receive this service, and/or students who received special services, such as speech and language assistance, were identified as being "at risk" for literacy-learning difficulties and comprised a second group (*n* = 14). The 39 students not receiving any special services formed the third, "no risk" group.

The student reports of perceived cognitive competence for each of the three groups at each testing were compared (Table 1). Examination of this table shows that the "high risk" students had the lowest mean cognitive competence scores (3.04) at the beginning of the year, followed by the "at risk" students (3.29), and the "no risk" students (3.37). This pattern persisted at the midyear testing with means of 3.29, 3.37, and 3.53 for the "high risk", "at risk", and "no risk" groups, respectively, and at the end of year

testing with means of 3.17, 3.64, and 3.68, for the "high risk", "at risk", and "no risk" groups, respectively. It should be noted that the "high risk" students' mean score rose to 3.29 at midyear, and then dropped to 3.17 at the end of the year. Although their mean end of year score was higher than their beginning of year score, the difference was not significant, $t(7) = .67$, $p = .52$, using a paired t test. Also not significant for this 'high risk" group were the comparisons of means from the beginning of the year to midyear, $t(7) = 1.43$, $p = .19$, and midyear to the end of the year, $t(7) = .96$, $p = .37$.

In contrast, the "at risk" students and "no risk" students demonstrated a slow, steady increase in cognitive competence over the school year. For both groups, the positive difference found between the beginning of the year and the end of the year mean scores was significant: the "at risk" group with $t(13) = 4.10$, $p = .00$ and the "no risk" group with $t(38) = 4.61$, $p = .00$. Also significant for these two groups was the increase in scores from midyear to end of the year: the "at risk" group with $t(13) = 3.84$, $p = .00$ and the "no risk" group with $t(38) = 2.49$, $p = .02$. The "no risk" students showed a significant increase from the beginning of the year to midyear, $t(38) = 2.45$, $p = .02$; however, for this time period the increase for the "at risk" students was not significant, $t(13) = .86$, $p = .41$.

The teachers' ratings of their students' cognitive competence were dramatically lower than the students' ratings of their competence (Table 2). A comparison of the students' ratings with the teachers' ratings shows that for each of the risk groups and for each testing time, the teacher ratings are substantially lower than the student ratings. As with the student ratings, the teacher ratings show the same pattern at each testing time

with the “high risk” group having the lowest mean, the “at risk” group having the middle mean, and the “no risk” group having the highest mean.

For all three groups of students, the teachers’ ratings of the cognitive competence of the children showed a steady increase from the beginning of the school year to midyear and from midyear to the end of the year. However, the difference was most dramatic for the “high risk” children where the difference between the beginning of the year mean and the end of year mean was 1.02; the corresponding differences for the “at risk group” was .82 and for the “no risk” group was .37. For each of the groups, paired *t* tests were done between the scores obtained on the beginning and the midyear testings, the midyear and the end of year testings, and the beginning of the year and the end of year testings. Every one of these *t* tests was significant at $p = .05$ (Table 3).

Next, the children’s reading achievement during the school year was examined. First, considering Text Reading Level (TRL) as the criterion measure, there were dramatic differences among the groups at the beginning of the school year, with the “high risk” group having a mean score of .63, the “at risk” group with a mean score of 1.66, and the “no risk” group with a mean score of 6.06 (Table 4). At midyear the “high risk” group was reading almost as well as the “no risk” group” with TRL means of 11.13 and 11.92, respectively; the same was true at the end of the year, when the “high risk” group had a mean of 18.25 and the “no risk” group had a mean of 18.49. Whereas, the “at risk” group did not perform as well and had a mid year TRL mean of 7.86 and an end of year mean of 15.86. As is obvious from inspection of Table 4, each of the three risk groups demonstrated significant gains in reading achievement as measured by TRL from the beginning of the year to the end of the year as shown by the paired *t* test results: the

"high risk" group with $t(7) = 16.01$, $p = .00$; the "at risk" group with $t(13) = 9.90$, $p = .00$; the "no risk" group with $t(38) = 18.90$, $p = .00$.

Using teacher grades in reading as the criterion for achievement, at both midyear and end of year, the 'high risk" students had the lowest mean grade, the "at risk" students had the middle mean grade, and the "no risk" students had the highest mean grade (Table 5). For each group, there was a significant increase in mean grade from midyear to end of year: the "high risk" group with $t(7) = 9.00$, $p = .00$; the "at risk" group with $t(13) = 5.08$, $p = .00$; the "no risk" group with $t(38) = 8.72$, $p = .00$. Note that the "high risk" group had teacher-assigned grade means that were considerably lower than the grade means for the "no risk" group, whereas when TRL was the criterion, the midyear and end of year mean scores for these two groups were quite similar. A possible reason for this is discussed in the following.

Although the Peer Acceptance scale data were analyzed, the analyses yielded no particularly interesting or significant information as related to school achievement or the development of self-esteem.

Discussion

Early elementary school children do not have a fully developed concept of their self as measured by self-report measures. However, this concept of self develops as they mature in elementary school. Harter (1990) notes that correlations between children's perceived cognitive competence and their teacher's evaluations of their perceived cognitive competence increase from grades three to six. Consequently, the teachers' evaluations of their first grade students' cognitive competencies are probably more accurate reflections of the students' competencies than are the students' own evaluations.

The teachers' evaluations of the "high risk" (M =1.65) and "at risk" (M = 1.99) students were low at the beginning of the school year. Although there are different descriptors for each item, across the 24 items on the cognitive competence scale, a score of 1 generally indicates not very good in the area, a score of 2 indicates sort of good in the area, a score of 3 indicates pretty good in the area, and a score of 4 indicates really good in the area. The teacher evaluations showed significant student improvement for all students over the school year, but the end-of-school means for the "high risk" (M = 2.67) and "at risk" (M = 2.81) groups were still below 3.0. Whereas, the teachers' mean for the "no risk" or nonintervention students was 3.02 at the beginning of the year and 3.39 at the end of the year. The larger gains made by the "high risk" and "at risk" groups over the school year, as compared to the "no risk" group, suggests that the interventions received by these two groups had an impact on their cognitive competencies. While the "high risk" and "at risk" students are not believed by their teachers to have high cognitive competence, one can only wonder how these students would have fared if they had not had some form of intervention. The "no risk" students with no intervention showed an increase of .37 points over the school year. If the "high risk" and "at risk" children had had this same increase and no intervention, their mean cognitive competency scores would have been 2.02 and 2.36, respectfully, or "sort of good", at the end of the school year.

It is noteworthy that the "high risk" and "at risk" students perceive themselves as considerably higher than their teachers perceive them. These researchers noted in the testing of these students that they often showed reluctance to evaluate themselves in the cognitive area or very rapidly evaluated themselves. Possibly they did not want to confront the reality of their academic performance and abilities. Additionally, at each

testing time the “high risk” students had the lowest perceived cognitive competency mean score and the “no risk” students had the highest mean score, with the “at risk” group in the middle. Thus, while these students may not have been as accurate as their teachers in evaluating their cognitive competency, the patterns were the same as their teachers, except for the small peak in the midyear scores of the “high risk” group’s own evaluations.

While all the children demonstrated significant improvement over the year on both the reading achievement criterion measures, the patterns were somewhat different depending on the measure. At the beginning of the school year, the “high risk” and “at risk” groups were at the reading readiness level and the “no risk” group was at the pre-primer level on the TRL. At the end of first grade the “high risk” and “no risk” groups had similar TRL scores that were on the second grade level and the “at risk” group had TRL scores that were at the end of first grade level. In contrast, the mean teacher reading grades followed the same pattern as the perceived cognitive competency scores. At both midyear and end of year, the “high risk” group had the lowest mean reading grade, the “no risk” group had the highest mean reading grade, and the “at risk” group mean was in the middle. A possible explanation for different patterns for the TRL scores and the teacher grades is that the teacher grade is based on broader reading behaviors than just ability to read and comprehend a book, as are TRL scores The teacher reading grade encompasses additional behaviors, such as selecting appropriate books, reading independently, making inferences, and actually reading during silent reading period.

In summary, based on the teacher evaluations of their students’ cognitive competencies, it appears that children’s participation in a literacy intervention serves to

strengthen their cognitive competencies, which are the precursors to their developing self-esteem. Neither Reading Recovery nor the other interventions that the students received claim to affect anything other than student achievement. Most academic interventions at the elementary school level focus only on strengthening a cognitive weakness and do not allege that they have any other impact. Yet this study indicates that when teacher judgments are taken into consideration, these interventions probably have a psychological, as well as an academic, outcome and this non-targeted outcome needs to be factored into the educational decision-making process. Educational evaluations must move away from targeting only student achievement test scores and begin to examine the broad spectrum of outcomes that might accompany the introduction of an educational program.

References

Clay, M. M. (1993). *An observation survey of early literacy achievement*. Portland, NH: Heinemann.

Harter, S. (1993). Causes and consequences of low self-esteem in children and adolescents. In R. F. Baumeister (Ed.), *Self-Esteem: The puzzle of low self-regard* (pp. 87-116). New York: Plenum Press.

Harter, S. (1990). Issues in the assessment of the self-concept of children and adolescents. In S. M. LaGreca (Ed.), *Through the eyes of the child: Obtaining self-reports from children and adolescents* (pp. 292-325). Boston: Allyn & Bacon.

Harter, S. (1988). Developmental processes in the construction of the self. In T. D. Yawkey & J. E. Johnson (Eds.), *Integrative processes and socialization: Early to middle childhood* (pp. 45-78). Hillsdale, NJ: Lawrence Erlbaum Associates.

Harter, S., & Pike, R. (1983). *The Pictorial Scale of Perceived Competence and Social Acceptance for Young Children [procedural manual]*. Denver, CO: University of Denver.

Harter, S., & Pike, R. (1984). The Pictorial Scale of Perceived Competence and Social Acceptance for Young Children, *Child Development, 55,* 1969-1982.

Hummel-Rossi, B. & Ashdown, J. (2002). The state of cost-benefit and cost-effectiveness analysis in education. *Review of Educational Research,72,* 1-30.

Levin, H. M. (1988). Cost-effectiveness and policy implications. *Educational Evaluation and Policy Analysis, 10*(1), 51-69.

Levin, H. M. (1995). Cost-effectiveness analysis. In M. Carnoy (Ed.), *International encyclopedia of economics of education* (2nd ed., pp. 360-364). New York: Pergamon Press.

Levin, H. M., & McEwan, P. J. (2001). *Cost-effectiveness analysis: Methods and applications* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Marsh, H. W. (1990). Causal ordering of academic self-concept and academic achievement: A multiwave, longitudinal panel analysis. *Journal of Educational Psychology, 82,* 646-656.

Rice, J. K. (1997). Cost analysis in education: Paradox and possibility. *Educational Evaluation and Policy Analysis, 19,* 309-317.

Rumbaugh, W., & Brown, C. (2000). The impact of Reading Recovery participation on students' self-concepts. *Reading Psychology, 21,* 13-30.

Table 1

Student Perceptions of their Cognitive Competence by Time

	Beginning of Year		Midyear		End of Year	
Group	M	SD	M	SD	M	SD
Hi Risk	3.04	.58	3.29	.29	3.17	.30
At Risk	3.29	.37	3.37	.28	3.64	.36
No Risk	3.37	.44	3.53	.45	3.68	.43

Note. n = 8 for Hi Risk, n = 14 for At Risk, n = 39 for No Risk

Table 2

Teacher Evaluation of Student Cognitive Competence by Time of Year

	Beginning of Year		Midyear		End of Year	
Group	M	SD	M	SD	M	SD
Hi Risk	1.65	.39	2.21	.32	2.67	.40
At Risk	1.99	.63	2.44	.81	2.81	.71
No Risk	3.02	.83	3.15	.71	3.39	.59

Note. n = 8 for Hi Risk, n = 14 for At Risk, n = 39 for No Risk

Table 3

Comparisons of Teacher Evaluations of Student Cognitive Competency by Time of Year

	Beginning - Midyear		Midyear- End Year		Beginning – End Year	
Group	t(7df)	p	t(13df)	p	t(38df)	p
Hi Risk	4.98	.00	4.95	.00	13.98	.00
At Risk	3.01	.01	4.11	.00	5.92	.00
No Risk	2.08	.04	3.85	.00	4.45	.00

Note. n = 8 for Hi Risk, n = 14 for At Risk, n = 39 for No Risk

Table 4

Text Reading Level (TRL) Scores by Time of Year

	Beginning of Year		Midyear		End of Year	
Group	M	SD	M	SD	M	SD
Hi Risk	.63	.92	11.13	4.05	18.25	3.11
At Risk	1.86	1.61	7.86	3.63	15.86	5.42
No Risk	6.06	4.50	11.92	5.94	18.49	4.44

Note. n = 8 for Hi Risk, n = 14 for At Risk, n = 39 for No Risk

Table 5

Teacher Reading Grades by Time of Year

	Midyear		End of Year	
Group	M	SD	M	SD
Hi Risk	2.13	.35	2.87	.64
At Risk	2.86	.54	3.64	.75
No Risk	3.62	1.02	4.28	.83

Note: n = 8 for Hi Risk, n = 14 for At Risk, n = 39 for No Risk