

The Integration of Renaissance Programs into an Urban Title I Elementary School, and its  
Effect on School-wide Improvement.

2001-2002 Annual Report to Renaissance Learning, Inc.

Linda A. Sadusky & Sarah K. Brem

Arizona State University

19 September 2002

## Table of Contents

<b>TABLE OF CONTENTS .....</b>	<b>2</b>
<b>I. EXECUTIVE SUMMARY.....</b>	<b>3</b>
<b>II. METHODOLOGY &amp; SCHOOL NARRATIVES .....</b>	<b>4</b>
METHODOLOGY.....	4
SCHOOL A NARRATIVE .....	4
SCHOOL B NARRATIVE.....	5
<b>III. QUANTITATIVE ANALYSIS .....</b>	<b>6</b>
HAS THERE BEEN AN INCREASE IN READING SAT9S SINCE THE INCEPTION OF READING RENAISSANCE AT SCHOOL A?.....	6
DO THE RESULTS OF STAR TESTING CORRELATE WITH SAT9 READING SCORES? ....	10
ARE SAT9 READING SCORES CORRELATED WITH ACTIVITY IN THE READING RENAISSANCE PROGRAM?.....	12
<b>IV: PERCEPTIONS OF AR: TEACHERS, PARENTS, AND STUDENTS .....</b>	<b>13</b>
GENERAL SATISFACTION .....	13
OPPORTUNITIES FOR IMPROVEMENT: WIDENING THE ZONE .....	17
OPPORTUNITIES FOR IMPROVEMENT: ALTERING PERCEPTIONS OF ASSESSMENT .....	20
<b>V: FUTURE RESEARCH.....</b>	<b>31</b>

## I. Executive Summary

This study was undertaken to determine the effects of implementing Reading Renaissance (RR) at a Title I school in the metropolitan Phoenix area, from its inception (1997) to present (2001). Drawing upon five years of SAT9 and Accelerated Reader data for approximately 1150 students, survey data from approximately 300 current students, 150 parents, and 20 teachers, and focus groups with parents, teachers, and students, we examine whether and how reading and school culture have been influenced by RR.

We found that since the inception of RR, there has been an 18% gain in students' SAT9 scores. Students' scores on Renaissance STAR testing are moderately to strongly correlated with their scores on the SAT9, and students who perform well in the AR program tend to perform well on the SAT9. However, we found little evidence linking specific AR benchmarks (e.g., points earned, number of books read, and gain in level) to improvement on the SAT9. Thus, while there is evidence for "macro-changes," overall gains corresponding to the introduction and use of RR, we have found little evidence for "micro-changes," gains associated with specific AR activities.

In reviewing responses from both surveys and focus groups, we find that RR has been favorably received by parents, teachers, and students in the school. The school has made a strong investment in the Renaissance program and philosophy. A strong majority believes that RR improves reading. Parents appear to be knowledgeable about the program, and they interact with their children about reading and RR. They are pleased by the amount of reading their children were doing, and the motivation the program provided. Many of the children enjoy the reading and the sense of success that comes with their achievements. However, there are some elements of RR about which parents, teachers, and students are ambivalent; these conflicted feelings center on competition and pressure. All three groups feel that while setting goals and striving for these goals are a positive and important aspect of RR, this can lead to competition and anxiety among students, as they compare themselves to others on reading level and number of points earned, as well as pressure among teachers, as they worry about whether they will make model or master classroom.

While some of these concerns may be inherent to the RR program (or any other form of assessment), we believe the situation is aggravated by fundamental misunderstandings regarding the purpose of RR and RR assessment. These misconceptions are not the responsibility of any one group or aspect of the program. They are fueled by some program design elements, political pressure over accountability, and misconceptions held by parents, teachers and students. Based upon our surveys, interviews, and focus groups, we believe that the dominant model of assessment at the school is what we term the "high-stakes testing" model, which stresses summative evaluation, striving for perfect performance, little feedback and contextualization, and substantial, long-lasting consequences. This model is quite different from that which underpins RR, which we term the "challenge" model. The challenge model stresses formative evaluation, striving for an intermediate level of performance indicating that the task is neither too easy nor too hard, rich feedback, and relatively minor, short-term consequences. Given the current political climate in Arizona and other states, the high-stakes model is a pervasive influence. We believe that if the alternative challenge model was more effectively taught to teachers, students and parents, and was understood as a valuable tool to complement the high-stakes model, we would see a more effective implementation of RR, larger gains in reading performance, and greater comfort and satisfaction among all members of the school community.

## II. Methodology & School Narratives

### Methodology

This is a retrospective longitudinal study exploring the impact of the implementation of the RR program model at a K-6 urban, Title 1 elementary school. This site was chosen because of its choice, five years ago, to implement the RR model school-wide, and its continued involvement with RR strategies. Collecting quantitative academic data allows us to track the time before training, during initial phases of use, and over years of continued implementation. Collecting qualitative data allows us to explore personal experiences with AR and RR throughout these phases of program adoption and implementation.

In order to answer the question, “Has implementation of the RR program model made a difference in academic growth?” SAT9 reading scores at the RR school will be used to look for changes within that school. In addition, the quantitative results of this implementation will be compared to a matched site that has not fully implemented AR or RR, to control for any additional changes in the district or community. The second part of this project addresses what takes place within a school community when a program model addressing school-wide reform is embraced. What evidence is there that the school environment has changed as a result of the RR implementation?

#### *Instruments and Data Collection*

- SAT9 reading scores, 1997 to 2002.
- Demographic information including ethnicity, gender, and Title I eligibility.
- Pre- and post-STAR test results, 1997 to 2001.
- Additional AR data including number of points earned, book levels and number of books read, 1997 to 2001.
- Surveys of parents, teachers, and students addressing reading attitudes and impressions of AR/RR.
- Focus groups with students, teachers, and parents focusing on program use and effects.
- Reading specialist and librarian interviews focusing on program implementation.

### School A Narrative

This site is a K-6 urban, Title I elementary school with 36% of its 625 students qualifying for free or reduced lunch. Twenty-four percent of the school’s population comes from minority backgrounds. On-site programs for gifted students English as a second language learners, special education students, an all-day kindergarten program, a pre-school, before school math labs, an after school homework club and the services of a full-time social worker allow the school to best meet the diverse needs of its community. Before and after school day care is also available on site. Parent involvement includes volunteer readers, site council membership, a parent-teacher organization and Active Parenting classes.

School A has supplemented the district adopted reading series with other programs designed to meet the needs of at-risk students. Their first grade teachers are trained to use CLIP strategies, and that training is currently being implemented at the kindergarten level. They purchased Accelerated Reader (AR) in 1992 with 200 books/tests and had their first Reading Renaissance (RR) training in the Fall of 1998. The reading specialist at the time

remembers that as a school they earned about 3,000 points a year prior to the training. During the 1998-99 school year, they earned over 30,000 points. School A has received Model School recognition in each year since 1999. In the spring of 2000 Master School certification was awarded. This site was chosen because of this extended experience with Accelerated Reader both before and after the Renaissance program model was introduced.

### **School B Narrative**

A matched site was desired in order to report comparative data. This site was selected because it is also a Title I elementary school in the same district as School A. It has used Accelerated Reader since 1996, but has not chosen to be involved in the Reading Renaissance program model. It is this distinction that makes it an interesting comparative site. The distinguishing variable becomes the actual implementation of the Renaissance model. This study examines the impact of Reading Renaissance strategies and the potential acceptance of the program model as an acceptable Comprehensive School Reform Demonstration (CSRD) program.

Of the 617 K-6 students, 18% qualify for free and reduced lunch. School B offers a variety of special programs designed to meet the needs of its community including programs for gifted, English Language Learners (ELL) and special education students. School B also offers an all day kindergarten program, pre-school and on site day care. A full-time social worker is also on site daily. Parents take an active role in the school's site council and parent-teacher organization, and volunteer in classrooms. Active Parenting classes are also offered in the evenings.

In addition to the district adopted reading series, School B has always been involved in literature studies and promoting the use of trade books in the classroom. Prior to AR they were very involved in a program called Quest. Quest was a collection of about 300 multiple choice, true false tests on trade books. Each test was assigned a number of points, but books were not leveled. There has been only sporadic use of AR, based primarily upon each individual teacher's desire to use the program, but they have seen a noticeable increase in library circulation.

We were not permitted access to the parents or students of School B, and only the librarian and reading specialist are data sources for this study.. Thus, we do not have information about the implementation of AR, but it appears to be quite haphazard, and RR is not present.

### III. Quantitative Analysis

Six years of SAT9 data were collected from Schools A and B, 1997 to 2002, as well as data on gender, ethnicity, and Title I eligibility. For School A, we also have AR data from 1997 to 2001, though STAR test results were available only beginning in 1998.

The primary questions to be addressed through quantitative analysis were:

- Has there been an increase in reading SAT9s since the inception of Reading Renaissance at School A?
- Do the results of STAR testing correlate with SAT9 reading performance?
- Are SAT9 reading scores correlated with activities in the Reading Renaissance program?
- Are increases or decreases in SAT9 reading scores correlated with activities in the Reading Renaissance program?

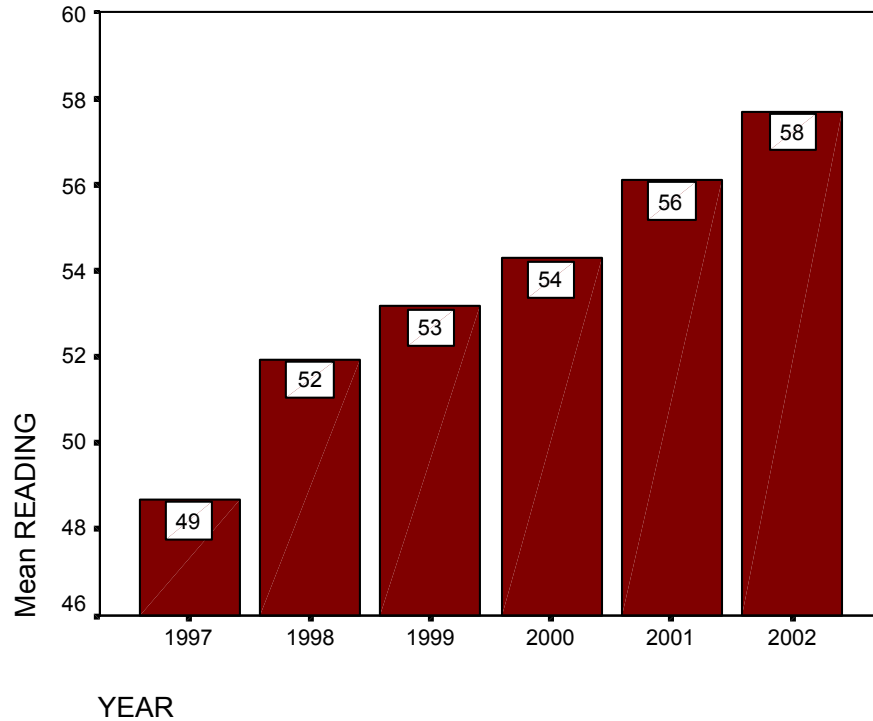
#### **Has there been an increase in SAT9 reading performance since the inception of Reading Renaissance at School A?**

From 1997 to 2002, School A showed a 9-point increase in SAT9 reading scores, an 18% increase. This increase is significant by the nonparametric Kruskal-Wallis test,  $X^2(5) = 35.9$ ,  $p < 0.001$ . There were no significant gains in the control school during that time.

At School A, there has been a significant improvement in SAT9 reading scores since Reading Renaissance and AR were first deployed. Normal curve equivalent (NCE) scores rose from 49 to 58, an 18% gain. The increase held for both Title I eligible and non-eligible students (Figure 2), and both boys and girls (Figure 3). Looking at ethnicity (Figure 4), however, we see that the gains were stronger for White students than for Hispanic students (other minorities represent such a small proportion of the school population that statistical analysis is inappropriate).

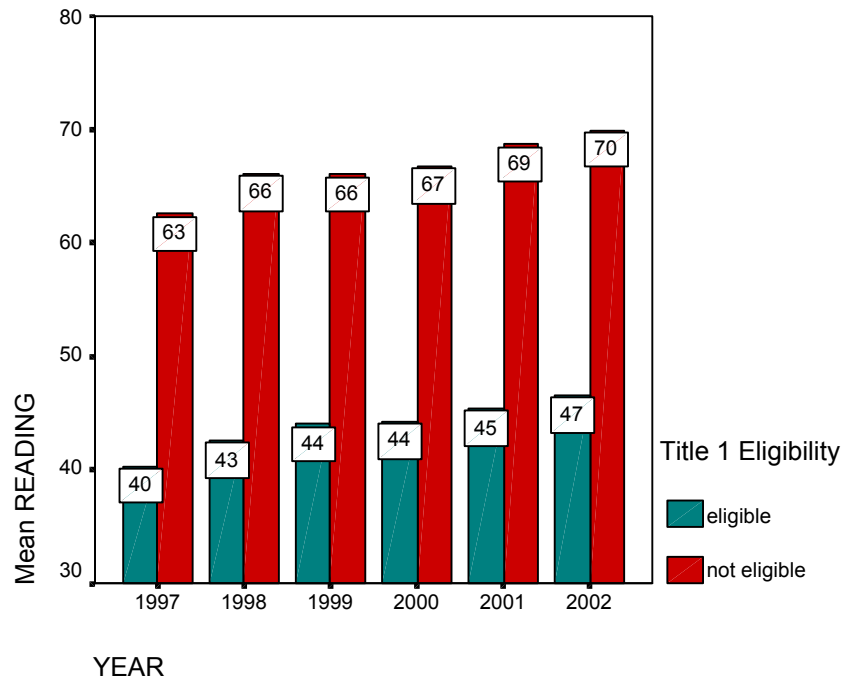
**Figure 1. SAT9 total reading scores for School A, 1997 to 2002.**

*Includes grades 3 and 4 in 1997, grades 3 through 6 in 1998, and grades 2 through 6 for 1999 to 2002.*



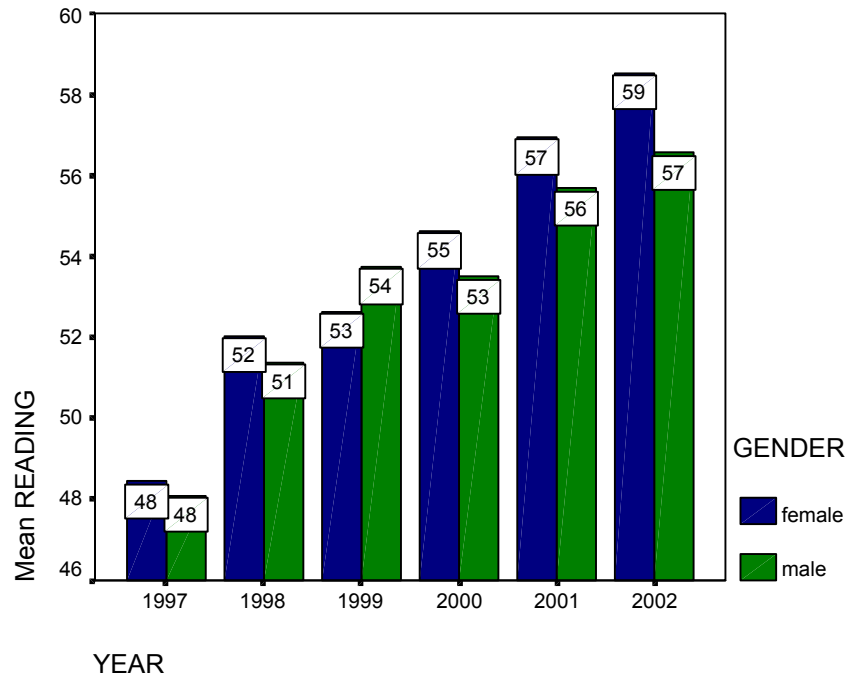
**Figure 2. School A, SAT9 scores broken down by Title I eligibility.**

*Both Title I eligible and ineligible students improved an average of 7 points. Gains for both groups were significant. Title I eligible:  $X^2(5) = 23.0, p < 0.001$ , Title I ineligible:  $X^2(5) = 21.9, p < 0.01$ .*



**Figure 3. School A, SAT9 scores broken down by Gender.**

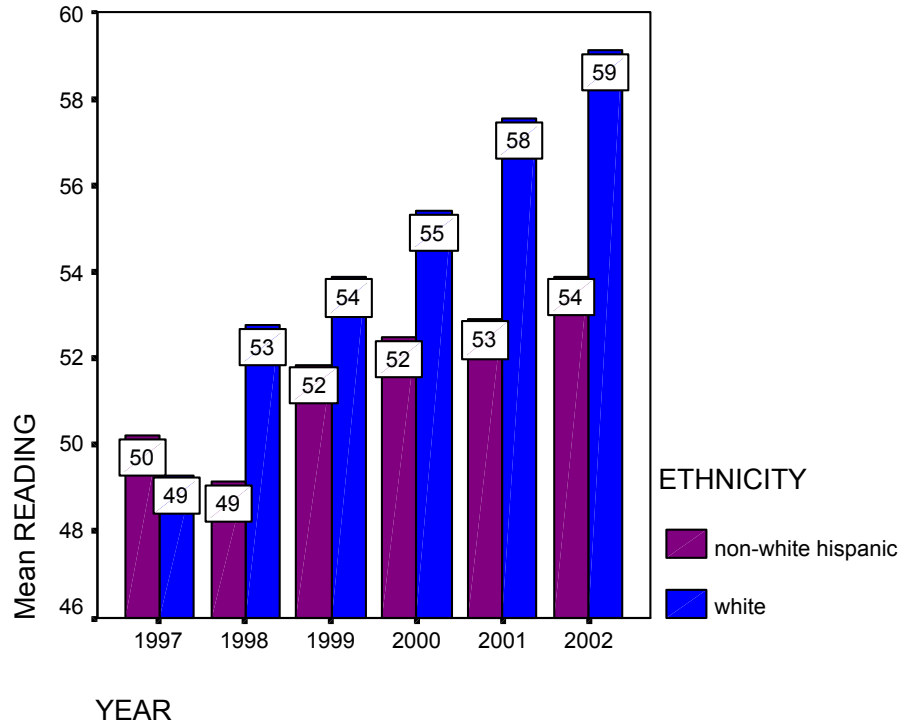
*Girls improved 11 points,  $X^2(5) = 29.1, p < 0.001$ , boys improved 9 points,  $X^2(5) = 12.0, p < 0.05$ .*





**Figure 4. School A, SAT9 scores broken down by Ethnicity.**

*Gains for white students averaged 10 points, and were significant,  $X^2(5) = 32.3, p < 0.001$ , but gains for non-white Hispanic students were smaller, and not significant,  $X^2(5) = 3.1, p > 0.50$ .*



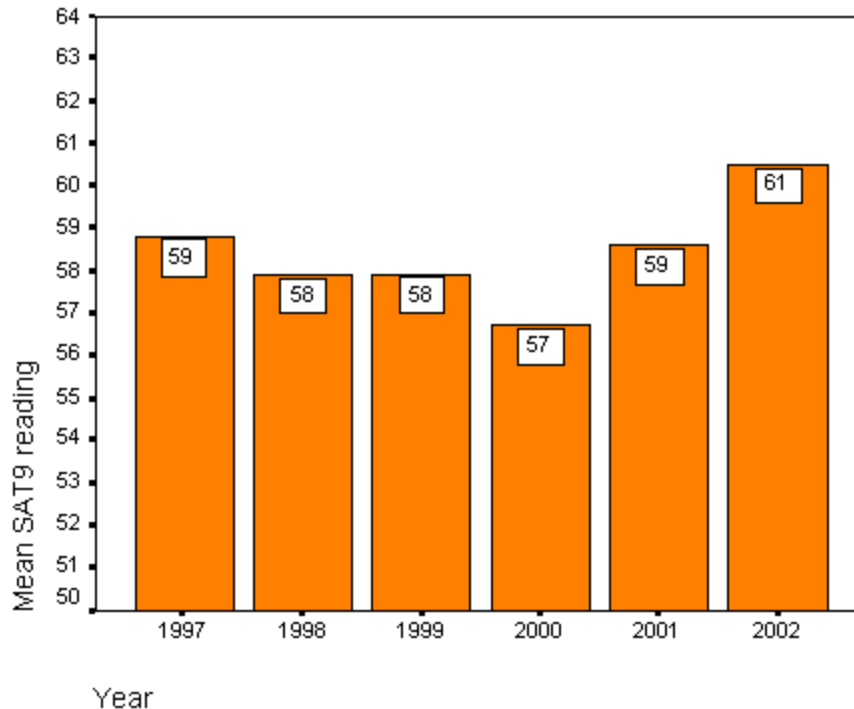
The advantage of a retrospective study is that the school did not enter this program as a “showcase school,” knowing that its progress was being charted by researchers. The changes cannot be due to being the subject of research, but rather reflect the natural process at School A.

It is possible that other changes in the school or district are responsible for these gains. To address this, we examined a second school that had not implemented Reading Renaissance (Figure 5). The school is in the same district, is geographically close to School A, and uses a similar curriculum, with the exception of RR. The demographics are similar, but School B has slightly fewer Title I students, and a slightly smaller minority population. Therefore, if anything, the situation is biased in favor of School B. School B did not show gains paralleling those of School A. Breaking the results down by gender, ethnicity, and Title I eligibility produced no significant gains; there was a marginally significant gain for Title I eligible students, from 41.4 to 46.0 ( $X^2(5)=9.57, p<0.10$ ).

Thus, it would appear that the emphasis placed on reading at School A has produced results. Of course, we cannot be certain what element of the reading program at School A is responsible for these gains. However, based on the history provided by the librarian and reading specialist, we believe that the changes to reading instruction are largely, if not solely, the result of implementing Reading Renaissance.

**Figure 5. SAT9 total reading scores for School B, 1997 to 2002.**

*Includes grades 3 and 4 for 1997, grade 3 for 1998, grades 2 through 4 for 1999, and grades 2 through 5 for 2000, and grades 2 through 6 for 2001 and 2002.*



### Do the results of STAR testing correlate with SAT9 reading scores?

Correlations between STAR scores and SAT9 scores are between 0.65 and 0.75. Thus, it appears that STAR results can serve as a moderately good predictor of SAT9 performance in reading.

Given that the only significant gains were at School A, for the remainder of the report we focus on this school. A potential asset of the Reading Renaissance program is its ability to provide schools with a sense of how their students will perform on standardized tests. To function as a good predictor of SAT9 scores, the RR summative testing instrument, the posttest STAR, should correlate well with SAT9 performance. We tested whether this was the case at School A by examining the correlations between STAR9 and SAT9 performance from 1998 to 2001 (the years for which we had STAR data). As shown in Table 1, the correlations are significant, and moderate in size.

*Table 1. Correlations of the STAR posttest with the SAT9 total reading scores, School A only, 1998-2002. All correlations significant,  $p < 0.001$ .*

	SAT9 total reading, 1998	SAT9 total reading, 1999	SAT9 total reading, 2000	SAT9 total reading, 2001
STAR posttest, 1998	0.66 (n=44)			
STAR posttest, 1999		0.69 (n=234)		
STAR posttest, 2000			0.67 (n=389)	
STAR posttest, 2001				0.73 (n=361)

The correlations hold at these levels when the School A population is broken down by ethnicity, Title I eligibility, and gender. This suggests that for most students, their STAR scores will be moderately predictive of their SAT9 scores in reading.

*Table 2. Correlations of the STAR posttest with the SAT9 total reading scores, School A only, 1998-2001, broken down by ethnicity. All correlations are significant,  $p < 0.001$ , unless otherwise noted.*

Ethnicity		SAT9 total reading, 1998	SAT9 total reading, 1999	SAT9 total reading, 2000	SAT9 total reading, 2001
Hispanic	STAR, 1998	.55 (N=7) <i>n.s.</i>			
	STAR, 1999		.64 (N=42)		
	STAR, 2000			.74 (N=67)	
	STAR, 2001				.71 (N=76)
White	STAR, 1998	.69 (N=35)			
	STAR, 1999		.75 (N=179)		
	STAR, 2000			.71(N=287)	
	STAR, 2001				.73 (N=255)

*Table 3. Correlations of the STAR posttest with the SAT9 total reading scores, School A only, 1998-2001, broken down by Title I eligibility. All correlations are significant,  $p < 0.05$ , unless otherwise noted.*

Title I eligibility		SAT9 total reading, 1998	SAT9 total reading, 1999	SAT9 total reading, 2000	SAT9 total reading, 2001
Not eligible	STAR, 1998	0.29 (N=13) <i>n.s.</i>			
	STAR, 1999		0.50 (N=106)		
	STAR, 2000			0.51 (N=185)	
	STAR, 2001				0.52 (N=175)
Title I eligible	STAR, 1998	0.37 (N=31)			
	STAR, 1999		0.37 (N=128)		
	STAR, 2000			0.45 (N=204)	
	STAR, 2001				0.52 (N=186)

*Table 4. Correlations of the STAR posttest with the SAT9 total reading scores, School A only, 1998-2001, broken down by gender. All correlations are significant,  $p < 0.01$ .*

Gender		SAT9 total reading, 1998	SAT9 total reading, 1999	SAT9 total reading, 2000	SAT9 total reading, 2001
Girls	STAR, 1998	0.81 (n=22)			
	STAR, 1999		0.71 (n=118)		
	STAR, 2000			0.67 (n=183)	
	STAR, 2001				0.73 (n=172)
Boys	STAR, 1998	0.67 (n=19)			
	STAR, 1999		0.67 (n=107)		
	STAR, 2000			0.68 (n=184)	
	STAR, 2001				0.73 (n=170)

### Are SAT9 reading scores correlated with activity in Reading Renaissance?

While students who perform well on AR activities also tend to perform well on the SAT9, we did not find a relationship between AR activity and SAT9 improvement. There is a macro-level association between the presence of the program and a rise in scores, but no micro-level association between specific activities and a rise in scores.

School A students who perform well on the SAT9 tend to be those who show the greatest success in and interaction with AR; not surprisingly, good reading and good reading test scores tend to go together. Students with a higher average AR reading level tended to achieve higher SAT9 reading scores, as shown in Table 5. Likewise, the number of points they earn is also associated with test-taking ability, since the number of points received depends on the length of the book, the difficulty level and the score they receive on the corresponding quiz. Consistent with this, we see in Table 6 that the number of points earned in a year is moderately correlated with SAT9 reading scores.

Table 5. Correlations between average reading level in AR and SAT9 reading scores, 1999 to 2001. All correlations are significant at the 0.001 level.

	AR level in 1999	AR level in 2000	AR level in 2001
SAT9 total reading, 1999	.44 (N=422)		
SAT9 total reading, 2000		.41 (N=369)	
SAT9 total reading, 2001			.49 (N=382)

Table 6. Correlations between AR points earned by students, and their SAT9 total reading scores, 1999 to 2001. Correlations are significant at 0.001, unless otherwise noted.

	Points earned, 1999	Points earned, 2000	Points earned, 2001
SAT9 total reading, 1999	.50 (N=419)		
SAT9 total reading, 2000		.47 (N=366)	
SAT9 total reading, 2001			.52 (N=381)

However, the number of books read *did not correlate* with SAT9 performance, with no correlation greater than 0.10. In addition, we found little association between *change* in students' performance on the SAT9 reading and their AR activity. Increases and decreases in their SAT9 reading scores were not associated with the change in their AR level, the number of points earned, or the number of AR books read that year, except from 2000 to 2001, there was a weak correlation between change in SAT9 readings scores and change in AR level ( $r=0.24$ ), and books read ( $r=0.27$ ).

In short, while the presence of the AR program is associated with a rise in reading scores, this rise cannot readily be attributed to particular activities or aspects of the Reading Renaissance program. There is a macro-level association between the presence of the program and a rise in scores, but no micro-level association between the specific activities of AR and a rise in scores.

#### IV: Perceptions of AR: Teachers, Parents, and Students

Working from focus groups, interviews, and surveys of parents, teachers, and students, we have developed a portrait of School A, and how AR has affected not only reading education, but also the school environment more generally. In general, the implementation of the Reading Renaissance program has lived up to the school's expectations. The focus on reading and its value to learning has truly become part of the school's culture. It was quite evident in all interviews that teachers and students know what Reading Renaissance is, its goals and purpose, and why the school is so heavily involved in the program. It is equally evident that all participants 'speak' the Reading Renaissance language, and most readily acknowledge the benefits of the program. Students intelligently discussed reading strategies concerning choosing books, reading for understanding, and what it means to be a good reader. Conversations about specific books and book series, types of books, issues related to reading fiction and non-fiction, how to deal with difficult words, the differences between reading and skimming, and even a recognition that they were reading with the purpose of getting better at reading all flowed easily and naturally. One group even considered the idea that sometimes, based upon one's interest in a book, even a good reader could be a 'not so good' reader because "*when you're reading and you don't really like your book, you're not concentrating as hard as if you really like your book.*" This is not to say that there weren't dissenters, and there were important concerns that should not be overlooked. But often, elements that were seen as problematic were also viewed as having positive effects. Still, there are opportunities for fuller implementation of the program.

#### General satisfaction

The primary reason for becoming involved in Reading Renaissance was to improve reading ability. The program was attractive to teachers and administrators because it offered a clear and comprehensive set of components that were purported to motivate students to read and provide the tools necessary for teachers, parents and students to monitor reading progress.

The program is seen as living up to these claims. In focus group interviews, parents, teachers and students held the program in high regard, especially its ability to motivate students, recognize achievement, create independence and self-reliance, and give students' the sense that they "owned" their reading progress. A parent noted:

*R is actually writing in her journal now. I mean all on her own...I think it's kind of neat. She's reading and she's writing. I have to go tell her to turn off the light at 10:20 at night, but I don't mind (laughs)...I'm pleased. I think the Reading Renaissance program gave that, the ability and desire to read.*

One teacher describes how the elements of the program work together to create student independence.

*And the independent time on the computer, without the teacher standing over there telling them what to do...They feel totally responsible for the book they read and their performance on the test. And the teacher is not guiding that, they're monitoring from the background, but they're not standing over them...I think they like that. That feeling of success. (They're) fully responsible for the selection of the book, how well they read it, their performance on the test.*

Parents also noticed that their child's reading skills were improving, that the goal of reading for understanding was being internalized and that the program provided evidence of growth for their children.

*Oh, I want to read that book, but I can't get there yet. I've gotta get to that point and then all of a sudden, you realize they're a better reader and they didn't know it.*

When asked to comment on changes seen in their children since before the school's involvement with Reading Renaissance, one fifth grader's parent reflects,

*Basically it was his excitement. It was how he was handling things. When he had to do book reports in 2<sup>nd</sup> grade, it was 'Okay, I have to read this many books. It wasn't necessarily the content of the book; it was just the book itself. And, when we started AR it was more 'Okay, now I have to read at a certain level, and comprehend it, beyond what I got out of it. I have to understand what they're going to ask me.'*

The elements of 'doing AR' and achieving Reading Renaissance goals seem to have become ingrained in 'who' the school is. Everybody is doing it. New students and new teachers quickly see that this is 'how it is around here'. One teacher comments,

*Kidwise, the big difference that I saw immediately was they were talking about books and reading them. I mean before they'd check books out, but they never read them. An, I think that probably, because we had that expectation, the library did, I did, for the classroom. If they went to (another teacher's room), she did. (The principal) supported it. Everywhere around the kids knew, oh god, we have to read.*

And other teachers comment on their recognition of change,

*K: I think that when AR started, I think it intimidated everybody, kind of quite a bit, but I also think it gave an entire faculty a focus that wasn't there before. Everybody had a purpose. We knew, I mean everybody's on the same page.*

*D: We were learning it together at the same time...helping each other.*

*K: At that point everybody had a common goal and it felt like that.*

Furthermore, all stakeholders appear to understand the feedback that Reading Renaissance provides and, for the most part, use this information appropriately. Students and teachers were able to describe how tests, points, zones and goals are integrated into the program model. All understood the concept of being 'at-risk' and how to 'get out of being at-risk'. Though not all parents were able to correctly interpret the data provided on the student reports and records, all were able to verbalize a general understanding of how the program worked. Some parents had a different use for the program than was perhaps planned, creating a "bonding moment" around the Renaissance reports, rather than using them as a form of reading evaluation.

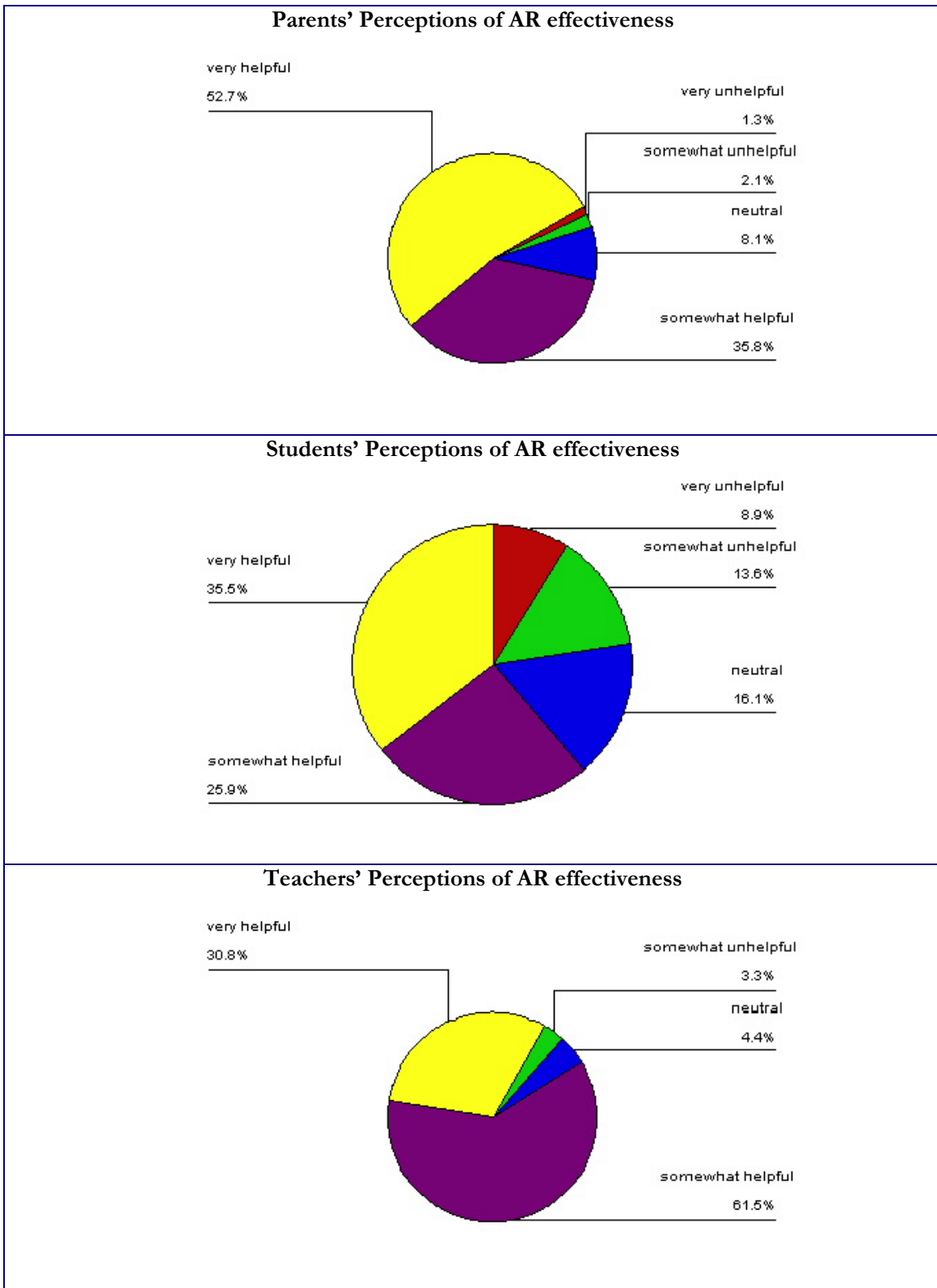
*I'll tell you how I use this. And it's not for reading. How I use it is it gives me that chance to put my arm around her and say, 'way to go, kiddo'. You know, 'keep it up, that's neat.' 'So how many points you got now?' and 'Are you going to the store?' And she'll tell me all about it...It's not that I take a lot of time deciphering the scores, but it gives me a chance to talk to her about it. A chance to reinforce.*

This satisfaction is borne out in the surveys distributed to the broader school population. The majority of teachers, parents and students believe that the use of AR and RR in their school has made a positive contribution to reading (Figure 7). *Participants were also asked to name the one most positive aspect of Reading Renaissance. The type and frequency of their responses is listed in Table 7.*

Table 7.

<b>Participant</b>	<b>Category</b>	<b>Themes</b>	<b>Frequency</b>	<b>Percent</b>
<b>Parents</b>	Positive aspects of RR and reading at School A	Amount of reading & performance	31	27.2
		AR generally	19	16.7
		Providing positive motivation to read	15	13.2
		Positive environment/staff	14	12.3
		Enthusiasm about & focus on reading	13	11.4
		Accountability as a positive influence	9	7.9
		Good reading resources	7	6.1
		<i>Other</i>	6	5.4
		<b>Total</b>	<b>114</b>	<b>100.0</b>
		<b>Students</b>	Positive aspects of RR and reading at School A	The act of reading itself
Accountability as a positive influence	54			14.5
Positive environment/staff	43			11.6
Rewards for reading	42			11.3
AR generally	40			10.8
Good reading resources	28			7.5
Amount of reading & performance	26			7.0
Enjoyment of reading activities	21			5.6
Found nothing positive about AR	20			5.4
<i>Other</i>	15			4.0
<b>Total</b>	<b>372</b>	<b>100.0</b>		
<b>Teachers</b>	Positive aspects of RR and reading at School A	Amount of reading & performance	9	32.1
		AR generally	4	14.3
		Enthusiasm about & focus on reading	4	14.3
		Providing positive motivation to read	2	7.1
		Good reading resources	2	7.1
		<i>Other</i>	7	25.0
		<b>Total</b>	<b>28</b>	<b>100.0</b>

Figure 7. When asked how helpful AR and RR were in building reading skills, the majority of parents, teachers and students believed the program was very or somewhat helpful.





## Opportunities for Improvement: Widening the Zone

The school has implemented the concept of “ZPD,” or “zone of proximal development” in a way that is different than that prescribed by Renaissance. They have chosen to narrow students’ zones, a process they call “zoning the zone.” They reduce the breadth of the recommended zones and require each student to read a certain number of books within each zone, at increasing point levels.

For example, the recommended range given at the time of the STAR testing may be 3.2 – 5.0. The Reading Renaissance program fully expects the child to read up and down within this entire range. The school, however, breaks this range down into four colors: red (3.1-3.5), pink (3.6-4.0), white (4.1-4.5) and orange (4.6-5.0), and restricts the reader to only “red book” before he is allowed to move to “pink books”, and so on. The school uses both the number designations and a color-coded system to help students recognize books in their zone. Any book in the school library that is an AR book (the school owns the corresponding test) is labeled with a colored dot on the binding, and a label indicating the book’s title, author, level and point value on the front cover. The school also publishes lists of all the AR books available in the library, sorted by color.

Teachers and administrators explain their decision as motivated by several factors: they want each reader to take their time in each zone, reduce the chance that readers will miss “great books” in each zone, and avoid younger readers accessing “age-inappropriate” books.

Students understand that individuals are placed at their own ability level and see movement to more difficult zones as progress. They metaphorically see this progression through the zones as a ladder which one climbs. They also believe that it is “fair” to have students read in their zone. In fact, most of the fifth and sixth graders remember that when AR first came to the school, the books were not leveled, but simply coded as an AR book and given a red dot.

*K: I didn't like it when it was all red 'cause we didn't know (if it) was easy or hard.*

*L: You'd just get a book.*

*LS: So the red dot just meant it was an AR book...but didn't say what level?*

*L: No, and like if it'd be hard or easy...Like I'm in light blue and she's in like pink, so that's an easier book for her and it would be way easy for me.*

*K: (answers L) So that would be unfair. If she took a test on the easier book and I couldn't take a test on those because they were too hard for me.*

*L: So we have to stay in the color we're in.*

*LS: So by not being fair, it would not be fair because of what?*

*K: Because we wouldn't know which books to check out, which ones are harder.*

*L: and which are easy.*

However, narrowing the zone in which a student can read may have some unexpected and undesired side effects. Narrowing the range of books a child can read at any particular time causes issues related to access. There simply has to be enough books available within a student's zone if the concept of choice is to be supported. Parents seemed most aware of the gaps.

*There is a shortage of certain books at certain levels and certain ladders. It's almost like a gap thing. I mean sometimes you'll hit a level and they'll be lots and lots of books, but then you'll go check one rung and suddenly you're scrambling just to try to find a book.*

Financially, this is quite burdensome. Not only are more books needed, but also each new book must have its companion test or it's not 'available' as an AR book. As one parent summarized, "If you're going to buy into it (the program), then you have to buy into it (spend the money)."

Even though they acknowledge that reading at one's individualized level is desirable, they do not like the confinements of the program. One parent voices concern that her child never has time to read books of his own choosing.

*The sad thing is when your child doesn't have time to do that. He only has time during the summer because he's a slow reader. And he only has so much time in class to do it. And that's an awful point.*

These zone lists are sent home to help parents know which books are in their children's zones. But for some the list is a physical reminder of how limiting the concept of zones can be. Parents felt frustrated when they heard their children say that they would not read a book because it was not on the list.

When a school adopts the RR program, it also accepts a new definition of choice. A student has choice, but only within his zone. Sometimes books that a student wanted to read were not in this zone. Sometimes, even within a series of books, like the Little House on the Prairie series, books jumped from zone to zone without regard for the actual progression within the series. So a student might be able to read the first two books and the fourth, but not the third.

Some teachers have begun to deal with this problem by allowing some out-of-level reading, and the RR program does support this, but the effects of this option raise other concerns as illustrated by the following dialogue between one group of teachers.

*A: (They should) be allowed, if they meet a goal, to go and get whatever color they want.*

*B: And the program does suggest things like that (for) the kids that want something different. You make a deal, you do this, and I'm going to let you read whatever you want for a while.*

*C: So they can get one that's in their level and one that's not?*

*B: And you know, that's not for every kid, you don't make that general, but that's for the kid who wants to read Harry Potter, but it's not in their level.*

*A: But then they don't get to take the test on it.*

*B: Yeah, they do.*

*A: They do?*

This comment raised a lot of conversation and confusion. Finally, A asks:

*A: But if it's below their level then that puts them at risk, doesn't it?*

*B: If they do it too many times*

*D: You (have to) limit doing that.*

The color coding system has become a way that children label not only their choice in books, but one another. Both teachers and students refer to individuals and groups of students by their color. Teachers describe the range of reading abilities in their classrooms as 'going from yellow to black'. Other teachers understand this to mean that this particular classroom has readers from pre-primer to well into junior high reading. Students refer to themselves, and their reading ability, as their 'color'. "I'm pink. She's blue."

Though knowing your zone and knowing that you have the goal of progressing through the zones is one of the elements that generated the most enthusiasm (since students are motivated by seeing their progress), the zone itself still publicly marks the children. Knowing your zone and knowing that you have the opportunity, through your own efforts, to improve your reading and move through the zones is one of the great motivators in this program. One parent notices that his child measures success by watching zone movement.

*(She) will comment about 'I've gone to pink' or white or whatever it is. It's a big deal what ever it is. 'I'm in chapter books.' You know they're improving when you see that progress.*

On the other hand, knowing that your peers are reading in 'higher' zones is frustrating. The following comment nicely sums up the dichotomy.

*It was definitely a change. And I know he was disappointed with the color he was in, and it was a matter of pushing himself. So, I saw a change in him that way. He was pushing himself up to the next color level and the exuberance when he came home. He MADE IT to the next color level.*

Interestingly, these 'labels' are not limited to one teacher's classroom, but follow students at all grades throughout the entire school. You can be 'yellow' in first grade, but you can also be 'yellow' in fourth grade. Parents of students with learning disabilities were most frustrated with how the program 'labels' their children.

### **Recommendations.**

Widening the zone would require more careful monitoring and adjusting by the teachers. The lock-step progression designed by the school is much easier to 'control'. In a way it manages the teachers' decision making as much as it does the children's. Though the school has made strides towards measurable improvement in reading, as evidenced by the SAT9 and STAR data, what if even more progress could be made by following the Reading Renaissance recommendation? Within that decision also lies the possible reduction of some of the negative implementation effects. Keeping colors in place, but widening a student's

zone would mean he could read red, pink, white and orange. It would be more difficult for anyone but the child to know his zone. The overlap of colors into many zones would further ‘disguise’ the categorization. In addition, there would be a broader selection of books for each child. As is the program has had to sacrifice some of its intent to motivate pleasure reading.

When asked on surveys to name the one most negative aspect of Reading Renaissance, the responses were heavily weighted toward concerns about competition, pressure, and inappropriate forms of motivation (Table 8).

Table 8.

Participant	Category	Themes	Frequency	Percent
<b>Parents</b>	Negative aspects of RR and reading at School A	Accountability as a negative influence	25	29.4
		Found nothing negative to say about RR	24	28.2
		Inadequate resources	17	20.0
		Undermining self-confidence/self-esteem	7	8.2
		Lack of parental involvement in RR	4	4.7
		Inappropriate use of extrinsic motivation	3	3.5
		Poor environment/staff	3	3.5
		<i>Other</i>	2	2.4
		<b>Total</b>	<b>85</b>	<b>100.0</b>
		<b>Students</b>	Negative aspects of RR and reading at School A	Accountability as a negative influence
Found nothing negative to say about RR	71			24.3
Negative feelings from difficulty reading	26			8.9
Poor environment/staff	19			6.5
Problems with technology	16			5.5
Unhappy with reading resources	11			3.8
Inadequate resources	10			3.4
Reading is an unpleasant activity	10			3.4
Teasing because of poor performance	9			3.1
Too much work (reading or other areas)	7			2.4
<b>Other</b>	<b>16</b>			<b>5.4</b>
<b>Total</b>	<b>292</b>	<b>100.0</b>		
<b>Teacher</b>	Negative aspects of RR and reading at School A	Accountability as a negative influence	10	35.7
		Found nothing negative about AR	4	14.3
		Too time-consuming	4	14.3
		Inadequate resources	2	7.1
		Problems with technology	2	7.1
		Too much emphasis on AR	2	7.1
		<b>Other</b>	<b>4</b>	<b>14.4</b>
<b>Total</b>	<b>28</b>	<b>100.0</b>		

### Opportunities for Improvement: Altering Perceptions of Assessment

Narrowing the zone not only makes the program more internally manageable, it is also less risky from an assessment perspective. Students are carefully reined in; they cannot make huge leaps forward, but they are also unlikely to backslide. In the current climate of Arizona,

where high-stakes testing is highly visible, tied to school funding, and quantitative measures are often seen as the *only* legitimate measures of a school's performance, there is a tendency among all stakeholders to treat every assessment as if it were a high-stakes situation. Since Reading Renaissance data is intended for formative assessment, as "low-stakes" testing that merely provides all concerned with feedback to provide the best instruction, the pressure to treat it as one would summative, high-stakes evaluation can have an undesirable effect.

Progress is identified in several ways: individual test scores, aggregate reports (e.g., at-risk reports), point accumulation, goal achievement, and the achievement of Model or Master status at the classroom, school and library levels. With the exception of Model and Master status, which is clearly a public designation, the rest are intended as sources of information about individual and classroom progress. Ideally, poor performance would not be used as a trigger of recriminations against parents, teachers or students, but simply as information indicating that a different approach was needed. Unfortunately, the atmosphere is one in which anything less than strong performance can lead to anxiety and a sense of failure.

**Individual Performance.** The school uses the point system as prescribed by Reading Renaissance, and the concept of points seems well understood and accepted. Children clearly understand that a book worth 5 points is both longer and harder than a book worth 1 point, and can further distinguish that a 5 point book in a lower level is easier than a 5 point book in a higher level. Points appear to be an easily understood, easily measured program component. Furthermore, the attainment of goals is made more readily understandable through points, and there is no question that goal-setting has had some positive effects. There seems to be an active focus on moving oneself along some continuum toward a specific outcome. There are many instances of recognition and celebration as individual and group goals are met. During many morning announcements you will hear the principal happily congratulating a number of students who have recently met their goals. Within those classrooms many students applaud their peers. Special assemblies gather the entire school population, parents, district personnel and board members to honor students who meet their quarterly and yearly goals, earn top points, read the most books and make most progress.

However, the use of points has also had a downside. The same pressure to increase availability of books in each zone exists within each zone. There must be a wide range of books at all point values within each zone in order for students to benefit from the full range of difficulty before moving 'up the ladder'. This need has positively influenced the number of books available in the library. In most classrooms you will also find the teacher's private collection of books labeled with 'AR dots'. This is a wonderful outcome of the school's focus on reading, but at the same time, the school never seems to be able to satisfy its needs.

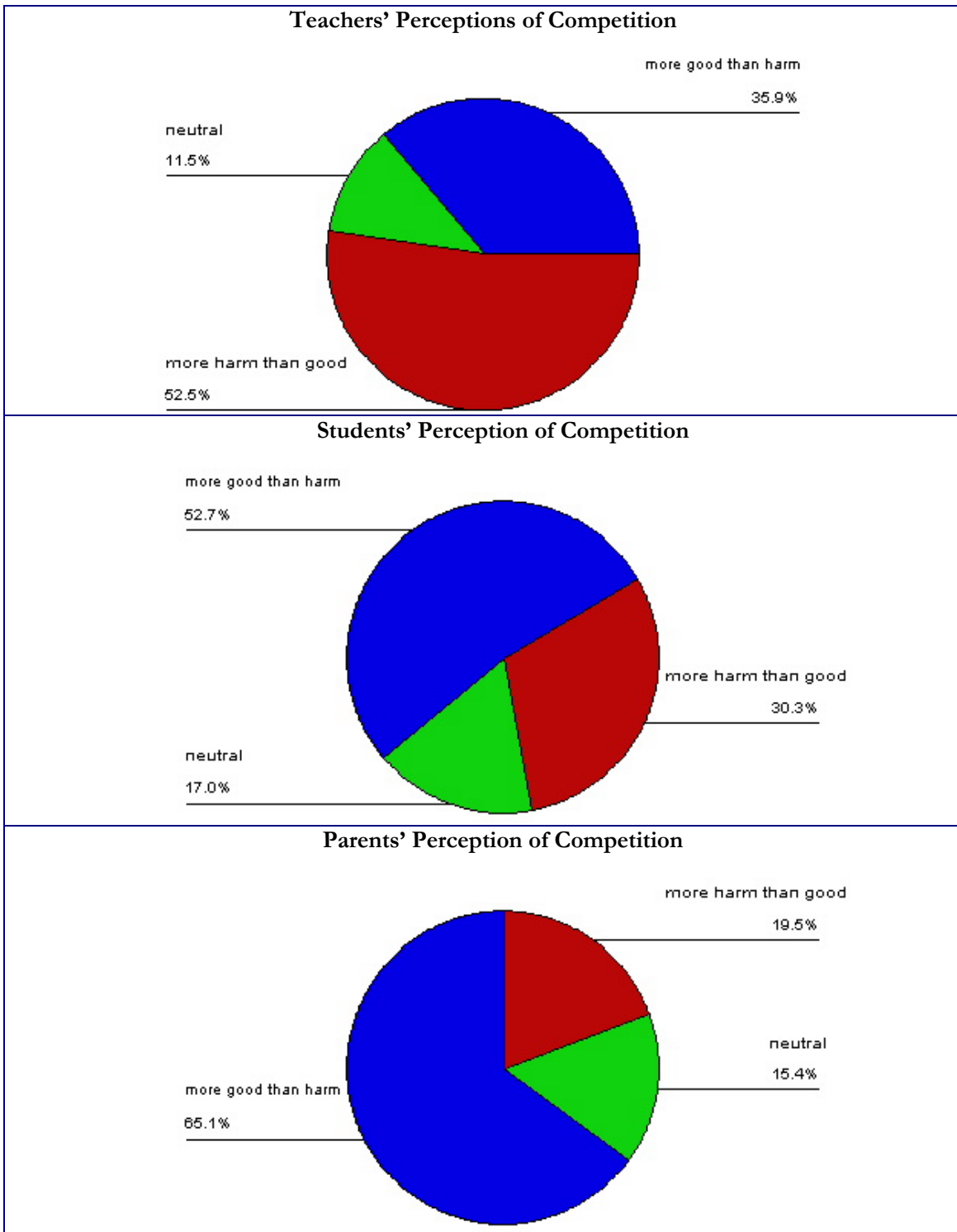
The effect of points on the issue of meeting individualized needs while safeguarding against categorizing students is similar to those raised about zoning, though to a lesser extent. Students know how many points books are worth. They can usually tell just by size alone, so no one has to announce that someone is reading an easier book. And in the midst of all the celebrations of achievement, there are always a few who have not met their goals. And here is where all of the 'highs' of recognition sit side by side with the 'lows' of unfulfilled expectations. Students feel time slipping by as they realize they have fallen too far behind to achieve their goals. Parents agonize over the dilemma that goal setting, motivation and recognition are all wonderful aspects of the world except when your child doesn't see himself as someone who has something to celebrate. The problems that labeling creates are

simply compounded when points enter into the equation. Kids boast about how many points their current book is worth; they compare how many points each person has earned, and they complain when they feel a particular books has not been assigned a “fair” point value.

In addition, children know that reading books with more points not only helps them meet their goals, but also gives them more ‘buying’ power. When asked if the number of points influences which book might be chosen, one student responds, “Yeah, I just get higher ones so I can get more points.” Another comments, “You just want to get it done,” when asked if there are any negative aspects of points. There is a sense that this element of the AR program has become an essential part of how students view independent reading. It is now an assignment. Students have a designated reading level, and each book they read is valued by points not content. That is not to say that students aren’t also saying, ‘You should read this great book.’ But a comment like that is often followed by two questions: ‘What color is it?’ and ‘How many points is it worth?’

Based on the survey results, it appears that children, parents, and teachers are all divided over the effect of points and competition in the AR program. All three groups have a strong showing of those who believe rewards and competition are healthy ways to promote reading, and of those who find these elements of the program worrisome. In Figure 8, teachers express great concern regarding competition; over half fear that it does more harm than good. Students and parents are show, but still substantial, levels of concern.

Figure 8. Teachers were the most concerned about competition, the majority believing that it does more harm than good. Parents and students also expressed concern, but at far lower levels.



In Figure 9, again teachers have the most pessimistic view; almost 70% believe that the use of rewards has a negative effect on reading motivation, causing student to value rewards

over reading. Students and parents viewed the rewards far more positively, though approximately 20% expressed concern.

Students have learned to cope with the pressure and competition, but in ways that may be undesirable. Students manipulate their book selection, based upon point value and ease of readability, to increase their chances of making their goals. Teachers, too, learn to “beat the system.” More than a few teachers have been known to choose a ‘big point read-aloud’ so that their classes have a better chance of meeting their goals. Students are not unaware of these manipulations, as the following excerpt from one of the sixth grade focus groups illustrates. This comment comes out of a conversation about the idea that sometimes teachers advise ‘at-risk’ students to read quick, easy books in their zone (in order to get out of being ‘at-risk’).

*N: What I understand is that (when) we read the little books they are worth half the points (than) if you read chapter books. But if you read two little ones you get the points. So it's faster to read two little ones.*

*LS: Why do you think that? Is there a reason why they'd want you to do that?*

*N: Cause the chapter books have bigger words and stuff and take your time to read. Otherwise you're reading little books and they give you faster points, but that's not what really matters.*

Later in the same section another student adds,

*Some people even in our grade will just try to get half point books. They say, 'Oh, that's easier.' You can probably get like five points reading half point books. That's not like reading a five point large book. It's just easier for them. It's the easier way to do it. So they are taking the easy way through."*

Even parents reinforce the role of the point system as an extrinsic motivator. For example, one parent mentioned a ‘points as money’ system used at home.

*I said, "I'll pay you a dollar a point." So they went and they read and when June came around my money was real and \$50.00 seemed like a lot of money. I could see their eyes going 'ka-ching'. So the next year when S came home with 220 points, I'm like, it's working!*

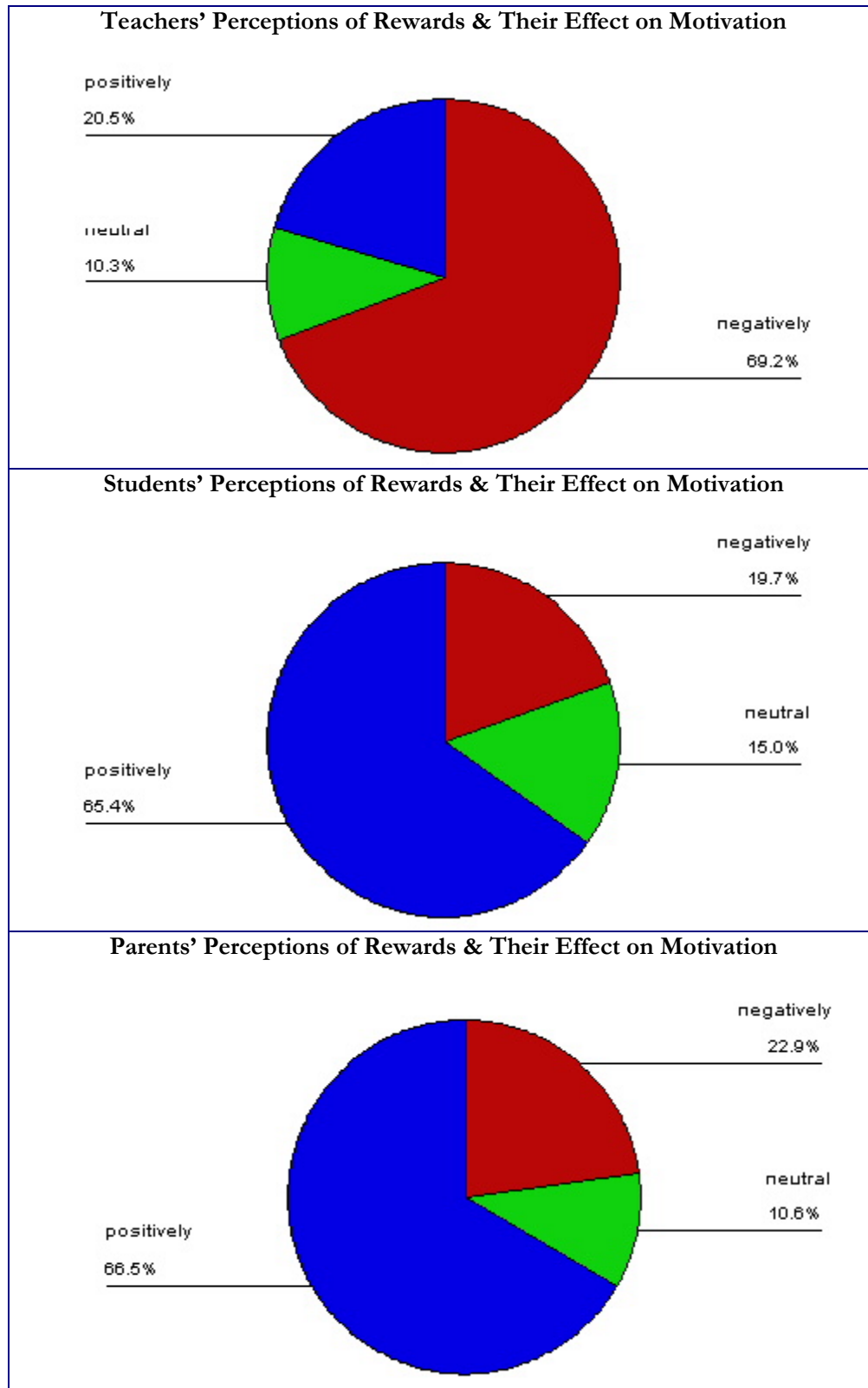
Some parents saw points and the store as an opportunity to learn about spending habits. One comments, “It gives them a chance to manage money skills. You know, am I going to (shop) this quarter or am I going to save?” The ‘dollar a point’ parent adjusted the money payout to a period payoff rather than paying all at the end of the year. Another parent, when discussing the AR store, also commented on the reality of an external rewards system.

*The rewards have been nice. On the face of it we're all on a rewards system. You get a paycheck that (indicates) that you're worth whatever you've done. So it really does do better if there's some kind of reward or some kind of goal or award to be working towards.*

Figure 9 summarizes the survey results regarding the effects of rewards on student motivation. While teachers are very concerned that rewards for reading will have a negative effect on student motivation, parents and students are more optimistic.



Figure 9.



Goals and the program have become so much a part of the students' lives that any reading that doesn't get them closer to their goal may be considered a waste of time. One student discovered that the book she had read didn't have an AR test. She comments,

*A: So I read that whole thing for nothing.*

*LS: So, since you couldn't take a test on it, you feel you read it for nothing.*

*A: Because it wasn't even an AR book. There were no questions about it. It wasn't that I didn't like it, but it wasn't an AR book.*

It's possible that the teachers reinforce this attitude. One child explains, "You have to read a book, and you have to pay good attention. You're gonna take the test. And some books aren't AR and maybe you want to read them, but your teacher's saying, 'No, that's a waste of time.'"

These students' voices echo the conflict between the positives and negatives of goal setting. Recognition and reward doesn't come without competition and an ever-increasing pressure to succeed. This pressure comes from their peers as well as their teachers. One girl describes a classmate, "(She'll say) 'I got this much and you only have five.' I actually know a person like that. And I absolutely hate them for it." Another adds, "Well, like they're saying, if people can look at your reading log or people watch you take a test. They're like, 'I read that same book and I got a hundred.' Well, that's you. I'm like a slower reader... They like to brag about how you messed up."

And of course, there's pressure at home. Students who have to spend time reading outside of class in order to reach their goals can disturb family schedules and routines. They have to 'fit in' sports, family, homework, etc. For some children, divorce or daycares further complicate their lives. Books are left in one place and needed in the other.

Parents are well aware of the pressures of meeting goals. Although there is a general feeling of confidence in most teachers, there are some who are perceived as putting the group goal ahead of the individual's goal.

*I've had some teachers who allow the kids to still read for pleasure and others who are so concerned about points that they want to go ahead with that ladder exactly where it's supposed to be. I think the teacher has to be flexible and in tune enough with kids, so they don't dampen the joy of reading by making it a chore to earn points, you know?*

Another parent explains that competition is two sided.

*The dark side of competition is you make winners and then you make losers. There is no, it isn't always a win. So when you set it up as a goal and you tried to reach your own goal, it doesn't matter. Then you look at the competition that is (within) the class as in how many have reached their goal, it takes on a different meaning. When you have competition in that way, in that way it's a matter of who's reaching the goal.*

Parents and students recognized that privacy may be an issue. When recognition is so public, and goals are so well shared by all members of the group, peers seem to feel entitled to knowing how others are doing. One parent recalls overhearing kids comparing and

boasting, 'What color are you in? Well, I'm in this.' 'It can be a big ego booster', she says, "but also a very big esteem loss."

The following excerpt from a sixth grade group's discussion about using reading logs clearly addresses the issue of privacy and comparisons.

*J: Well, if you get a 20% or 40% then you don't want anybody to know, but your teachers still want to look at your reading log.*

*LS: So it could be a not so good thing if you have to write down the unsuccessful experience.*

*K: Well, I think it could be harmful in a way. Cause if you're carrying it around and let's say, you're a 5<sup>th</sup> grader and you are only in yellow.*

*JG: What's yellow?*

*J and the others: lowest color, lowest range*

*K: And then you took it and you dropped it. When somebody finds it, they could make fun of you...Or, you're in black (the highest level) and you got nothing on a test. And you dropped it and somebody found it. And they would make fun of you because your name is in it.*

*JG: Does that happen a lot?*

*J and B: Yeah.*

*B: I've been made fun of just because I got like a 20%.*

*JG: By people in your class?*

*B: No, by people that I had take my log. Like to the library before and I left it. They found it and opened it up and looked in it before.*

*J: Yeah, that happened to me, too. Like right here (points to log entry) I got a 40% on my first non-fiction of the year. My teacher said I'm border level and I read it in a light green and I got a 40%. I took that back to get another book and like a 5<sup>th</sup> grader or a 4<sup>th</sup> grader, they go "Nah, nah (sing-song) You got a 40%. You stink at reading.*

*LS: So that would be a negative thing, huh K? I also got the feeling that you thought your log was private.*

*J and K: like a diary almost. Or a journal.*

But, as one student suggests, using logs is still a good thing. "Even though it's embarrassing. It's better because when you didn't have a log it was hard to keep (all your reading) in your head." Consideration should be given to how the school can balance the tools of recognition and the negative effects of competition and pressure.

Setting point goals redefines the purpose of taking the test. The student is not focused on assessing a level of understanding. He is focused on earning points. Even while reading, some students seem to be focusing on taking the test. When asked what they think about while reading, students said:

*I'm thinking about getting done.*

*It (the test) helps you memorize the book because you want to get a good score.*

*It sounds like something I did last year because it took a long time to read it and so I had like two chapters left and so I just skipped a chapter. A whole chapter. And it was the last one. I took the test. I got like an 80. Not that bad.*

Parents suspect that things haven't changed much since they were in school. Kids will figure out ways to circumvent learning and reach the goal.

*I remember being in school. I remember sometimes you do things cause you want to finish it right and other times you're trying to get to the goal. The sense of, you know, how it ends. Let's say you were to read a certain series, the author has a style. You can almost guarantee that you could read certain parts of it, if you've read several of their books; you know what will be coming. And so, that might be a good thing, too. He's learning how to skim. Figure out what's going on. But you can learn how to answer tests.*

**Classroom Performance.** In the Reading Renaissance model there are three levels of goal integration. Individual student goals are set based upon each student's initial STAR test results. The program provides guidance for teachers to then set point and level goals for each student. The second tier of goal setting involves the classroom as a group. Following prescribed calculations, teachers are able to determine yearly point and level goals for the classroom. When classrooms attain a Model Classroom banner you can hear the group cheering gleefully, and it doesn't take long for that banner to be displayed proudly outside the classroom door. However, with this comes teachers who feel pressure when their classes don't progress as expected. Attaining these goals along with evidence that they are using other program components permits the class to apply for recognition from Reading Renaissance as a Model and then Master Classroom. Schools can set yearly (and then quarterly) point goals for the school at large, and they can apply to Reading Renaissance for recognitions as Model and Master School, and Model Library.

While the concepts of Model and Master status were familiar to all, the actual criteria for achieving Model and Master status at the classroom, school, and library levels were unfamiliar to most. All groups were familiar with the terms and the desirability of achieving these designations, but parents, teachers and students were not able to state specifically how these designations were attained. Additionally and perhaps in part fueled by this missing knowledge, there was some concern that organizational goals were overshadowing individual goals. Recently teachers have begun to feel more pressure about maintaining the school's successful pattern of goal attainment. They, like students, can feel the pressure of comparison. The announcements, banners and hallway displays say 'Mrs. X's class made Model classroom.' The absence of these recognitions also point out those whose classes are not 'progressing' as desired. Furthermore, receiving Master School recognition one year seems to increase pressure the next year. In the final year of this study, the school reached Model School status by the end of the first quarter of the school year; since nine weeks of

data are required to attain this goal, this represents the fastest possible attainment. Some teachers see the school's current focus on making Model and Master School as a deterrent to maximizing individual growth.

*We do class goals. But you can't even say that that's what your kids need to focus on. You may have little Suzie who's going to read 100 points in nine weeks just so she can help make the class goal. And then there's little Sally who's just going to read 2 points because 'oh, look, we're getting to our goal.' You know, I think we really need to focus more on individualized goals.*

The students in our focus groups seemed well aware of these pressures. One comments, "I think some of the time they push you to read SO MUCH and in so little time. It can make you a stronger reader, but for some people it messes them up." Another adds, "Well, if you don't do good. Like if we're really slow on our books or if we don't do so good. They sort of scold us."

### **Recommendations**

Students, parents and teachers appear to understand and use the basic elements of the Renaissance program effectively. They understand points and levels, and can establish and reach goals at the individual, classroom, and school level. Indeed, given the rate at which they achieve Model and Master levels, they clearly can use these concepts with great success. However, there does appear to be a price for these achievements; students, parents and teachers all recognize that pressure and competition frequently come with accountability and acknowledgement of achievement.

Some of these feelings may arise because all parties involved are highly aware of the stakes associated with evaluation and the public announcement of the results. Some of this is perhaps inevitable; as many participants pointed out, if there are high-achievers, there are also bound to be low-achievers, and they will be identified if only by their lack of awards. In addition, however, these problems may be aggravated by the tendency to see every quiz and every book as a test with consequences, as if each one is a direct measure of a child's reading ability, rather than one piece of evidence, or one datum. A formative approach would stress the collection of a large body of evidence before making assessments or decisions, while a summative approach emphasizes a branchpoint at which final decisions and rankings are made. Thus, rather than attempting to provide the appropriate level of challenge, the goal becomes to achieve success on every quiz.

Additionally, if test results were viewed as a fluid concept of progress and the widening of the zones were combined, students' feelings of being limited would be lessened. A struggling reader who really wants to read that 'black' Harry Potter book might be encouraged to do so. And if he tests well on it, then he's demonstrated his ability to read a book assumed to be too difficult. And, if he doesn't do well on it, then he has evidence that though he may have enjoyed the book, he didn't really fully understand it. Or, what if after reading just a bit of the book, the child realizes for himself, "I'm not getting this." By allowing students to have these experiences, teachers would truly be fostering independence, ownership and self-reliance.

It is important to note that in some ways, the design of the Renaissance program may encourage this mindset. By design, the program recognizes quiz scores between 85 –92 % as indicating that the student is reading books that are appropriately challenging. However,

maximum points are only assigned when students receive a 100%. This seems to discourage risk-taking and works contrary to the notion of reading at your challenge level. If points were assigned on the basis of whether students average scores remained in the challenge range, it would allow for the occasional lower grade and also discourage a stream of 100's. This supports the idea that the testing system should be viewed as mapping a process, a peek at one's progression towards becoming a better reader. Unfortunately, under the current point scheme, students are encouraged to think that 100's are best because only perfect score are rewarded with full points.

## V: Future Research

School A has shown substantial gains in SAT9 reading scores since 1997. After a close examination of how RR and AR are implemented and perceived within the school, we have hypothesized that the effects of the program may have been attenuated by the way it has been used at the school. Moreover, there are signs of considerable worry or discomfort associated with RR and AR, centering primarily on competition and the results of competition.

Increasingly, schools in Arizona and especially this district are being forced to focus on high-stakes testing. Testing is occurring more frequently and more broadly than ever before, and there are serious financial consequences for both schools and individual teachers if their students do not perform as demanded. Thus, the dominant model for assessment is competitive, norm-referenced, and high-stakes. RR is none of these things; the focus is rather on individual mastery, with quizzes as formative assessments to help teachers and students set goals, identify problems, and determine a student's instructional level (ZPD).

From a cognitive standpoint, we describe this situation as a problem of *competing mental models*. Mental models are, effectively, the world recreated in our own heads. When we think about how our car works, the best route to work, or what a test is for, we are drawing upon mental models. Mental models are responsible for our ability to make sense of the world, to make decisions and interpret events.

People are cognitively conservative. Our tendency is to try to fit new situations to an existing mental model; if the fit is less than perfect, we will change our picture of the new situation before we will try changing the established mental model. Thus, if our dominant mental model of assessment was built on the SAT9 and similar sorts of testing, our tendency should be to try to assimilate all new forms of assessment to that model, including RR.

This assimilation would account for a number of problems we have identified with the RR implementation. The high-stakes testing model has an element of competition; thus, it is natural to look for an element of competition in RR. Teachers and students find it by comparing levels, worrying about how many points they have, and making model classroom. In high-stakes testing, higher scores are always better, so we stress high scores and perfect performance; in RR, the ideal score is between 85 and 92%, but teachers and students consistently worry about less than perfect scores. Because of the tremendous consequences attached to high-stakes testing, even a single percentile can make a difference; in RR, the key is whether or not students are in "in their zone" as measured by average quiz performance. A few percentage points are irrelevant, as is the occasional low score, but the dominant high-stakes testing model makes this an extremely difficult position to accept.

When the dominant mental model leads to inappropriate or inaccurate outcomes, the solution is to actively introduce an alternative mental model, and help learners to draw new boundaries. Thus, the solution is a two-step process: they need to not only develop a new mental model that can sit alongside the old one, but they also have to learn to identify which model is needed when. In this case, we believe that what is needed is a new mental model for parents, teachers and students that embodies the notion of ZPD, formative assessment, and criterion-referenced testing. Once that model is available in their minds, they need to learn when to apply this new model, and when it will be more fruitful to rely on the high-stakes testing model.

We hypothesize that this training will allow the school to implement RR more effectively, and will alleviate some of the stresses that are currently associated with its use. In particular, we would expect the following results

- **The acceptance of a broader ZPD:** Currently, students' ZPD as used in School A is no more than 0.5 levels wide in most zones. This ensures that students do not “regress,” by reading books that are lower than the last one they read, or risk a low score by reading a challenging book. It also gives the appearance of constant progress, since students move through these zones fairly quickly. Rapid progress, high scores, and no “backsliding” are important features of a high-stakes model, but are incompatible with the RR program as it was meant to be implemented. Once the school community understands that the high-stakes model is inappropriate, and perhaps even counterproductive, we expect that they would be willing to accept a broader ZPD as recommended by RR, and provide more opportunities for challenge and growth.
- **Less concern about performance relative to other students and classrooms:** Students and teachers feel pressured to “be better” than everyone else. In a norm-referenced system, to which the SAT9 belongs, for each student who is above the mean, there has to be someone below the mean. This is not true for criterion-referenced systems, in which everyone can achieve mastery. Competition is not particularly relevant in this model, and we would expect this concern to decrease as the new model grew more familiar. This could also result in performance gains, if it allowed for greater risk-taking and therefore greater opportunities for challenge.
- **Less concern about individual quiz scores, and more accurate interpretations of quiz performance:** High-stakes testing allows for no mistakes, and an entire year's work rests on one score. In keeping with this, students and teachers at School A are overly aware of students' exact performance on each quiz, and will avoid challenge to ensure high scores. We predict the new mental model, which stresses that scores can be *too high*, indicating a lack of challenge, will result in a greater willingness to try a more difficult book, increasing the chances for learning and improvement.

The overall result of this switch to a more accurate mental model should be: (a) improved reading skills, as measured on SAT9 tests, (b) greater satisfaction with the RR program among parents, teachers and students, as measured through focus groups and surveys, and (c) a better understanding of the psychological principles underlying the RR program, as measured through focus groups and surveys.