THE DEVELOPMENT OF READING FLUENCY AND MOTIVATION TO READ: AN EXAMINATION OF THE CAUSAL RELATIONSHIP

by

MATTHEW PIERCE QUIRK

(Under the direction of Paula J. Schwanenflugel)

ABSTRACT

The current investigation is comprised of two studies examining the validity of an adapted version of the Motivation to Read Profile (MRP-A) and the causal relationship between motivation to read and reading fluency skill development. Both studies were conducted with 168 second grade students from a rural school district in northeast Georgia. The first study scrutinized the validity of the MRP-A along five aspects of Messick’s (1989) framework for unitary construct validity. In this study, the MRP-A functioned well along the content, substantive, structural, and generalizability aspects of the framework, yet questions still remained regarding its relation to external criterion variables. The second study examined the causal relationship between motivation to read and reading fluency skill development using structural equation modeling techniques in a three wave longitudinal design. The three subscales of the MRP-A (reading self-concept, value for reading, and goals for reading) were used as observed indicators of a motivation to read latent variable and the reading fluency skill latent variable was defined by three fluency subscales (the TOWRE sight word efficiency subtest, the TOWRE phonemic decoding efficiency subtest, and the DIBELS oral reading fluency scale). Previous research conducted on the relationship between academic self-concept and academic achievement has identified four possible models to explain their causal relationship which were used as hypothesized models to explain the relationship between motivation to read and reading fluency skill development. Surprisingly, the SEM analyses supported a self-enhancement model, which asserted that motivation to read influenced motivation to read...
and reading fluency skill at the subsequent time point, but that reading fluency skill only influenced reading fluency skill level at the subsequent time point.

INDEX WORDS: motivation to read, reading fluency skill, self-enhancement relationship
THE DEVELOPMENT OF READING FLUENCY AND MOTIVATION TO READ:
AN EXAMINATION OF THE CAUSAL RELATIONSHIP

by

MATTHEW PIERCE QUIRK
B.S., Pennsylvania State University, 1999
M.A., The University of Georgia, 2002

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial
Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2005
THE DEVELOPMENT OF READING FLUENCY AND MOTIVATION TO READ:
AN EXAMINATION OF THE CAUSAL RELATIONSHIP

by

MATTHEW PIERCE QUIRK

Major Professor: Paula Schwanenflugel
Committee Members: Deborah Bandalos
                  James Baumann
                  Nancy Knapp
                  Paul Schutz

Approval for Electronic Submission:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
August 2005
DEDICATION

This project was supported by the Interagency Education Research Initiative, a program of research managed jointly by the National Science Foundation, the Institute of Education Sciences in the U.S. Department of Education, and the National Institute of Child Health and Human Development in the National Institutes of Health. Funding for the project was provided by the National Institute for Child Health and Human Development HD407-03 and the National Science Foundation.
ACKNOWLEDGMENTS

Thanks to my wonderful committee for all of their help and support: Paula Schwanenflugel, Deborah Bandalos, James Baumann, Nancy Knapp, and Paul Schutz.
TABLE OF CONTENTS

ACKNOWLEDGMENTS........................................................................................................v

CHAPTER

1  INTRODUCTION..................................................................................................1
   Key Motivational Constructs and Their Influence on Skill Development........3
   Existing Research on the Motivation/ Skill Development Relationship...11
   Limitations of Prior Research.....................................................................17
   Measurement of Motivation to Read...........................................................19
   Implications for the Current Investigation................................................24
   Models Regarding the Relationship between the Development
   of Reading Skill and Motivation to Read..............................................27

2  STUDY 1.........................................................................................................32
   Method........................................................................................................35
   Results........................................................................................................41
   Discussion...................................................................................................57

3  STUDY 2.........................................................................................................66
   Method........................................................................................................70
   Results........................................................................................................75
   Discussion...................................................................................................81

4  GENERAL DISCUSSION...............................................................................83

REFERENCES...........................................................................................................91
### APPENDIX

<table>
<thead>
<tr>
<th></th>
<th>DESCRIPTION</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>ORIGINAL 28 ITEMS FROM THE MRP-A</td>
<td>103</td>
</tr>
<tr>
<td>B</td>
<td>TEACHER RATINGS OF STUDENT MOTIVATION TO READ (TRSM)</td>
<td>110</td>
</tr>
</tbody>
</table>
CHAPTER 1
INTRODUCTION

In the past, much of the research on motivation related to education examined the notion of achievement or academic motivation as a broad construct generalizing across all domains in a child’s academic experience. However, more recently research in the area of academic motivation has shown evidence that motivation is much more specific than was once thought (Wigfield, 1997). These findings have shifted research in the area of academic motivation to a more domain specific approach. Instead of examining a student’s overall motivation to learn, a more specific approach is taken examining their motivation in many different domains (i.e. math, science, reading, social studies, etc.). In fact, research has shown that as early as Kindergarten, young children have domain specific beliefs about their abilities in math and reading that can have a profound effect on their motivation to learn in these respective academic domains (Eccles, Adler, Futterman, Goff, Kaczala, Meece, & Midgley, 1983; Marsh, Craven, & Debus, 1991). The most influential experiences a young student has in school are related to reading. In the primary grades, approximately half of each day is spent working on literacy based activities (Ficken & McClellan, 2002). Therefore, the motivational nature of early reading instruction has a significant impact on every young student’s view of their ability to succeed in school, which can impact their school experiences for the remainder of their academic careers (Poskiparta, Niemi, Lepola, Ahtola, & Laine, 2003). Enhancing motivation to read early in a student’s academic career is important for several reasons: First, children who are motivated to read are more likely to spend more time reading, which has been directly linked to improved reading achievement (Taylor, Frye, &
Maruyama, 1990). Second, scales of reading motivation in upper elementary school children account for approximately 10% of the variance in reading performance measures (Wigfield, Wilde, Baker, Fernandez-Fein, & Scher, 1996). Thus, students who experience motivating instruction in reading relatively early in their academic careers will be more likely to develop into good readers and are also more likely to hold a positive overall academic self-concept.

The impact that motivation to read can have on reading achievement and skill development is best illustrated by the well-known Matthew effects in reading (Stanovich, 1986), which is the finding that the gap between good and poor readers widens as struggling readers avoid reading and skilled readers seek out additional opportunities to read. As good readers continue to gain reading skill through extra practice, the poor readers’ skill level remains stagnant because they avoid engagement in reading tasks. Studies have suggested that good readers read approximately five times as many minutes per day as average readers and nearly 200 times as much as poor readers (Anderson, Wilson, & Fielding, 1988). Beyond the effect that amount of reading has on skill development, research has also demonstrated that motivated readers tend to choose more challenging reading materials, persevere when reading is difficult, cognitively process reading materials more deeply, and show higher levels of reading comprehension (Allington, 1986, 1991; Anderson, Wilson, & Fielding, 1988; Grolnick & Ryan, 1987; Hidi, 1990; Morrow, 1992; Schiefele, 1991; Smith-Burke, 1989; Taylor, Frye, & Maruyama, 1990; Tobias, 1994).

The impact that student motivation to read has on the development of reading skill is not a topic that just interests researchers but one that interests teachers too. As noted by Gambrell (1996), “The results of a national survey conducted by the National Reading Research Center revealed that this is a question of great interest to teachers (O’Flahavan, Gambrell, Guthrie, Stahl, & Alvermann, 1992). Out of 84 reading topics, teachers identified “creating interest in reading” as the top priority for reading research.
Three other topics related to motivation appeared in the top 10: increasing the amount and breadth of children’s reading; developing intrinsic desire for reading; and exploring the roles teachers, peers, and parents play in increasing children’s motivation to read (p. 15).”

Researching motivation to read to better understand the effect that it has on the development of skilled and strategic readers is a topic of interest for both reading researchers and teachers and could have significant implications for students in their classrooms.

Key Motivational Constructs and Their Influence on Skill Development

Wigfield and Guthrie (1997) have demonstrated that motivation to read is not a single construct, but rather it is a conglomerate of multiple constructs that interact within individuals at many different levels. Their research has identified eleven distinct constructs that they believe comprise an individual’s motivation to read including reading self-efficacy, reading challenge, reading curiosity, reading involvement, importance of reading, reading work avoidance, competition in reading, recognition for reading, reading for grades, social reasons for reading, and compliance (reading because of an external goal or requirement). This collection of constructs represents the broadest conceptualization of reading motivation. Other researchers have developed a narrower conceptualization of reading motivation that is comprised of fewer constructs or have focused their attention on a single construct that they view as central to impacting reading achievement (Baker & Scher, 2002; Chapman & Tunmer, 1995; Gambrell, Palmer, Codling, & Mazzoni, 1996; McKenna & Kear, 1990). For the purposes of the current research, I have narrowed my focus on three constructs that I view as central to the relationship between motivation to read and reading skill development: reading self-concept, value for reading, and reading goals.

Self-efficacy is defined as, “People’s judgments of their capabilities to organize and execute courses of action required to attain designated types of performances (Bandura, 1986, p.391).” In other words, self-efficacy is a personal assessment of one’s
ability to successfully accomplish a task at a specified level of proficiency, or can I do it? A person’s self-efficacy for a given task has been shown to influence behavioral instigation, direction, effort, and persistence, which hold obvious implications for skill development (Bandura, 1986; Locke & Latham, 1990; Weiner, 1985). People that have low self-efficacy for a given task tend to put forth little effort or avoid it altogether, while those that feel capable are more likely to actively participate and persist in the face of difficulty (Schunk, 1991). Thus, a person’s self-efficacy in a given domain, such as reading, plays a significant role in their motivation to actively engage in activities in that domain, which is a prerequisite to skill development.

Even as early as kindergarten, students have domain specific ability beliefs (Marsh et al., 1991). Although these early beliefs may or may not apply specifically to reading, the early emphasis on learning to read in school and at home would surely focus these beliefs on reading by the end of the 1st grade. The fact that students develop domain specific ability beliefs in their early elementary years makes early experiences with reading critical to the development of healthy reading-efficacy beliefs.

A domain specific derivative of Bandura’s (1977) self-efficacy construct, Wigfield and Guthrie (1997) have proposed that an influential aspect of students’ motivation to read is reading self-efficacy. They define reading efficacy as a person’s belief that they can be successful in reading. This definition of self-efficacy is slightly different from Bandura’s in that it focuses on broader expectations of future success in reading rather than task-specific beliefs pertaining to a person’s judgment of their ability to complete a specific task at a designated level of performance, which is more “past” oriented. As a result of this difference, Wigfield and Guthrie’s reading efficacy bears more of a resemblance to reading self-concept than self-efficacy as it was originally conceptualized by Bandura.

In their research, Wigfield and Guthrie (1997) found children’s level of reading efficacy, combined with reading curiosity and involvement, to be a determinant of the
amount and breadth of reading done by students outside of school. They also found a
strong correlation (ranging from .49 - .51) between students’ reading efficacy and their
willingness to read challenging texts. This is important when considering the impact that
choosing challenging texts can have on a child’s skill development and overall reading
achievement. Guthrie and Wigfield (2000) discuss the reciprocal relationship that
challenging texts can have on cognitive skill development in children. Basically, they
argue that motivation guides cognition and language use, where a motivated reader is
more likely to persist when reading difficult texts, which further develops their cognitive
skill; whereas, a less motivated reader would likely give up prior to the development of
those same cognitive skills. Schunk (1991) gives a good conceptualization of what is
meant by the cyclical nature of motivation and achievement, relating it to self-efficacy
beliefs.

“As they (students) work at the task, they engage in activities they believe will
lead to goal attainment: attend to instruction, rehearse information to be remembered,
expend effort, and persist. Self-efficacy is substantiated as learners observe goal progress,
which conveys they are becoming skillful… Heightened self-efficacy sustains motivation
and improves skill development (pp. 213).”

Research has also demonstrated that students with higher levels of reading
efficacy also tend to be more strategic when they encounter difficulties with reading
(Guthrie et al., 1996). Chapman and Tunmer (2003) explain this reciprocal relationship
by proposing that, “just as self-efficacy beliefs are said to influence the nature and extent
of strategic learning behavior, so too are self-efficacy beliefs themselves influenced by
the extent to which strategic behavior and specific strategies are fruitful in contributing to
desirable task outcomes (p.12).” Because becoming skillful in reading includes becoming
more strategic, this holds significant implications on the relationship between reading-
efficacy and reading skill development.
There are also developmental changes in reading efficacy as students progress through their elementary school years. Students’ reading efficacy increases over time; however, some students decrease in how much they believe they can do when compared with their peers (Wigfield & Eccles, 2000). This increase in reading efficacy is most likely the result of the students’ awareness of their own increases in skill level as they get older, but it seems that at the same time they are beginning to believe that they are not progressing as fast as their peers. Eccles et al. (1998) found that students’ self-efficacy beliefs became more closely tied to indicators of actual performance as students get older because students become more sophisticated at cognitively processing and assessing the evaluative feedback they receive from teachers, which is consistent with the literature on academic self-concept (Skaalvik and Hagtvet, 1990; Wigfield, et al., 1997).

Because most of the existing research on the relationship between self-beliefs and achievement have focused on self-concept rather than self-efficacy, it is useful to point out some of the key similarities and differences between the two. Bong and Skaalvik (2003) reviewed the literature on self-efficacy and self-concept pointing out key similarities and differences between the two constructs. Both self-efficacy and self-concept treat perceived competence as the central element in construct definition and assessment. They both are influenced by prior mastery experience, social comparisons with peers, and evaluative feedback from significant persons (i.e. teachers and parents). Self-efficacy and self-concept are both domain specific constructs and have been found to predict motivation, emotion, and performance in various academic domains. One key difference between self-efficacy and self-concept is the specificity of a person’s judgments about themselves. Self-efficacy is conceptualized at a task specific level (“I can read ten pages of this book today”); whereas, self-concept is viewed at a more global level (“I am a good reader”). Therefore, self-concept tends to be past-oriented, stable over time, and more resistant to change. Because of the task specific nature of self-efficacy, it tends to be more malleable and is influenced by individual experiences of
successes and failures more than self-concept. Judgments of self-concept rely more heavily on external social comparisons of competence; whereas, judgments of self-efficacy are more influenced internally by achievement of goals and past mastery experiences. It is important to note that many of the published instruments designed to assess motivation to read contain items that essentially assess reading self-concept rather than more task specific items that would assess self-efficacy. For example, as previously mentioned, Wigfield and Guthrie’s (1997) reading efficacy subscale on the MRQ is a better indicator of reading self-concept than self-efficacy due to the fact that the items target global beliefs about how a person views themselves as a reader instead of targeting a person’s beliefs regarding confidence in their ability to execute specific reading tasks. This is important because self-concept is less malleable and may not reflect subtle changes in the self-system resulting from positive experiences in mastering specific skills that self-efficacy items may better detect.

Confidence in one’s ability to succeed alone does not guarantee a person will be motivated to engage in a given task. In addition to high self-efficacy or confidence in one’s ability to succeed, it is also necessary for the person to perceive some value in participating in the activity. Expectancy-value theory relates personal beliefs that one will succeed at a given task to how much the person values success at the given task (Eccles et al., 1983). Expectancy-value theorists would argue that the extent to which a task is valued is key to explaining an individual’s choice, persistence, and performance on it (Wigfield & Eccles, 2000). Because gains in reading skill are directly related to the amount of reading activity in which students choose to engage (Anderson, Wilson and Fielding, 1988), it is necessary for students to perceive some personal value in becoming a skillful reader.

The idea of subjective task value has been broken down to include multiple levels including interest value, attainment value, and utility value (Eccles, Adler, Futterman, Goff, Kaczala, Meece, & Midgley, 1983). Interest value is defined as how much an
individual likes or is interested in the activity. Attainment value pertains to the importance of doing well at a given activity. Finally, utility value is the perceived usefulness of the activity. Thus, value for a given activity, such as reading, can be established at many different levels, which is reflected on scales designed to assess the value that students’ place on reading. Eccles et al. (1983) claim that a high subjective task-value in combination with a high personal expectation for success will result in activity choice and engagement in a given task, which is vital to skill development.

A student’s achievement goals are related to their skill development and self-efficacy in that they are a benchmark for evaluating whether one has succeeded or failed at a given task (Turner, Husman, and Schallert, 2002). The literature on achievement goals exemplifies a variety of ways in which goals can be conceptualized. In the current study, goals are conceptualized in terms of the goal orientation of the students, examining whether students set goals for their reading and distinguishing what the driving forces are behind the goals that they set. Goals have been shown to direct activity toward actions, regulate effort expenditure, and affect the persistence of action, which all hold implications for the development of reading skill (Locke & Latham, 1994).

The theory on goal orientation is focused on what the driving motives are behind setting achievement goal(s). Most of the literature on goal orientation deals with the concepts of performance and mastery goals. Students who hold a performance goal orientation focus on demonstrating competence in comparison with others, whereas students who hold a mastery goal orientation are focused on the development of personal competence or task mastery. Elliot (1999) has summarized much of the theory resulting from past research on these two types of students. Performance goals are hypothesized to be linked to a negative set of processes and outcomes including withdrawal of effort in the face of failure, surface processing of study material, and decreased task enjoyment; whereas, mastery goals are hypothesized to be linked to a positive set of processes and outcomes including persistence in the face of failure, deep processing of study material,
and enhanced task enjoyment (Ames, 1992; Dweck & Leggett, 1988; Nicholls, 1989; Nolen, 1988). Ames (1992) also links mastery goals with the belief that effort leads to success (Weiner, 1979), a preference for challenging work and higher levels of risk taking (Ames & Archer, 1988; Elliot & Dweck, 1988), and intrinsic interest in learning. In contrast, performance goal orientations have been associated with avoidance of challenging tasks, negative affect following failure, and a judgment that one lacks ability.

In the past, performance and mastery goal orientations have been conceptualized as a dichotomy; however, more recently researchers have argued that a student could be motivated to reach a goal for both mastery and performance based reasons (Morrone & Schutz, 2000). For example, a student may be predominantly motivated to do well in her reading class because she generally enjoys reading and sees the utility in becoming a good reader, but at the same time may try to outperform her best friend by reading more Accelerated Reader books throughout the school year. Studies by Zimmerman and Kitsantas (1997, 1999) have shown that students who were initially motivated to acquire writing revision skills for mastery reasons (termed process goals) and then later shifted their focus to performance goals (termed outcome goals) for utilizing these new skills to attain a given outcome outperformed groups of students who were exclusively using mastery goals. For this reason, both performance and mastery goals for reading will be interpreted as positive indicators of a student’s motivation for reading in the current investigation.

This recent reconceptualization positing that performance and mastery goals can work together is predicated on another recent distinction in the literature on achievement goals. Elliot (1999) and other goal theorists have proposed that the performance/mastery goal dichotomy be revised to include approach and avoidance motives. This trichotomous framework includes mastery goals, as they were traditionally viewed; however, the traditional conceptualization of performance goals is split to include performance-approach and performance-avoidance goals. The distinction between
performance-approach and performance-avoidance motives lies in the valence behind these motives. Persons with performance goals and high perceptions of personal competence try to demonstrate their competence relative to others, which would constitute behavior consistent with performance-approach goals, and persons with low perceptions of competence try to avoid demonstrating their lack of competence relative to others, which is consistent with performance-avoidance goals. Presumably, if performance and mastery goals were to work in combination with one another, the performance goals would have to be performance-approach goals. Because the current study is interpreting both performance and mastery goals to be indicators of positive motivated behavior, all of the items on the goal subscale dealing performance goals were designed to be indicators of performance-approach goals, which are viewed as more positive motivators than performance-avoidance goals in the literature.

Research on motivation to read over the past decade has demonstrated that it is multidimensional, domain specific, and changes developmentally with cognitive maturation and influential experiences in academic settings (Eccles et al., 1998; Marsh et al., 1991; Wigfield, 1997). When considering the impact that skill development has on motivation, there seems to be a few key motivational constructs that are most influential. The development of skill would theoretically have the largest impact on competence beliefs, whether it is at a task specific level (self-efficacy) or more general beliefs about one’s ability in a given domain (self-concept). In addition, task-value is theoretically influential to the development of skill due to the finding that people are more likely to actively engage and persist in activities that they see some personally relevant value in completing (Eccles et al., 1983). Finally, the goals that people set and their goal orientation impact skill development because they are used as a benchmark for evaluating one’s performance at a given task or in a given domain and help to explain the driving force behind a person’s motivation in a given domain (Ames, 1992; Dweck & Leggett,
1988; Nicholls, 1989; Nolen, 1988). As a result of these theoretical relationships with skill development, these three areas of motivation will be the focus of this investigation.

**Existing Research on the Motivation/ Skill Development Relationship**

Although there has been considerable interest in how self-beliefs and motivation are related to achievement in reading (Chapman, 1988; Chapman & Tunmer, 1997; Cox & Guthrie, 2001; Schunk & Rice, 1993; Shell, Brunning, & Colvin, 1995), to my knowledge, no empirical research has investigated the causal relationship between the development of specific reading skills and *motivation to read*. However, there is an extensive body of research that has investigated the causal relationship between academic self-concept and academic achievement, some of which is specific to language arts (Guay, Marsh, & Boivin, 2003; Helmke & van Aken, 1995; Marsh & Koller, 2004; Marsh & Yeung, 1997; Mujis, 1997; Skaalvik & Hagtvet, 1990; Skaalvik & Valas, 1999).

Research in the areas of academic self-concept and academic achievement have established a clear relationship between the two, with correlational evidence positively and moderately correlated (Hansford & Hattie, 1982). However, as Byrne (1996) has pointed out, simply establishing a correlational relationship between the two does not shed light on the causal relationship between them, which has theoretical and practical implications. Beyond the studies that have established a correlational relationship between academic self-concept and academic achievement, there have been a series of studies that address the issue of causality through the use of path analysis and structural equation modeling (Guay, Marsh, & Boivin, 2003; Helmke & van Aken, 1995; Marsh & Koller, 2004; Marsh & Yeung, 1997; Skaalvik & Hagtvet, 1990; Skaalvik & Valas, 1999). These studies have yielded mixed results with three different patterns of relationships emerging: a) increases in achievement "causing" increases in self-concept (skill development model), b) increases in self-concept "causing" increases in achievement (self-enhancement model), and c) a cross-lagged relationship where
increases in achievement impact increases in self-concept and increases in self-concept impacting increases in achievement (reciprocal model).

Proponents of the skill development model of the achievement/ self-concept relationship argue that changes in self-concept variables are primarily the result of academic achievement (Calsyn & Kenny, 1977). To test this model, Skaalvik and Valas (1999) conducted a two-wave (two time points), three variable (self-concept, achievement, and motivation) longitudinal study with 3 cohorts of Norwegian students in the 3<sup>rd</sup>, 6<sup>th</sup>, and 8<sup>th</sup> grades. They examined the three variables in a domain specific manner looking at each variable in both math and reading. They defined motivation in terms of investment (persistence in the face of difficulty) and interest (how much they liked the given subject matter and related activities). They defined self-concept in terms of the general feelings of doing well or poorly in mathematics and language arts. Achievement was measured through the use of achievement tests containing developmentally appropriate items and teacher ratings of student ability in math and language arts respectively. The data was analyzed by conducting path analyses opening all possible paths between all three variables from time 1 to time 2 in each age group (separate analyses for both math and language arts). Then the cross-lagged path coefficients that were not statistically significant were eliminated one by one. The results showed support for a skill development model with the students in the 3<sup>rd</sup>/ 4<sup>th</sup> grade cohort, where achievement in 3<sup>rd</sup> grade significantly impacted self-concept in 4<sup>th</sup> grade but no relationship was found between self-concept at 3<sup>rd</sup> grade with achievement in 4<sup>th</sup> grade. This may suggest that self-concept is still relatively unstable in the primary grades, which would limit the possibility of self-concept contributing to future achievement in a given domain. Also, these results suggest that achievement, which was relatively stable across time points, is a major contributor to students’ subsequent self-concept and may be the primary indicator students use to form their self-concept in a particular domain. These results were similar to other studies of elementary aged children that support the skill
development model of the self-concept/achievement relationship (Chapman & Tunmer, 1997; Helmke & van Aken, 1995). However, the results of this study also showed similar effects (skill development effects) for both of the older cohorts of students, which is contradictory to past results of similar studies that showed a developmental trend between self-concept and achievement as self-concept became more stable and influential in late elementary, middle, and high school students (Marsh, 1990; Skaalvik & Hagtvet, 1990).

The developmental perspective on the self-concept/achievement relationship posits that young children have very positive academic self-concepts that may appear to be biased in relation to external indicators of academic achievement (Marsh, 1990; Wigfield & Karpathian, 1991). As children move into late elementary and early middle school, their self-concepts become more closely tied to indicators of their actual performance level. Harter (1999) proposed a possible explanation for this developmental trend of changing self-concept. First, as children mature, they become more cognitively sophisticated which enables them to make more realistic judgments about their ability by comparing their performance with that of their peers. Second, older children are better able to interpret evaluative standards of others (teachers and parents), which leads to a less egocentric view of the self. This developmental trend in the self-concept/achievement relationship was supported by a study conducted by Skaalvik and Hagtvet (1990). In this study, they examined the relationship between general academic self-concept, defined as the general feeling of doing well in school, and general academic achievement, defined as a composite of teacher rankings of students in reading comprehension, mathematics, and social/natural sciences for two cohorts of children across two school years. They found that the model for the younger cohort of students (grades 3/4) supported a skill development model with achievement at time 1 making a significant contribution to self-concept at time 2 and self-concept at time 1 not making a significant contribution to achievement at time 2. In addition, evidence from this study supported a reciprocal effects model for the older cohort of students (grades 6/7) with
self-concept at time 1 significantly contributing to achievement at time 2 and
achievement at time 1 significantly contributing to self-concept at time 2. Although this
study suggests that as students get older their self-concepts become more stable and begin
to influence subsequent achievement, these results should be interpreted with caution for
a few reasons. First, this study examined the relationship between global academic self-
concept and a composite of academic achievement across multiple subjects, which does
not take into account the widely accepted notion that self-concept is domain specific
(Eccles, Wigfield, Harold, & Blumenfeld, 1993; Marsh et al., 1991; Marsh & Hattie,
1996; Wigfield et al., 1997). Second, using a composite of teacher ratings across three
academic subjects as the only measure of academic achievement is problematic in that it
also overlooks the possible impact of domain specific academic achievement on this
relationship. Also, using teacher ratings as the sole indicator of academic achievement
does not take into account any achievement related experiences encountered by students
such as tests or grades, which are major contributors to a person’s self-concept (Bong &
Skaalvik, 2003).

Studies conducted by Guay, Marsh, and Boivin (2003) and Kurtz-Costes and
Schneider (1994) found no evidence to support this developmental trend between
academic self-concept and achievement. In both studies, the evidence supported a
reciprocal model (bidirectional relationship) for students in the 2nd, 3rd, and 4th grades,
with achievement impacting self-concept and vice versa in all age groups. Kurtz-Costes
and Schneider (1994) examined the relationship between general academic self-concept
and school grades (as an indicator of academic achievement) in German students age 8 at
time 1 and age 10 at time 2. They found a bidirectional relationship with academic self-
concept at time 1 significantly influencing achievement at time 2 and achievement at time
1 significantly impacting academic self-concept at time 2.

A study conducted by Guay, Marsh, and Boivin (2003) found similar results.
Their study was a multiwave (multiple time points) multicohort design that examined the
relationship between general academic self-concept and academic achievement longitudinally with three different cohorts of students ranging in age from 2\textsuperscript{nd} grade (cohort 1), 3\textsuperscript{rd} grade (cohort 2), and 4\textsuperscript{th} grade (cohort 3). Self-concept and achievement for each cohort was measured in three waves at one year increments, with the students in 4\textsuperscript{th} grade (cohort 1), 5\textsuperscript{th} grade (cohort 2), and 6\textsuperscript{th} grade (cohort 3) at the time of the final wave of data collection. Their initial analyses showed evidence to support a reciprocal effects model over the first two waves of data collection for all three cohorts. That is, they found achievement at time 1 significantly impacted self-concept at time 2 and self-concept at time 1 significantly impacted achievement at time 2 for all three cohorts of students. Interestingly, this relationship was not the same for the data at time 2 with time 3 in the full model (examining all three time points at once). In all three cohorts, the data from time 2 and time 3 supported a self-enhancement model, which posits that self-concept will influence subsequent achievement but that achievement will not influence subsequent self-concept. However, when they took out the data from time 1 and only analyzed data from times 2 and 3, the reciprocal relationship re-emerged with self-concept contributing significantly to subsequent achievement and achievement contributing significantly to subsequent self-concept. It should be noted that in both of the aforementioned studies self-concept and achievement were measured across multiple domains, which may impact the relationships found in these studies compared to those that examined these relationships at a domain specific level.

The logic of the reciprocal effects model for older children argues that as self-concept evaluations become more accurate with age, they become better predictors of actual academic achievement (Marsh & Koller, 2004; Marsh & Yeung, 1997). Studies conducted by Marsh and Koller (2004) and Marsh and Yeung (1997) on middle and high school students provided support for this logic. Both studies provided evidence supporting a reciprocal effects model for the self-concept/achievement relationship in older students with self-concept contributing to subsequent achievement and vice versa.
When reviewing all of the studies on the self-concept/achievement relationship, I was only able to find one study that dealt with causation related specifically to reading self-concept and reading achievement. Chapman and Tunmer (1997) conducted a longitudinal study that examined the relationship between reading self-concept and reading achievement for students in their 1st-3rd years of schooling (equivalent to K-2). They found that there was a low to moderate concurrent correlation between reading achievement and reading self-concept \( (r = .11, .21, .35) \) respectively over the three time periods, suggesting that a significant relationship between the two begins to form in the 1st grade and becomes stronger in 2nd, which is sooner than previous research has suggested (Nicholls, 1978; Stipek, 1993). They also found that reading self-concept becomes more stable in the 2nd and 3rd years of schooling \( (r = .49) \), but that there is still some instability at this age. Reading achievement showed considerably higher levels of stability across all three time points \( (r = .69 \text{ from T1 - T3}, r = .91 \text{ from T2-T3}) \). Using path analysis, they found no relationship between RSC and reading achievement from T1-T2, but they did find a relationship supporting the skill development model from T2-T3, showing that reading achievement at time 2 significantly contributed to self-concept at time 3 \( (\text{path loading of .27}) \). This study clearly indicated that the relationship between reading achievement and reading self-concept begins to emerge between the 2nd and 3rd years of schooling and that evidence provided support for an emerging skill development relationship. This was contradictory to findings from other studies which have suggested that self-concept does not begin to influence achievement until somewhere between grades 3 - 5 (Kurtz-Costes & Schneider, 1994). Taken together, the results of these studies provide contradictory evidence regarding the self-concept/achievement relationship.
Limitations of Prior Research

While there is more evidence to support a skill development model with children in the primary grades (Chapman & Tunmer, 1997; Helmke & van Aken, 1995; Skaalvik & Hagtvet, 1990; Skaalvik & Valas, 1999), the studies conducted by Guay, Marsh, and Boivin (2003), Mujis (1997), and Kurtz-Costes and Schneider (1994) support a reciprocal model for primary aged children. Theoretical work on the self-concept/achievement relationship, as well as limitations of previous empirical studies may help to explain this contradictory evidence and provide guidance for future research.

One limitation to previous research on the self-concept/achievement relationship pertains to the rationale behind using self-concept as the single indicator of academic motivation. These studies have operated under the assumption that self-system variables (self-concept, self-efficacy, self-worth, and attributions) are thought to influence achievement mainly by means of their effect on motivation (Bandura, 1986; Chapman & Tunmer, 1997; Schunk, 1991). However, as previously discussed, research on motivation has shown that it is multifaceted (Wigfield & Guthrie, 1997), which leaves open the possibility that some influential aspects of a student’s motivation on their achievement have been ignored. Eccles et al. (1983) have pointed out that it is possible for a person to hold high expectations for success in a given domain, yet still be unmotivated to actively engage in activities within that domain because they are not personally valued. Moreover, research in the area of goal theory has shown that the use of academic goals can be a powerful influence on directing student behavior and is a key element to the hallmark of motivated student behavior, self-regulated learning (Yowell & Smylie, 1999). Wolters (2003) states that self-regulated learners tend to be highly self-efficacious, hold a mastery goal orientation for learning, and view material to be learned in school to be valuable for them to learn. If research is being conducted to examine the relationship of self-concept and achievement via increased motivation for learning, it would make sense to examine the influence of self-concept as well as other motivational variables, such as
task-value and goal orientation that should theoretically contribute to this relationship. Thus, it is possible that assessing motivation more broadly (i.e. including more than just self-concept or self-efficacy as indicators of motivation) may increase the accuracy with which we are able to assess student motivation as well as our ability to examine the relationship between motivation and achievement.

Another limitation with many of these studies pertains to an issue of specificity. Many of the studies reviewed used broad measures of achievement, such as grades or teacher ratings across a variety of academic domains. Others used broad achievement tests that assessed a wide range of developmentally appropriate knowledge in a variety of domains. Marsh (1992) conducted a study that provided evidence for the importance of considering content specificity when examining the self-concept/achievement relationship. They found correlations of .45 - .70 between content specific self-concepts and achievement, which are higher than previously reported correlations between general academic self-concept and general academic achievement. In a meta-analysis of achievement/ self-belief studies, Valentine, DuBois, and Cooper (2004) concluded that examining these relationships at the domain specific level has shown to be more effective than looking at the achievement/ self-concept relationship at the global level. Chapman and Tunmer (1997) discussed how their measures of achievement were very stable across measurement time points, which has statistical implications when examining constructs that have theoretically modest effects on one another. Because these broad measures of achievement are very stable, there is not much variance in achievement scores. Skaalvik and Valas (1999) point out that restricted variance in achievement measures may mask the influence that self-concept has on subsequent achievement. Thus, it is possible that increasing the specificity of measurement from broad achievement scores or grades to that of specific skills may introduce more variance in measurement that may help to “unmask” these relationships.
Examining the development of domain specific skills as a more detailed indicator of achievement within a particular domain presents a second problem of previous research worthy of consideration. Many of the previous studies of the self-concept/achievement relationship have examined these constructs across multiple grade levels for different cohorts of students; however, the skills necessary to achieve in a given domain at different grade levels changes qualitatively. The skills necessary to achieve at a high level in reading at the second grade, where decoding skills heavily influence “learning to read”, are qualitatively different from the skills necessary to achieve at a high level in reading at the third grade level, where the emphasis shifts to comprehension for “reading to learn”. Thus, if an investigation were to increase its level of specificity to examine the relationship between domain specific skill development and self-beliefs, it would have to do so over the course of a single year where the skills necessary to achieve remain qualitatively similar. Such an investigation would be of particular significance during “critical periods” in skill development where students are making large gains in a particular skill over the course of a single year in a single academic domain.

Measurement of Motivation to Read

Part of the issue in discerning the relationship between motivation to read and the development of reading skill is the problem of measuring reading motivation. Ironically, despite the amount of research that has been done on students’ motivation for reading, there are very few published instruments available to assess motivation to read. Of the measurement tools that have been published, all of them are designed as self-report measures for which there is limited validity evidence to support their use with young children. How can a topic that has generated this much interest in teachers and researchers and that has such serious implications on the development of avid and skillful readers remain so difficult to assess? I will address these issues and suggest some possible solutions in the next section of this paper.
The major problem with instruments used to assess student *motivation to read* is a theoretical one, as there seems to be no clear definition of what *motivation to read* actually consists. All of the measures reviewed propose to measure, at least partially, young children’s motivation for reading.

Wigfield and Guthrie’s (1997) Motivations for Reading Questionnaire (MRQ) is a self-report Likert-type scale and is the broadest instrument designed to assess *motivation to read* for students in the 3rd grade or older. It is comprised of eleven subscales proposed to assess a conglomerate of motivational constructs related to reading. However, Watkins and Coffey (2004) dispute Wigfield & Guthrie’s MRQ research, stating that there is evidence that reading motivation is multidimensional but that there is no empirical support for the proposed eleven dimensions underlying the MRQ. They were unable to replicate the factor structure from past research on the MRQ and suggest that it be revised. One of their major criticisms is that the MRQ is too broad, attempting to cover a lot of different constructs with too few items. Many of the subscales in the MRQ consist of only three or four items. Fabrigar, MacCallum, Wegener, and Strahan (1999) suggest that subscales should be comprised of at least five items. This becomes problematic when you are attempting to use eleven subscales to assess young children’s motivation for reading because it would require 55 items, which would be difficult for young children to reliably complete. Other measures of *motivation to read* propose fewer numbers of constructs that make up a student’s reading motivation, which may allow for more reliable factor structure analyses but also may lead to construct underrepresentation. Another unrelated criticism of the MRQ is that it requires the students to read the items themselves, which can be problematic for students who are early in the reading developmental process or are having difficulties learning to read. This problem could cause unreliability resulting from problems with decoding and reading comprehension that are unrelated to factors of *motivation to read*. 
The Reading Self-Concept Scale (RSCS) was developed by Chapman and Tunmer (1999) to address motivational issues specific to children’s self-concepts as readers. The RSCS specifically examines three subcomponents within the domain specific area of academic reading self-concept: perceptions of competence, perceptions of difficulty, and attitudes toward reading (Chapman & Tunmer, 1995). Perceptions of competence refer to beliefs regarding ability and proficiency in reading tasks (very similar to reading self-efficacy). Items from this scale use questions like “Can you work out hard words by yourself when you read?” that are aimed at assessing whether students are confident in their reading abilities. The perceptions of difficulty subscale refers to the student’s beliefs that reading tasks are challenging or problematic. “Are the books you read in school hard for you?” is an example question from this subscale, which also gets at a student’s beliefs in their ability to read. Attitude toward reading uses questions like “Do you look forward to reading?” to examine the affective component addressing issues of value, usefulness, and the general enjoyment (or lack thereof) of reading related activities. However, the questions from this subscale don’t really address specific reasons that students may or may not feel that reading is important. They only ask very general questions that assess if they value reading, not why they value reading, which may be important. One of the major problems with this scale is that the questions in all subscales are very general and do not provide much detailed information about why a student is motivated or unmotivated to read. Another major problem with this scale as a measure of a student’s motivation to read is that it is only designed to address issues related to self-concept and does not take into consideration other motivationally relevant constructs such as reading goals or intrinsic/extrinsic reasons for reading. Again, many of the constructs that this measure is designed to assess are similar to those measured by the MRQ; however, many differences remain between the definitions used by each.

The Motivation to Read Profile (MRP; Gambrell, Palmer, Codling, & Mazzoni, 1996) is made up of two basic instruments. The MRP consists of The Reading Survey, a
self-report instrument similar to the MRQ and the Conversational Interview, which is a qualitative measure designed for individual administration. It is designed to assess two specific dimensions of motivation to read: self-concept as a reader and value of reading (10 items each). A strength of this scale is that it focuses attention on the two most important parts of a student’s motivation to read, self-concept and value. It also provides information about intrinsic and extrinsic motivation to read with value questions like, “People who read a lot are... a) very interesting, b) interesting, c) not very interesting, d) boring”. Some questions on the self-concept scale also address these issues. For example, “I worry about what other kids think about my reading... a) never, b) once in a while, c) almost every day, and d) every day”. These types of questions are useful because they provide information regarding a student’s self-concept and value for reading but also provide information regarding intrinsic and extrinsic motivations for reading, essentially maximizing the amount of information that you can get out of a 20-question format. One limitation of the scale is that it doesn’t ask any questions that would provide information about whether students set goals related to their reading activity and whether these goals are driven by a mastery or performance orientation, which has implications for both motivation and skill development. This instrument takes a much more focused approach to defining motivation to read than the MRQ, but also manages to capture more aspects of a student’s motivation to read that the RSCS.

Despite the fact that all of these measures take a different approach to defining what makes up a student’s motivation to read, all of these instruments have similar intended uses when drawing interpretations from test scores, which creates serious validity issues for these measures. These uses include providing classroom teachers with information about their students’ reading motivation to inform classroom practices that might target problem areas that are highlighted by the scales, developing interventions designed to improve motivation for reading, and tracking changes in motivation over the course of the school year. Because the uses of these instruments are so similar, it would
make sense that they should be measuring similar constructs to inform these decisions, but they do not. Some instruments, like the MRQ, take a very broad approach to defining what makes up motivation to read, while others, like the RSCS and MRP, have a much narrower focus looking only at specific component(s) of overall reading motivation. This would not be as much of a problem if the developers of these instruments had made their rationale clear for choosing the aspects of motivation that their respective instruments are designed to measure.

A second problem with the existing measures of motivation to read is that they almost exclusively rely on self-report. This is problematic for a few reasons. First, it has been well established by research that self-report measures are not always reliable and valid, especially with young children (Furlong & Smith, 1994). Young children have an exceptionally difficult time making assessments of their own abilities and feelings, which requires a solid understanding of one’s self-concept (Furlong & Smith, 1994; Harter, 1983). Young children tend to be overly optimistic, equate effort with ability and more likely to respond in a socially desirable way to please adults administering the test (Lepola, 2004). Consequently, other forms of evidence are needed to provide validity evidence for these self-report measures. This evidence should show that students’ responses to these self-report instruments are consistent with motivated behaviors exhibited by those same students in various situations where they are reading.

Further, because these measures all take the same approach methodologically, it is impossible to determine how much of the variance detected in motivation to read can be attributed to method specific variance. Kane (1992) argues that triangulation, or multiple sources of information that converge on similar findings, is needed for validity evidence. There needs to be a way to look at the behavioral manifestations of motivated readers that would allow comparisons between information collected by self-report with observations of actual reading behaviors. This would be a key addition to the growing body of validity evidence for measurements of motivation to read.
Finally, the MRQ requires students to read the items on the questionnaire, which is problematic. This leaves open the possibility that the instrument could be assessing constructs other than motivation for reading, such as decoding, reading comprehension, and reading fluency. These skills would contribute non-trait specific variance to interpretation of these test scores. Even though the MRQ is designed for older populations of students (3rd grade and above), it does not provide a way to gauge reading motivation for those students who continue to struggle with reading past the third or fourth grade. Further, because of the relative focus in this literature on older children, there is currently a dearth of information regarding motivation to read in younger children who are currently in the process of formulating a motivational stance to reading.

Implications for the Current Investigation

This study will examine the relationship between changes in reading skill and motivation to read in 2nd grade students; therefore, it is necessary to discuss which reading skills are most important to achievement in reading at this level. Skilled word reading provides the necessary foundation for development of subsequent reading comprehension skills (Jenkins, Fuchs, van den Broek, Espin, & Deno, 2003) and is a major determinant of reading ability in the elementary grades (Juel, 1988; Stanovich, 1991). As students become more automatic at decoding individual words, they free up cognitive resources that can then focus on comprehending connected text (LaBerge & Samuels, 1974). Support for this theory of verbal efficiency has been demonstrated through research findings showing a positive relationship between word-reading speed and reading comprehension (McCormick & Samuels, 1979; Perfetti & Hogaboam, 1975). Thus, during the primary years of elementary school, much of the emphasis on early reading instruction is focused on decoding skill. In order to develop decoding skill, instruction is focused on phonics (letter-sound correspondences) as well as other word level decoding skills such as rhyming, language rules (ie. the silent “e” rule), the blending of individual sounds into complete words, and vocabulary building activities so that
words can be recognized and understood. Once decoding skills and strategies have been mastered, reading instruction in the middle and late elementary school years begins to shift from “learning to read” to “reading to learn”, where the emphasis is placed on pulling and integrating ideas from connected text.

Many of the students that are in the 2nd grade are beginning this transition from “learning to read” to “reading to learn”. One area in which these students advance rapidly during their second grade year to aid them with this transition is in the area of reading fluency. Reading fluency has traditionally been defined in reference to “the ability of readers to read quickly, effortlessly, and efficiently with good, meaningful expression” (Rasinski, 2003). Previous research studying reading fluency has defined it in many different ways, which holds implications for how it is assessed. Definitions of reading fluency in previous research has emphasized speed of accurate reading (Torgesen, et al., 2001), others including context-free word reading latencies (Stanovich, 1980; Torgesen et al., 1997), and others emphasizing speed and accuracy of word reading along with intonation or prosodic features of reading (Allington, 1983; Rasinski, 1990). Other definitions have included accuracy, speed, expression, and simultaneous understanding of text (Jenkins et al, 2003). The current investigation defines reading fluency in terms of word reading speed and accuracy both in and out of context. Measurements of word level reading are commonly conducted in fluency research for words in and out of context; however, the two tasks are not identical (Jenkins, et al. 2003). Context-free reading tasks are done in the form of timed reading of word lists, such as the Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 1997). Because the words in these lists are isolated, they are read without the aid of expectancy based context effects such as semantic activation and syntactic rules that aid in word recognition (Posner & Snyder, 1975). Other instruments such as the Dynamic Indicators of Basic Early Literacy Skills or DIBELS (Kaminski & Good, 1998) assess word reading accuracy and rate within the context of coherent passages that allow for these expectancy based effects to aid in word
recognition. Regardless of how fluency is assessed, it seems that a critical period of
development from word-by-word to fluent reading occurs throughout the 2nd grade year
(Schwanenflugel, Hamilton, Kuhn, Wisenbaker, & Stahl, 2004). How well a student is
able to navigate this transitional year and develop the skills necessary to become a fluent
reader has implications for how motivated they are and will continue to be in the area of
reading. Conversely, it is possible that initial motivation to read at the beginning of the
2nd grade year has implications for how engaged they are in tasks designed to develop
these skills. These are the relationships the current investigation will attempt to address.

There are also implications for the current research on the measurement of
motivation to read. These implications are twofold. First, the current research will
contribute much needed criterion validity evidence for the measurement of motivation to
read by examining the relationship between the MRP and teacher ratings of directly
observable behaviors of engaged reading as an indicator of motivated readers. Due to the
fact that many of the existing measures of motivation to read were developed out of an
engagement perspective on reading (Baker & Wigfield, 1999), it seems reasonable to
expect that a good criterion for assessing whether readers are motivated is to have
teachers rate their observations of individual students based on indicators of their level of
engagement when involved in reading related tasks. Second, the current investigation
will attempt to address the issue of construct underrepresentation in measures of
motivation to read by examining the effect of adding items to the MRP that are related to
goals that students set for their reading as well as their goal orientation. This “reading
goals” subscale will be added to the existing items on the MRP that focus solely on
reading self-concept and value for reading. It is hypothesized that these added items will
improve the reliability of the MRP, as well as improve the explanatory power of the MRP
when attempting to characterize a student’s overall level of motivation to read.
Models Regarding the Relationship between
the Development of Reading Skill and Motivation to Read

Through a review of the literature on motivation to read and reading skill development, as well as the existing studies on the academic self-concept/academic achievement relationship, I have identified four models of the reading motivation/fluency skill development relationship worthy of testing. When considering this relationship, there are four possible relationships/hypotheses that emerge.

The Skill Development Model: The model most heavily supported by previous research is the skill development model, which asserts reading skill level will impact subsequent motivation to read with no effect of motivation to read on subsequent reading skill level. Studies supporting a skill development model for students in the primary grades found that achievement level at time 1 (T1) contributed significantly to self-concept at time 2 (T2), but they found no significant contribution of T1 self-concept to T2 achievement level (Helmke & van Aken, 1995; Skaalvik & Valas, 1999; Skaalvik & Hagtvet, 1990). All of these studies tested models over two time points with students ranging in age from 2nd - 4th grade. However, the most convincing evidence to support the skill development model for the current study came from Chapman and Tunmer’s (1997) study of the relationship between reading self-concept and reading achievement, with the evidence showing a significant contribution of reading achievement level in 1st grade to reading self-concept in the 2nd grade and no significant contribution of reading self-concept in 1st grade to reading achievement level in 2nd grade. The evidence from this study suggested that young children have relatively unstable self-concepts that are still changing in the 1st and 2nd grades. These findings are consistent with theoretical work on achievement related self-perceptions that suggest changes in self-concept are primarily due to early experiences (successes and failures) in reading situations (Schunk, 1991). Thus, this body of research supports my first hypothesis that reading skill development
will contribute to subsequent *motivation to read*, but that *motivation to read* will have no significant impact on subsequent reading skill development.

*The Reciprocal Model:* The second model to be tested is a reciprocal effects model that asserts a cross-lagged effect where reading skill level will impact subsequent *motivation to read* and the level of student *motivation to read* will impact subsequent skill development. Despite the aforementioned evidence supporting the likelihood of a skill development model for the skill development/ motivation relationship, there is also evidence to suggest the possibility of a reciprocal relationship between these variables in primary aged readers. Studies supporting the reciprocal model show similar skill development effects from T1 to self-concept at T2 as the previously reviewed studies; however, the studies supporting the reciprocal effects model also show evidence that self-concept at T1 makes a significant contribution to achievement at T2 (Guay, Marsh, & Boivin, 2003; Kurtz-Costes & Schneider, 1994; Mujis, 1997). Of these studies, the most convincing evidence comes from the study conducted by Guay, Marsh, & Boivin (2003) who examined the relationship between academic self-concept and academic achievement with 2nd, 3rd, and 4th grade students. They pointed out that the studies by Skaalvik and Hagtvet (1990) and Skaalvik and Valas (1999) were conducted on Norwegian students where grades are not given until students reach the 7th grade, which may cause a delay in the development of stable self-concepts due to the fact that these students are not given grade-based feedback on their academic performance in the early elementary years. In addition, the study by Chapman and Tunmer (1997) showed that self-concept became increasingly stable from K-2nd grade, which leaves open the possibility that its influence on achievement may become more significant as it continues to stabilize. Therefore, there seems to be enough evidence to suggest the possibility of a reciprocal effect of motivation on skill development and vice versa. Thus, my second hypothesis will test the model showing a reciprocal effect between *motivation to read* and reading skill development over the course of students’ 2nd grade year.
The Self-Enhancement Model: The third model to be tested is a self-enhancement model which asserts motivation to read will impact subsequent fluency skill development; however, fluency skill level will have no impact on subsequent motivation to read. While none of the aforementioned studies found a clear-cut self-enhancement effect, the study conducted by Guay, Marsh, and Boivin (2003) did find a self-enhancement effect in one portion of their model. The results of their study suggested that there was a reciprocal effect between academic self-concept and academic achievement from T1 to T2; however, they found a self-enhancement effect from T2 to T3. While subsequent analyses in Guay et al.’s (2003) study found that the reciprocal effect re-emerged from T2 to T3 when the first time point was removed from the model, their initial findings combined with research that has suggested the possibility that reading self-concept stabilizes earlier than was once thought (Chapman & Tunmer, 1997) leave open the possibility that a self-enhancement model could emerge when examining the relationship between motivation to read and the development of reading fluency skill in 2nd grade readers. Thus, my third hypothesis will test the possibility of a self-enhancement effect.

The Independence Model: The independence model asserts that changes in reading skill development and motivation to read operate independently from one another, which is similar to the null model proposed by Skaalvik and Hagtvet (1990). Previous research related to the current study has focused on either the relationship between general academic self-concept and general academic achievement level or the relationship between reading self-concept and reading achievement across multiple school years. Thus, this is the first investigation to my knowledge of the causal relationship between a more broadly defined motivation to read (including value for reading and goals for reading as well as reading self-concept) and specific reading skill development over the course of a single school year. Therefore, it seems prudent to consider the possibility of an independence model in addition to the two models that have
been supported by related yet distinctly different empirical research. The null hypothesis implied by the independence model states that motivation to read and skill development operate independently of one another, where reading skill at each time point will only influence subsequent reading skill at later time points and motivation to read at each time point will only impact subsequent motivation to read at later time points.

The purpose of this study is to investigate the relationship between second grade students’ motivation to read and the development of a specific reading skill, reading fluency. A better understanding of this relationship has practical and theoretical implications. If motivation to read has a subsequent effect on reading skill development, this would indicate a need to integrate more techniques into early reading instruction that are focused on improving student motivation to read as well as techniques important to the development of specific reading skills. If reading skill development is found to significantly impact subsequent motivation to read, this would indicate a need to focus more on improving specific reading skills without as much emphasis on explicitly building student motivation to read, as this would be a by-product of developing skill.

This study is intentionally focused on the second grade year, as research has shown this to be a critical period of rapid skill development that can take students from word-by-word reading in the beginning of the year to fluent speech-like reading by the end of their second grade year.

A second purpose of the current investigation is to test the validity of an adapted version of The Motivation to Read Profile (Gambrell et al., 1996). One possible reason for the relative lack of research on the relationship between motivation to read and other academic variables is that instruments used to measure motivation to read have very little validity evidence to support interpretations made from scores on these measures. The current investigation will attempt to address some of these validity concerns by examining the relationship between self-reported motivation to read with two criterion variables. First, I will examine the relationship between students’ scores on the adapted
Motivation to Read Profile (MRP-A) with teacher ratings of individual student’s motivation to read based on observable behaviors during reading related activity. Items from the teacher rating scale to be used were selected from an existing instrument, The Teacher Questionnaire on Student Motivation to Read (3rd Edition) by Sweet, Guthrie, and Ng (1998). Second, I will examine the relationship between students’ scores on the MRP-A and their observed on-task behavior during formal reading instructional time. It is assumed that more motivated students will exhibit a higher frequency of on-task behavior during reading instructional time than their less motivated peers.

The present research consists of two studies, a validation study of the MRP-A and a study examining the relationship between second grade students’ motivation to read and the development of reading fluency. Study 1, the validity study of the MRP-A, will consist of two parts. In addition to the aforementioned criterion validity analyses, this study will examine the internal structure of the MRP-A, specifically investigating the validity of each subscale using Messick’s (1989) framework for unified construct validity. Study 2 will be an examination of the relationship between student motivation to read and the development of reading fluency skill during students’ second grade year. Specifically, I will be investigating four possible models of this relationship. First, the skill development model, where students’ level of reading skill at each time point will significantly influence subsequent motivation to read at later time points, with no influence of motivation to read on subsequent reading skill. Second, the reciprocal model, which asserts a bidirectional relationship where motivation to read influences subsequent reading fluency skill level and reading fluency skill significantly impacts subsequent motivation to read. Third, the self-enhancement model, which asserts that motivation to read will impact subsequent reading fluency skill but that reading fluency skill level will not impact subsequent motivation to read. Finally I will test the independence model, where motivation to read and reading fluency skill are hypothesized to operate independently of one another.
CHAPTER 2
STUDY 1

The purpose of this study was to examine the validity of an adapted version of Gambrell et al.’s (1996) Motivation to Read Profile (MRP-A) for identifying second grade students’ motivation to read using Messick’s (1989) framework for unitary construct validity. Messick’s framework incorporates a scale’s domain content relevance, representativeness, and internal structure with its criterion-relatedness. Messick (1989, 1995) has suggested evidence to support a scale’s construct validity should address the following six aspects:

1. **Content**: This aspect refers to the theoretical and empirical basis for specifying the boundaries and structure of the construct domain being assessed;

2. **Substantive**: This aspect incorporates domain content and performance regularity in assessment (Loevinger, 1957);

3. **Structural**: This aspect assesses relationships among items and subscales, the degree of homogeneity within tests or subscales, and the dimensionality of the inter-item structure reflecting the construct’s domain;

4. **Generalizability**: This aspect refers to the consistency or ability to replicate assessment results across population groups, assessment settings, and time;


5. **External**: This aspect involves the degree with which a test score relates to other measures, and whether those relationships are consistent with the theory of the construct being assessed (Loevinger, 1957); and

6. **Consequential**: This aspect is concerned with any possible negative implications for individuals or groups resulting from measurement invalidity. If interpretations made from test scores do not result in high stakes decisions for the individuals or groups being assessed, then this aspect of validity need not be addressed.

These six aspects of construct validity represent a framework that is interrelated; therefore, a combination of evidence from all six areas is needed to substantiate a validity argument for a given measure. Thus, decisions regarding a measure’s validity should not be based on any single aspect; rather, all applicable aspects should be taken into consideration holistically to make an overall validity judgment for a given measure. The analyses conducted in Study 1 were designed to assess the construct validity of the MRP-A, at least in part, along five of these six dimensions. The consequential aspect of construct validity was excluded from consideration since no high stakes decisions are made as a result of a student’s score on the MRP-A.

The validation study of the MRP-A was carried out in two phases. The first phase of Study 1 was designed to revise the 28 item MRP-A along six revision criteria (the full 28 item scale can be found in Appendix A). Each item was evaluated in terms of: 1) their content validity as it relates to motivational theory from its respective domain (reading self-concept, subjective task-value, and academic goals), 2) frequency distributions of responses to each item, 3) inter-item correlations with other items on the same subscale, 4) their impact on the subscale’s internal consistency, 5) how well each item loaded on the primary factor extracted during an exploratory factor analyses of each individual subscale, and 6) whether items were consistent across gender. The first phase of Study 1 was designed to refine the MRP-A through a systematic items-analysis in order to
optimize each subscale’s performance before any external validity examination was carried out in the second phase of this study.

The second phase of Study 1 was designed to evaluate the validity of the revised MRP-A using Messick’s (1989) framework for unified construct validity. Evaluations along the five aspects in this framework will be done for both the full 28-item scale as well as the revised 20-item scale to ensure that revisions made in phase 1 of these analyses does not impact validity judgments made on the MRP-A in phase 2. Beyond the item and subscale level analyses, confirmatory factor-analyses were conducted on the full and revised scales to determine its factor structure. These analyses evaluated whether the items best reflect a three-factor model, representing the three subscales on the MRP-A, or a one-factor model, representing a single motivation to read factor. Because revision decisions were made from time 1 data, all confirmatory factor analyses were conducted using time-2 data. All of the CFA models to be tested were nested models; therefore, the chi-square difference ($\chi^2_{\text{diff}}$) statistic was used to determine whether the three-factor model was an improvement over the one factor model for both the full and revised scales. In addition to the $\chi^2_{\text{diff}}$ statistic, model fit for each model tested was evaluated using the root mean square error of approximation (RMSEA) and comparative fit index (CFI), based on recommendations made by Hu and Bentler (1996). All CFA’s were conducted using the LISREL 8 (Joreskog & Sorbom, 1993) program.

Phase 2 of Study 1 also addressed a major problem facing existing measures of motivation to read, a lack of criterion-related validity evidence. In the past, criterion validity evidence for self-report measures of motivation to read or attitude toward reading scales has come in the form of time watching television or ownership of a library card (McKenna & Kear, 1990). While this type of validity evidence could theoretically be related to motivation to read, they are both weak and indirect criteria to use as indicators of a student’s motivation to read. Because almost all existing measures of motivation to read utilize a self-report format, it is imperative to build criterion-related validity
evidence for these measures showing their ability to predict actual motivated behaviors in reading-related situations before interpretations from scores on these measures can be made with any level of confidence. This study examined criterion-related validity evidence for the MRP-A in relation to two different criterion variables. First, I examined how well student’s scores on the MRP-A were able to predict teacher ratings of individual student’s motivation to read. Teacher ratings were chosen as a criterion variable based on the assumption that teachers would be the best judges of each student’s motivation to read due to the fact that they are able to observe each student in reading related situations over the course of the student’s second grade year. Second, I examined how well student’s scores on the MRP-A were able to predict on-task behavior during formal reading instruction. On-task behavior was chosen as a second criterion variable based on the assumption that students who were motivated readers would be intrinsically involved in reading related instruction.

Method

Participants

Participants in Study 1 were 168 second grade students from four elementary schools located in a rural area of the southeastern United States. Based on the 2003-2004 state report cards for participating schools, 41% of enrolled students were eligible for free/reduced lunches (Georgia Department of Education, 2004). The demographic make-up of the schools is as follows: approximately 72.5% of the students were Caucasian, 12% African American, 7.5% Hispanic, 6% Asian, and 2% multiracial. Of the participating students, 54% were female. Participating students ranged in age from 7-9 years old, with a mean age of 7 years 8 months (7.66) and a standard deviation of 5 months (.43). The 168 participating students were randomly selected from a group of 337 students participating in a larger study on the development of reading fluency.
Measures

Self-Reported Motivation to Read. An adapted version of The Reading Survey portion of Gambrell et al.’s (1996) Motivation to Read Profile (MRP-A) was used to assess students’ motivation to read. The MRP-A is comprised of 28 items that use a 4-point Likert-type response scale and is administered to students individually. The survey is read aloud to students to avoid any issues where reading skill could confound young or lower skilled students’ responses. The original two subscales are designed to assess two specific dimensions of motivation to read: self-concept as a reader and value for reading (10 items each). More specifically, the self-concept items are designed to assess general attitudes about themselves as readers and how efficacious they feel in various reading situations. The value items assess the value students place on reading in terms of how important they feel reading is, how frequently they would like to engage in reading related activity, and how useful they feel reading will be in their future lives. A third experimenter-constructed subscale containing eight additional items was added to assess student’s goals for reading. All items were designed to meet the following criteria: a) applicability to grades 2 through 6, b) applicability to all teaching approaches and materials, c) suitability for group administration (even though it will be administered individually in this study), and d) accuracy in reflecting the appropriate dimension of motivation (self-concept, value, or goals). Some items were designed to be reverse-scored to avoid the possibility of particular response set. Field test data from a study conducted by Gambrell (1996) reported acceptable reliability for the original MRP (20 items), with an alpha of .75 for the self-concept subscale and .82 for the value subscale. Following the items-analysis conducted in the current study and the resulting revisions, the scale was reduced to 17 items (6 self-concept items, 4 value items, and 7 goal items). The revised 17 item scale was used for all validity analyses done in the second phase of data analysis for Study 1 of the current investigation. A summary of all of the items from both the original and revised MRP-A can be found in Appendix A.
Teacher Ratings of Students’ Motivation to Read. The items for the Teacher Ratings of Student Motivation to Read (TRSM) scale used in this study were selected from a previously published instrument, the Teacher Questionnaire on Student Motivation to Read-3rd Edition (Sweet, Guthrie, & Ng, 1998). This previously published scale was designed to assess teacher’s perceptions of motivated readers; therefore, many of the items were designed for teachers to rate students based on observable behaviors that are theoretically indicative of motivated readers, which is what the items selected from this scale were used for in the current investigation. Specifically, the TRSM includes the 13 items from the individual, topic, and social subscales of the original questionnaire. Items from the individual subscale ask teachers to rate students’ behaviors that would be consistent with intrinsic motivation for reading. The topic subscale asks teachers to rate students’ behaviors that would indicate their interest in reading as it relates to favorite topics or genres, which is also closely related to intrinsic reasons for reading. The social subscale asks teachers to rate students’ behaviors that would indicate their willingness to share ideas about books in social situations and to work on reading activities in groups. The decision to select these particular subscales was based on three criteria. First, the items from these subscales ask teachers to rate behaviors they have observed in students that are consistent with the motivational dimensions assessed by the MRP-A. Second, these particular subscales have shown adequate psychometrics in a study conducted by Sweet et al. (1998), with alphas of .89 for the individual subscale, .84 for the topic subscale, and .64 for the social subscale. Finally, all three subscales were related to one another, with intercorrelations amongst these subscales ranging from .69-.81, which is consistent with the theoretical interrelations of these constructs in the motivation literature. A copy of the TRSM can be found in Appendix B at the end of this manuscript. Reliability of the TRSM in the current study was .87 alpha for the total scale was calculated due to the fact that only the composite total score were used in the final analyses.
Student’s On-Task Behavior during Formal Reading Instruction. An experimenter-constructed protocol was used to record second grade students’ on-task behavior during formal reading instructional time. This protocol included each participating student’s picture next to a grid with eight columns and three rows (a total of 24 on/off-task cells for each student). Observations were conducted in 8 five-minute intervals (representing the 8 columns on each student’s grid) over the course of an hour-long reading instructional block. Each classroom was observed 3 times (representing the 3 rows on each student’s grid) over the course of the academic year, with observations in each classroom approximately corresponding with the 3 time points of MRP-A data collection. At the end of each 5-minute interval of an observation, the research assistants recorded any student who was observed to be off-task by marking an “x” next to that student in the corresponding box on their grid. To improve inter-rater reliability, the two observers met prior to beginning observations to determine what would constitute off-task behavior. At the end of each hour-long observation, the total number of x’s marked next to each individual student was added together and subtracted from 8, representing the total number of time-points that the student was observed to be on-task during the reading instructional block. Observations were collected as part of a larger observational study on reading fluency instruction. Inter-rater reliability between the two research assistants was established through concurrent observations of the same classroom, with inter-rater agreement of 95%.

Reading Fluency Skill. Reading fluency skill was measured using two separate instruments, one that assesses isolated word list fluency and one that assesses oral reading fluency in context.

a) Isolated Word List Fluency. To assess students’ ability to fluently read lists of isolated words I used the Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 1999), which measures children’s ability to fluently (i.e. quickly and accurately) read both real
words and non-words. The TOWRE is comprised of two subtests, one subtest is comprised of items designed to assess sight word reading efficiency and the other subtest is designed to assess phonemic decoding efficiency of reading non-words. Both subtests on the TOWRE consist of lists of words (sight word efficiency), or non-words (phonemic decoding efficiency) for which children are asked to read as quickly and accurately as they can for 45 seconds. The number of words and non-words read correctly in each 45 seconds on each subtest is recorded. The TOWRE reports test-retest reliabilities between .90 and .97, and validity estimates with other decoding measures between .91 and .94. The two subtests of the TOWRE were used as criterion variables to examine the relationship between the MRP-A and decoding efficiency.

b) Oral reading fluency in context. The DIBELS Oral Reading Fluency is a standardized, criterion referenced test of oral reading fluency. Children are asked to read three grade-level reading passages, of similar difficulty, aloud for one minute. The examiner records all the oral reading errors, similar in fashion to a running record, including words omitted, substituted, and hesitations of more than three seconds. Misread words that are self-corrected within three seconds, repeated words, and inserted words are scored as accurate. The number of words read correctly in one minute (cwpm) for each passage is calculated by counting the number words correctly read by the child. The median cwpm of the three passages was used to identify children’s benchmark level at each time point. The DIBELS Oral Reading Fluency scale is often used by schools to identify children who are at-risk of failing end of the year high-stakes tests and to monitor students reading progress throughout the year. The DIBELS Oral Reading Fluency reports alternate-form reliability between .89 and .96 and
criterion-related validity estimates ranging from .91 to .96 for second grade students. Students’ median score on the DIBELS at each time point was used as a criterion variable to examine the relationship between the MRP-A and oral reading fluency.

Procedure

The MRP-A was administered to all participating students three times over the course of their second grade year, with testing dates at each school spaced approximately 60 instructional days apart to allow adequate time for skill and motivational growth. The MRP-A was administered to students individually by specially trained research assistants as part of a larger assessment battery. The children were also administered the TOWRE and the DIBELS. This was counterbalanced with the MRP-A so that if one student received the MRP-A first, the next received the reading fluency skill assessments first. Completion of the MRP-A and the reading fluency skill assessment battery took approximately twenty minutes per student. Verbal assent was obtained from each individual student in addition to the written consent provided by the parents before participating. Each participant was given a small token or sticker as a sign of appreciation for their time and effort.

The TRSM was given to teachers immediately prior to the third data collection time point, so that their responses on the questionnaire could be analyzed concurrently with MRP-A data from the third time-point. Teachers were given directions for completing the questionnaire at the time they were passed out and were instructed to have them completed by the final date of student data collection. Also, the written directions on the TRSM include a brief description of what constitutes behaviors that they have observed “rarely”, “seldom”, “sometimes”, or “often”, which is designed to make the teacher’s responses more reliable across classrooms. The TRSM took teachers approximately five minutes to complete per student.
Results

Phase 1: Items-Analysis and Scale Revision

The items analysis of the MRP-A was conducted by creating a grid that tracked how all 28 items functioned based on six performance criteria. These criteria were as follows: 1) Items were judged on how well they reflected motivational theory relevant to the subscale’s construct of interest (self-concept, value, or goals). 2) Frequency distributions were examined to flag items that were not normally distributed; using skew and kurtosis statistics exceeding an absolute value of 2.0 as the cutoff criteria. 3) Items that did not correlate at least moderately (> .20) with a majority of the other items on the subscale were flagged. 4) Items that had a negative impact on the subscale’s internal consistency (alpha) were flagged. 5) Exploratory factor analyses were conducted on each of the three subscales using maximum likelihood estimation and a Promax rotation, since factors were assumed to correlate. Items with factor loadings less than .40 on the primary factor extracted or items with cross-loadings greater than .30 on a secondary factor were flagged as being potentially problematic. 6) An ANOVA was conducted on all three subscales using gender as a categorical variable to determine if any items exhibited a gender bias. Items with significantly different means for boys versus girls ($p < .05$) were flagged. Any items that were flagged 3 or more times based on these six criteria were removed from their respective subscale and any subsequent analyses.

The first criterion for revision of the 28 item MRP-A was to examine how well each item on the three subscales (self-concept for reading, value for reading, and goals for reading) reflected the motivational theory from its respective domain. This criterion is aligned with the content aspect of Messick’s (1989) framework for unitary construct validity. Items were examined by three independent reviewers, consisting of one expert in early reading and two graduate students with doctoral level training in academic motivation. Any item that was identified as being problematic by two or more of the
reviewers was flagged for potential content validity problems. Of the original 28 items, 4 were identified as being theoretically problematic (1 self-concept item and 3 value items).

Chapman and Tunmer’s (1997) definition of reading self-concept was used as a reference point for evaluating items from the self-concept subscale. They defined reading self-concept as the combination of three interrelated components: “1) perceptions of competence in performing reading tasks; 2) perceptions that reading activities are generally either easy or difficult, and 3) attitudes felt towards reading (p. 280).” This definition is similar to other definitions of academic self-concept, yet it is specific to reading. Based on this definition, two of the three independent reviewers identified item 17 as being potentially problematic. All of the other items from the self-concept subscale were judged to fall within the boundary of Chapman and Tunmer’s (1997) definition.

The content validity of items from the value subscale was evaluated in terms of Eccles et al.’s (1983) definition of subjective task-value as it pertained specifically to reading. Thus, value for reading was also defined along three interrelated dimensions: 1) how much an individual likes or is interested in reading; 2) the perceived importance of doing reading well, and 3) the perceived usefulness of reading. Based on this definition, 3 items from the value for reading subscale were flagged by the independent reviewers for potential content validity problems. Items 6 and 20 were identified by two of the three reviewers and item 10 was identified by all three reviewers as being potentially problematic based on this definition.

Items from the goals for reading subscale were evaluated based on theory from the achievement goal literature. Achievement goals are broadly defined as the purposes for engaging in an activity (Maehr, 1989); therefore, items from the goals subscale were designed to assess whether students had a purpose for engaging in reading activity and what those purposes might be. Based on this definition none of the items were flagged for potential content validity problems. This is not surprising given that the items were constructed with this definition in mind.
The second criterion for revision of the MRP-A was an examination of each item’s frequency distribution. Item response frequencies were evaluated in terms of the skewness and kurtosis statistics for each individual item. Items with skewness or kurtosis statistics that exceeded an absolute value of 2 were flagged as being potentially problematic. Of the original 28 items on the MRP-A, 4 were identified as having problematic response frequencies. Item 8 had a kurtosis statistic of 2.34, item 12 had a kurtosis statistic of 3.78, item 20 had a skewness statistic of -2.78 and a kurtosis statistic of 7.50, and item 27 had a kurtosis statistic of 2.84. Each of the flagged items had a positive kurtosis statistic greater than 2.0 indicating that their response distributions were more peaked than would be expected in a normal distribution (i.e., students responded to the item the same way frequently). For example, 147 of the 168 participating students responded that they would be “very happy” to receive a book for a present; therefore, this item did little to differentiate students who valued reading from those who did not.

The third criterion for revision of the MRP-A aligned with the structural aspect of Messick’s (1989) unitary construct validity. It involved an examination of the inter-item correlations to see how well each item correlated with other items from its respective subscale. A requirement that each item correlated at least .20 with a majority of the other items from its respective subscale was used as the cutoff-criterion. Of the 28 original items on the MRP-A, 14 of them failed to meet the cutoff-criterion. Items 5, 7, and 11 had particularly low inter-item correlations with a majority of the other items from the self-concept subscale. All of the items except for item 8 had low inter-item correlations on the value subscale, indicating that this subscale was particularly problematic in this area. Finally, items 21 and 26 were flagged for low inter-item correlations with a majority of the items from the goals for reading subscale.

The fourth criterion for revision of the MRP-A aligned with the substantive aspect of construct validity. It involved an examination of the impact each item had on its respective subscale’s internal consistency (alpha). Four items were flagged for having a
negative impact of its subscale’s alpha statistic. Alpha for the 10-item self-concept subscale was .63; however, the alpha statistic would increase to .73 if item 11 were deleted from the subscale, indicating that responses to this item were not consistent with responses to other items on the subscale. Alpha for the 10-item value subscale was .57, yet alpha would increase to .58 if either item 12 or item 18 were eliminated from the subscale, again indicating that these items were negatively impacting the internal consistency of the value subscale. Finally, alpha for the 8-item goals subscale was .69; however, if item 26 were eliminated alpha would increase to .70.

The fifth criterion for revision of the MRP-A aligned with the substantive and structural aspects of construct validity. To examine items along this criterion, exploratory factor analyses were conducted on each of the three subscales using maximum likelihood estimation and a Promax rotation, since factors were assumed to correlate. Items with factor loadings less than .40 on the primary factor extracted or items with cross-loadings greater than .30 on a secondary factor were flagged as being potentially problematic. Of the original 10 items on the self-concept subscale, 6 were flagged for loadings less than .40 on the primary factor extracted. Item 3 had a loading of .09, item 5 had a loading of .28, item 7 had a loading of .37, item 11 had a loading of .36, item 13 had a loading of .35, and item 17 had a loading of .15 on the primary factor extracted. Of the original 10 items on the value subscale, only item eight had a factor loading greater than .40 on the primary factor extracted. Item 2 had a loading of .29, item 4 loaded .22, item 6 loaded .08, item 10 loaded .01, item 12 loaded -.06, item 14 loaded .04, item 16 loaded -.10, item 18 loaded .12, and item 20 loaded .22 on the primary factor extracted. Of the original 8 items on the goals subscale, 5 had loadings less than .40 on the primary factor extracted. Item 22 had a loading of .35, item 23 had a loading of .23, item 25 had a loading of -.02, item 26 had a loading of .30, and item 28 had a loading of -.06 on the primary factor extracted. The fact that this criterion was particularly
problematic for a majority of the items is not surprising considering the low inter-item correlations for most of these items reported earlier.

The sixth and final criterion used to revise the original MRP-A aligned with the generalizability aspect of Messick’s (1989) framework. This criterion involved examining for the possibility of gender bias in each of the items. To test this possibility all items were subjected to an ANOVA using gender as a categorical variable. Any item with a significant difference between groups \( (p < .05) \) was flagged as being potentially problematic. Of the 28 original items, only 3 were flagged as being potentially biased by gender. Item 3 had an \( F(1) = 2.95, p = .03 \); item 13 had an \( F(1) = 2.05, p = .03 \); and item 16 had an \( F(1) = 4.57, p = .03 \). All other items had non-significant between group values indicating that they were not biased across gender groups at time 1.

Following a review of how each of the 28 original items on the MRP-A functioned along the six criteria in the items-analysis, 8 items were removed for failing to meet the criterion in three or more of the six areas. A summary of how each item performed along all six revision criteria can be found in Table 2.1. From the self-concept as a reader subscale, only item 11 was removed from further analyses leaving 9 of the original 10 items on the subscale. From the value for reading subscale, items 6, 10, 12, 16, 18, and 20 failed in at least three of the six areas and were removed leaving 4 of the original 10 items on the subscale. Of the three subscales, the value subscale was by far the worst functioning with over half of the items failing to meet enough requirements to be retained in the final revised MRP-A. From the goals for reading subscale, item 26 failed in three of the six areas and was removed from further analyses leaving 7 of the original 8 items on the subscale. Evidence to support the improved validity of the revised subscales will be presented in the next section.

Phase 2: Validation of the Revised MRP-A

Because analyses were guided by Messick’s (1989) framework for unitary construct validity, results will be presented along the five aspects previously discussed.
When possible, results will be reported for both the full and revised MRP-A to show how revision improved the validity of each respective subscale.

Table 2.1.  

<table>
<thead>
<tr>
<th>Item</th>
<th>Theory</th>
<th>Skew/Kurtosis</th>
<th>Low Int-Item correlation</th>
<th>Negatively Impacts Alpha</th>
<th>EFA (Loads on scale factor)</th>
<th>Bias (Gender)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>26</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>27</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

*Note.* Three or more categories failed (X) resulted in the item being deleted from its respective subscale; Theory decisions were based on recommendations made by three independent reviewers; reviewer’s content validity recommendations were based on the theory from each subscale’s domain within the motivation literature. Skew and Kurtosis decisions were based on a cut-off of 2.0. Item intercorrelation decisions were based on the number of low/moderate correlations (.20 or higher) each respective item had with all other items on the subscale (a majority was required). Items that had a negative influence on alpha were deleted. Exploratory factor analyses were conducted on each subscale using a Promax rotation (oblique), since factors were assumed to correlate. Items with factor loadings less than .40 on the primary factor extracted or items with cross-loadings greater than .30 on a secondary factor were flagged as being potentially problematic. An ANOVA was conducted using gender as the categorical variable to detect item bias, *p*-values less than .05 were indicative of biased items.
Content Aspect

Content validity for each item on the full and revised subscales was initially examined during the items-analysis of the MRP-A. Three of the four items that were flagged for potential theoretical problems during the items-analysis failed to meet at least two other revision criteria, which resulted in the deletion of those items from their respective subscales. The only item that was flagged for potential theoretical problems yet remained following revision of the three subscales was item 17. While this item still looks to have theoretical problems, its correlation with the self-concept scale was .65 at time 2 and .57 at time 3, which were both statistically significant ($p < .05$). These high correlations indicate that this item is functioning well despite the potential content validity concerns. All other items that remained in the MRP-A following revision were evaluated to be theoretically relevant during the items-analysis. The fact that three of the four items flagged for theoretical problems were eventually discarded due to other problems is evidence that the initial content validity review was effective in identifying potentially problematic items on solely theoretical grounds.

Substantive Aspect

Examination of the substantive aspect of the MRP-A’s construct validity focused on the suitability of the items for second grade students and the impact that revision had on each subscale’s internal consistency. Gambrell et al. (1996) explained that items for the original version of the MRP were designed to be applicable to students ranging in age from grades 2 through 6. Despite the fact that the instrument was originally designed to be administered to groups of students, the MRP-A was administered to students individually. This was done to ensure that each item was read aloud and understood before the student responded to the item. This change in administration procedures was viewed as a safeguard against students with low reading ability or poor reading comprehension skill, since participants were only in the second grade and had limited reading ability. The second area addressed relating to the substantive validity of the
MRP-A was an examination of the internal consistency for each subscale following revision. Alpha for the revised subscales at time 2 were as follows (alpha for the full subscales are in parentheses): self-concept = .76 (.68), value = .55 (.62), goals = .67 (.67).

No items on the revised subscales had a detrimental effect on alpha when examining the “alpha when item deleted” statistic at time 2. Alpha actually increased in the self-concept subscale and remained the same in the goals subscale, indicating that removal of the edited items has made them more reliable. Alpha decreased slightly in the value subscale following revision; however, this slight decrease in alpha would be expected due to the fact that six of the original ten items were removed from the subscale during revision.

Structural Aspect

Items analysis. Item-total correlations were recalculated after revision of the MRP-A, using time 2 data, to examine how well the remaining items on each subscale correlated with one another. Inter-item correlations of the remaining nine items on the self-concept subscale ranged from .06 to .49 at time 2 with 12 correlations falling below the .20 cut-off used to revise the scale at time 1. In contrast, inter-item correlations of the full 10-item self-concept subscale ranged from -.28 to .49 at time 2 with 21 inter-item correlations falling below the .20 cutoff. This indicates that the inter-item structure of the self-concept subscale was improved by the removal of item 11, yet there were still a number of low inter-item correlations on the self-concept subscale following revision. Inter-item correlations for the revised value subscale ranged from .13 - .34 at time 2 with one inter-item correlation falling below the .20 cutoff. Inter-item correlations for the full 10-item value subscale ranged from -.02 - .34 at time 2 with 31 of the 45 inter-item correlations falling below the .20 cut-off used for revision. This is evidence that the inter-item relationships on the value subscale were significantly improved by revision. Inter-item correlations for the revised 7-item goals subscale ranged from .05 - .51 at time 2 with 7 correlations falling below the .20 cut-off. In contrast, inter-item correlations for the full 8-item goals subscale ranged from .05 - .51 at time 2 with 14 correlations falling
below the .20 cut-off. While the improvement of this subscale was less evident than in the value subscale, this was not surprising given that only one item was removed during revision.

**Factor analysis.** Confirmatory factor analyses were conducted on the MRP-A using time 2 data to examine the factor structure of both the full and revised scales. Two CFA models were tested for both the full and revised scales. The first set of models looked at how well items on the MRP-A loaded on a single motivation to read factor. The second set of models tested how well those same items loaded on a three-factor model, representing the three subscales of the MRP-A. All analyses were conducted on LISREL 8 (Joreskog & Sorbom, 1993) using maximum likelihood estimation. In the first set of models tested, one-factor models for both the full 28 item MRP-A and revised 20 item MRP-A, fit indices indicated marginal fit of both models to the observed time 2 data. The full 28-item MRP-A had a $\chi^2$ (350) = 562.62 ($p < .001$), an RMSEA = .058, and a CFI = .90. Following Hu and Bentler’s (1996) guidelines for interpreting model fit, the RMSEA of .058 indicated good model fit; however, the CFI of .90 fell well below the suggested .95 cutoff, indicating potential problems with this model. The revised 20-item MRP-A had a $\chi^2$ (170) = 282.54 ($p < .001$), an RMSEA = .07 and a CFI = .92, indicating marginal fit to a one-factor model.

The second set of models tested examined how well data from time 2 for the full and revised MRP-A’s fit to a three-factor model. In this model the self-concept items were hypothesized to load on a single latent factor, the value items were hypothesized to load on a second latent factor, and the goal items were hypothesized to load on a third latent factor for both scales. The three factor model showed an improved, yet marginal fit to the full MRP-A’s time 2 data, with the $\chi^2$ (347) = 522.25 ($p < .01$), RMSEA = .050, and CFI = .92. Estimated correlations between the three factors ranged from .69-.88, showing a moderate to high correlation between factors which reflect expected relationships from the motivational literature (all correlations reported on the upper-
The three factor model for the revised 20 item MRP-A showed good fit to the time 2 data, with the $\chi^2 (167) = 241.81$ ($p < .01$), RMSEA = .049 and CFI = .95. Estimated correlations between the three factors from this model ranged from .56-.82 indicating that they were distinct yet related as would be expected by motivational theory (all correlations reported on the lower diagonal of Table 2).

### Table 2.2. Factor Correlations from the three factor CFA on the MRP-A at Time 2

<table>
<thead>
<tr>
<th>Factor</th>
<th>self-concept</th>
<th>value</th>
<th>goals</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. self-concept</td>
<td>1.00</td>
<td>.69</td>
<td>.83</td>
</tr>
<tr>
<td>2. value</td>
<td>.56</td>
<td>1.00</td>
<td>.88</td>
</tr>
<tr>
<td>3. goals</td>
<td>.82</td>
<td>.69</td>
<td>1.00</td>
</tr>
</tbody>
</table>

*Note.* Correlations from the full MRP-A are reported on the upper diagonal and correlations from the revised MRP-A are reported on the lower diagonal.

Finally, the $\chi^2_{\text{diff}}$ statistic was calculated to test whether the three-factor models were significantly better than the single-factor models in explaining the factor structure from time 2 data for both the full and revised MRP-A. For the full 28 item scale, the one-factor model had a $\chi^2 (350) = 562.62$ and the three factor model had a $\chi^2 (347) = 522.25$; therefore, the $\chi^2_{\text{diff}} (3) = 40.37$, which was statistically significant ($p < .001$). The significant $\chi^2_{\text{diff}}$ statistic indicated that the addition of two factors to the one-factor model significantly improved the model’s fit to the time 2 data. For the revised 20 item MRP-A, the one-factor model had a $\chi^2 (170) = 282.54$ and the three factor model had a $\chi^2 (167) = 241.81$; therefore, the $\chi^2_{\text{diff}} (3) = 40.73$, which was also statistically significant ($p < .001$). These results clearly indicate that the three-factor model fit significantly better than the one-factor model for both the full and revised MRP-A at time 2. A summary of the fit indices and model comparisons can be found in Table 2.3.
Table 2.3. **CFA Fit Indices and Model Comparisons of the Full and Revised MRP-A**

<table>
<thead>
<tr>
<th>Scale</th>
<th>Model</th>
<th>$\chi^2$</th>
<th>df</th>
<th>RMSEA</th>
<th>CFI</th>
<th>$\chi^2_{diff}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full 2</td>
<td>One-factor</td>
<td>562.62</td>
<td>350</td>
<td>.058</td>
<td>.90</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Three-factor</td>
<td>522.25</td>
<td>347</td>
<td>.050</td>
<td>.92</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Difference</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>40.37*</td>
</tr>
<tr>
<td>Revised 2</td>
<td>One-factor</td>
<td>282.54</td>
<td>170</td>
<td>.065</td>
<td>.92</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Three-factor</td>
<td>241.81</td>
<td>167</td>
<td>.049</td>
<td>.95</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Difference</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>40.73*</td>
</tr>
</tbody>
</table>

*Note.* Full 2 = full MRP-A at time 2; Revised 2 = revised MRP-A at time 2; RMSEA = root mean square error of approximation; CFI = comparative fit index; $\chi^2$ = chi-square difference statistic; * indicates significance at the .001 level.

**Generalizability Aspect**

The generalizability aspect of the revised MRP-A was previously addressed by examining how each of the items functioned across gender in the items-analysis portion of the study. One of the three items that was identified as functioning different by gender at time 1 was eliminated during revision for failing to meet at least two other criteria. To test whether any of the remaining items were biased across gender, a one-way ANOVA was run on the items from the revised MRP-A at times 2 and 3. Of the remaining 20 items, only item 7 at time 3 and item 23 at time 3 had significant differences in means across groups. Item 7 at time 3 had an $F(1) = 4.51, p < .05$ and item 23 at time 3 had an $F(1) = 3.98, p < .05$ signifying that they functioned differently across gender at the third time point. None of the items showed significant differences across gender at time 2. In addition to the gender bias analysis, total scores on the three revised subscales at time points 2 and 3 were correlated to assess test-retest reliability (generalizability across time). Test-retest reliabilities for the three revised subscales of the MRP-A were all significant ($p < .001$): .70 (self-concept), .52 (value), and .69 (goals). The fact that these were moderate to high correlations would be expected since assessments were spaced
approximately 60 instructional days apart, giving ample time for motivational constructs
to change slightly.

External Aspect

To assess the external aspect of the MRP-A’s validity, two separate sets of
analyses were conducted to evaluate the relationship between the MRP-A with two
criterion variables, teacher ratings of students’ motivation to read and time on-task during
formal reading instructional time.

Relationship to Teacher Ratings. The first set of analyses examined the
relationship between students’ scores on the MRP-A (both full and revised) at time 3 and
teacher ratings of students’ motivation to read, as measured by the TRSM. Time 3 data
was collected concurrently with the teacher ratings; therefore, only MRP-A data from the
third time point was used in these analyses. Table 2.4 summarizes the means, standard
deviations, and correlations to the TRSM for composite scores from the teacher rating
scale, as well as the self-concept, value, and goals subscales of the MRP-A from time 3.
Correlations between the TRSM and the MRP-A (subscale composites and total scores
for full and revised scales) ranged from -.05 to .12 and were all nonsignificant (p > .10).
The lack of any significant correlations indicates that scores on both the full and revised
MRP-A had no relationship with teacher ratings of student’s motivation to read.

Table 2.4. Means and Standard Deviations of the TRSM and MRP-A at Time 3

<table>
<thead>
<tr>
<th>Scale</th>
<th>M</th>
<th>SD</th>
<th>r (TRSM)</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRSM</td>
<td>35.70</td>
<td>7.04</td>
<td></td>
<td>168</td>
</tr>
<tr>
<td>Full Self-Concept</td>
<td>31.04</td>
<td>3.76</td>
<td>.12</td>
<td>168</td>
</tr>
<tr>
<td>Full Value</td>
<td>33.73</td>
<td>4.48</td>
<td>.03</td>
<td>168</td>
</tr>
<tr>
<td>Full Goals</td>
<td>24.37</td>
<td>4.15</td>
<td>-.02</td>
<td>168</td>
</tr>
<tr>
<td>Full Total MRP-A</td>
<td>89.14</td>
<td>10.01</td>
<td>.05</td>
<td>168</td>
</tr>
<tr>
<td>Revised Self-Concept</td>
<td>28.11</td>
<td>3.81</td>
<td>.04</td>
<td>168</td>
</tr>
<tr>
<td>Revised Value</td>
<td>12.93</td>
<td>2.36</td>
<td>.10</td>
<td>168</td>
</tr>
<tr>
<td>Revised Goals</td>
<td>20.96</td>
<td>3.80</td>
<td>-.05</td>
<td>168</td>
</tr>
<tr>
<td>Revised MRP-A Total</td>
<td>51.88</td>
<td>7.42</td>
<td>.01</td>
<td>168</td>
</tr>
</tbody>
</table>
To follow up these results, all items on the MRP-A were correlated with the TRSM total score to see if any items had a relationship to teacher ratings of student’s motivation to read. Of all 28 items on the MRP-A, only items 8 and 11 had statistically significant correlations to the TRSM total score. Item 8 had the highest correlation with the TRSM total score ($r = .17, p < .05$) and item 11’s correlation was barely significant ($r = .15, p < .05$). These results clearly indicate that student’s scores on the MRP-A are not significantly related to teacher ratings of students’ motivation to read, as measured by the TRSM, so no further analyses were needed.

To test the possible influence of reading skill on teacher’s ratings of individual students, scores on the TRSM were correlated with students’ scores on the DIBELs at time 3. A correlation of .45 ($p < .001$) was found between teacher ratings and students’ oral reading fluency at the spring time point, indicating the possibility that teacher’s ratings were more influenced by students’ reading skill level than by their motivation level. Given this problematic finding, it is unclear whether teacher ratings of motivation are a good indicator of children’s “true” motivation or whether teachers just assume that children with higher ability are more motivated.

**Relationship to Time-on-Task.** In light of the nonsignificant relationship between the MRP-A and the TRSM, an examination of the relationship between the MRP-A (full and revised scales) and a second criterion variable, time-on-task during formal reading instructional time, was conducted. Data for time-on-task was collected at all three time points; therefore, correlations were calculated between time-on-task with composite scores on the MRP-A (both full and revised) at corresponding time points. Means, standard deviations, and correlations between the MRP-A and time-on-task for each respective time point are reported in Table 2.5. Correlations between time on task and the MRP-A (subscale composites and total scores for full and revised scales) at all three time points ranged from -.14 to .09 and were all nonsignificant ($p > .05$). The lack of significant correlations between time on-task and any of the MRP-A composite scores at
each time point indicates that there was no relationship between time on-task during formal reading instruction and students’ scores on the MRP-A at all three time points.

Given the lack of relationship between the MRP-A and either of the validity benchmark measures, the TRSM and time on-task, it was important to discern whether there was a relationship between any of the validity measures. In fact, the correlation between the TRSM and time on-task was also nonsignificant ($r = -.04, p > .05$). Thus, it

Table 2.5. *Means, Standard Deviations, and Correlations for Time on Task and the MRP-A at all 3 Time Points*

<table>
<thead>
<tr>
<th>Scale</th>
<th>M</th>
<th>SD</th>
<th>$r$ (TOT)</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time on Task 1</td>
<td>7.67</td>
<td>.80</td>
<td>TOT1</td>
<td>168</td>
</tr>
<tr>
<td>Full Self-Concept 1</td>
<td>31.46</td>
<td>3.88</td>
<td>-.01</td>
<td>168</td>
</tr>
<tr>
<td>Full Value 1</td>
<td>34.21</td>
<td>3.36</td>
<td>.02</td>
<td>168</td>
</tr>
<tr>
<td>Full Goals 1</td>
<td>25.50</td>
<td>4.05</td>
<td>.02</td>
<td>168</td>
</tr>
<tr>
<td>Full Total MRP-A 1</td>
<td>91.18</td>
<td>9.37</td>
<td>.01</td>
<td>168</td>
</tr>
<tr>
<td>Revised Self-Concept 1</td>
<td>28.80</td>
<td>3.88</td>
<td>.03</td>
<td>168</td>
</tr>
<tr>
<td>Revised Value 1</td>
<td>13.26</td>
<td>1.94</td>
<td>.00</td>
<td>168</td>
</tr>
<tr>
<td>Revised Goals 1</td>
<td>22.18</td>
<td>3.79</td>
<td>.00</td>
<td>168</td>
</tr>
<tr>
<td>Revised Total MRP-A 1</td>
<td>64.24</td>
<td>6.46</td>
<td>.01</td>
<td>168</td>
</tr>
<tr>
<td>Time on Task 2</td>
<td>7.65</td>
<td>1.0</td>
<td>TOT2</td>
<td>168</td>
</tr>
<tr>
<td>Full Self-Concept 2</td>
<td>31.47</td>
<td>4.15</td>
<td>-.12</td>
<td>168</td>
</tr>
<tr>
<td>Full Value 2</td>
<td>34.18</td>
<td>3.62</td>
<td>.00</td>
<td>168</td>
</tr>
<tr>
<td>Full Goals 2</td>
<td>25.51</td>
<td>3.85</td>
<td>.01</td>
<td>168</td>
</tr>
<tr>
<td>Full Total MRP-A 2</td>
<td>91.17</td>
<td>9.49</td>
<td>-.05</td>
<td>168</td>
</tr>
<tr>
<td>Revised Self-Concept 2</td>
<td>28.67</td>
<td>4.14</td>
<td>-.13</td>
<td>168</td>
</tr>
<tr>
<td>Revised Value 2</td>
<td>13.16</td>
<td>2.10</td>
<td>.01</td>
<td>168</td>
</tr>
<tr>
<td>Revised Goals 2</td>
<td>22.04</td>
<td>3.62</td>
<td>.01</td>
<td>168</td>
</tr>
<tr>
<td>Revised Total MRP-A 2</td>
<td>63.88</td>
<td>5.81</td>
<td>-.05</td>
<td>168</td>
</tr>
<tr>
<td>Time on Task 3</td>
<td>7.66</td>
<td>1.10</td>
<td>TOT3</td>
<td>168</td>
</tr>
<tr>
<td>Full Self-Concept 3</td>
<td>31.04</td>
<td>3.76</td>
<td>-.14</td>
<td>168</td>
</tr>
<tr>
<td>Full Value 3</td>
<td>33.73</td>
<td>4.48</td>
<td>.02</td>
<td>168</td>
</tr>
<tr>
<td>Full Goals 3</td>
<td>24.37</td>
<td>4.15</td>
<td>-.02</td>
<td>168</td>
</tr>
<tr>
<td>Full Total MRP-A 3</td>
<td>89.14</td>
<td>10.01</td>
<td>-.06</td>
<td>168</td>
</tr>
<tr>
<td>Revised Self-Concept 3</td>
<td>28.11</td>
<td>3.81</td>
<td>-.14</td>
<td>168</td>
</tr>
<tr>
<td>Revised Value 3</td>
<td>12.93</td>
<td>2.36</td>
<td>.09</td>
<td>168</td>
</tr>
<tr>
<td>Revised Goals 3</td>
<td>20.96</td>
<td>3.80</td>
<td>-.03</td>
<td>168</td>
</tr>
<tr>
<td>Revised Total MRP-A 3</td>
<td>62.00</td>
<td>6.29</td>
<td>-.05</td>
<td>168</td>
</tr>
</tbody>
</table>

*Note.* TOT1 = Time on task at time point 1, TOT2 = Time on task at time point 2, TOT3 = Time on task at time point 3
is unclear which, if any, measure of reading motivation represents a valid one or whether each merely assesses a different aspect of reading motivation.

*Relationship to “Flat” Reading Skill.* Because no relationship was found between scores on the MRP-A, the TRSM, or time-on-task, analyses were conducted to examine the relationship between scores on the MRP-A and flat reading skill at each of the three time points. Scores from each of the three subscales were correlated with each of the three reading skill scores from corresponding time points. None of the correlations were significant indicating that none of the subscales were related to flat reading skill at any of the three time points.

*Summary of validity results.* In summary, 8 of the original 28 items on the MRP-A were removed from the scale during revision. Item 11 was dropped from the self-concept subscale due to low correlations with other items on the subscale, its negative impact on the subscale’s internal consistency, and because it failed to load sufficiently on the primary factor extracted in the EFA. Items 6, 10, 12, 16, 18, and 20 were dropped from the value subscale for a multitude of reasons. All of the items dropped had extremely low correlations with other items on the value subscale and failed to sufficiently load on the primary factor extracted during the EFA. Items 6, 10, and 20 were judged to have theoretical problems based on Eccles et al.’s (1983) definition of subjective task value as it pertained specifically to reading. Items 10 and 12 had a negative impact on the internal consistency of the value subscale. Items 12 and 20 had problems with their response distributions and items 10 and 16 exhibited a gender bias. The value subscale was by far the most problematic scale in both the revision and validity evaluation phases of these analyses. Only item 26 was dropped from the goals for reading subscale due to low correlations with the other items on the subscale, its negative impact on the subscale’s internal consistency, and an insufficient loading on the primary factor extracted during the EFA. A summary of the final items retained following revision can be found in Table 2.6. The results of this investigation clearly indicated that
the revised scale functioned better along all five aspects of Messick’s (1989) framework for construct validity. These will be discussed in detail in the following section.

Table 2.6.  **Final Items Retained Following Revision of the MRP-A**

<table>
<thead>
<tr>
<th>Subscale</th>
<th>Items</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reading Self-Concept</strong></td>
<td>1. My friends think that I am _____________.</td>
</tr>
<tr>
<td></td>
<td>A very good reader, A good reader,</td>
</tr>
<tr>
<td></td>
<td>An OK reader, A poor reader</td>
</tr>
<tr>
<td></td>
<td>3. I read _____________.</td>
</tr>
<tr>
<td></td>
<td>Not as well as my friends, About the same as my friends,</td>
</tr>
<tr>
<td></td>
<td>A little better than my friends, A lot better than my friends</td>
</tr>
<tr>
<td></td>
<td>5. When I come to a word I don’t know, I can _____________.</td>
</tr>
<tr>
<td></td>
<td>Almost always figure it out, Sometimes figure it out,</td>
</tr>
<tr>
<td></td>
<td>Almost never figure it out, Never figure it out</td>
</tr>
<tr>
<td></td>
<td>7. When I am reading by myself, I understand _____________.</td>
</tr>
<tr>
<td></td>
<td>Almost everything that I read, Some of what I read,</td>
</tr>
<tr>
<td></td>
<td>Almost none of what I read, None of what I read</td>
</tr>
<tr>
<td></td>
<td>9. I am _____________.</td>
</tr>
<tr>
<td></td>
<td>A poor reader, An OK reader,</td>
</tr>
<tr>
<td></td>
<td>A good reader, A very good reader</td>
</tr>
<tr>
<td></td>
<td>13. When my teacher asks me a question about what I have read, I _________.</td>
</tr>
<tr>
<td></td>
<td>Can never think of an answer, Have trouble thinking of an answer,</td>
</tr>
<tr>
<td></td>
<td>Sometimes think of an answer, Always think of an answer</td>
</tr>
<tr>
<td></td>
<td>15. Reading is ____________.</td>
</tr>
<tr>
<td></td>
<td>Very easy for me, Kind of easy for me,</td>
</tr>
<tr>
<td></td>
<td>Kind of hard for me, Very hard for me</td>
</tr>
<tr>
<td></td>
<td>17. When I am in a group talking about stories, I _____________.</td>
</tr>
<tr>
<td></td>
<td>Almost never talk about my ideas, Sometimes talk about my ideas,</td>
</tr>
<tr>
<td></td>
<td>Almost always talk about my ideas, Always talk about my ideas</td>
</tr>
<tr>
<td></td>
<td>19. When I read out loud I am _____________.</td>
</tr>
<tr>
<td></td>
<td>A poor reader, An OK reader,</td>
</tr>
<tr>
<td></td>
<td>A good reader, A very good reader</td>
</tr>
<tr>
<td><strong>Value for Reading</strong></td>
<td>2. Reading a book is something I like to do.</td>
</tr>
<tr>
<td></td>
<td>Never, Not very often,</td>
</tr>
<tr>
<td></td>
<td>Sometimes, Often</td>
</tr>
<tr>
<td></td>
<td>4. My best friends think reading is _____________.</td>
</tr>
<tr>
<td></td>
<td>Really fun, Fun,</td>
</tr>
<tr>
<td></td>
<td>OK to do, No fun at all</td>
</tr>
</tbody>
</table>
8. People who read a lot are _____________.
   Very interesting, Interesting,
   Not very interesting, Boring
14. I think that reading is _________________.
   A boring way to spend time, An OK way to spend time,
   An interesting way to spend time, A great way to spend time

Goals for Reading

21. I feel like I am getting better at reading ________________.
   Never, Sometimes,
   Almost everyday, Everyday
22. Finishing a book that I have started reading ________________.
   Is very important to me, Is important to me,
   Is sometimes important to me, Is never important to me
23. I like to try reading books that older kids read ________________.
   Never, Every once in a while,
   A lot of the time, Always
24. When I am in reading class, I do the work to get better at reading ________________.
   Everyday, Most of the time,
   Sometimes, Never
25. When I am in reading class, I do the work to get better at reading than my friends ________________.
   Everyday, Most of the time,
   Sometimes, Never
27. The reason that I want to get better at reading is so I can read more interesting books.
   A lot like me, A little like me,
   Not much like me, Not at all like me
28. The reason that I want to get better at reading is so I can read better than my friends.
   A lot like me, A little like me,
   Not much like me, Not at all like me

Note. Items were scored on a 1-4 point scale. Items 1, 4, 5, 7, 8, 15, 22, 24, 25, 27, and 28 were reverse coded.

Discussion

The purpose of Study 1 was twofold. First, Study 1 was designed to analyze all of the 28 items on the MRP-A and make revisions to the scale based on six revision criteria applied to the data collected at time 1. Second, after revisions were made to the MRP-A, a detailed examination of the revised scale’s validity was conducted using Messick’s (1989) framework for unitary construct validity. These analyses were conducted using the full 28-item scale as a reference point for evaluating the effectiveness of revision decisions made in phase 1 of the study. All of the analyses conducted in phase 2 of the
study used data collected at times 2 and 3 to confirm revision decisions that were based on time 1 data. All of the information from phases 1 and 2 were then integrated to make an overall judgment of the MRP-A’s validity.

**Revision of the MRP-A**

In the first phase of Study 1, all of the 28 items from the MRP-A were scrutinized along six revision criteria. Each item’s performance was systematically tracked and any item that failed in three or more areas was subsequently deleted from the scale. These six criteria were chosen for two reasons. First, they are commonly used criteria when conducting an items-analysis on a newly developed scale. Second, they were roughly aligned with the five aspects of Messick’s (1989) framework that were used in phase 2 to evaluate the validity of the MRP-A; therefore, they were designed to improve the revised scale’s performance along the 5 aspects for unified construct validity. The findings in the second phase of these analyses provided support for the revisions made, as the revised MRP-A performed better than the full scale along all of the five aspects of construct validity examined. The specific improvements of the revised MRP-A over the full MRP-A are discussed in the following sections of this paper.

**Content Aspect of Validity**

The content aspect of validity examines the theoretical rationale behind an item’s purpose for being on a given scale. In general, it seemed that the original MRP was developed using a broader definition of reading self-concept and value for reading; therefore, there were a few items that fell outside of the boundaries of the definitions used here. Specifically, there was 1 item from the original *self-concept as a reader* subscale and 3 items from the *value for reading* subscale that fell outside of the boundary set from the definitions for reading self-concept and subjective task-value used in this study. The fact that 3 of these 4 items were eventually deleted from the scale for failing in other areas of item-functioning is evidence that the content validity judgments made were accurate.
Beyond the possibility that items could fall outside of a construct’s defined boundaries, a second possibility is that the construct is underrepresented, in that potentially important aspects of a construct are not addressed by the items that make up a given scale. In this study, reading self-concept was defined in terms of three interrelated aspects: 1) perceptions of competence in performing reading tasks; 2) perceptions that reading activities are generally either easy or difficult, and 3) attitudes felt towards reading. A review of the items remaining on the revised MRP-A showed that at least one item from the scale addressed each of these three areas, providing support for the claim that reading self-concept was not underrepresented.

Similarly, value for reading was defined along three interrelated aspects: 1) how much an individual likes or is interested in reading; 2) the perceived importance of doing well reading, and 3) the perceived usefulness of reading. While the 4 remaining items on the revised value subscale address issues related to individual interest in reading and the perceived usefulness of reading, the perceived importance of reading was not represented. This aspect of value for reading was represented by an item that was eventually deleted from the scale (item 12); however, the fact that this item was deleted was not surprising given that most of the emphasis in early schooling is placed on reading, thus almost all of the students surveyed responded that they felt reading was very important. It may be difficult to assess this aspect of value for reading in a society that places such a heavy emphasis on literacy, which would make such an item subject to a socially desirable response-set, especially when assessing young children (Lepola, 2004).

Finally, an examination of the goals for reading subscale revealed that there is a distinct possibility of construct underrepresentation for this subscale. However, when you consider the definition of reading goals used, the purposes for engaging in reading, there are a limitless number of possible reasons why a person may engage in reading. Therefore, a decision was made to only include items that represented the most common purposes for reading in school related contexts. With that in mind, the current set of
items seems to adequately represent the most common goals that students have for reading in schools: extrinsic reasons (i.e. to read better than their friends) and intrinsic reasons (i.e. to become a better reader to because they are genuinely interested in reading).

Overall, the remaining items on the revised MRP-A were evaluated to have content validity, in that they adequately represent each construct’s domain as defined in the current investigation. While the addition of new items might increase the content validity of each subscale, the developmental immaturity of young readers places a limit on the number of items that can be feasibly included on a scale directed at this age group.

**Substantive Aspect of Validity**

An evaluation of the substantive aspect of the MRP-A’s construct validity focused on an examination of the administration procedures and the effects that revision had on the reliability of the subscales, as measured by coefficient alpha. As was previously discussed, the MRP was originally designed to be administered to children from the 2nd to 6th grades. To increase the probability of response consistency across students, it was a priority to make sure that each student understood what each item was asking before they responded. Because the participants in this study were at the lower-bound edge of the appropriate age range for the instrument, the MRP-A was administered individually rather than to groups of students. This allowed the testers to monitor each student’s comprehension of each item, which should have improved response consistency. Also, potentially confusing items would elicit varied responses from individual students resulting in lower internal consistencies for the subscales. The fact that revision of the MRP-A took into consideration the impact that each item had on its subscale’s internal consistency would result in the deletion of any items that were confusing across respondents because of the negative impact that the item would have had on coefficient alpha for its subscale. Coefficient alpha for the revised self-concept subscale increased with item 11 removed from the original 10-item scale. Alpha decreased slightly for the
revised value subscale; however, this small decrease was expected since 6 of the original 10 items were deleted from the subscale for various reasons and because all 10 items had low inter-item correlations to begin with. Finally, coefficient alpha remained the same for the goals subscale despite the fact that item 26 was deleted during revision, reducing the number of items on the revised subscale from 8 to 7.

Another issue related to the substantive aspect of the MRP-A’s validity is the issue of using self-report measures with young children. Gambrell et al. (1996) reported that the items for the original MRP were designed on a 4-point scale based on findings in previous research that found young children have difficulty simultaneously differentiating among more than five discrete categories (Case & Khanna, 1981). Despite taking this into account when designing items for the MRP-A, there are still concerns related to how reliably young children respond to self-report measures, especially when you consider the abstract nature of motivation related questions. Previous research has shown that the reliability of self-report measures of self-concept used with young children tend to be lower than when used with older populations of students (Van Den Bergh & Rycke, 2003). This finding is consistent with the low internal consistencies found for each subscales on the MRP-A in the current investigation. Overall, the substantive aspect of construct validity was improved by the change in administration procedures and the revisions made to the scale.

**Structural Aspect**

The structural aspect of the MRP-A’s validity was tested by examining inter-item correlations within each subscale and through CFA’s that were conducted to confirm the factor structure of both the full and revised scales. The inter-item correlations for items in each of the three subscales were clearly improved by the revisions made to the MRP-A. Before revisions were made there were a number of items on each subscale that had extremely low or negative correlations with the other items on their respective subscales, indicating that they were not functioning well to measure the unified construct that they
were designed to assess. However, many of the items that were problematic in the full scale were deleted during revision because they failed other areas of item functioning, which was not surprising. Yet, an examination of the inter-item correlations for the revised MRP-A indicated that there were still problems despite removal of the extremely problematic items, especially for the value subscale. The problems encountered by the value subscale were not surprising given that these assessments were conducted on a young population of students, who are prone to giving socially desirable responses (Lepola, 2004), especially when asked questions about reading, which is heavily emphasized early in children’s academic careers. Another possible explanation for the low inter-item correlations is the inherently elusive nature of measuring motivation. Like motivation as a whole, self-concept, task-value, and achievement goals are complex individually, and can be influenced by many different interrelated factors. This complexity makes it very difficult to assess all of the possible factors and/or events that might influence a person’s motivation on a scale with a limited number of items.

The confirmatory factor analyses conducted on the MRP-A to test its factor structure following revision provided evidence to support a three factor structure for the scale representing its three subscales, reading self-concept, value for reading and goals for reading. At time 2, the three-factor solution fit significantly better than the one-factor model for both the full and revised scales. The results of these confirmed the hypothesis that the MRP-A is made up of three correlated factors, as the results clearly indicated that the MRP-A is best represented by a three factor model.

*Generalizability Aspect*

The generalizability aspect of validity is concerned with measurement consistency across population groups, measurement settings, and time. An analysis of the items from the revised MRP-A at time 3 indicated that two items functioned differently across gender. However, those same items showed no indication of bias at time 2 and they were
close to being insignificant at time 3. Thus, it was determined that no serious gender bias was present in any of the items on the MRP-A.

Another set of analyses examined the test-retest reliability of the three subscales on the revised MRP-A to test generalizability across time. Test-retest reliabilities for all three subscales were significant and consistent with those found in previous research on the MRP conducted by Gambrell et al. (1996). The moderate to high correlations between scores on the three subscales over time reflected the moderately stable nature of self-concept, task-value, and achievement goals that would be expected. Overall, the MRP-A functioned well with regards to the generalizability aspect of its construct validity.

External Aspect

Perhaps the most important piece of validity evidence for the MRP-A was criterion validity evidence, which the external aspect of Messick’s (1989) framework for unitary construct validity is designed to address. It was hypothesized that teachers’ ratings on the TRSM would be positively related to students’ scores on the MRP-A. It was assumed that because the TRSM was designed for teachers to rate student’s motivation to read based on the frequency of directly observable reading behaviors, that it would be the best possible criterion to test the validity of students’ responses on the MRP-A. A positive relationship between the MRP-A and TRSM would indicate that student’s responses on the MRP-A accurately reflect the frequency of directly observable behaviors indicative of their motivation level. However, the results from these analyses showed no relationship between scores on the MRP-A and TRSM. One possible reason for the lack of a relationship between teachers’ ratings of individual student’s motivation to read with student’s individual scores on the MRP-A could be that teacher’s ratings are being influenced by student’s reading skill level rather than rating students purely on behaviors that would indicate their level of motivation to read. The statistically significant correlation found between students’ scores on the TRSM and their spring
scores on the DIBELS indicated ratings on the TRSM may have been influenced by reading skill level. However, even with the possibility of TRSM ratings being influenced by students’ reading skill level, there should still be at least a minimal relationship between the TRSM and MRP-A, which was not the case.

Given that no relationship was found between the TRSM and the MRP-A, a second set of analyses tested the relationship between student’s scores on the MRP-A and their on-task behavior during formal reading instructional time. It was hypothesized that students who were more motivated to read would spend more time on-task during reading instructional time than their less motivated peers. Results of these analyses showed no relationship between student’s scores on the MRP-A and their on-task behavior during formal reading instructional time. One possible reason for the lack of a relationship could be that very few students were found to be off-task during reading instruction. Because so few students were observed to be off-task, there is a possibility that there was not enough variance in the time-on-task data to detect a relationship with scores on the MRP-A. However, like the analyses with the TRSM, the fact that there was no correlation at all between any of the items on the MRP-A with the time-on-task data at any time point suggests that there was no real meaningful relationship present.

Finally, because no relationship was found between the MRP-A and the TRSM or the time-on-task data, the relationship between scores on the MRP-A and reading fluency skill was tested at each time point. The results showed a significant relationship between the self-concept subscale of the MRP-A with some of the reading skill subtests at corresponding time points; however, the value and goals subscales had no significant correlations with reading skill subtests at any time point. The correlations found between the self-concept subscale with reading skill are consistent with findings from a study conducted by Chapman and Tunmer (1997) examining the development of the relationship between reading self-concept and reading achievement in K-2 students. Yet,
it is still problematic that no relationship was found between the other two subscales of the MRP-A with flat reading skill.

The lack of a relationship between the MRP-A with any of the criterion variables tested in the current investigation leaves questions regarding the criterion related validity of the MRP-A. Yet, the fact that none of the criterion variables correlated with one another suggests that they may have been poor indicators of students’ actual motivation to read. Based on the results of the current investigation, it remains unclear whether the lack of a relationship between the MRP-A with each of the criterion variables was due to problems with the MRP-A or the criterion variables themselves. These results suggest that future research needs to focus on identifying more reliable and valid indicators of students’ actual motivation to read that could be used to test the criterion validity of self-report measures like the MRP-A.
CHAPTER 3
STUDY 2

Study 2 was an examination of the “causal” relationship between motivation to read and the development of reading fluency skill over the course of students’ second grade academic year. Previous research in this area has focused on the relationship between academic self-concept and academic achievement, most of which has been general to a conglomerate of school subjects (i.e., reading, math, and social/natural sciences) and a few others that focused on a specific domain such as math or language arts. However, regardless of the level of specificity, all of the previous research has focused on this relationship over the course of multiple school years. One reason for examining these relationships over the course of multiple years is that achievement, which has been traditionally assessed using grades or teacher ratings of student ability, is relatively stable over the course of a single year, which statistically limits the explanatory power of such a study. However, the current investigation has taken a more specific approach by examining the relationship between motivation to read and the development of a specific reading skill (reading fluency), which has been shown to develop rapidly during students’ second grade year (Schwanenflugel et al., 2004).

On the basis of previous research examining the achievement/self-concept relationship, I formulated four possible models of the motivation to read/fluency skill development relationship to test for second grade readers which represent the four hypothesized relationships between motivation to read and reading fluency skill in the current investigation. The first model tested was the reciprocal model, which was based on previous research that has suggested a cross-lagged relationship may exist between a
student’s academic self-concept and their academic achievement. A cross-lagged relationship would be evident in the current investigation if motivation to read at each time point impacted reading fluency skill at the subsequent time point and reading fluency skill had a significant impact on motivation to read at the subsequent time point. This type of model has been supported by research that has argued students’ self-concepts become well established earlier than previous research has suggested and that students as young as the second grade have self-concepts that significantly impact their subsequent achievement (Guay, Marsh, & Boivin, 2003; Kurtz-Costes & Schneider, 1994; Mujis, 1997).

The second model tested was a skill development model, which implied that students’ fluency skill level at each time point will contribute significantly to fluency skill and motivation to read at the subsequent time point but it eliminates the paths from motivation to reading fluency. This model is based on the findings from previous research done by Chapman and Tunmer (1997) that suggested reading self-concept is influenced by previous reading achievement in primary aged children but that reading self-concept did not have a significant impact on later reading achievement. The skill development relationship emerges as a result of students using achievement related experiences in reading (both successes and failures) as a primary indicator when establishing their self-concept as a reader.

The third model tested was the self-enhancement model, which asserted that motivation to read will have an impact on reading fluency skill level each subsequent time point, but that reading fluency skill level will have no impact on motivation to read at subsequent time points. Traditionally it was believed that children’s academic self-concepts were still developing until later in their elementary school careers (Stipek, 1993) which limited its ability to impact subsequent achievement; however, more recent research has shown that children’s reading self-concepts begin to stabilize during their second grade year (Chapman & Tunmer, 1997). Although previous research on the
causal relationship between academic self-concept and academic achievement has not supported the self-enhancement model with young children, the finding that reading self-concept begins to stabilize during student’s second grade year leaves open the possibility that a self-enhancement relationship may emerge in the current investigation.

The final model tested was the independence model, which assumed that no cross-lagged effects will be found between motivation to read and fluency skill development at subsequent time points. This model implies that motivation to read will only contribute to motivation to read at each subsequent time point and fluency skill level will only contribute to fluency skill level at each subsequent time point. Path diagrams of all of the models to be tested can be found in Figure 1.

Based on the findings of previous research examining the relationship between academic self-concept and academic achievement level, it was hypothesized that the skill development model would best explain the data from the current investigation. This hypothesis was based on the assumption that young children’s reading self-concepts are still emerging, which could limit the ability of a student’s motivation to read to impact subsequent reading fluency skill. However, because two of the largest influences on a child’s academic self-concept are evaluative feedback from significant others (i.e. teachers) and successes/ failures experienced in academic situations, it is assumed that fluency skill level will have a significant impact on subsequent motivation to read.

Figure 3.1  Structural Models Tested to Examine the Relationship between Reading Fluency Skill and Motivation to Read
Model 2: Skill Development model
Time 1
MIR1
FLU1
Time 2
MIR2
FLU2
Time 3
MIR3
FLU3

Model 3: Self-Enhancement Model
Time 1
MIR1
FLU1
Time 2
MIR2
FLU2
Time 3
MIR3
FLU3

Model 4: Independence Model
Time 1
MIR1
FLU1
Time 2
MIR2
FLU2
Time 3
MIR3
FLU3

Note. MTR = motivation to read, FLU = reading fluency skill; numbers following variable names represent the time point of data collection; observed variables are not included in the structural models (see measurement model in Figure 1)
Method

Participants

Participants in this study were the same as those from Study 1.

Measures

Self-Reported Motivation to Read. The three revised subscales of MRP-A from Study 1 were used as observed indicators of the motivation to read latent variable.

Reading Fluency Skills. The reading skill assessments were the same as reported for Study 1. The TOWRE sight word efficiency, TOWRE phonemic decoding efficiency, and DIBELS Oral Reading Fluency subtests represent the three observed indicators for the reading fluency skill latent variable.

Procedure

The procedures for data collection were the same as reported for Study 1. The three waves of data collection represent the three time points used in the longitudinal analyses in this study.

Statistical Analyses

All of the analyses were conducted using the LISREL 8 (Joreskog & Sorbom, 1993) statistical software program. Fit of the various models run was determined using maximum likelihood estimation. Maximum likelihood estimation was chosen for two reasons: 1) it is the most widely used estimation procedure in SEM studies, and 2) it is the most accurate procedure for estimating model fit for normally distributed data. The overall fit of each model tested was evaluated in terms of the root mean square error of approximation (RMSEA) and comparative fit index (CFI), which are both fit indices reported when a model is tested on LISREL 8. The RMSEA is a standardized measure of the lack of fit of the population data to the model being tested; therefore, a small RMSEA is desirable. The CFI statistic represents the proportion of increased fit the hypothesized model shows over the null baseline model (a model with all of the parameters forced to equal zero) incorporating the expected values of the chi-square value for a model under
the non-central chi-square distribution. The RMSEA and CFI were chosen due to the fact that they take two different perspectives on assessing model fit; therefore, they compliment one another well when interpreting a model’s overall fit to the observed data. Hu and Bentler’s (1996) guidelines for evaluating model fit were used, with an RMSEA < .06 and a CFI > .95 indicating good model fit to the observed data.

Each of the four models tested contained two latent variables at each of the three time points of data collection: motivation to read and reading fluency skill. Motivation to read was defined by three indicators or observed subscales: reading self-concept, value for reading, and goals for reading at each time point. Reading fluency skill was also defined by three indicators or observed variables: the TOWRE sight word efficiency subscale, the TOWRE phonemic decoding efficiency subscale, and the DIBELS oral reading fluency subscale at each time point (means and standard deviations for each observed variable are reported in Table 3.1). Before any analyses were conducted the data was screened for outliers through an examination of scatter plots for each observed variable. Tests of univariate and multivariate normality were also conducted due to the fact that they are both basic assumptions of data when using maximum likelihood estimation to determine model fit (Kline, 1998). Tests of each variable’s univariate normality were all good with skewness and kurtosis statistics ranging from -1.26 - 2.02. A summary of each variable’s univariate normality statistics can be found in Table 3.1. The relative multivariate kurtosis statistic was 1.13 for the observed fluency variables and 1.10 for the observed motivation to read variables, indicating that the requirement of multivariate normality was met by all observed variables used in these analyses (a cut-off of < 2.0 was used). Because attrition was relatively low (9% of cases, or 17 out of 185 cases), listwise deletion methods were used for cases of missing data. The covariance matrix used in these analyses can be found in Table 3.2, with each scale’s variance on the diagonal.
Table 3.1  
*Descriptive Statistics for Fluency Skill and Motivation to Read Variables*

<table>
<thead>
<tr>
<th>Scale</th>
<th>M</th>
<th>SD</th>
<th>Skew</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSC 1</td>
<td>28.80</td>
<td>3.88</td>
<td>-.05</td>
<td>-.85</td>
</tr>
<tr>
<td>VR 1</td>
<td>13.26</td>
<td>1.94</td>
<td>-.93</td>
<td>.72</td>
</tr>
<tr>
<td>GR 1</td>
<td>22.18</td>
<td>3.79</td>
<td>-.11</td>
<td>-.99</td>
</tr>
<tr>
<td>DIBELS 1</td>
<td>65.20</td>
<td>26.59</td>
<td>.75</td>
<td>.67</td>
</tr>
<tr>
<td>TOWRE SWE 1</td>
<td>48.35</td>
<td>12.22</td>
<td>-.10</td>
<td>-.81</td>
</tr>
<tr>
<td>TOWRE PDE 1</td>
<td>21.20</td>
<td>9.78</td>
<td>.27</td>
<td>-.48</td>
</tr>
<tr>
<td>RSC 2</td>
<td>28.67</td>
<td>4.14</td>
<td>-.23</td>
<td>-.75</td>
</tr>
<tr>
<td>VR 2</td>
<td>13.16</td>
<td>2.10</td>
<td>-.66</td>
<td>-.08</td>
</tr>
<tr>
<td>GR 2</td>
<td>22.04</td>
<td>3.62</td>
<td>-.22</td>
<td>-.61</td>
</tr>
<tr>
<td>DIBELS 2</td>
<td>89.54</td>
<td>29.64</td>
<td>.45</td>
<td>-.06</td>
</tr>
<tr>
<td>TOWRE SWE 2</td>
<td>54.44</td>
<td>11.10</td>
<td>-.22</td>
<td>-.45</td>
</tr>
<tr>
<td>TOWRE PDE 2</td>
<td>24.09</td>
<td>11.22</td>
<td>.20</td>
<td>-.89</td>
</tr>
<tr>
<td>RSC 3</td>
<td>28.11</td>
<td>3.81</td>
<td>-.06</td>
<td>-.65</td>
</tr>
<tr>
<td>VR 3</td>
<td>12.93</td>
<td>2.36</td>
<td>-1.26</td>
<td>2.02</td>
</tr>
<tr>
<td>GR 3</td>
<td>20.96</td>
<td>3.80</td>
<td>-.28</td>
<td>-.08</td>
</tr>
<tr>
<td>DIBELS 3</td>
<td>99.68</td>
<td>29.81</td>
<td>.30</td>
<td>.23</td>
</tr>
<tr>
<td>TOWRE SWE 3</td>
<td>58.35</td>
<td>9.81</td>
<td>-.54</td>
<td>.24</td>
</tr>
<tr>
<td>TOWRE PDE 3</td>
<td>28.55</td>
<td>10.66</td>
<td>.20</td>
<td>-.26</td>
</tr>
</tbody>
</table>

*Note.* RSC = reading self-concept, VR = value for reading, GR = goals for reading (time point represented by the number following the test name); N = 168 for all variables except DIBELS 1, which had an N of 167

Table 3.2  
*Variances and Covariances of Reading Fluency and Motivation to Read*

<table>
<thead>
<tr>
<th>Scale</th>
<th>RSC 1</th>
<th>VR 1</th>
<th>GR 1</th>
<th>RSC 2</th>
<th>VR 2</th>
<th>GR 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSC 1</td>
<td>15.16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VR 1</td>
<td>2.62</td>
<td>3.78</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GR 1</td>
<td>8.11</td>
<td>3.03</td>
<td>14.22</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RSC 2</td>
<td>8.72</td>
<td>1.75</td>
<td>5.23</td>
<td>2.07</td>
<td>2.79</td>
<td>2.45</td>
</tr>
<tr>
<td>VR 2</td>
<td>1.29</td>
<td>1.78</td>
<td>7.17</td>
<td>4.39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GR 2</td>
<td>5.91</td>
<td>2.25</td>
<td>8.05</td>
<td>8.53</td>
<td>2.82</td>
<td>7.29</td>
</tr>
<tr>
<td>RSC 3</td>
<td>7.46</td>
<td>1.70</td>
<td>5.32</td>
<td>10.61</td>
<td>2.82</td>
<td>2.15</td>
</tr>
<tr>
<td>VR 3</td>
<td>1.50</td>
<td>2.14</td>
<td>2.57</td>
<td>1.90</td>
<td>2.56</td>
<td>2.15</td>
</tr>
<tr>
<td>GR 3</td>
<td>5.51</td>
<td>3.13</td>
<td>7.78</td>
<td>5.63</td>
<td>2.89</td>
<td>9.25</td>
</tr>
<tr>
<td>DIB 1</td>
<td>14.87</td>
<td>0.71</td>
<td>3.31</td>
<td>27.36</td>
<td>-0.31</td>
<td>1.68</td>
</tr>
<tr>
<td>TSWE 1</td>
<td>6.71</td>
<td>0.40</td>
<td>1.31</td>
<td>11.91</td>
<td>-0.30</td>
<td>-0.66</td>
</tr>
<tr>
<td>TPDE 1</td>
<td>3.19</td>
<td>-0.26</td>
<td>-3.12</td>
<td>6.09</td>
<td>-0.47</td>
<td>-1.07</td>
</tr>
<tr>
<td>DIB 2</td>
<td>18.66</td>
<td>2.71</td>
<td>2.68</td>
<td>26.27</td>
<td>0.22</td>
<td>1.88</td>
</tr>
<tr>
<td>TSWE 2</td>
<td>5.22</td>
<td>-0.25</td>
<td>-0.34</td>
<td>9.95</td>
<td>-0.94</td>
<td>.58</td>
</tr>
<tr>
<td>TPDE 2</td>
<td>4.55</td>
<td>-0.91</td>
<td>-1.98</td>
<td>8.39</td>
<td>-1.01</td>
<td>-1.34</td>
</tr>
<tr>
<td>DIB 3</td>
<td>21.67</td>
<td>4.76</td>
<td>9.69</td>
<td>28.43</td>
<td>4.57</td>
<td>5.57</td>
</tr>
<tr>
<td>TSWE 3</td>
<td>5.57</td>
<td>1.01</td>
<td>1.99</td>
<td>10.76</td>
<td>0.13</td>
<td>1.85</td>
</tr>
<tr>
<td>TPDE 3</td>
<td>5.87</td>
<td>0.45</td>
<td>1.08</td>
<td>9.71</td>
<td>-0.28</td>
<td>.64</td>
</tr>
</tbody>
</table>

---
<table>
<thead>
<tr>
<th>RSC 3</th>
<th>14.30</th>
</tr>
</thead>
<tbody>
<tr>
<td>VR 3</td>
<td>2.40</td>
</tr>
<tr>
<td>GR 3</td>
<td>7.47</td>
</tr>
<tr>
<td>DIB 1</td>
<td>24.67</td>
</tr>
<tr>
<td>TSWE 1</td>
<td>8.48</td>
</tr>
<tr>
<td>TPDE 1</td>
<td>4.05</td>
</tr>
<tr>
<td>DIB 2</td>
<td>24.14</td>
</tr>
<tr>
<td>TSWE 2</td>
<td>6.83</td>
</tr>
<tr>
<td>TPDE 2</td>
<td>5.98</td>
</tr>
<tr>
<td>DIB 3</td>
<td>26.52</td>
</tr>
<tr>
<td>TSWE 3</td>
<td>7.42</td>
</tr>
<tr>
<td>TPDE 3</td>
<td>6.45</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DIB 2</th>
<th>TSWE 2</th>
<th>TPDE 2</th>
<th>DIB 3</th>
<th>TSWE 3</th>
<th>TPDE 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIB 2</td>
<td>880.71</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TSWE 2</td>
<td></td>
<td>123.84</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TPDE 2</td>
<td></td>
<td>101.78</td>
<td>126.63</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIB 3</td>
<td></td>
<td>287.80</td>
<td>257.37</td>
<td>888.26</td>
<td></td>
</tr>
<tr>
<td>TSWE 3</td>
<td></td>
<td>96.88</td>
<td>83.33</td>
<td>245.47</td>
<td>96.67</td>
</tr>
<tr>
<td>TPDE 3</td>
<td></td>
<td>93.17</td>
<td>104.49</td>
<td>245.65</td>
<td>81.74</td>
</tr>
</tbody>
</table>

Note. RSC = reading self-concept, VR = value for reading, GR = goals for reading, TSWE = TOWRE sight word efficiency, TPDE = TOWRE phonemic decoding efficiency, DIB = DIBELS (number following each subtest represents the time point of data collection); variances for each subtest are on the diagonal.

The measurement model was tested first to examine how well the six observed variables measured the two latent variables over all three time points (see Figure 2). Once the measurement model was tested, the four hypothesized structural models were tested. All five models tested include correlated measurement errors for each subscale at adjacent time points, which assumes that the measurement error for repeated measures of each subtest at adjacent time-points were correlated for reasons outside of the model (Kline, 1998). The structural models were tested individually, and because they are nested models, the chi-square difference statistic ($\chi^2_{\text{diff}}$) was then calculated to make comparisons between models.

In the current investigation, the structural portion of the least restrictive model in Figure 1 (model 1), which assumes a cross-lagged effect between motivation to read and fluency skill level, was compared with the measurement model to determine how much replacing the 15 correlations among the six constructs with the 9 structural paths worsened the fit of the model. Then, $\chi^2_{\text{diff}}$ tests were used to compare the fit of the independence model, which was the most restrictive model representing the null
hypothesis, with the fit of the less restrictive skill development and self-enhancement models. Finally, the fit of the skill development, self-enhancement, and independence models were compared to the fit of the reciprocal model, which was the least restrictive model tested. All comparisons used the chi-square difference statistic ($\chi^2_{\text{diff}}$) to determine if each model became significantly better by adding paths to the most restrictive model.

Figure 3.2

Full Measurement Model of Fluency Skill and Motivation to Read

Note. MTR = motivation to read, FLU = reading fluency skill; RSC = reading self-concept, VR = value for reading, GR = goals for reading, DIB = DIBELS, TSWE = TOWRE sight word efficiency subtest, TPDE = TOWRE phonemic decoding efficiency subtest; $\epsilon$ = error term for each variable; the number following each abbreviated variable’s name represents the time point of data collection
Results

Measurement Model. The first model tested was the full measurement model, which examined the relationship between the two latent variables (motivation to read and reading fluency skill) and their observed indicator variables (the three subscales on the MRP-A and the three reading fluency subtests) at all three time points. The full measurement model had a $\chi^2 (108) = 248.12 \ (p < .001)$, an RMSEA = .082, and a CFI = .97. The RMSEA of .082 was slightly larger than the cut off of .06 indicating marginally good model fit; however, the CFI of .97 indicated that the model was significantly better than the baseline model with all of the loadings forced to equal zero. The factor loadings in the measurement model were all statistically significant ($p < .05$), indicating that all of the subtests were significantly related to the factor they were designed to assess. The variance explained in each of the fluency subtests by the reading fluency latent factor were substantial, with $R^2$ ranging from .72 to .91, indicating that almost all of the variance in these subtests was explained by the reading fluency skill factor. However, this was not the case for two of the three subscales on the MRP-A. Variance explained in the goals subscale by the latent motivation to read factor over all three time points was good, with $R^2$ ranging from .65 - .76. Yet the variance explained in the self-concept subscale by the motivation to read factor was low with $R^2$ ranging from .38 - .47, indicating that a majority of variance in the self-concept subscale was explained by factors outside of the model (factors other than the motivation to read latent factor). The value for reading subscale performed even worse, with $R^2$ values ranging from .22 to .30, again indicating that the majority of variance in the value subscale was not explained by the motivation to read factor. A summary of all the $R^2$ values, unstandardized factor loadings, their standard errors and subsequent t-values for the full measurement model are presented in Table 3.3. Also, correlations between all of the latent variables in the measurement model are reported in Table 3.4. Overall, the measurement model fit
marginally well to the observed data; however, the low $R^2$s of the three observed subscales on the MRP-A indicate that they are not correlated very highly.

Table 3.3  
*Parameter Estimates for the Measurement Model of Motivation to Read and Reading Fluency Skill*

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Time 1</th>
<th>Time 2</th>
<th>Time 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Factor Loadings (SE)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTR ---&gt; Self-concept</td>
<td>1.00$_{nt}$</td>
<td>1.00$_{nt}$</td>
<td>1.00$_{nt}$</td>
</tr>
<tr>
<td>MTR ---&gt; Value</td>
<td>.41* (.07)</td>
<td>.37* (.07)</td>
<td>.54* (.09)</td>
</tr>
<tr>
<td>MTR ---&gt; Goals</td>
<td>1.20* (.16)</td>
<td>1.10* (.15)</td>
<td>1.40* (.19)</td>
</tr>
<tr>
<td>FLU ---&gt; DIBELS</td>
<td>1.00$_{nt}$</td>
<td>1.00$_{nt}$</td>
<td>1.00$_{nt}$</td>
</tr>
<tr>
<td>FLU ---&gt; TOWRE SWE</td>
<td>.48* (.02)</td>
<td>.38* (.02)</td>
<td>.32* (.02)</td>
</tr>
<tr>
<td>FLU ---&gt; TOWRE PDE</td>
<td>.33* (.02)</td>
<td>.34* (.02)</td>
<td>.33* (.02)</td>
</tr>
<tr>
<td><strong>$R^2$ of Observed Variables</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTR ---&gt; Self-concept</td>
<td>.43</td>
<td>.47</td>
<td>.38</td>
</tr>
<tr>
<td>MTR ---&gt; Value</td>
<td>.30</td>
<td>.22</td>
<td>.30</td>
</tr>
<tr>
<td>MTR ---&gt; Goals</td>
<td>.66</td>
<td>.65</td>
<td>.76</td>
</tr>
<tr>
<td>FLU ---&gt; DIBELS</td>
<td>.85</td>
<td>.89</td>
<td>.87</td>
</tr>
<tr>
<td>FLU ---&gt; TOWRE SWE</td>
<td>.91</td>
<td>.90</td>
<td>.81</td>
</tr>
<tr>
<td>FLU ---&gt; TOWRE PDE</td>
<td>.71</td>
<td>.77</td>
<td>.72</td>
</tr>
</tbody>
</table>

*Note.  MTR = motivation to read, FLU = reading fluency skill; SE = standard error of the unstandardized path loading; nt indicates that the path loading was not tested for significance since it was set to equal 1.0; *$p<.01$*
Table 3.4  
*Correlations between Motivation to Read and Reading Fluency Skill*

<table>
<thead>
<tr>
<th>Variable</th>
<th>MTR 1</th>
<th>FLU 1</th>
<th>MTR 2</th>
<th>FLU 2</th>
<th>MTR 3</th>
<th>FLU 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>MTR 1</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FLU 1</td>
<td>.07</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTR 2</td>
<td>.82</td>
<td>.10</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FLU 2</td>
<td>.06</td>
<td>.97</td>
<td>.09</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTR 3</td>
<td>.77</td>
<td>.03</td>
<td>.77</td>
<td>.03</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>FLU 3</td>
<td>.14</td>
<td>.97</td>
<td>.16</td>
<td>.97</td>
<td>.11</td>
<td>1.00</td>
</tr>
</tbody>
</table>

*Note.* MTR = motivation to read, FLU = reading fluency skill, the number following each abbreviated variable represents the time point of data collection; all boldface correlations were collected at the same time; all underlined correlations were cross-lagged in the reciprocal effects model.

*Independence Model.* The first structural model tested examined the possibility that motivation to read and reading fluency skill development operate independently of one another. This model was tested first because it represents the null hypothesis and was the benchmark used to evaluate all of the other structural models tested. Not surprisingly, all of the paths in this model were significant (all *p* > .05), indicating that MTR at each time point had a significant impact on MTR at the subsequent time point and FLU at each time point also had a significant impact on FLU at the subsequent time point. All parameter estimates, their standard errors, and t-values for the independence model are summarized in Table 3.5. Overall, the independence model showed marginally good fit to the observed data with a $\chi^2 (118) = 268.29$, an RMSEA = .081, and a CFI = .97.

*Skill Development Model.* The second structural model examined was the skill development model, which tested the hypothesis that reading fluency skill (FLU) would impact *motivation to read* (MTR) at subsequent time points, but that MTR would have no effect on FLU at subsequent time points. The results from this model showed no
indication that FLU impacted MTR at adjacent time points, with all of the skill development parameters insignificant \( p > .05 \). A summary of all the parameter estimates, their standard errors, and t-values for this model can be found in Table 3.5. Similar to the independence model, the overall fit of the skill development model was acceptable, \( \chi^2 (116) = 267.75 \ (p < .01) \), an RMSEA = .082, and a CFI = .97. However, the fact that the structural equation analysis found that reading fluency skill had little impact on subsequent motivation suggests little support for this model. A comparison between the independence and skill development models confirmed this, with a \( \chi^2_{\text{diff}} (2) = .54 \ (p = .76) \), indicating that the added skill development paths did not significantly improve the model over the independence model given the difference in degrees of freedom.

**Self-Enhancement Model.** The third model tested was the self-enhancement model, which tested the hypothesis that motivation to read (MTR) would impact reading fluency skill (FLU) at subsequent time points, but that FLU would have no impact on MTR at subsequent time points. In this model, MTR at time 1 had no impact on FLU at time 2; however, MTR at time 2 did have a significant impact on FLU at time 3 with a parameter estimate of .83 \( (p < .05) \). This result indicated the possibility that students’ motivation to read may have begun to stabilize in the middle of their second grade year, which allowed it to have a significant impact of reading fluency skill development toward the end of the year. A summary of all the parameter estimates, their standard errors, and t-values for this model can be found in Table 3.5. Overall, this model showed a slight improvement in fit to the observed data, with a \( \chi^2 (116) = 261.73 \ (p < .01) \), an RMSEA = .081, and a CFI = .97. Because the independence model fit the observed data better than the skill development model in the previous comparison, the self-enhancement model was only compared to the independence model. The comparison between the independence and self-enhancement models revealed a significant difference, with the \( \chi^2_{\text{diff}} (2) = 6.56 \ (p < .05) \), which indicated that the added self-enhancement paths were a
significant improvement over the independence model given the difference in degrees of freedom.

**Reciprocal Effects Model.** The final structural model tested examined the possibility of a cross-lagged effect between *motivation to read* (MTR) and reading fluency skill (FLU) at the subsequent time point, which reflects the hypothesis implied by the reciprocal effects model. This model was tested last because it was the least restrictive model to be tested. In this model, the paths between MTR at time 1 and MTR at time 2, MTR at time 2 and MTR at time 3, FLU at time 1 and FLU at time 2, and FLU at time 2 and FLU at time 3 were all significant ($p < .05$). However, only the cross-lagged parameter estimate from MTR at time 2 on FLU at time 3 was significant, indicating that MTR only influenced FLU toward the middle/end of students’ second grade year. A summary of all the parameter estimates, their standard errors, and t-values for the reciprocal model can be found in Table 3.5. Like the previous models tested, this model fit marginally well, with a $\chi^2 (114) = 261.10$ ($p < .01$), an RMSEA = .082 and a CFI = .97, but the finding that only one of the cross-lagged parameter estimate was significant indicated little support for the autoregressive view of skill and motivation taken by this model. This model was compared with both the independence and self-enhancement models. The comparison with the independence model revealed no significant difference, with a $\chi^2_{\text{diff}} (4) = 7.19$ ($p = .13$), indicating that the reciprocal model did not fit significantly better than the null model, given the cost in degrees of freedom. The comparison with the self-enhancement model also showed no significant difference between models, with a $\chi^2_{\text{diff}} (2) = .63$ ($p = .73$). Taken together, the results of these comparisons and those conducted previously clearly indicated that the self-enhancement model fit the observed data the best out of the four hypothesized models. A summary of all model comparisons can be found in Table 3.6.
Table 3.5  Parameter Estimates of the Structural Models Tested Examining the Relationship between Motivation to Read and Reading Fluency Skill

<table>
<thead>
<tr>
<th>Model</th>
<th>Path</th>
<th>Parameter Estimate</th>
<th>SE</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reciprocal</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MTR 1→→MTR 2</td>
<td>.86**</td>
<td>.130</td>
<td>.73</td>
</tr>
<tr>
<td></td>
<td>FLU 1→→MTR 2</td>
<td>.003</td>
<td>.007</td>
<td>.73</td>
</tr>
<tr>
<td></td>
<td>MTR 1→→FLU 2</td>
<td>-.04</td>
<td>.340</td>
<td>.73</td>
</tr>
<tr>
<td></td>
<td>FLU 1→→FLU 2</td>
<td>1.10**</td>
<td>.050</td>
<td>.94</td>
</tr>
<tr>
<td></td>
<td>MTR 2→→MTR 3</td>
<td>.75**</td>
<td>.098</td>
<td>.94</td>
</tr>
<tr>
<td></td>
<td>FLU 2→→MTR 3</td>
<td>-.004</td>
<td>.005</td>
<td>.65</td>
</tr>
<tr>
<td></td>
<td>MTR 2→→FLU 3</td>
<td>.84*</td>
<td>.330</td>
<td>.95</td>
</tr>
<tr>
<td></td>
<td>FLU 2→→FLU 3</td>
<td>.97**</td>
<td>.036</td>
<td>.96</td>
</tr>
<tr>
<td><strong>Skill Development</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MTR 1→→MTR 2</td>
<td>.85**</td>
<td>.130</td>
<td>.72</td>
</tr>
<tr>
<td></td>
<td>FLU 1→→MTR 2</td>
<td>.003</td>
<td>.007</td>
<td>.72</td>
</tr>
<tr>
<td></td>
<td>FLU 1→→FLU 2</td>
<td>1.10**</td>
<td>.050</td>
<td>.94</td>
</tr>
<tr>
<td></td>
<td>MTR 2→→MTR 3</td>
<td>.75**</td>
<td>.097</td>
<td>.65</td>
</tr>
<tr>
<td></td>
<td>FLU 2→→MTR 3</td>
<td>-.004</td>
<td>.005</td>
<td>.64</td>
</tr>
<tr>
<td></td>
<td>FLU 2→→FLU 3</td>
<td>.97**</td>
<td>.036</td>
<td>.95</td>
</tr>
<tr>
<td><strong>Self-Enhancement</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MTR 1→→MTR 2</td>
<td>.86**</td>
<td>.130</td>
<td>.73</td>
</tr>
<tr>
<td></td>
<td>MTR 1→→FLU 2</td>
<td>-.04</td>
<td>.340</td>
<td>.73</td>
</tr>
<tr>
<td></td>
<td>FLU 1→→FLU 2</td>
<td>1.10**</td>
<td>.050</td>
<td>.94</td>
</tr>
<tr>
<td></td>
<td>MTR 2→→MTR 3</td>
<td>.78**</td>
<td>.097</td>
<td>.65</td>
</tr>
<tr>
<td></td>
<td>MTR 2→→FLU 3</td>
<td>.83*</td>
<td>.330</td>
<td>.95</td>
</tr>
<tr>
<td></td>
<td>FLU 2→→FLU 3</td>
<td>.97**</td>
<td>.036</td>
<td>.96</td>
</tr>
<tr>
<td><strong>Independence</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MTR 1→→MTR 2</td>
<td>.85**</td>
<td>.130</td>
<td>.72</td>
</tr>
<tr>
<td></td>
<td>FLU 1→→FLU 2</td>
<td>1.10**</td>
<td>.050</td>
<td>.94</td>
</tr>
<tr>
<td></td>
<td>MTR 2→→MTR 3</td>
<td>.75**</td>
<td>.100</td>
<td>.65</td>
</tr>
<tr>
<td></td>
<td>FLU 2→→FLU 3</td>
<td>.97**</td>
<td>.036</td>
<td>.95</td>
</tr>
</tbody>
</table>

*Note.* MTR = motivation to read, FLU = reading fluency skill, each abbreviated variable is followed by a number signifying the time point for that variable; R² values represent the amount of variance explained in a variable by the combination of all variables with a path leading to it; *p < .05   **p < .01
Table 3.6  
Fit Indices and Comparisons between Models

<table>
<thead>
<tr>
<th>Model</th>
<th>$\chi^2$</th>
<th>df</th>
<th>RMSEA</th>
<th>CFI</th>
<th>$\chi^2_{\text{diff}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independence</td>
<td>268.29</td>
<td>118</td>
<td>.081</td>
<td>.97</td>
<td></td>
</tr>
<tr>
<td>Skill Development</td>
<td>267.75</td>
<td>116</td>
<td>.082</td>
<td>.97</td>
<td></td>
</tr>
<tr>
<td>Comparison between models</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td>.54</td>
</tr>
<tr>
<td>Independence</td>
<td>268.29</td>
<td>118</td>
<td>.081</td>
<td>.97</td>
<td></td>
</tr>
<tr>
<td>Self-Enhancement</td>
<td>261.73</td>
<td>116</td>
<td>.081</td>
<td>.97</td>
<td></td>
</tr>
<tr>
<td>Comparison between models</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td>6.56*</td>
</tr>
<tr>
<td>Independence</td>
<td>268.29</td>
<td>118</td>
<td>.081</td>
<td>.97</td>
<td></td>
</tr>
<tr>
<td>Reciprocal Effects</td>
<td>261.10</td>
<td>114</td>
<td>.082</td>
<td>.97</td>
<td></td>
</tr>
<tr>
<td>Comparison between models</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>7.19</td>
</tr>
<tr>
<td>Self-Enhancement</td>
<td>261.73</td>
<td>116</td>
<td>.081</td>
<td>.97</td>
<td></td>
</tr>
<tr>
<td>Reciprocal Effects</td>
<td>261.10</td>
<td>114</td>
<td>.082</td>
<td>.97</td>
<td></td>
</tr>
<tr>
<td>Comparison between models</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td>.63</td>
</tr>
</tbody>
</table>

Note. $\chi^2$ = chi-square statistic, df = degrees of freedom, RMSEA = root mean square error of approximation, CFI = comparative fit index, $\chi^2_{\text{diff}}$ = chi-square difference statistic; *$p < .05$

Discussion

The results of this study provided surprising evidence for a self-enhancement relationship between motivation to read and the development of reading fluency skill for second grade students. Contrary to previous studies on the relationship between academic self-concept and academic achievement in young children, which have mostly found achievement to be predominant over self-beliefs (Chapman & Tunmer, 1997; Helmke & van Aken, 1995), the current investigation found no influence of reading fluency skill on motivation to read at subsequent time points. Similarly, the results of this study do not fully support previous findings of a possible reciprocal effect where academic self-concept effects academic achievement at subsequent time points and academic achievement impacts academic self-concept at subsequent time points (Kurtz-Costes & Schneider, 1994). However, the results of this study were partially consistent with the study conducted by Guay, Marsh, and Boivin (2003), which found a partial self-enhancement effect where academic self-concept influenced academic achievement in the latter portion of their model (from time 2 to time 3); however, their results also suggested a reciprocal effect in the earlier portion of the model (from time 1 to time 2) which was
not found to be the case in the current investigation. The fact that results from this study were not consistent with previous studies on the academic self-concept/academic achievement relationship is not surprising given the many differences regarding the level of specificity in their respective designs. Yet, the fact that the self-enhancement model fit better than each of the other three hypothesized models was unexpected, even taking these differences into account.
CHAPTER 4
GENERAL DISCUSSION

The purpose of the present research was twofold. First, it examined the validity of the MRP-A as a self-report measure designed to assess 2nd grade students’ motivation to read. The validity of the MRP-A was systematically scrutinized along five aspects of Messick’s (1989) framework for unitary construct validity. Second, the present research explored the causal relationship between motivation to read and fluency skill development during students’ second grade year. Because it was the first study of this kind to be conducted, research findings on the causal relationship between academic self-concept and academic achievement were used as a reference point. Based on those findings, four possible models were hypothesized as possible explanations for the relationship between motivation to read and reading fluency skill development.

The results of the validation study of the MRP-A yielded mixed results. As part of this validation study, the original 28-item scale was revised to improve the content validity and psychometric performance of items comprising each of the three subscales on the MRP-A: reading self-concept, value for reading, and goals for reading. To this end, the revisions proved to be effective at significantly improving the construct validity of the MRP-A along four of the five aspects of Messick’s (1989) framework. When compared with the full 28-item scale, the 20-item revised scale showed significant improvements along the content, substantive, structural, and generalizability aspects of construct validity. However, the question of the MRP-A’s criterion related validity remained unanswered, with both the full and revised scales failing to show any significant relationship to the external criterion variables they were tested against. Follow up
analyses also showed no relationship amongst the criterion variables, which brought into question their utility as benchmarks for assessing the criterion validity of the MRP-A. More work still needs to be done to conceptualize better criterion variables before a reliable judgment of the MRP-A’s criterion related validity can be made.

Overall, based on the findings of the current validation study, the MRP-A was generally successful along the content, substantive, and generalizability aspects of construct validity. The self-concept and goals subscales also functioned well along the structural aspect; however, the 4-item value subscale was plagued with low inter-item correlations and marginal internal consistencies. Yet, despite the structural problems with the value subscale, confirmatory factor analysis provided evidence to support a three-factor structure for the scale, which was important given that they were used as three separate indicators of *motivation to read* in Study 2. While there is definitely room for improvement in measuring students’ *motivation to read*, the MRP-A functioned well enough to warrant its use in the second study of this investigation.

The results of the second study provided surprising evidence supporting a self-enhancement model to explain the causal relationship between *motivation to read* and reading fluency skill development. Specifically, the findings from the current investigation show that *motivation to read* and reading fluency skill act independently of one another from time 1 to time 2, but that *motivation to read* at time 2 has an impact on reading fluency skill at time 3 (see Figure 4.1). This finding was different from findings in previous research that have suggested a skill development model best describes the relationship between academic self-concept and academic achievement in young children due to the possibility that self-system variables do not begin to stabilize until the late stages of a student’s elementary school experience (Kurtz-Costes & Schneider, 1994; Skaalvik, 1997; Skaalvik & Hagvet, 1990). However, research conducted by Chapman and Tunmer (1997) found that reading self-concept began to stabilize sometime in student’s third year of schooling (ie. their second grade year). The findings of the current
investigation support Chapman and Tunmer’s findings, in that student’s scores on the MRP-A became more stable as they progressed through their second grade year, which could have influenced the emergence of a self-enhancement effect over the second and third time points.

Figure 4.1 The Self-Enhancement Model: Unstandardized Path Weights and Parameter Estimates

Note. Solid lines represent significant paths; dashed lines represent nonsignificant paths. MTR = motivation to read, FLU = reading fluency skill; numbers following variable names represent the time point of data collection; rsc = reading self-concept, vr = value for reading, gr = goals for reading, tsw = TOWRE sight word efficiency, tpd = TOWRE phonemic decoding efficiency, dib = DIBELS oral reading fluency.

In order to understand how this self-enhancement effect emerged halfway through the students’ second grade year, it is important to consider previous research on the development of academic self-concept. Evidence supporting the development of self-concept during student’s early elementary years was provided by a study conducted by Van den Bergh and Rycke (2003). They explored the multidimensionality of students’ academic self-concepts through an examination of its factor structure with first, second,
and third grade students. Interestingly, they found that the factor structure changed as students progressed through the primary grades. Specifically, they found that the factor structure changed for first and second graders and remained the same for second and third grade students signifying that their academic self-concept became more differentiated and began to stabilize sometime during students’ second grade year. They also found that the measurement error of self-concept was larger for second graders than it was for third graders, signifying that the reliability of measuring academic self-concept increased as students progressed through second grade and into third grade. Both of these findings provide supportive evidence for the partial self-enhancement effect found later in students’ second grade year in the current investigation. Because the first time point of data was collected during the first month of students’ second grade year, it seems likely that this stabilization had not yet occurred, which helps to explain why this effect emerged in the data from the second and third time points, which were collected at the middle and end of the students’ second grade year.

Another possible explanation for the differences between the findings of the current investigation and those of previous research conducted on the self-concept/achievement relationship has to do with the level of specificity of measurement in those studies. Previous research has conducted longitudinal studies over the course of multiple school years using broad measures of achievement such as grades or achievement test scores. It was hypothesized that by increasing the level of specificity of measurement in the current investigation to examine the development of motivation to read rather than general academic self-concept and a specific reading skill (reading fluency) rather than general academic achievement would result in a more sensitive test of the causal relationship than was achieved in previous research. This sensitivity was important given that the measures of reading fluency skill accounted for almost all of the variance in the reading fluency latent variable, leaving little added variance for the motivation variable to explain. In fact, it seems that this was the case given that motivation to read at time 2
was able to explain a statistically significant amount of added variance in reading fluency skill at time 3 in addition to the 95% explained by reading fluency skill at time 2.

A third unrelated, yet relevant possible explanation for the differences between the findings in the current investigation and some of the previous research that has been conducted on the causal relationship between academic self-concept and academic achievement was described by Guay et al. (2003). They pointed out that much of the research on the causal relationship between academic self-concept and academic achievement had been conducted in Norwegian countries where children are not given grades until they reach middle school. This is an important because grades are one of a few important sources of external feedback that children use to form their academic self-concept. In contrast, students in the United States are given evaluative feedback at increasingly young ages, with high stakes tests starting as early as the first grade, which could have a significant impact on the development of academic self-concept. Harter’s (1996) claim that socialization experiences with caregivers have a significant impact on the content of self-evaluations, which implies that the aforementioned differences between social climates could have a significant effect on developmental differences between cultures.

**Implications for Educational Practice.** The results of the current investigation hold implications for educational practice. Previous research has suggested that academic achievement, or the development of specific academic skills, predominate over self-system variables that influence academic motivation (Chapman & Tunmer, 1997; Helmke & van Aken, 1995; Skaalvik & Valas, 1999). From this view, instruction should focus primarily on developing students’ academic skills, based on the assumption that as they become more skilled they will also become more motivated to achieve. The lack of a relationship between reading fluency skill and subsequent motivation in the current investigation contradicts this belief. Thus, these results indicate that early reading instruction needs to address more than simply developing specific reading skills in young
readers. Rather, instruction should be designed in a way that develops students’ motivation for reading while developing specific reading skills at the same time. This type of instruction would imbed motivationally relevant techniques within instruction designed to develop specific reading skills. Motivationally relevant techniques might include things like helping students to begin setting goals for their reading development, emphasizing the value in becoming a better reader by showing students how their developing skills can be utilized to achieve personally relevant goals, and by fostering a belief in students that with effort, they can improve their reading skill. A specific example of this would be a teacher who holds individual conferences with each student once a month to discuss individual progress and set goals for the upcoming month relevant to the specific reading skills that are being targeted during formal reading instructional time. While these conferences would only take approximately five minutes per student each month, they could have a profound influence on focusing students’ attention on their own personal progress and teach them the utility in setting specific academic goals to monitor their progression.

These results also hold implications for early interventions designed to remediate early reading problems. Many of the existing remedial reading interventions focus solely on accelerating reading skill growth, so that students can “catch up” with their peers. In order to achieve this goal, many programs take a “skill and drill” approach to instruction, ignoring the possible impact that motivation plays in remediating reading difficulties. For example, popular programs such as SRA’s Reading Mastery (Englemann & Bruner, 1988) devote almost all of their instructional time to working on phoneme identification and blending as well as explicit phonics instruction focused on letter-sound correspondences. While these techniques have been shown to be effective in accelerating decoding skill growth, they ignore the possible influence of motivating struggling readers through the use of authentic texts and discussions regarding personally relevant uses for their new found skills. Similar to the implications for reading instruction in the regular
classroom, the results of this study indicate that early reading interventions should also take into account the possible contribution that motivational variables could make in accelerating reading growth.

Limitations. Although the results of present investigation supported a self-enhancement model to explain the relationship between motivation to read and reading fluency skill development, several limitations of this study should be taken into consideration when interpreting these findings. First, the results of Study 1 suggested possible validity concerns for using the MRP-A to assess motivation to read with young children. While SEM techniques help to mitigate some of the reliability issues plaguing the MRP-A, the finding that the three motivational subscales were not highly correlated with the motivation to read latent variable may indicate the possibility that some important relationships were not identified. Second, previous studies with similar designs have suggested the use of multiple cohorts of students to allow comparisons between cross-sectional results with those from the true longitudinal design (Marsh, Byrne, & Yeung, 1999). While this type of design would have clearly been desirable to check the replicability of the current results, it was not feasible given the resources available for data collection in the current investigation. Third, the fact that the three reading fluency skill subscales accounted for almost all of the variability in the reading fluency skill latent variable across time points may have masked the true ability of the motivation to read variable to explain any added variance. Therefore, it is possible that motivation to read had more of an influence of reading fluency skill than was found due to the fact that it was masked by the strong relationships found between the fluency measures. Finally, because it was the first study conducted specifically on the relationship between motivation to read and reading fluency skill development, it was exploratory at times, based on the fact that there was not a significant body of previous research to guide the current investigation.
In conclusion, the current study was designed to examine the relationship between motivation to read and reading fluency skill development. While numerous studies have been conducted examining the relationship between general academic self-concept and academic achievement (Guay, Marsh, & Boivin, 2003; Helmke & van Aken, 1995; Marsh & Koller, 2004; Marsh & Yeung, 1997; Mujis, 1997; Skaalvik & Hagtvet, 1990; Skaalvik & Valas, 1999), this is the first study to specifically examine motivation to read and reading fluency skill development at this specific level. With an increasing body of research providing evidence that achievement and motivational variables function at a domain specific level (Harter, 1982; Eccles et al., 1983; Marsh et al., 1991; Wigfield, et al., 1997), future research on these relationships must take domain specificity into account. In contrast to previous research conducted on the relationship between reading self-concept and reading achievement, the current investigation found self-system variables to predominate over achievement variables with young children. In summary, this study suggests that children’s motivation to read at the middle of their second grade year has an influence on the development of their reading fluency skills toward the end of their second grade year.
REFERENCES


APPENDIX A

ORIGINAL 28 ITEMS FROM THE MRP-A
Appendix A

Original 28 Items from the MRP-A

1. My friends think that I am _________________.
   A very good reader
   A good reader
   An OK reader
   A poor reader

2. Reading a book is something I like to do.
   Never
   Not very often
   Sometimes
   Often

3. I read _________________.
   Not as well as my friends
   About the same as my friends
   A little better than my friends
   A lot better than my friends

4. My best friends think reading is __________.
   Really fun
   Fun
   OK to do
   No fun at all

5. When I come to a word I don’t know, I can _________________.
   Almost always figure it out
   Sometimes figure it out
   Almost never figure it out
Never figure it out

6. I tell my friends about good books that I read.
   Never
   Almost never
   Some of the time
   A lot

7. When I am reading by myself, I understand ______________.
   Almost everything that I read
   Some of what I read
   Almost none of what I read
   None of what I read

8. People who read a lot are ________.
   Very interesting
   Interesting
   Not very interesting
   Boring

9. I am ________________.
   A poor reader
   An OK reader
   A good reader
   A very good reader

10. I think libraries are ____________.
    A great place to spend time
    An interesting place to spend time
    An OK place to spend time
    A boring place to spend time

11. I worry about what other kids think about my reading ____________.
Never
Once in a while
Almost everyday
Everyday

12. Knowing how to read well is _________.
Not very important
Sort of important
Important
Very important

13. When my teacher asks me a question about what I have read, I
______________.
Can never think of an answer
Have trouble thinking of an answer
Sometimes think of an answer
Always think of an answer

14. I think that reading is _____________.
A boring way to spend time
An OK way to spend time
An interesting way to spend time
A great way to spend time

15. Reading is ______________.
Very easy for me
Kind of easy for me
Kind of hard for me
Very hard for me

16. When I grow up I will spend _____________.
None of my time reading
Very little of my time reading
Some of my time reading
A lot of my time reading

17. When I am in a group talking about stories, I ____________.
   Almost never talk about my ideas
   Sometimes talk about my ideas
   Almost always talk about my ideas
   Always talk about my ideas

18. I would like for my teacher to read books out loud to the class
   ____________.
   Every day
   Almost every day
   Once in a while
   Never

19. When I read out loud I am a _____________.
   Poor reader
   OK reader
   Good reader
   Very good reader

20. When someone gives me a book for a present, I feel _____________.
   Very happy
   Sort of happy
   Sort of unhappy
   Unhappy

21. I feel like I am getting better at reading _____________.
   Never
   Sometimes
Almost every day
Every day

22. Finishing a book that I have started reading ______________.
Is very important to me
Is important to me
Is sometimes important to me
Is never important to me

23. I like to try reading books that older kids read__________.
Never
Every once in a while
A lot of the time
Always

24. When I am in reading class, I do the work to get better at
reading__________.
Every day
Most of the time
Sometimes
Never

25. When I am in reading class, I do the work to get better at reading than my
friends__________.
Every day
Most of the time
Sometimes
Never

26. If a book I want to read is too hard for me, I ____________.
Will definitely try to read it again later
Might try to read it again later
Probably won’t try to read it again later
Will never try to read it again later

27. The reason that I want to get better at reading is so I can read more interesting books.
   A lot like me
   A little like me
   Not much like me
   Not at all like me

28. The reason that I want to get better at reading is so I can read better than my friends.
   A lot like me
   A little like me
   Not much like me
   Not at all like me

*Note*. All items removed during revision of the MRP-A are italicized
APPENDIX B

TEACHER RATINGS OF STUDENT MOTIVATION TO READ (TRSM)
Appendix B

*Teacher Ratings of Student Motivation to Read (TRSM)*

Teacher Name: ___________________________________ Date: __________
School: __________________________________________
Student Name: ____________________________________

Directions: For each item, circle the choice that best describes the frequency with which the student displays the behavior described. Use these frequency descriptions to guide your choice:

Response choices: Rarely (1)  Seldom (2)  Sometimes (3)  Often (4)

**Rarely:** You have observed the student exhibit this behavior only once, twice, or never.
**Seldom:** You have observed the student exhibit this behavior several times over a two month period.
**Sometimes:** The student exhibits this behavior once or twice a week.
**Often:** The student exhibits this behavior nearly every day for substantial amounts of time.

For example: Chooses to go to the library (1) (2) (3) (4)  Meaning: He/she sometimes chooses to go to the library

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Is easily distracted while reading</td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>2. Talks about his/her feelings related to a book or story</td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>3. Has definite preferences for favorite topics or authors</td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>4. Is a voracious reader</td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>5. Avoids participating in reading group activities</td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>6. Has a specialized reading interest</td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>7. Gets “lost” in books (lost meaning “consumed in”)</td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>8. Discussions about books with teacher and other students are complex—including motivations, plot, and personal response</td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>9. Spends a long time reading about topics he/she likes</td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>10. Is easily discouraged when he/she encounters</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
difficult text

11. Does better on reading activities when working with peers

12. Chooses to read about favorite topics

13. Is enthusiastic about reading