A PSYCHOMETRIC STUDY OF THE SELF-REGULATION STRATEGY INVENTORY – TEACHER RATING SCALE

A DISSERTATION

SUBMITTED TO THE FACULTY

OF

THE GRADUATE SCHOOL OF APPLIED AND PROFESSIONAL PSYCHOLOGY

OF

RUTGERS,

THE STATE UNIVERSITY OF NEW JERSEY

BY

BRACHA SCHNAIDMAN

IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE

OF

DOCTOR OF PSYCHOLOGY

NEW BRUNSWICK, NEW JERSEY AUGUST 2018

APPROVED: ____________________________
Timothy J. Cleary, Ph.D.

____
Ryan J. Kettler, Ph.D.

DEAN: ____________________________
Francine Conway, Ph.D.

ABSTRACT

The current study examined the reliability and validity of the Self-Regulation Strategy Inventory – Teacher Rating Scale (SRSI-TRS), a measure used to assess teacher perceptions of students' use of self-regulated learning (SRL) in a classroom context. The SRSI-TRS is part of the larger SRSI assessment system that also includes a student self-report questionnaire (SRSI-SR) and a parent rating scale (SRSI-PRS). Data from 343 seventh- and eighth-grade students was used for this study. The data was collected as part of a larger longitudinal study examining the relations between students' SRL, motivation, background variables, and academic performance. The measures in the current study included the SRSI-TRS, the SRSI-SR, and a student version of the SRSI-TRS (STRS). The STRS had the same items as the teacher rating scale, but was reworded in first person to reflect students' perspective. Construct validity of the SRSI-TRS was examined used principal axis factoring analysis. Results yielded a two factor structure with subscales which paralleled subscales from the SRSI-SR and the SRSI-PRS. Interrater reliability was examined using data from a subsample of students who had ratings completed independently by two teachers. Pearson correlations and mean differences between scores indicated high levels of agreement between teachers. Finally, correlations were used to examine convergent validity between the SRSI-TRS and the two student self-report measures. The SRSI-TRS was found to have statistically significant small to medium correlations with the SRSI-SR and STRS. The SRSI-TRS did not have a significantly higher correlation with the STRS, indicating that teachers and students do not show high levels of agreement even when both are rating behaviors in the same context. The results of this study provide preliminary support for use of the SRSI-TRS as a valid and reliable measure of teacher perceptions of student SRL. The study also highlights areas for future research for the SRSI-TRS and SRL assessment in general.

ACKNOWLEDGMENTS

As I reach this milestone of culminating my dissertation and graduate schooling, I would like to take the opportunity to thank those who have helped me along this journey. Firstly, my utmost appreciation goes to my dissertation chairman, Dr. Timothy Cleary, who introduced me to the world of research in a collaborative and enriching environment. I am grateful for his mentorship in research and many other areas of professional development, and for his support throughout my schooling and during the dissertation process in particular. I am greatly appreciative to Dr. Ryan Kettler, who served on my dissertation committee and has been an exemplar of academic excellence since I began at GSAPP. I also extend a heartfelt thank you to my faculty mentor, Dr. Anne Gregory, for her support and guidance during my time in graduate school.

I would like to use this opportunity to thank my family, who have stood with me throughout this journey and continue to be my endless source of strength and support. Thank you to my parents, for believing in what I do and providing encouragement and help at every step of the way. Thank you to my parents-in-law for their constant giving and devotion. Thank you also to my sister Leah and her family for being excellent role models of what can be achieved. My deepest gratitude, of course, goes to my husband and children. This accomplishment is as much theirs as it is mine.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Introduction

Self-regulated learning (SRL) has been identified as a critical factor for students' academic success (Dent & Koenka, 2016; Pintrich, 2000; Wigfield, Klauda, & Cambria, 2011; Zimmerman, 2001). SRL refers to a variety of behavioral, metacognitive, and motivational processes students use to achieve their self-set learning goals (Zimmerman & Schunk, 1989). Over the past several decades, the theoretical bases of SRL have been explored and explicated. As researchers gained an understanding of the importance of SRL for academic functioning, they have increased their focus on developing interventions to promote students' SRL (Dignath & Buttner, 2008; Dignath van Ewijk, 2011).

A key area of research that is central to both theory and intervention is SRL assessment. Historically, the measures used to assess SRL have mirrored researchers' conceptualizations of SRL as a construct and have been a tool to further inform the theories (Ben-Eliyahu & Bernacki, 2015; Boekaerts & Corno, 2005; Perry, 2002). While a prevalent purpose of SRL assessment is description in basic research, assessment plays an integral role in applied research as well (Winne & Perry, 2000). Practitioners and educators need to assess how students apply regulatory strategies and behavior in learning contexts in order to teach skills to those who lack them (Greene, Robertson, & Croker Costa, 2011). Assessment is also important for establishing the effectiveness of programs targeting students' SRL (Boekarts & Corno, 2005). Assessment, therefore, is not an end in and of itself, but a powerful means to shape theory and practice.

SRL assessment encompasses broad-based measures like ratings scales and interviews as well as more fine-grained measures, such as think-alouds, diaries, and microanalysis (Winne & Perry, 2000). Self-report questionnaires are the most commonly used form of measurement (Dinsmore, Alexander, & Loughlin, 2008); however, these types of measures have been criticized in the literature, with evidence showing that students are not reliable reporters of their own behavior (Winne & Jamieson-Noel, 2002) and that broad questionnaires do not capture the processual nature of SRL (Boekaerts, Pintrich, & Zeidner, 2000). Researchers have shifted focus to a variety of other measures, with an understanding that SRL, in both basic and applied research, is best understood through the use of multiple methods and measures (Winne & Perry, 2000). The current study focuses on teacher ratings of student SRL, a potentially valuable assessment method that to date has not been explored extensively in the literature. Teacher ratings have been widely used to assess many areas of student functioning, and are an important piece of a multi-informant comprehensive assessment (Merrell, 2008; Sattler, 2008).

The Self-Regulation Strategy Inventory – Teacher Rating Scale (SRSI-TRS) is an existing scale designed to assess SRL behaviors that students exhibit in the classroom setting (Cleary & Callan, 2014). The SRSI-TRS was developed as part of a system of assessment that also includes a self-report questionnaire and a parent rating scale. Although several recent studies have shown the SRSI-TRS to be a promising tool in assessing student SRL (Callan & Cleary, 2017; Cleary & Callan, 2014; Cleary, Dembitzer, & Kettler, 2015), there are many issues that were not explored in prior research. The current study will examine the factor structure and interrater reliability of the scale, two psychometric properties that have not been addressed. In addition, the study will build on existing research by examining the scale's convergence with two student self-report measures. It is expected that the results generated from this study will underscore the value of the SRSI-TRS as a method for assessing students' self-regulated behaviors in the classroom, and more generally, will add to the broad literature base examining the effectiveness of teacher ratings for various types of student behavior.

Overview of Self-Regulated Learning

Self-regulated learning (SRL) is broadly described as "the degree to which students are metacognitively, motivationally, and behaviorally active participants in their own learning process" (Zimmerman, 1986). This overarching definition includes several key dimensions of SRL. Effective self-regulated learners use metacognition before, during, and after learning to check their understanding of the task, develop plans or set goals, and monitor and evaluate their performance (Winne, 2011). Motivation beliefs, such as self-efficacy and task interest, play an integral role in the initiation and maintenance of regulatory behavior (Eccles & Wigfield, 2002; Pajares, 2008; Schunk & Usher, 2011; Wolters & Rosenthal, 2000). Finally, self-regulation includes the use of behavioral strategies to select and structure environments in order to create an optimal learning experience (Alexander, Graham, & Harris, 1998; Broadbent, 2017; Corno, 2011).

Although many students will use regulatory strategies from time to time, a key distinguishing feature of self-regulated learners in school contexts is their deliberate and thoughtful use of strategies to improve their academic performance. Thus, they understand the link between regulatory processes and learning outcomes and purposefully apply strategies to achieve their goals (Zimmerman, 1990). As Zimmerman (1998) aptly explains, self-regulated learners "are distinguished by their view of academic learning as something that they do for themselves rather than as something that is done to or for them" (p. 1).

Zimmerman (2000) elaborated a model of SRL based on social-cognitive theory. According to this model, SRL is a cyclical process that occurs in three phases: forethought, performance, and reflection (Zimmerman, 2000). The forethought phase sets the stage for learning. During this phase, effective self-regulated learners engage in the processes of task analysis, goal setting, and strategic planning (Butler & Cartier, 2004; Schunk, 2001; Zimmerman, 2011). During the performance phase, students apply regulatory strategies, such as environmental structuring, use of self-consequences, and seeking information from others, as well as task strategies, such as making predictions while reading or drawing pictures to solve algebra problems (Cleary, 2018; Pressley & Harris, 2008; Zimmerman & Martinez-Pons, 1986). Self-observation is another key performance phase process in which students monitor the effectiveness of their learning and strategy use while engaged in the task (Chen & Rossi, 2013). The final phase, reflection, occurs after the learning task. The key processes for this phase are self-judgements and self-reactions; that is, students assess their performance, make causal attributions about the reasons they performed the way they did, and react affectively based on their judgement (Schunk & Usher, 2013; Zimmerman, 2011). The feedback generated during the reflection phase is used for a subsequent forethought phase, thus beginning the cycle again (Zimmerman, 2000). The feedback loop is an integral part of the SRL cycle. Students use feedback from their own monitoring and attributions as well as external feedback from others to continuously regulate their behavior, affect, and environment (Zimmerman, 2001; Zimmerman & Schunk, 2011).

Importance of Self-Regulated Learning

Self-regulated learning has been identified as an important factor for students' academic success (Pintrich, 2000; McClelland & Cameron, 2011; Zimmerman, 2001). Research has shown that students with good SRL skills perform better academically than other students (Nota, Soresi, & Zimmerman, 2004; Perry & Rahim, 2011; Weinstein & Acee, 2013), and they also have more optimistic views of their future (Zimmerman, 2002). Ertmer and Newby (1996) explain that the difference between expert and novice learners lies not in the actual amount of knowledge they have in a domain, but in their ability to be aware of their knowledge and deficits, control their learning, and reflect – essentially, their SRL skills. SRL is a key to compensating for individual differences in ability and allowing all students to access success (Zimmerman, 2002). Beyond the classroom, self-regulation is important for success in the workforce and in recreational settings, where it is necessary to be proactive about learning and refining skills. A goal of education is to prepare students for lifelong learning, and SRL skills are an important component of this (Weinstein & Acee, 2013; Zimmerman, 2002).

Researchers have documented the importance of SRL skills across a variety of academic domains, such as reading, writing, mathematics, and science (Cleary & Platten, 2013; Fuchs et al., 2003; Graham & Harris, 2009; Tonks & Taboada, 2011). To this end, several intervention programs that enhance students' regulatory skills and academic performance in different content areas have been developed over the past decade (Butler, Beckingham, & Lauscher, 2005; Cleary, Velardi, & Schnaidman, 2017; Graham & Harris, 2009). In addition, numerous studies have shown how SRL skills impact academic performance for students across all grade levels, from preschool through college (Adagideli, Sarac, & Ader, 2015; Hofer, Yu, & Pintrich, 1998; Perry, VandeKamp, Mercer, & Nordby, 2002).

Although SRL is important across educational levels, these skills are particularly crucial as students transition to middle school. Middle school students typically experience a decrease in motivation and self-efficacy and a decline in academic performance (Cleary & Chen, 2009; Dembo & Eaton, 2000; Eccles et al., 1993). In middle school, students are often faced with multiple assignments given by many different teachers, and they have less personal connection with teachers and principals (Barber & Olsen, 2004). Students are expected to be more self-motivated and take responsibility for their learning (Rudolph, Lambert, Clark, & Kurlakowsky, 2001). When combined with other changes such as increased complexity in peer relationships and the "storm and stress" of adolescence, this becomes a particularly challenging time for students. In the classroom, self-regulated learners can employ strategies and behaviors, such as self-monitoring the number of math problems completed during class, planning the format for a writing assignment, and seeking help before a test, in order to engage in their learning in a meaningful and productive way (Moos & Ringdal, 2012; Pape, Bell, & Yetkin, 2003). As middle school students navigate a rigorous academic environment and cope with other challenges in their personal lives, SRL skills can help put them on the path to success.

Assessment of Self-Regulated Learning

Assessment of SRL occupies a central place in research and practice. Without assessment, it is difficult to apply and evaluate interventions geared to improving students' SRL (Boekarts & Corno, 2005). As the conceptual understanding of SRL evolved and expanded, research on SRL measurement progressed as well (Perry, 2002). Historically, ratings scales – particularly student self-reports – were the predominant method of measurement (Dinsmore et al., 2008; Greene et al., 2011.) Because of inherent limitations with this method, researchers have begun to shift their focus to other assessment tools and methodologies.

The measures designed to evaluate students' SRL behaviors and strategies are grounded in researchers' understanding of the construct they seek to assess. Winne and Perry (2000) posit that SRL has characteristics of both an aptitude and an event. An aptitude is a relatively stable attribute or trait that predicts future behavior. An event, on the other hand, is a discrete entity with a beginning and end that exists within a larger series of events unfolding over time. Using this framework, Winne and Perry (2000) grouped SRL measures into two broad categories: event measures and aptitude measures. The two types of measures differ in their purpose and scope, and each yields important information regarding students' SRL.

Event measures, such as microanalysis, behavioral traces, and think-alouds, are contextualized forms of assessment that seek to capture information about regulatory processes as they occur during a specific learning task (Cleary & Callan, 2017; Zimmerman, 2008). These types of assessment tools emphasize SRL as a process that involves applying domain-specific skills in an actual learning situation (Boekaerts & Corno, 2005; Schunk, 2001). Event measures are useful for diagnosing and remediating deficits in SRL, particularly as they relate to specific tasks and situations (Zimmerman, 2008). Event measures vary in their formats and the types of data they generate; for example, think-alouds require students to verbalize their thoughts while engaged in a task (Greene & Azevedo, 2007), online trace logs provide a record of strategies used by the student (Winne et al., 2006), and microanalysis uses structured protocols to interview students on the processes they use before, during, and after engaging in a learning task (Cleary & Callan, 2017).

In contrast, aptitude measures, such as self-report questionnaires, rating scales, and some types of structured interviews, assess a student's overall use of SRL skills and strategies (Zimmerman, 2008; Winne & Perry, 2000). These tools often provide a global measure of the frequency or likelihood with which students apply SRL concepts in their school and/or homework. Typically, aptitude measures ask respondents to retrospectively rate or record their use of specific strategies and behaviors. The ratings are then aggregated to yield a single score that represents a global picture of the student's use of the strategies and behaviors (Winne & Perry, 2000). Two commonly used aptitude measures that are relevant to the study, self-report questionnaires and teacher rating scales, are discussed in greater detail next.

Self-report questionnaires. Self-report questionnaires, such as the *Learning and Study Strategies Inventory* (LASSI; Weinstein, Schulte, & Palmer, 1987), and the *Motivated Strategies for Learning Questionnaire* (MLSQ; Pintrich, Smith, Garcia, & McKeachie, 1991), have traditionally been the most frequently used measures of SRL (Dinsmore et al., 2008; Winne & Perry, 2000). Self-report questionnaires have many benefits. They are relatively easy and cost effective to design, administer, and score (Winne & Perry, 2000). In addition, self-report questionnaires allow students to report on motivation beliefs that are not observable to others (Patrick & Middleton, 2002). However, there are disadvantages to using student self-reports, as they are susceptible to response biases and may not accurately represent what students actually do while engaged in learning tasks (Cleary & Chen, 2009; Winne & Jamieson-Noel, 2002). Also, because they capture SRL as a global construct, self-report questionnaires may not be sensitive to small, nuanced shifts in students' SRL, such as following an intervention (Cleary, et al., 2017). Furthermore, research has shown that students' reports of their behavior often have a low level of correspondence with adults' reports, and are less predictive of future outcomes (Loeber, Green, & Lahey, 1990).

Teacher rating scales. Extensive research has documented the value of gathering information about children from multiple informants, such as parents and teachers, because each informant can contribute a unique perspective based on the context and setting in which they observe and interact with the child (Achenbach, McConaughty, & Howell, 1987; De Los Reyes et al. 2015). Teachers can provide valuable information about students' functioning, particularly with regard to academic skills and behaviors (Frick, Barry, & Kamphaus, 2010; Kettler et al., 2014; Perry & Meisels, 1996). Because teachers interact with students for many hours over several months, most will have a rich and varied sample of student behavior and performance on which to base their judgements (Gerber & Semmel, 1984). Critics have raised concerns that rating scales are susceptible to response bias, and teacher ratings may be influenced by students' abilities or cultural factors (Darling-Hammond, 1995; Hoge & Butcher, 1984; Merrell, 2008). However, a relatively large literature base shows that teacher ratings are a valuable and accurate measure of students' functioning across a range of academic and behavioral outcomes (Hoge & Colodarci; 1989; Loeber, Green, Lahey, & Stouthamer-Loeber, 1989; Perry & Meisels, 1996). Teacher rating scales are also a cost-effective and efficient method of measurement (Kenny & Chekaluk, 1993; Perry & Meisels, 1996). Many SRL measures, such as self-report scales, microanalysis, and think-alouds, rely on student reports; thus, gathering data from teachers is important because it can be used to corroborate students' perspectives (Achenbach et al., 1987; Loeber et al., 1989).

Teachers are well positioned to provide information about students' SRL because they interact frequently with the students and can observe them engaging in various learning activities over a long period of time (Callan & Cleary, 2017). Several studies have examined the reliability and validity of teacher ratings of student SRL. The measures used in prior research are the *Rating Student Self-Regulated Learning Outcomes: A Teacher Scale* (RSSRL; Zimmerman & Martinez-Pons, 1988) and the SRSI-TRS (Cleary & Callan, 2014), which is the target measure that is the focus of this study.

Zimmerman and Martinez-Pons (1988) developed the RSSRL to measure students' use of SRL strategies that are observable in school, as well as outcomes of strategy use. The purpose of their initial study was to establish the validity of a student self-report interview (Self-Regulated Learning Interview Schedule; Zimmerman & Martinez-Pons, 1986) by using teacher ratings as a criterion measure. Using a sample of 80 tenth grade students, the authors found that a single underlying construct best represented the RSSRL and the construct was distinct from student achievement. Many (but not all) of the strategies endorsed by students on a self-report interview were significantly correlated to the underlying SRL construct of the teacher measure (Zimmerman & Martinez-Pons, 1988). In a later study, the RSSRL was shown to be predictive of students' achievement on a science task; it accounted for 26% of the variance in scores (DiBenedetto & Zimmerman, 2013).

The SRSI-TRS was developed by Cleary and Callan (2014) and is designed to capture teacher perceptions' of students' use of regulatory behavior in the classroom. Several studies have examined the predictive validity of the SRSI-TRS and its relations with student self-report measures (Callan & Cleary, 2017; Cleary & Callan, 2014). Using a sample of 128 ninth grade students, Cleary and Callan (2014) found that the SRSI-TRS accounted for a medium amount of variance (9.4%) in students' mathematics test scores, after controlling for prior achievement and students' self-reported regulatory behaviors and motivation beliefs. More recently, Callan and Cleary (2017) conducted a study using a sample of 100 eighth grade students and found that the SRSI-TRS accounted for a medium amount of variance in two of the three outcome measures examined (students' posttest mathematics scores and standardized test scores). The SRSI-TRS has also been shown to have significant, medium correlations with the Self-Regulation Strategy Inventory – Self-Report (SRSI-SR; Callan & Cleary, 2017; Cleary et al., 2015).

The studies reviewed provide promising results about the use of teacher ratings of student SRL, and specifically the SRSI-TRS, although further work remains regarding the reliability of scores and ensuing inferences from the measure. It should be emphasized that both event (i.e., contextualized) and aptitude (i.e., broad, aggregate) measures, and the multiple types of measure included within each category, have the potential to generate unique and important information about a student's SRL (Winne & Perry, 2000). Thus, best practice indicates that a combination of measures is most useful for assessing SRL processes (Boekaerts & Corno, 2005; Patrick & Middleton, 2002), and teacher ratings can be a valuable piece of a multi-method assessment (Callan & Cleary, 2017). High quality measurement requires tools yielding reliabile scores from which valid inferences can be drawn; therefore, researchers need to continue to focus on the psychometric properties of the SRL measures, including triangulation and convergence across different measures (Callan & Cleary, 2017; Cleary, 2011; Winne & Perry, 2000).

SRL Assessment Practices in the Schools

Although SRL assessment has been an emerging focus of researchers, there exists a gap between the research findings and actual practice within the schools. While there is room for improvement with regard to assessment of SRL in practice, there is evidence to indicate that educators and practitioners perceive assessment of SRL to be valuable and useful (Cleary, Gubi, & Prescott, 2010; Cleary & Zimmerman, 2006), and school psychologists are interested in learning more about SRL assessment (Cleary et al., 2010). Developing and researching measures that can be used within school systems can help bridge the gap between research and practice.

Historically, assessment has occupied a large role in school psychology practice. Estimates of the percentage of time spent on assessment have remained stable over the last four decades, with studies indicating that practitioners spend roughly half of their time on assessment activities (Castillo, Curtis, & Gelley, 2012; Farling & Hoedt, 1971; Goh, Teslow, & Fuller, 1981; Hosp & Reschly, 2002; Hutton, Dubes, & Muir, 1992). However, there have been shifts in the focus and scope of assessments by school psychologists over the years. The need to assess skills and competencies beyond intelligence and academic ability has become apparent as researchers and practitioners gain a better understanding of the factors that influence students' learning.

Various studies have stressed the importance of assessing academic enablers, including motivation, interpersonal skills, and self-regulation, as part of an effort to promote student achievement (Cleary, 2006; DiPerna, Volpe, & Elliot, 2002; Linnenbrink & Pintrich, 2002).

The trends in school psychology assessment practice form an important backdrop for incorporating SRL assessment in the schools. Furthermore, best practice in school psychology mandates a multi-method assessment approach that evaluates multiple areas of functioning and makes use of different measures, methods, and informants (Fagan & Wise, 2007; Sattler, 2008; Whitcomb & Merrell, 2013). SRL skills are an important factor for students' success at school, and can be a crucial point of intervention for trying to improve academic achievement (Cleary et al., 2017). In addition, self-regulation and motivation deficits are involved in about a quarter of all student referrals for special education evaluations (Bramlett, Murphy, Johnson, & Wallingsford, 2002; Cleary, 2009). Thus, SRL clearly has a place in a comprehensive, multi-method assessment.

Despite the prevalence of SRL-related referrals and increased understanding of the importance of SRL skills for academic success, school psychologists do not routinely conduct evaluations of students' SRL and motivation skills. Additionally, studies have found that school psychologists report little to no familiarity with published, available self-report and ratings scales designed to measure SRL and motivation (Cleary, 2009; Cleary et al., 2010). School psychologists surveyed about SRL assessment methods indicated a preference for self-report/rating scales, informal interviews, and classroom observations. About 88% of the respondents included self-report/rating scales as one of their top three assessment techniques (Cleary, 2009, p. 82). These findings may indicate that self-report/rating scale measures of SRL would be acceptable to school psychologists, and would fit well within a multi-method assessment framework.

Rationale for the Current Study

The literature reviewed thus far supports the premise that teacher ratings can be a valuable source of information about students' SRL. Further, teacher ratings scales are cost-efficient and easy to administer, and can be linked to intervention strategies. School psychologists often rely on teachers to provide information about students' academic functioning, because teachers interact with students across a wide variety of learning situations. Accordingly, teacher rating scales of SRL may be a particularly acceptable measure for school psychologists to incorporate into their assessment repertoire.

This study focuses on examining the reliability and validity evidence pertaining to an exisiting SRL teacher rating scale, the SRSI-TRS, in order to evaluate its potential as a psychometrically sound measure of SRL. Prior studies have found that the SRSI-TRS has high internal consistency, and have provided evidence of predictive and convergent validity (Callan & Cleary, 2017; Cleary & Callan, 2014). The current study seeks to further explore evidence of reliability and validity of the scale in a comprehensive way. In terms of validity, factor analysis will be used to investigate construct validity, and additional evidence of convergent validity will be gathered by examining the relations between the SRSI-TRS and two student self-report measures. In terms of reliability, internal consistency and interrater reliability will be examined. Specifically, this study seeks to answer the following questions:

1. a. What is the factor structure of the SRSI-TRS?

   b. What is the internal consistency of the scale (and subscales)?

2. What is the level of interrater agreement between two teachers in the same classroom using the SRSI-TRS to rate the same students?
3. What is the level of convergence between the SRSI-TRS and two student self-report measures?

The first question addresses the factor structure of the SRSI-TRS. Factor analysis is an essential tool for establishing construct validity of a measure (Brown, 2015). To date, there is no research examining the factor structure of the SRSI-TRS. The factor structures of the other versions of the SRSI (the self-report questionnaire and parent rating scale) have been examined in prior research; both have a three-factor structure. The factors in the Self-Regulation Strategy Inventory – Self-Report (SRSI-SR) are *Managing Environment and Behavior*, *Seeking and Learning Information*, and *Maladaptive Regulatory Behaviors* (Cleary, 2006). The factors in the Self-Regulation Strategy Inventory – Parent Rating Scale (SRSI-PRS) include *Managing Behavior and Learning*, *Maladaptive Regulatory Behaviors*, and *Managing Environment* (Chen, Cleary, & Lui, 2015). The factor structures of these two measures are highly similar; both include two adaptive scales and one maladaptive scale, and there is overlap in the item wording and content (Chen et al., 2015). This is not surprising considering that the SRSI-SR and SRSI-PRS both assess students' use of SRL behaviors and strategies while studying and doing homework; that is, outside of the classroom context. The SRSI-TRS, on the other hand, captures teachers' perceptions of student SRL behaviors in the classroom. In addition, the SRSI-TRS does not include items reflecting maladaptive behaviors. Although the SRSI-TRS was developed from the SRSI-SR, and some items are similar, the underlying factor structure may be different than that of the other measures. Because the SRSI-TRS was not developed based on a specific factor structure, exploratory factor analysis will be used for this study, with no a priori hypothesis about the factor structure.

The second question examines interrater agreement for teacher ratings of student SRL, a topic that has not been explored in the literature. Interrater reliability is important to calculate in studies involving and subjective measurement, an external variable is measured and there is subjectivity within the measurement, such as student behaviors measured by teacher ratings. High levels of agreement between raters indicate that the measure is reliable and will yield stable results across respondents (Litwin, 1995; Thomas, 2017). In their meta-analysis of cross-informant agreement, Achenbach et al. (1987) found large correlations for interrater agreement between two teachers (r = .64); indeed, pairs of teachers had the highest level of agreement among all pairs of informants. Thus, it is hypothesized that similar results will be found with regard to teacher ratings of SRL; that is, large correlations between ratings given by two teachers in the same classroom.

The final question focuses on convergence between the SRSI-TRS and two student self-report measures. One student measure is the previously established SRSI-SR, which asks students to rate their use of SRL strategies and behaviors at home while studying and doing homework. The other student measure used in this project is identical to the SRSI-TRS, with items reworded in first person (e.g., “I monitor how well I learn class material”) and is referred to as the STRS (i.e., student version of the TRS). As is the case with the SRSI-TRS, the STRS asks students to rate their use of SRL strategies and behaviors in the classroom.

Previous studies found small to medium correlations between the SRSI-TRS and the SRSI-SR. Callan and Cleary (2017) reported a significant, medium correlation (r = .30) between the two scales. Cleary and colleagues (2015) reported small correlations between the SRSI-TRS and SRSI-SR adaptive subscales (r = .23 and r = .24) and a medium correlation between the SRSI-TRS and the SRSI-SR maladaptive subscale (r = -.39). These findings are generally consistent with research on cross-informant agreement for a variety of emotional and behavioral problems, which has shown small correlations between teacher and student reports (Achenbach et al., 1987; De Los Reyes et al., 2015; Phares, Compas, & Howell, 1989). In addition, the SRSI-TRS and SRSI-SR measure SRL in different contexts, which may explain the reason larger correlations were not found (Cleary et al., 2015). It is expected that the current study will find similar small to medium correlations between the SRSI-TRS and SRSI-SR. There is evidence that higher agreement can be achieved when the teacher and student questionnaires used for comparison are similar (Kettler et al., 2014). Additionally, the SRSI-TRs and STRS assess SRL in the same context. Thus, it is hypothesized that significant, medium correlations, at minimum, will be found for the SRSI-TRS and the STRS.

Methods

The data for this project was collected as part of a larger longitudinal study that aimed to examine the relations among background variables (e.g., prior achievement, socioeconomic status), motivation beliefs, self-regulated learning behaviors, and academic performance. A number of data sources were utilized in the larger study, including student self-report questionnaires, teacher rating scales, and school records. Data was collected at three time points: Spring of 2013, Fall of 2013, and Spring of 2014. For the purposes of this study, data from the first and second collection phases (Spring and Fall of 2013) was used. Data from the first collection phase was used for interrater reliability analyses since that phase had the largest sample of students with two ratings. Data from the second collection phase was used for factor analysis and convergent validity analyses because that phase included all of the measures that were needed for this study.

Sample

School. A middle school located in a Northeastern suburban school district participated in the longitudinal study. The school had a total student population of approximately 1,200 students. About 29% of the students in the school qualified for Free/Reduced Lunch.

Students. The sample for this study includes 343 middle school students. Two hundred twenty students (61.2%) were in seventh grade and 123 students (35.9%) were in eighth grade during phase two of data collection. Of the sample, 196 students (57.1%) were female, and 86 students (25.1%) were eligible for Free/Reduced Lunch. Table 1 describes the demographic characteristics of the overall sample. Data for a subsample of 40 students from phase one was used to calculate interrater agreement.

Teachers. A total of 14 math teachers provided ratings for the 343 students. Eleven teachers were female (78.6%) and three were male (21.4%). Ten of the teachers identified as white (71.4%) and two as Hispanic or Latino (14.3%); the ethnicity of the last two teachers is unknown. Their years of teaching experience ranged from three to 38 years, with a mean of 13 years. For the interrater reliability analysis, each student was rated by a general education teacher and a special education teacher who worked in the same math class.

Table 1

Demographic Characteristics of Participating Students

Measure	*n*	%
Gender		
Male	147	42.9
Female	196	57.1
Grade		
Seventh	220	64.1
Eighth	123	35.9
Ethnicity		
White (non-Hispanic)	148	43.1
Hispanic or Latino	77	22.4
Asian/Pacific Islander	72	21.0
Black or African American	18	5.2
Interracial	16	4.7
Native American	2	0.6
Free/Reduced Lunch	86	25.1

Measures

Self-Regulation Strategy Inventory – Teacher Rating Scale (SRSI-TRS). The SRSI-TRS is an SRL measurement tool designed to capture teacher perceptions of students' adaptive SRL behaviors in the classroom context (Cleary & Callan, 2014). The SRSI-TRS is part of a multidimensional assessment system that also includes a self-report questionnaire and a parent rating scale. The SRSI-TRS was developed to parallel the self-report version, but only includes items that would be observable to teachers in the classroom. The SRSI-TRS includes thirteen items to which teachers respond using a five point Likert scale from 1 *(almost never) to 5 (almost always.)* Sample items on the SRSI-TRS are, "The student monitors how well he or she learns class material" and "The student is prepared for class." The final score is calculated by taking the mean of all item responses. Prior research has shown high internal consistency for the SRSI-TRS (α = .97; Cleary & Callan, 2014; and α = .99; Callan & Cleary, 2017).

Self-Regulation Strategy Inventory – Teacher Rating Scale, Student Version (STRS). The STRS was a measure administered during the second and third data collection phases of the longitudinal study. This measure was identical to the TRS, except all items were reworded to reflect the students' perspective, such as "I monitor how well I learn class material." This scale aimed to capture students' perceptions of their behaviors and strategy use in the classroom, and was developed to be another student self-report measure in addition to the SRSI-SR. Because the STRS was a new measure, no prior reliability or validity information is available.

Self-Regulation Strategy Inventory – Self-Report (SRSI-SR). The SRSI-SR is a self-report questionnaire designed to assess the frequency with which students engage in adaptive and maladaptive regulatory behaviors while studying and doing homework (Cleary, 2006). It includes 28 items to which students respond using a five point Likert scale from 1 *(almost never)* to *5 (almost always.)* Sample items on the SRSI-SR are, "I try to study in a quiet place" and "I rely on my math class notes to study." The items on the SRSI-SR were developed based on general categories of SRL strategies (Zimmerman & Martinez-Pons, 1988) and address the three major dimensions of regulation – motivation, strategy use, and metacognition. The SRSI-SR has a three factor structure: *Managing Environment and Behavior* (MEB), *Seeking and Learning Information* (SLI), and *Maladaptive Regulatory Behavior* (MRB; Cleary et al., 2015). The items on the MRB scale are negatively worded; for example, "I try to forget about the topics that I have trouble learning," and are reverse scored when calculated a composite score. Prior research has shown high internal consistency for the overall SRSI-SR (α = .92) and for the subscales (α = .76 to .87; Cleary et al., 2015).

Demographic Information. Demographic information for the participating students and teachers was provided by the school district.

Procedures

The SRSI-TRS, SRSI-SR, and STRS were all included in the longitudinal study, although the STRS was only administered at the second and third data collection phases. At each phase, trained graduate research assistants administered the measures to the student participants. The research assistants read the instructions aloud and answered questions as needed. Students completed their measures in one 20-25 minute testing session. All student measures were collected over a three week period. Mathematics teachers completed the teacher rating scale for their respective students within two weeks after the students completed their measures. For a subset of students, a second SRSI-TRS was completed by another teacher who worked in the classroom. Data was entered into an SPSS database by trained graduate students.

Data Analysis

Several quantitative techniques were used to analyze the psychometric properties of the SRSI-TRS. Table 2 outlines the research questions and analytic techniques.

The factor structure of the SRSI-TRS was examined using principal axis factoring (PAF) analysis. This is an exploratory process that identifies a number of factors underlying a larger set of variables (Meyers, Gamst, & Guarino, 2017). Because there is no prior research on the factor structure of the SRSI-TRS, exploratory factor analysis is an appropriate technique to observe the possible underlying factors without imposing a preconceived structure (Brown, 2015). After the factor structure was determined, internal consistency of the overall scale and subscales were calculated using Cronbach's alpha. Item-total correlations were examined as well.

Interrater agreement between teachers was calculated for the subset of students (n = 40) who enrolled in classrooms with two full-time teachers. Both teachers independently rated the students using the SRSI-TRS. Pearson correlations were the primary measure of interrater reliability, and were computed for the total scores as well as the subscale scores. Pearson correlations can reveal the level of consistency between raters (Geisinger, 2017). Because they are commonly used to calculate interrater and cross-informant agreement (Achenbach et al., 1987), they allow for comparison of the level of agreement with other similar studies. However, Pearson correlations only provide a measure of the linear relationship and do not capture the level of absolute agreement between informants (Stolarova, Wolf, Rinker, & Brielmann, 2014). Therefore, the average difference in scores between pairs of ratings was computed to give additional information about agreement.

Pearson correlations were used to examine the level of convergence between the SRSI-TRS, the STRS, and the SRSI-SR. Correlations were computed for the total (mean) scores, for subscale scores, and for each individual item score on the SRSI-TRS and STRS. The correlations for the individual items were examined qualitatively to determine whether students and teachers have greater agreement for certain SRL behaviors.

Table 2

Data Analyses

Research Questions	Data Used	Data Analytic Techniques
1. a. What is the factor structure of the SRSI- TRS? **b. What is the internal consistency of the scale (and subscales)?**	a. Item level scores for all students in the sample b. Item level scores for total scale and subscales	a. Principle Axis Factoring (PAF) b. Cronbach's alpha
2. What is the level of interrater agreement for the SRSI-TRS?	Mean scores for total scale and subscales on two sets of ratings for 40 students	Pearson correlations, Descriptive analysis and t-test for mean difference in scores
3. What is the level of convergence between the SRSI-TRS and two student self-report scales (the SRSI-SR and the STRS?	Mean scores for total scales (and subscales), Item level scores for the SRSI-TRS and STRS	Pearson correlations, Qualitative

Results

This chapter examines the results from the data analyses performed. Preliminary analyses were first conducted to check statistical assumptions and to examine missing data. Following data screening and cleaning procedures, a variety of statistical techniques were employed to address the three primary research questions. Principal axis factoring (PAF) analysis was used to examine the factor structure of the SRSI-TRS and Cronbach's alpha was computed to determine the internal consistency of the scale and subscales. Pearson correlations and mean differences in scores were used to examine interrater reliability for the SRSI-TRS, and Pearson correlations were used to examine convergent validity. All statistical procedures were performed using IBM SPSS Statistics Version 25.

Screening Procedures

Of the original 343 participants in the study, two were removed because their SRSI-SR data was missing, yielding a sample of 341 students. Missing data was examined for all three measures (SRSI-SR, STRS, and SRSI-TRS.) Missing data was minimal for both the SRSI-SR and the STRS. On the SRSI-SR, no item was missing more than one data point (0.3%), and no case was missing more than one data point (3.6%), except one case, which was missing two (7.1%). On the STRS, two items were each missing two data points (0.6%).

The SRSI-TRS included an option for teachers to rate "don't know." All "don't know" responses were treated as missing data, and there was no other missing data aside from the "don't know" responses. Most items had at least one case with a missing value, and one item (Question 8) had 36 missing values. Additionally, six students had missing values for more than two items on the scale, which includes a total of 13 items. Downey & King (1998) recommend removing cases that are missing more than 20% of the data, so these six cases were deleted, leaving a sample of 335 students. Missing data was analyzed again after deleting the six cases; three items were missing four, three, and one values (1.2%, 0.9%, and 0.3%, respectively), and Item 8 was missing 30 values (9.0%).

Since factor analysis uses item level data, missing data on the SRSI-TRS were not replaced prior to the analysis. Cases with missing data were deleted listwise, yielding an *n* of 303. Skewness and kurtosis were examined for the 13 items on the scale. All values were within normal limits (between 2 and -2; Ferguson & Cox, 1993), except one kurtosis value of 4.14 (see Table 3). Ferguson and Cox (1993) suggest that data is acceptable for factor analysis if less than 25% of the items exceed acceptable limits for skewness and kurtosis; thus, the one large value did not pose a problem. Furthermore, individual sampling adequacies were examined using Kaiser-Meyer-Olkin (KMO) tests (Ferguson & Cox, 1993). The KMO Measure of Sampling Adequacy examines the partial correlations between variables. Small partial correlations indicate a high level of shared variance due to common underlying factors, indicating that the variable is suitable for factor analysis. Generally, variables with KMO values about .60 or higher are considered to be suitable for factor analysis. All KMO values for the variables in this study exceeded .80, and can be characterized as “meritorious” according to Kaiser’s (1974) original guidelines for sampling adequacy.

Table 3

Descriptive Statistics of SRSI-TRS Items

Item	Mean	SD	Skewness	Kurtosis
1. The student asks about topics that might appear on upcoming tests.	2.76	1.47	0.23	-1.30
2. The student keeps his or her class materials very organized.	4.25	0.90	-1.01	0.25
3. The student asks insightful questions in class.	3.09	1.27	0.05	-0.96
4. The student asks questions about errors he or she makes on tests or assignments.	3.31	1.24	-0.11	-0.99
5. The student seeks help or attends extra help sessions.	2.87	1.35	0.38	-1.11
6. The student asks questions in class when he or she does not understand something.	3.55	1.16	-0.34	-0.75
7. The student keeps himself or herself motivated even when they struggle to learn something.	3.89	0.95	-0.43	-0.68
8. The student monitors how well he or she learns class material.	3.85	0.98	-0.53	-0.53
9. The student asks about the format of upcoming tests (short-answer, multiple choice)	2.06	1.47	1.09	-0.33
10. The student pushes himself or herself to understand the details of the topics presented in class.	3.92	0.98	-0.60	-0.34
11. The student is enthusiastic about learning.	3.83	1.09	-0.47	-0.87
12. The student makes excellent use of class time.	4.15	1.01	-0.98	-0.08
13. The student is prepared for class.	4.53	0.79	-1.96	4.14

Note: $n = 303$.

After the PAF analysis was complete, missing data on all measures were replaced using multiple imputations (MI). This method was chosen because a Missing Value Analysis indicated that the date was not missing completely at random (Little's MCAR test was not significant.) MI procedures are less biased than traditional estimation methods because they incorporate random error, and are therefore recommended for data missing not at random (Meyers et al., 2017; Schlomer, Bauman, & Card, 2010). The automatic imputation method (fully conditional specification) in SPSS was used, and 20 imputed data sets were created, as recommended by Baraldi and Enders (2010). After the MI were completed, composite (mean) scores were computed for the three measures, and skewness and kurtosis was examined for all variables. Descriptive statistics are included in Table 4.

Table 4

Descriptive Statistics of Composite Measures for Overall Sample

Measure	Mean	SD	Skewness	Kurtosis	Alpha
SRSI-TRS	3.45	.84	.14	-.80	.927
SRSI-SR	3.75	.67	-.45	.06	.883
STRS	3.75	.70	-.38	-.21	.929

Note: n = 335. SRSI-TRS = Self-Regulation Strategy Inventory – Teacher Rating Scale, SRSI-SR = Self-Regulation Strategy Inventory – Self-Report, STRS = Student version of Teacher Rating Scale. Pooled results of multiple imputation procedures were used to obtain means. SD, skewness, kurtosis, and alpha were calculated based on the original data.

Data from two teachers who separately rated a subsample of 40 students were used for interrater reliability analyses. Again, the measure (SRSI-TRS) offered an option for teachers to rate "don't know," and all "don't know" ratings were considered missing. Across teacher ratings, three variables were each missing one data point (2.5%). One variable (Item 8) was missing four data points (10%) within one set of teacher ratings. One case was deleted because it was missing three values (23.1%), yielding an n of 39. MI procedures were used to address the remaining missing data. Composite (mean) scores were computed for each teacher's ratings. Descriptive statistics for the composite scores are included on Table 5.

Table 5

Descriptive Statistics of SRSI-TRS Scores for Interrater Reliability Sample

Rater	Mean	SD	Skewness	Kurtosis	Alpha
Teacher 1	3.36	.78	.05	-1.10	.919
Teacher 2	3.28	.86	.35	-1.13	.925

Note: n = 39. Teacher 1 is the general education classroom teacher, and Teacher 2 is the special education in-class support teacher.

Research Question 1: Factor Structure and Internal Consistency

Exploratory factor analysis (EFA) was used to examine the factor structure of the SRSI-TRS. Prior to conducting the EFA, the data was assessed to determine whether it was adequate for factor analysis. The Kaiser-Meyer-Olkin Measure of Sampling Adequacy indicated that the strength of the relationships among variables was high (KMO = .919). As noted previously, the KMO measure examines shared variance between variables. A high level of shared variance indicates that the variables are measuring a common factor and are therefore suitable for factor analysis. Bartlett's test of Sphericity was also significant (χ^2 [435] = 4751.73, $p < .05$). The null hypothesis for this test is that the correlation matrix is an identity matrix, with no collinearity between variable. A significant result rejects the null hypothesis and indicates sufficient correlation between the variables to proceed with the analysis.

Principal axis factoring (PAF) analysis was chosen because it aligns conceptually with the purpose of the investigation; that is, to identify a latent construct underlying the measured variables (Meyers et al., 2017). A preliminary PAF analysis was conducted without rotation to obtain the scree plot and to determine the amount of variance explained by each of the factors. In the preliminary model, two factors had an eigenvalue greater than 1.0. The first factor had an eigenvalue of 7.22, and accounted for 55.53% of the variance. The second factor had an eigenvalue of 2.03, and accounted for 15.63% of the variance. Thus, both factors cumulatively accounted for 71.16% of the variance on the scale. A scree plot (see Figure 1) also provided support for a two-factor model. All items on the scale had communality values greater than .50, and were therefore retained for further analyses (Meyer et al., 2017).

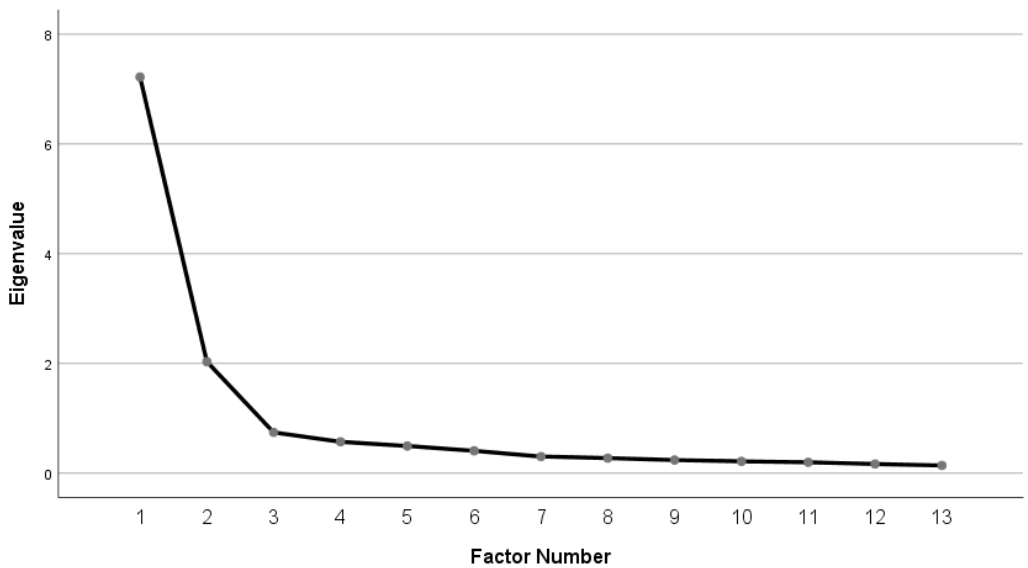

Figure 1. EFA Scree Plot.

The PAF analysis was conducted again using an oblique strategy with promax rotation. The model was constrained to two factors. An oblique strategy was chosen because SRL theory supports correlations between constructs in SRL (Chen, Cleary, & Lui, 2014). Promax rotation was used, as recommended by Meyers and colleagues (2017). The analysis was also done using a direct oblimin rotation, and both rotation methods yielded similar results. The promax was chosen because it gave a cleaner solution; that is, the items has higher loadings for their factors and lower loadings for the other factor, thus more clearly demonstrating the two-factor solution.

Factor loadings for items on the SRSI-TRS are presented on Table 6. Eigenvalues and percent of variance explained are also included for each factor. Factor 1 consisted of six items. The highest loading was .931, and the lowest loading was .708. These items reflected teacher perceptions of students' use of help-seeking and information-seeking behaviors, such as asking the format of upcoming tests or attending extra help sessions (see Table 6; items 1, 3, 4, 5, 6, 9). Thus, this factor was labeled *Seeking Help and Information*. Factor 2 also consisted of six items. The highest loading was .827, and the lowest loading was .647. These items reflected teacher perceptions of students' management of learning through organizational, motivation, and self-control strategies (items 2, 7, 8, 10, 12, 13). This factor was labeled *Managing Behavior*, which is consistent with other scales from the SRSI-SR and SRSI-PRS (Chen, Cleary, & Lui, 2015; Cleary, 2006). Item 11 (The student is enthusiastic about learning) cross-loaded with loadings above .40 on both factors (Meyer et al., 2017). This item was dropped from further analysis.

Table 6

Factor Loadings of Self-Regulation Strategy Inventory – Teacher Ratings Scale (SRSI-TRS)

	Factors	
Item	Factor 1	Factor 2
1. The student asks about topics that might appear on upcoming tests.	.931	-.097
3. The student asks insightful questions in class.	.871	-.005
4. The student asks questions about errors he or she makes on tests or assignments.	.864	-.012
6. The student asks questions in class when he or she does not understand something.	.821	-.011
9. The student asks about the format of upcoming tests (short-answer, multiple choice)	.818	-.187
5. The student seeks help or attends extra help sessions.	.708	.042
12. The student makes excellent use of class time.	-.335	.975
13. The student is prepared for class.	-.115	.827
2. The student keeps his or her class materials very organized.	-.045	.715
8. The student monitors how well he or she learns class material.	.216	.700
10. The student pushes himself or herself to understand the details of the topics presented in class.	.306	.666
7. The student keeps himself or herself motivated even when they struggle to learn something.	.298	.647
11. The student is enthusiastic about learning.	.479	.404
Eigenvalues	7.22	2.03
Percent of Variance Explained	55.53%	15.63%
Cronbach's Alpha	.918	.904

Note: Extraction Method: Principal Axis Factoring. Rotation Method: Promax with Kaiser Normalization.

Cronbach's alpha was examined for each of the factors. Both demonstrated excellent internal reliability. Factor 1 had an alpha of .918, and Factor 2 had alpha of .904. Item-total statistics were also examined (see Table 7). All items demonstrated adequate correlations to the total factor score, which indicates that the items are measuring the same construct. Reliability of the subscales would not be improved by deleting any of the items.

Table 7

Item-total Statistics for SRSI-TRS Factors

Item	Item-total Correlation	Cronbach's Alpha if Deleted
Factor 1: Seeking Help and Information		
1. The student asks about topics that might appear on upcoming tests.	.826	.895
3. The student asks insightful questions in class.	.806	.898
4. The student asks questions about errors he or she makes on tests or assignments.	.826	.896
5. The student seeks help or attends extra help sessions.	.701	.912
6. The student asks questions in class when he or she does not understand something.	.786	.902
9. The student asks about the format of upcoming tests (short-answer, multiple choice)	.684	.916
Factor 2: Managing Behavior		
2. The student keeps his or her class materials very organized.	.654	.899
7. The student keeps himself or herself motivated even when they struggle to learn something.	.783	.880
8. The student monitors how well he or she learns class material.	.795	.878
10. The student pushes himself or herself to understand the details of the topics presented in class.	.803	.877
12. The student makes excellent use of class time.	.692	.895
13. The student is prepared for class.	.710	.892

Research Question 2: Interrater Reliability

Pearson correlations were used to examine the level of agreement between ratings of student regulatory behaviors provided by two teachers. As noted previously, one of the teacher raters was a general education teacher (Teacher 1) while the second rater was a special education in-class support teacher in the same classroom (Teacher 2). The purpose of asking two teachers to rate each student was to examine whether informants in the same context (i.e., mathematics class) had similar perceptions of students' SRL behaviors relating to that context. It was hypothesized that large correlations would be found, as is consistent with previous research (Achenbach, 1987).

Correlations between the two teacher ratings were examined for the overall scale and for each of the two subscales (see Table 8). Descriptors for the magnitude of the correlations are based on Hopkins (2001) expansion of Cohen's (1988) guidelines: Trivial = .00-.09; Small = .10-.29; Medium = .30-.49; Large = .50-.69; Very Large = .70-.89; Almost Perfect = .90-1.00. Thus, the observed correlations for interrater reliability ranged from large to very large. The ratings for the SRSI-TRS composite score exhibited very large relations ($r = .75$).

Table 8

Interrater Reliability

Scale/Subscale	Correlation	95% Confidence Interval
SRSI-TRS Composite Score	.75	.57 – .86
Seeking Help and Information	.63	.39 – .79
Managing Behavior	.71	.51 – .84

Note: $n = 39$; all correlation are significant at $p < .01$ (one-tailed)

Mean differences between teacher ratings for each student were also examined in order to obtain more information about interrater agreement. The difference was computed between each pair of ratings, and absolute values were used to find the mean difference. The mean difference of the overall SRSI-TRS score was .45. The mean difference of the *Seeking Help and Information* subscale scores was .75, while the mean difference of the *Managing Behavior* subscale scores was .50. It should be noted that these results are descriptive and do not provide information whether the level of agreement is high or low. A one sample t-test with zero as the Test Values was conducted to see whether the mean difference of the overall score was significantly different than zero. Results showed that the test was not significant (t = -.632(38), p = .531). This test uses actual values of differences (not absolute values) in order to determine whether there is bias in the differences; that is, whether one rater gave consistently higher or lower scores than the other rater. Mean differences that are not significantly different than zero, as was found, indicate no bias and are suggestive of good agreement (Bland & Altman, 2003).

Research Question 3: Convergent Validity

Pearson correlations were also used to examine the level of convergence between the composite and subscale scores of the SRSI-TRS with the composite of the STRS, and the composite and subscales of the SRSI-SR. Guidelines to interpret the strength of the observed relations were noted above (Hopkins, 2001).

It was hypothesized that small to medium to correlations would be found between the SRSI-TRS and the SRSI-SR. It was also hypothesized that medium correlations, at minimum, would be found between the SRSI-TRS and the STRS. As hypothesized, the SRSI-TRS composite and subscales exhibited significant small correlations with the SRSI-SR composite and subscales (r = .16 - .29). The SRSI-TRS composite and subscales also exhibited significant small to medium correlations with the STRS ($r = .24 - .30$). Overall, majority of the correlations were in the small range, with some approaching medium. Steiger's Z-test was used to determine whether the correlation between the SRSI-TRS and the STRS ($r = .30$) was significantly higher than the correlation of the SRSI-TRS and the SRSI-SR composite ($r = .25$). Results showed that the difference was not significant ($z = -1.586$, $p = .06$), despite the STRS being an identical measure to the original teacher rating scale.

Table 9

Correlations among Teacher and Student Measures

Measure	1	2	3	4	5	6	7	8
1. TRS Composite	--							
2. TRS Seeking Help and Information	.92	--						
3. TRS Managing Behavior	.83	.55	--					
4. STRS	.30	.24	.30	--				
5. SR Composite	.25	.18	.29	.82	--			
6. SR Managing Environment and Behavior	.24	.16	.28	.75	.95	--		
7. SR Seeking and Learning Information	.19	.16	.19	.79	.89	.79	--	
8. SR Maladaptive Regulatory Behavior[a]	.24	.16	.28	.64	.80	.65	.57	--

Note: $n = 335$. TRS = Teacher Rating Scale, SR = Self-Report, STRS = Student version of TRS. All correlations are significant at $p < .01$ (one-tailed). [a] = reverse coded.

Item-level correlations for the SRSI-TRS and STRS were examined qualitatively to determine whether student and teacher ratings had greater convergence for certain SRL behaviors. Correlations ranged from .04 to .35 (see Table 10). All correlations were significant, except Item 9. Means for the items are presented as well.

Several patterns emerged when examining the item-level correlations. The items with correlations above .20 appear to reflect overt behaviors that can be readily observed by teachers (Items 1, 2, 4, 6, 7, 12, and 13). Many of these items can be easily understood by raters and thus do not require subjective interpretation. The item with the largest correlation (Item 13) clearly reflects a discrete behavior that can be easily rated by both students and teachers. Conversely, qualitative or descriptive analysis of items with correlations below .20 revealed that they either reflect covert processes (Items 8 and 10) or may involve behaviors that do not frequently occur in a classroom and are therefore more difficult to rate accurately (Items 5 and 9). Other items include words that require interpretation. For example, Item 3 asks about "insightful questions" and Item 11 about student being "enthusiastic," both which may be difficult to operationalize.

Table 10

Item-level Correlations and Mean Scores for SRSI-TRS and STRS

Item	Correlation	Mean TRS	Mean STRS
1. The student asks about topics that might appear on upcoming tests.	.21*	2.71	3.50
2. The student keeps his or her class materials very organized.	.27*	4.21	4.20
3. The student asks insightful questions in class.	.16*	3.03	3.40
4. The student asks questions about errors he or she makes on tests or assignments.	.24*	3.21	3.63
5. The student seeks help or attends extra help sessions.	.18*	2.80	3.07
6. The student asks questions in class when he or she does not understand something.	.26*	3.46	3.89
7. The student keeps himself or herself motivated even when they struggle to learn something.	.26*	3.82	3.83
8. The student monitors how well he or she learns class material.	.16*	3.79	3.59
9. The student asks about the format of upcoming tests (short-answer, multiple choice)	.04	1.99	3.39
10. The student pushes himself or herself to understand the details of the topics presented in class.	.13*	3.85	4.01
11. The student is enthusiastic about learning.	.15*	3.77	3.67
12. The student makes excellent use of class time.	.25*	4.12	3.99
13. The student is prepared for class.	.35*	4.50	4.62

Note: * = significant at $p < .01$ (one-tailed)

TRS = teacher rating scale, STRS = student version of teacher rating scale. Pooled results were used to obtain mean scores.

Discussion

The purpose of this study was to examine the psychometric properties of the SRSI-TRS, a tool used to assess teachers' perceptions of student SRL. While some prior research exists for the SRSI-TRS, this study was unique because it examined the reliability of scores and validity inferences drawn from this measure in a comprehensive way. This study utilized factor analysis procedures and examined interrater reliability of the SRSI-TRS, two areas that have not been addressed in prior research. In addition, this study adds to the convergent validity literature for the SRSI-TRS by including a student self-report questionnaire with identical items to the teacher rating scale targeting student behavior within the classroom. The findings of this study provide evidence about the viability of the SRSI-TRS as a measure of student SRL, and add to the general literature on SRL assessment and teacher ratings of student behavior.

The factor structure of the SRSI-TRS was examined through principal axis factoring, which yielded a two factor model. Analyses of internal reliability using Cronbach's alpha indicated that the overall scale and subscales had excellent internal consistency. Interrater reliability was examined using Pearson correlations and mean differences in scores; both indicated high levels of agreement between raters. Convergent validity was assessed by examining Pearson correlations between the SRSI-TRS and two student self-report measures of SRL. The majority of the correlations between student and teacher ratings were in the small range, with a few in the medium range. The results for the three research questions are discussed in greater detail below.

Factor Structure

The first objective of this study was to examine the internal factor structure of the SRSI-TRS. Exploratory factor analysis was used because the SRSI-TRS was not designed with a specific a priori structure (Cleary & Callan, 2014). Principal axis factoring procedures identified a two factor model. Each factor had six items with high loadings; thus, twelve of the thirteen items on the scale were included in the solution. The first factor reflected students' use of help-seeking and information-seeking behaviors. Sample items are, "The student asks about topics that might appear on upcoming tests" and "The student seeks help or attends extra help sessions." This factor was labeled *Seeking Help and Information*. The second factor reflected students' use of a variety of organizational, motivational, and metacognitive strategies to manage their learning and performance. Sample items are, "The student keeps his or her class materials very organized" and "The student monitors how well he or she learns class material." This factor was labeled *Managing Behavior*.

One item on the SRSI-TRS demonstrated significant cross-loading, with loadings above .40 on both factors. This item was, "The student is enthusiastic about learning." Enthusiasm is related to motivation for learning, an important component of SRL (Zimmerman, 2011). The two factors identified reflect metacognitive and behavioral processes of SRL; thus, the item about students' enthusiasm does not fit on either factor. Another item related to motivation ("The student keeps himself or herself motivated even when they struggle to learn something") did load on the *Managing Behavior* subscale, perhaps because it reflects on students' use of strategies to sustain their motivation, whereas the first item asks about students' general state of enthusiasm for learning. The item with cross-loadings was dropped from the scale for subsequent analyses because it did not add to the understanding of the construct measured by the SRSI-TRS.

Another point to note with regard to the items on the SRSI-TRS is that the teachers who participated in this study were given an option to rate "don't know." Several items had a few "don't know" responses, but for one item ("The student monitors how well he or she learns class material") about 10% of the responses were "don't know." This high rate of "don't know" responses may suggest that teachers found it difficult to comment on students' metacognitive processes, or that teachers were not be familiar enough with self-monitoring strategies to be able to assess students' use of them. In general, research on teacher ratings of student functioning has shown greater convergence with students' or other informants' reports when externalizing behaviors are measured (Achenbach et al., 1987; De Los Reyes et al., 2015). Thus, it may be easier for teachers to rate students' overt SRL behaviors (e.g., asking questions, attending help sessions) as opposed to more covert processes (e.g., self-monitoring.)

It is interesting to compare the factor structure of the SRSI-TRS to that of other measures within the SRSI assessment system: the student self-report (SRSI-SR) and the parent rating scale (SRSI-PRS). The teacher rating scale assesses teacher perceptions of students' use of SRL strategies and behaviors in the classroom, while the student self-report and parent rating scale both target students' use of SRL strategies and behaviors at home during studying and homework activities. Prior studies have found that the SRSI-SR and SRSI-PRS both have a three-factor structure, and the structures are fairly similar (Chen et al., 2014; Cleary et al., 2015). The student and parent scales both have two subscales measuring adaptive regulatory behaviors and one subscale measuring maladaptive regulatory behavior, while the SRSI-TRS has two adaptive scales and does not include any items targeting maladaptive behavior.

The two subscales on the SRSI-TRS closely mirror the adaptive subscales on the SRSI-SR and the SRSI-PRS (see Table 11). The SRSI-SR has a subscale *Seeking and Learning Information*, with items similar to the *Seeking Help and Information* scale on the SRSI-TRS. The SRSI-PRS does not include items related to help-seeking. The SRSI-SR also includes a scale of *Managing Environment and Behavior*, while the two adaptive scales on the SRSI-PRS are *Managing Environment* and *Managing Behavior and Learning*. The items about managing the environment (e.g., studying in a quiet place) are not relevant in the classroom context and therefore are not found on the SRSI-TRS. The *Managing Behavior* scale on the SRSI-TRS is similar to the *Managing Environment and Behavior* scale on the SRSI-SR and the *Managing Behavior and Learning* scale on the SRSI-PRS. Thus, all three measures include items reflecting strategies related to managing behavior and learning during the forethought (e.g., organizing materials, planning goals) and performance control (e.g., self-monitoring, motivation) phases of SRL.

Table 11

Comparison of SRSI Measures

Subscales	SRSI-TRS	SRSI-SR	SRSI-PRS
Adaptive	1. Seeking Help and Information	1. Seeking and Learning Information	
	2. Managing Behavior	2. Managing Environment and Behavior	1. Managing Behavior and Learning
			2. Managing Environment
Maladaptive		3. Maladaptive Regulatory Behavior	3. Maladaptive Regulatory Behavior

The results of the exploratory factor analysis provide a preliminary understanding of the construct measured by the SRSI-TRS; that is, students' use of a variety of adaptive regulatory behaviors and strategies. As was noted, the SRSI-TRS does not provide information on the frequency of students' maladaptive regulatory behavior. SRL encompasses adaptive behaviors as well as maladaptive behaviors, such as procrastination or avoidance (Zimmerman, 2000). While these behaviors may be more likely to occur during independent learning (i.e., when students are doing homework or studying), and are therefore included on the other two SRSI scales, they can also be relevant in the classroom (Boekaerts & Corno, 2005). Additionally, the SRSI-TRS was originally created to parallel the SRSI-SR, which in turn was designed to reflect ten general categories of SRL identified by Zimmerman and Martinez Pons (1986). The final version of the SRSI-TRS does not include items about some of those categories, mainly because they are not relevant in a classroom context (Cleary & Callan, 2014). Thus, while the SRSI-TRS provides valuable information about students' use of SRL strategies in class, it is important to keep in mind that it assesses a modest array of strategies included within the broader SRL construct.

Interrater Reliability

The second objective of this study was to examine interrater reliability of the SRSI-TRS. Interrater reliability for the SRSI-TRS has not been explored in prior studies. Research on teacher ratings of student behavior in other domains of functioning tends to report high levels of agreement between two teachers; for example, Achenbach et al. (1987) found average correlations of .64 between ratings given by pairs of teachers. For the current study, two math teachers in the same classroom were asked to rate the same students. Thus, the raters were very familiar with the students because they interacted with them on a daily basis in the same context. This is important given the contextualized nature of SRL. While students' use of SRL strategies and behaviors may differ based on the class or subject matter, it was hypothesized that two teachers in the same classroom would provide ratings with a high level of agreement. High agreement in this case would serve as an indicator of a reliable measure that yields ratings that are stable across respondents.

The findings of the interrater reliability analyses supported the hypothesis. Pearson correlations of the SRSI-TRS composite scores showed very high relations (r = .75). The *Seeking Help and Information* subscale had a correlation of .63, while the *Managing Behavior* subscale had a correlation of .71. Because correlations reflect the linear relationship between two variables and do not measure absolute agreement, an additional measure of mean differences between scores was examined. A one-sample t-test showed that the mean difference between pairs of scores for each student was not significantly different than zero, indicating that there was no bias in the ratings and suggesting good agreement (Bland & Altman, 2003).

There are some limitations to note with regard to the interrater reliability analyses. The sample size was small; only 39 students were rated by the two teachers. In addition, the raters were not consistent across the sample. There were six pairs of teachers, with each pair rating a different number of students, ranging from one to 17 students. Separate analyses showed that when correlations were examined for each pair of teachers individually, the results varied greatly, with correlations ranging from .43 (nonsignificant) to .99 (significant at $p < .01$.) Additional research with more robust samples is needed to provide more information about the interrater reliability of this measure; however, the preliminary findings from the current study show positive evidence for stability of scores across raters.

Convergent Validity

The third objective of this study was to examine convergent validity of the SRSI-TRS. Convergence across assessment measures is an area of SRL research that has received increased attention in recent years. Researchers emphasized the need to focus on triangulation across measures to determine how different measures complement each other and capture different aspects of SRL (Callan & Cleary, 2017; Winne & Perry, 2000). In general, measures within the same group (i.e., aptitude or event measures) show larger correlations than measures in different groups (Callan & Cleary, 2017). This study examined three broad measures of SRL, a teacher rating scale and two student self-report questionnaires. Prior research found small to medium correlations for teacher-student agreement with regard to ratings of SRL (Cleary et al., 2015), which is consistent with the general literature on cross-informant agreement between students and teachers (Achenbach, 1987).

This study used two different student self-report questionnaires. One of those measures, the SRSI-SR, was designed to assess students' use of SRL strategies and behaviors at home while studying and doing homework. In contrast, the SRSI-TRS assesses student SRL in the classroom. SRL theory supports the premise that SRL is highly contextualized and may fluctuate and vary across contexts and situations (McCardle & Hadwin, 2015; Panadero, 2017). Furthermore, contextual differences are a large factor underlying cross-informant discrepancies in children's behavior in general (De Los Reyes et al., 2015). Thus, it is not surprising that in this study the majority of correlations between measures targeting SRL behaviors in different contexts were in the small range, with a few approaching the medium range. These findings are consistent with prior research (Cleary et al., 2015).

The results for convergent validity of the TRS with the student version (STRS) were somewhat more surprising. The STRS measure was virtually identical to the SRSI-TRS (i.e., except for first person wording for students) and thus targeted SRL in the same context (i.e., math class.). In contrast to the author's hypotheses, small to medium correlations between the TRS and STRS were observed. The correlation between the STRS composite and the SRSI-TRS was not significantly larger than that of the SRSI-SR composite and the SRSI-TRS. Even when given identical measures that assessed SRL in the same context, teacher and student reports showed small to medium correlations, as is consistent with other research on teacher-student agreement.

Prior studies have offered different hypotheses to explain the small relations found between teacher and student ratings of SRL. With regard to the SRSI-TRS and the SRSI-SR, one logical explanation is that the two measures targeted SRL behaviors in distinct contexts and thus may be capturing different aspects of SRL (Callan & Cleary, 2017). Another explanation is that students may not be accurate reporters, and their perceptions of their behaviors may not match those of others, such as teachers or parents (Cleary & Chen, 2009; Winne & Jamieson-Noel, 2002). The latter hypothesis suggests that parent or teacher ratings are more objectively "accurate" than student ratings, and is supported by research showing that parent and teachers measures are more robust predictors of outcomes than student self-reports (Chen, Cleary, & Lui, 2015; Cleary & Callan, 2014).

The current study sheds some light on this issue because of the inclusion of the STRS measure. Because the SRSI-TRS and STRS were targeting SRL in the same setting, contextual differences cannot be an explanation for the small correlation sizes. In their review of the literature on cross-informant agreement for child mental symptoms, De Los Reyes and colleagues (2015) offered alternate hypotheses for discrepancies across raters in the same context; for example, differences in informants' perspectives, rater bias, and measurement error may all be contributing factors. Rater bias and measurement error are inherent limitations of any rating scale, but cannot alone explain cross-informant discrepancies (De Los Reyes, 2013). Differences in perspectives between teachers and student may be a plausible explanation for the small correlations found in this study. As was shown in prior research on SRL assessment, students' perceptions of their own behavior may not match those of outside observers. The current study does not provide information about which informant is more accurate, but it underscores the need for multimethod, multisource assessment of SRL. It also raises the question of incremental validity; that is, whether student reports of SRL in the same classroom provides unique predictive value above that of the SRSI-TRS.

Item-level convergence was also examined for the SRSI-TRS and the STRS to see whether student and teacher ratings had higher correlations for certain SRL behaviors. All correlations except one were significant, and majority of them were in the small range. Some patterns can be observed when examining the correlations qualitatively, although further research would be needed to ascertain whether teacher-student agreement varies significantly based on the different types of SRL behaviors measured. In this study, generally the items with higher correlations reflected overt behaviors that are easily observed, such as asking about errors on tests or assignments and keeping class materials organized. Items with lower correlations appeared to have more subjectivity (e.g., students asking “insightful” questions), or were behaviors that may not be observed often in a classroom, such as asking about the format of upcoming tests. The item that received the most “don’t know” responses (“The student monitors how well he or she learns class material”) had one of the lower correlations, as did the item that did not load on either factor (“The student is enthusiastic about learning.”) This information is only descriptive, but it can provide suggestions of how to potentially revise and improve items on the SRSI-TRS.

Limitations and Areas for Future Research

There are some limitations to note when considering the results of the current study. Firstly, nested data was used for this study; that is, ratings on the SRSI-TRS during the second data collection phase (that was used for the factor analysis and convergent validity analyses) were provided by fourteen teachers in different classrooms. The statistical methods used operated under the assumption that students’ ratings were independent of each other, when in fact multiple students shared a common rater (i.e., the same teacher.) Students in the same class may also share experiences or characteristics that would influence the ratings. Furthermore, each teacher rated a different number of students (ranging from 4 to 56 students), which can introduce additional biases into the data. Future research done should use multilevel modeling techniques to account for the nesting (Braun, Jenkins, & Grigg, 2006). Additionally, the subsample used to examine interrater reliability was also composed of students in different classrooms, and was further limited in that different pairs of teacher rated different numbers of students, as was noted.

The external validity of the study is also limited because the sample consisted of only seventh and eighth graders from one middle school. Also, the measures were adapted to target SRL behaviors in students' math class. These factors limit the generalizability of the findings. Future research should continue to explore the use of SRL teacher ratings across different age groups, student populations, and content areas.

Another limitation pertains to the measures used. While the focus of this study was the SRSI-TRS, only student self-report measures were used to examine convergent validity. Both teacher ratings and student self-reports are aptitude measures; that is, broad, aggregate measures of SRL. It is important for additional research to be done comparing teacher ratings to other types of SRL measures, particularly event measures that capture regulatory behaviors as they are happening in authentic learning situations. A couple of studies have looked at convergence between teacher ratings and SRL microanalysis, and more work needs to be done to determine how teacher ratings best fit in a multidimensional assessment of SRL that includes both event and aptitude measures. In light of recent research showing the importance of event measures in pinpointing deficits in students' SRL, it is particularly relevant to understand how teacher ratings compare to these measures, if teacher ratings are to be used diagnostically for assessment and intervention. In addition, the STRS student self-report questionnaire was a new measure developed for the longitudinal study, and no prior information was available regarding its reliability and validity.

The findings also point to some potential item-level revisions that may strengthen the SRSI-TRS. As was noted previously, teachers had an option to rate “don’t know” on the SRSI-TRS. Although the use of “don’t know” response was very limited across most items, there was one item that had approximately 10% “don’t know” responses from teachers (“The student monitors how well he or she learns class material.”). It may be that teachers rated “don’t know” for behaviors they cannot observe, or they may have limited knowledge about the skills the items were targeting (e.g., students’ monitoring of learning). Further research can focus on identifying which SRL behaviors may be more easily observed and rated, as well as how SRL behaviors can be operationalized in the classroom context in order to adapt the items appropriately. In addition, the SRSI-TRS may be improved by revising items that included subjective words, such as “The student asks insightful questions.”

Finally, while the results of this study are informative and provide support for the SRSI-TRS as a reliable and valid measure, it is important to keep in mind that the current results are preliminary. Confirmatory factor analysis is needed to provide support for the factor structure identified in this study. Further, more information is needed on interrater reliability and convergent validity with other measures and related constructs. Divergent validity analyses, which were not included in this study, would also further enhance the findings.

Implications for School Psychologists

The results of this study provide support for use of the SRSI-TRS as a valid and reliable measure of student SRL. As such, it may be an effective screening tool that school psychologists can incorporate into their assessment repertoire. Researchers and practitioners have recognized the importance of assessing factors that influence students' academic achievement, like SRL and motivation (Cleary, 2006; DiPerna, Volpe, & Elliot, 2002; Linnenbrink & Pintrich, 2002). In addition, self-regulation and motivation deficits are a common referral concern at both the elementary and secondary school levels (Bramlett, Murphy, Johnson, & Wallingsford, 2002; Cleary, 2009). Thus, including an SRL measure in a psychoeducational testing battery or using it as a screener for all students would allow school psychologists to identify individuals who need intervention. School psychologists often rely on teacher ratings to assess various aspects of student functioning, and teachers are well-positioned to provide information about students' learning strategies and behaviors. The SRSI-TRS can help school psychologists obtain reliable and valid information from teachers in an efficient manner.

The findings of this study regarding convergent validity also hold implications for school psychologists. Similar to prior research, small correlations were found between teacher and student ratings. While this study cannot determine whether student or teacher ratings are objectively more accurate, it suggests the importance of considering both sources of information. Although student self-report measures have been criticized in the literature because they often do not align well with other more objective measures (e.g., traces, microanalysis), they are still important because students' perceptions and judgements directly influence their regulatory behaviors (McCardle & Hadwin, 2015). Therefore, it is important for school psychologists to assess both student and teacher perceptions of SRL in order to gain a better understanding of student functioning and determine areas in need of intervention.

Lastly, one of the key strategies highlighted on the SRSI-TRS is help-seeking. Assessing middle school students' use of adaptive help-seeking behaviors is important since students may be reluctant to seek help because of peer pressures (Ryan, Pintrich, & Midgley, 2001).

Additionally, advances in technology have made help-seeking more versatile, and also require that students have skills to successfully navigate electronic or web-based help-seeking interactions (Mahasneh, Sowan, & Nasser, 2012). Help-seeking is a complex and important process that is related to academic achievement (Karabenick, 1998), and it is important target for assessment, instruction, and intervention. The SRSI-TRS can help school psychologists and educators identify students with skill deficits in this area in order to provide appropriate intervention.

Conclusion

The current study expanded on prior research by examining various psychometric properties of the SRSI-TRS. Exploratory factor analysis yielded a two factor model, with twelve of the thirteen items on the scale included in the solution. The structure found was similar to the structures of the SRSI-SR and the SRSI-PRS. Internal consistency of the overall scale and subscales was satisfactory. Results of interrater reliability analyses suggested high agreement between two raters, and were consistent with prior research on cross-informant agreement between pairs of teachers. Convergent validity was examined by looking at the correlations between the SRSI-TRS and two student self-report questions. The small to medium correlations found indicate that the measures are assessing related constructs. However, students and teacher ratings revealed low levels of correspondence, even when the measures given were identical. It appears that differences in student and teacher perspective on students' use of SRL in the classroom contribute to the discrepancies between ratings. The results of this study provide further evidence that scores yielded from the SRSI-TRS are stable and consistent, and valid inferences can be drawn from them. Overall, these results suggest that the SRSI-TRS may be a useful tool for assessing student SRL in the classroom context.

References

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101*(2), 213.

Adagideli, F. H., Sarac, S., & Ader, E. (2015). Assessing preschool teachers' practices to promote self-regulated learning. *International Electronic Journal of Elementary Education, 7*(3), 423-440.

Alexander, P. A., Dinsmore, D. A., Parkinson, M. M., & Winters, F. I. (2011). Self-regulated learning in academic domains. In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 393–407). New York: Routledge.

Alexander, P. A., Graham, S., & Harris, K. R. (1998). A perspective on strategy research: Progress and prospects. *Educational Psychology Review, 10*(2), 129–154.

Andrzejewski, C.E., Davis, H.A., Bruening, P.S., & Poirier, R.R. (2016). Can a self-regulated strategy intervention close the achievement gap? Exploring a classroom-based intervention in 9th grade earth science. *Learning and Individual Differences, 49*, 85-99.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes, 50*(2), 248-287.

Baraldi, A. N., & Enders, C. K. (2010). An introduction to modern missing data analyses. *Journal of School Psychology, 48*, 5–37.

Barber, B. K., & Olsen, J. A. (2004). Assessing the transitions to middle and high school. *Journal of Adolescent Research, 19*(1), 3-30.

Bell, C. V., & Pape, S. J. (2014). Scaffolding the development of self-regulated learning in mathematics classrooms. *Middle School Journal, 45*(4), 23-32.

Ben-Eliyahu, A., & Bernacki, M. L. (2015). Addressing complexities in self-regulated learning: a focus on contextual factors, contingencies, and dynamic relations. *Metacognition and Learning*, *10*(1), 1-13.

Bland, J. M., & Altman, D. G. (2003).Applying the right statistics: Analyses of measurement studies. *Ultrasound in Obstetrics and Gynecology, 22*, 85–93.doi: .1002/uog.122

Boekaerts, M., & Cascallar, E. (2006). How far have we moved toward the integration of theory and practice in self-regulation? *Educational Psychology Review, 18*(3), 199–210. doi:10.1007/s10648006-9013

Boekaerts, M., & Corno, L. (2005). Self-regulation in the classroom: A perspective on assessment and intervention. *Applied Psychology*, *54*(2), 199-231.

Boekaerts, M., de Koning, E., & Vedder, P. (2006). Goal-directed behavior and contextual factors in the classroom: An innovative approach to the study of multiple goals. *Educational Psychologist*, *41*(1), 33-51.

Boekaerts, M., & Niemivirta, M. (2000). Self-regulated learning: Finding a balance between learning goals and ego-protective goals. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation*. San Diego, CA: Academic Press.

Boekaerts, M., Pintrich, P. R., & Zeidner, M. (Eds.). (2000). *Handbook of self-regulation*. San Diego, CA: Academic Press.

Bramlett, R. K., Murphy, J. J., Johnson, J., & Wallingsford, L. (2002). Contemporary practices in school psychology: A national survey of roles and referral problems. *Psychology in the Schools, 39*, 327–335.

Braun, H., Jenkins, F., and Grigg, W. (2006). Comparing private schools and public schools using hierarchical linear modeling (NCES 2006-461). U.S. Department of Education, National Center for Education Statistics, Institute of Education Sciences. Washington, DC: U.S. Government Printing Office.

Broadbent, J. (2017). Comparing online and blended learner's self-regulated learning strategies and academic performance. *Internet and Higher Education, 33*, 24-32. doi:10.1016/j.iheduc.2017.01.004

Brown, T.A. (2015). *Confirmatory factor analysis for applied research* (2nd ed). New York: Guilford Press.

Butler, D. L., Beckingham, B., & Lauscher, H. J. N. (2005). Promoting strategic learning by eighth-grade students struggling in mathematics: A report of three case studies. *Learning Disabilities Research and Practice, 20*(3), 156–174.

Butler, D. L., & Cartier, S. C. (2004). Promoting students’ active and productive interpretation of academic work: A key to successful teaching and learning. *Teachers College Record, 106*, 1729-1758.

Callan, G. L., & Cleary, T. J. (2017, March 30). Multidimensional assessment of self-regulated learning with middle school math students. *School Psychology Quarterly*. Advance online publication. http://dx.doi.org/10.1037/spq0000198.

Callan, G. L., Marchant, G. J., Finch, W. H., & German, R. L. (2016). Metacognition, strategies, achievement, and demographics: Relationships across countries. *Educational Sciences: Theory & Practice, 16*, 1485-1502. doi:10.12738/estp.2016.5.0137

Castillo, J. M., Curtis, M. J., & Gelley, C. D. (2012). School psychology 2010—part 2: School psychologists' professional practices and implications for the field. *Communique, 40*(8), 4–6.

Chen, P. P., Cleary, T. J., & Lui, A. M. (2015). Examining parents' ratings of middle-school students' academic self-regulation using principal axis factoring analysis. *School Psychology Quarterly, 30*(3), 385-397.

Chen, P. P., & Rossi, P. D. (2013). Utilizing calibration accuracy information with adolescents to improve academic learning and performance. In H. Bembenutty, T. J. Cleary, & A. Kitsantas (Eds.), *Applications of self-regulated learning across diverse disciplines: A tribute to Barry J. Zimmerman* (pp. 263-298). Charlotte, NC:gi Information Age.

Cleary, T. J. (2006). The development and validation of the self-regulation strategy inventory—self-report. *Journal of School Psychology, 44*(4), 307-322.

Cleary, T. J. (2009). School-based motivation and self-regulation assessments: An examination of school psychologist beliefs and practices. *Journal of Applied School Psychology, 25*(1), 71-94. doi: 10.1080/15377900802484190

Cleary, T. J. (2011). Emergence of self-regulated learning microanalysis: Historical overview, essential features, and implications for research and practice. In B. J. Zimmerman, & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 329–345). New York: Routledge.

Cleary, T. J. (2018). *The Self-Regulated Learning Guide: Teaching Students to Think in the Language of Strategies.* New York: Routledge.

Cleary, T. J., & Callan, G. L. (2014). Student self-regulated learning in an urban high school: Predictive validity and relations between teacher ratings and student self-reports. *Journal of Psychoeducational Assessment, 32*(4), 295-305. doi: 10.1177/0734282913507653.

Cleary, T. J., Callan, G. L., Malatesta, J., & Adams, T. (2015). Examining the level of convergence among self-regulated learning (SRL) microanalytic processes, achievement, and a self-report questionnaire. *Journal of Psychoeducational Assessment, 33*, 439-450. http://dx.doi.org/10.1177/0734282915594739

Cleary, T. J., & Chen, P. P. (2009). Self-regulation, motivation, and math achievement in middle school: Variations across grade level and math context. *Journal of School Psychology*, *47*(5), 291-314.

Cleary, T. J., Dembitzer, L., & Kettler, R. J. (2015). Internal factor structure and convergent validity evidence: The self-report version of self-regulation strategy inventory. *Psychology in the Schools, 52*(9), 829–844. http://dx.doi.org/10.1002/pits.21866.

Cleary, T. J., Gubi, A., & Prescott, M. V. (2010). Motivation and self-regulation assessments in urban and suburban schools: Professional practices and needs of school psychologists. *Psychology in the Schools, 47*(10), 985–1002

Cleary, T. J., & Kitsantas, A. (2017). Motivation and self-regulated learning influences on middle school mathematics achievement. School Psychology Review, 46(1), 88-107.

Cleary, T. J., & Labuhn, A. S. (2013). Application of cyclical self-regulation interventions in science-based contexts. In H. Bembenutty, T. J. Cleary, & A. Kitsantas (Eds.), *Applications of self-regulated learning across diverse disciplines: A tribute to Barry J. Zimmerman* (pp. 89-124). Charlotte, NC: Information Age.

Cleary, T. J., & Platten, P. (2013). Examining the correspondence between self-regulated learning and academic achievement: A case study analysis [special issue]. *Educational Research International*. http://dx.doi.org/10.1155/2013/272560.

Cleary, T. J., Velardi, B., & Schnaidman, B. (2017). Effects of the self-regulation empowerment program (SREP) on middle school students' strategic skills, self-efficacy, and mathematics achievement. *Journal of School Psychology*, 64, 28-42. doi:10.1016/j.jsp.2017.04.004

Cleary, T. J., & Zimmerman, B. J. (2004). Self-regulation empowerment program: A school-based program to enhance self-regulated and self-motivated cycles of student learning. *Psychology in the Schools, 41*, 537 – 550.

Cleary, T. J., & Zimmerman, B. J. (2006). Teachers' perceived usefulness of strategy microanalytic assessment information. *Psychology in the Schools, 43*, 149 – 155.

Corno, L. (2001). Volitional aspects of self-regulated learning. In B. J. Zimmerman & D. H. Schunk (Eds.), *Self-regulated learning and academic achievement: Theoretical perspectives* (pp. 191–226). Mahwah, NJ: Erlbaum.

Corno, L. (2011). Studying self-regulation habits. In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 361-375). New York: Routledge.

Darling-Hammond, L. (1995). Equity Issues in Performance-Based Assessment. In M. T. Nettles & A. L. Nettles (Eds.), *Equity and excellence in educational testing and assessment* (pp. 89-114.) Norwell, MA: Kluwer Academic Publishers.

De Corte, E., Mason, L., Depaepe, F., & Verschaffel, L. (2011). Self-regulation of mathematical knowledge and skills. In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 155-172). New York, NY: Routledge.

De Los Reyes, A. (2013). Strategic objectives for improving understanding of informant discrepancies in developmental psychopathology research. *Development and Psychopathology, 25*, 669–682. http://dx.doi.org/ 10.1017/S0954579413000096

De Los Reyes, A., Augenstein, T. M., Wang, M., Thomas, S. A., Drabick, D. A. G., Burgers, D. E., & Rabinowitz, J. (2015). The validity of the multi-informant approach to assessing child and adolescent mental health. *Psychological Bulletin, 141*(4), 858-900.

Demaray, M. K., & Elliott, S. N. (1998). Teachers' judgments of students' academic functioning: A comparison of actual and predicted performances. *School Psychology Quarterly, 13*(1), 8-24.

Dembo, M. H., & Eaton, M. J. (2000). Self-regulation of academic learning in middle level schools. *The Elementary School Journal*, *100*(5), 473-490.

Dent, A. L., & Koenka, A. C. (2016). The relation between self-regulated learning and academic achievement across childhood and adolescence: A meta-analysis. *Educational Psychology Review, 28*(3), 425-474. doi:10.1007/s10648-015-9320-8

DiBenedetto, M. K., & Zimmerman, B. J. (2013). Construct and predictive validity of microanalytic measures of students' self-regulation of science learning. *Learning and Individual Differences, 26*, 30-41. doi:10.1016/j.lindif.2013.04.004

Dicke, A.-L., Ludtke, O., Trautwein, U., Nagy, G., & Nagy, N. (2012). Judging students' achievement goal orientations: Are teacher ratings accurate? *Learning and Individual Differences, 22*, 844-849.

DiFrancesca, D., Nietfeld, J. L., & Cao, L. (2016). A comparison of high and low achieving students on self-regulated learning variables. *Learning and Individual Differences, 45*, 228-236. doi:10.1016/j.lindif.2015.11.010

Dignath, C., & Buttner, G. (2008). Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning, 3,* 231-264.

Dignath van Ewijk, C. (2011). Assessing students' acquisition of self-regulated learning skills using meta-analysis. In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 376-390). New York: Routledge.

Dinsmore, D. L., Alexander, P. A., & Loughlin, S. M. (2008). Focusing the conceptual lens on metacognition, self-regulation, and self-regulated learning. *Educational Psychology Review*, (4), 391.

DiPerna, J. C., Volpe, R. J., & Elliot, S. N. (2002). A model of academic enablers and elementary reading/language arts achievement. *School Psychology Review, 31*, 298 – 312.

Dörrenbächer, L., & Perels, F. (2016). Self-regulated learning profiles in college students: Their relationship to achievement, personality, and the effectiveness of an intervention to foster self-regulated learning. *Learning and Individual Differences*, 51, 229–241.

Downey, R. G., & King, C. V. (1998). Missing data in Likert ratings: A comparison of replacement methods. *The Journal of General Psychology, 125*(2), 175-191.

Eccles, J. S., Midgley, C., Wigfield, A., Buchanan, C. M., Reuman, D., Flanaga, C., & Mac Iver, D. (1993). Development during adolescence: The impact of stage-environment fit on young adolescents' experiences in schools and in families. *American Psychologist, 48*(2), 90-101.

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology,* 53, 109-132. doi:10.1146/annurev.psych.53.100901.135153

Ertmer, P. A., & Newby, T. J. (1996). The expert learner: Strategic, self-regulated, and reflective. *Instructional Science, 24*, 1–24.

Fagan, T., K. & Wise, P. S. (2007). *School psychology: Past, present, and future* (3rd ed.). Bethesda, MD: NASP.

Farling, W. H., & Hoedt, K. C. (1971). *National survey of school psychologists.* Washington, D.C.: NASP.

Ferguson, E., & Cox, T. (1993). Exploratory factor analysis: A users' guide. *International Journal of Selection and Assessment, 1*, 84-94.

Frick, P. J., Barry, C. T., & Kamphaus, R. W. (2010). *Clinical Assessment of Child and Adolescent Personality and Behavior* (3rd ed.). USA: Springer.

Fuchs, L. S., Fuchs, D., Prentice, K., Burch, M., Hamlett, C. L., Owen, R., & Schroeter, K. (2003). Enhancing third-grade students' mathematical problem solving with self-regulated learning strategies. *Journal of Educational Psychology, 95*(2), 306–315.

Geisinger, K. (2017). Interrater reliability. In A. Wenzel (Ed.), *The Sage Encyclopedia of Abnormal and Clinical Psychology* (Vol. 1, pp. 1891-1891). Thousand Oaks,, CA: SAGE Publications, Inc. doi: 10.4135/9781483365817.n759

Gerber, M. M., & Semmel, M. I. (1984). Teacher as imperfect test: Reconceptualizing the referral process. *Educational Psychologist*, *19*, 137–148.

Goh, D. S., Teslow, C. J., & Fuller, G. B. (1981). The practice of psychological assessment among school psychologists. *Professional Psychology, 12*, 696-706.

Graham, S., & Harris, K. R. (2009). Almost 30 years of writing research: Making sense of it all with the Wrath of Khan. *Learning Disabilities Research and Practice, 24*(2), 58-68.

Graham, J., Olchowski, A., & Gilreath, T. (2007). How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prevention Science, 8*(3), 206-213. doi:10.1007/s11121-007-0070-9

Greene, J. A., & Azevedo, R. (2007). Adolescents' use of self-regulatory processes and their relation to qualitative mental model shifts while using hypermedia. *Journal of Educational Computing Research*, *36*(2), 125-148.

Greene, J. A., & Azevedo, R. (2010). The measurement of learners' self-regulated cognitive and metacognitive processes while using computer-based learning environments. *Educational Psychologist, 45*(4), 203-209. doi:10.1080/00461520.2010.515935

Greene, J. A., Moos, D. C., & Azevedo, R. (2011). Self-regulation of learning with computer based environments. *New Directions for Teaching and Learning, 126*, 107-115. doi: 10.1002/tl.449

Greene, J. A., Robertson, J., & Croker Costa, L. J. (2011). Assessing self-regulated learning using think-aloud methods. In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 313-328). New York: Routledge.

Hadwin, A. F., Nesbit, J. C., Jamieson-Noel, D., Code, J., & Winne, P. H. (2007). Examining trace data to explore self-regulated learning. *Metacognition and Learning, 2*(2–3), 107–124. doi:10.1007/s11409-007-9016-7

Hofer, B. K., Yu, S. L., & Pintrich, P. R. (1998). Teaching college students to be self-regulated learning. In D. H. Schunk & B. J. Zimmerman (Eds.), *Self-regulated learning: From teaching to self-reflective practice* (pp. 57–85). New York, NY: Guilford Press.

Hoge, R. D., & Butcher, R. (1984). Analysis of teacher judgments of pupil achievement levels. *Journal of Educational Psychology, 76*, 777-781. doi:10.1037/0022-0663.76.5.777

Hoge, R. D., & Coladarci, T. (1989). Teacher-based judgments of academic achievement: A review of the literature. Review of Educational Research, 52, 297-313.

Hopkins, W.G. (2001). A scale of magnitudes for effect statistics. A New View of Statistics. Retrieved from http://sportsci.org/resource/stats/effectmag.html

Hosp, J. L., & Reschly, D. J. (2002). Regional differences in school psychology practice. *School Psychology Review, 31*, 11-29.

Hutton, J. B., Dubes, R., & Muir, S. (1992). Assessment practices of school psychologists: Ten years later. *School Psychology Review, 21*, 271 – 284.

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika, 39*, 31–36.

Karabenick, S. A. (2011). Methodological and assessment issues in research on help seeking. In In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 267-281). New York: Routledge.

Karabenick, S. A., & Berger, J. L. (2013). Help seeking as a self-regulated learning strategy. In H. Bembenutty, T. J. Cleary, & A. Kitsantas (Eds.), *Applications of self-regulated learning across diverse disciplines: A tribute to Barry J. Zimmerman* (pp. 237–261). Charlotte, NC: Information Age.

Karabenick, S & Zusho, A. (2015). Examining approaches to research on self-regulated learning: conceptual and methodological considerations. *Metacognition and Learning, 10*(1). doi:10.1007/s11409-015-9137-3.

Kenny, D. T., & Chekaluk, E. (1993). Early reading performance: A comparison of teacher-based and test-based assessments. *Journal of Learning Disabilities, 26*, 227-236.

Kettler, R. J., Elliott, S. N., DiPerna, J. C., Bolt, D. M., Reiser, D., & Resurreccion, L. (2014). Student and teacher ratings of academic competence: An examination of cross-informant agreement. *Journal of Applied School Psychology, 30*(4), 338-354. doi:10.1080/15377903.2014.950442

Linnenbrink, E. A., & Pintrich, P. R. (2002). Motivation as an enabler for academic success. *School Psychology Review, 31*, 313 – 327.

Litwin, M. S. (1995). *How to measure survey reliability and validity.* Thousand Oaks, CA: SAGE Publications.

Loeber, R., Green, S. M., & Lahey, B. B. (1990). Mental health professionals' perception of the utility of children, mothers, and teachers as informants on childhood psychopathology. *Journal of Clinical Child Psychology*, *19*, 136-143. doi:10.1207/s15374424jccp1902_5

Loeber, R., Green, S. M., Lahey, B. B., & Stouthamer-Loeber, M. (1989). Optimal informants on childhood disruptive behaviors. *Development and Psychopathology, 1,* 317–337. http://dx.doi.org/10.1017/S095457940000050X

Mahasneh, R. A., Sowan, A. K. & Nassar, Y. H. (2012). Academic help-seeking in online and face-to-face learning environments. *E-Learning and Digital Media, (9)*2, 196-210.

Malmberg, J., Jarvenoja, H., & Jarvela, S. (2010). Tracing elementary school students' study tactic use in gStudy by examining a strategic and self-regulated learning. *Computers in Human Behavior,* 26, 1034-1042. doi:10.1016/j.chb.2010.03.004

Matthews, G., Schwean, V. L., Campbell, S. E., Saklofske, D. H., & Mohamed, A. A. R. (2000). Personality, self-regulation, and adaptation: A cognitive-social framework. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 171-207). San Diego, CA: Academic Press.

McCardle, L., & Hadwin, A. F. (2015). Using multiple, contextualized data sources to measure learners' perceptions of their self-regulated learning. *Metacognition and Learning. 10,* 43-75. *doi*:10.1007/s11409-014-9132-0

McClelland, M. M., & Cameron, C. E. (2011). Self-regulation and academic achievement in elementary school children. In R. M. Lerner, J. V. Lerner, E. P. Bowers, S. Lewin-Bizan, S. Gestsdottir, & J. B. Urban (Eds.), Thriving in childhood and adolescence: The role of self-regulation processes. *New Directions for Child and Adolescent Development*, *133*, 29-44.

Meichenbaum, D. H., & Goodman, J. (1971). Training impulsive children to talk to themselves: A means of developing self-control. *Journal of Abnormal Psychology, 77*(2), 115-126.

Merrell, K. W. (2008). *Behavioral, social, and emotional assessment of children and adolescents* (3rd ed.). New York: Routledge.

Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G.G., Moreland, K. L., Dies, R. R., … Reed, G. M. (2001). Psychological testing and psychological assessment: A review of evidence and issues. *American Psychologist*, *56*(2), 128.

Meyers, L. A., Gamst, G.C., & Guarino, A. J. (2016). *Applied Multivariate Research: Design and Interpretation* (3rd ed.). LA: Sage Publications.

Montague, M. (2007). Self-regulation and mathematics instruction. *Learning Disabilities Research & Practice, 22*(1), 75-83.

Moos, D. C., & Ringdal, A. (2012). Self-regulated learning in the classroom: A literature review on the teacher's role. *Education Research International, 2012,* 1–15. http://dx.doi.org/10.1155/2012/423284.

Nickerson, A. B., & Nagle, R. J. (2001). Interrater reliability of the Devereux behavior rating scale-school form: The influence of teacher frame of reference. *Journal of Psychoeducational Assessment*, *19*(4), 299-316. doi:10.1177/073428290101900401

Nota, L., Soresi, S., & Zimmerman, B. J. (2004). Self-regulation and academic achievement and resilience: A longitudinal study. *International Journal of Educational Research*, *41*(3), 198-215. doi:10.1016/j.ijer.2005.07.001

Özcan, Z. C. (2016). The relationship between mathematical problem-solving skills and self-regulated learning through homework behaviours, motivation, and metacognition. *International Journal of Mathematical Education in Science and Technology*, *47*(3), 408-420. doi:10.1080/0020739X.2015.1080313

Pajares, F. (2008). Motivational role of self-efficacy beliefs in self-regulated learning. In B. J. Zimmerman & D. H. Schunk (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 111-140). New York: Erlbaum.

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology,* 8, 422. doi: 10.3389/fpsyg.2017.00422

Panadero, E., Jarvela, S., & Klug, J. (2016). Third wave of measurement in the self-regulated learning field: When measurement and intervention come hand in hand. *Scandinavian Journal of Educational Research*, *60*(6), 723-735. doi:10.1080/00313831.2015.1066436

Pape, S. J., Bell, C. V., & Yetkin, I. (2003). Developing mathematical thinking and self-regulated learning: A teaching experiment in a seventh-grade mathematics classroom. *Educational Studies in Mathematics*, *53*, 179-202. doi:10.1023/A:10260621218.

Patrick, H., & Middleton, M. J. (2002). Turning the kaleidoscope: What we see when self-regulated learning is viewed with a qualitative lens. *Educational Psychologist, 37,* 27–39.

Peeters, J., De Backer, F., Kindekens, A., Triquet, K., & Lombaerts, K. (2016). Teacher differences in promoting students' self-regulated learning: Exploring the role of student characteristics. *Learning and Individual Differences, 52*, 88-96. http://dx.doi.org/10.1016/j.lindif.2016.10.014

Perels, F., Dignath, C. & Schmitz, B. (2009). Is it possible to improve mathematical achievement by means of self-regulation strategies? Evaluation of an intervention in regular math classes. *European Journal of Psychology of Education, 24*(1), 17-32. http://dx.doi.org/10.1007/BF03173472

Perry, N. E. (2002). Introduction: Using qualitative methods to enrich understandings of self-regulated learning. *Educational Psychologist, 37*(1), 1-3.

Perry, N. E., Hutchinson, L., & Thauberger, C. (2007). Mentoring student teachers to design and implement literacy tasks that support self-regulated reading and writing. *Reading and Writing Quarterly, 23*(1), 27-50.

Perry, N. E., & Meisels, S. J. (1996). *How accurate are teacher judgments of students' academic performance?* Chicago, IL: National Opinion Research Center.

Perry, N. E., & Rahim, A. (2011). Studying self-regulated learning in classrooms. In In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 122-136). New York: Routledge.

Perry, N. E., & VandeKamp, K. O. (2000). Creating classroom contexts that support young children's development of self-regulated learning. *International Journal of Educational Research, 33*, 821–843.

Perry, N. E., VandeKamp, O., Mercer, L. K., & Nordby, C. J. (2002). Investigating teacher-student interactions that foster self-regulated learning. *Educational Psychologist, 37*, 5-15.

Phares, V., Compas, B. E., & Howell, D. C. (1989). Perspectives on child behavior problems: comparisons of children's self-reports with parent and teacher reports. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 1*(1), 68-71. http://dx.doi.org/10.1037/1040-3590.1.1.68

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 452-502). San Diego, California: Academic.

Pintrich, P. R., Smith, D. A., Garcia, T., & McKeachie, W. J. (1993). Reliability and predictive validity of the Motivated Strategies for Learning Questionnaire (MSLQ). *Educational and Psychological Measurement, 53*(3), 801-813.

Pressley, M. (1982). Elaboration and memory development. *Child Development, 53*, 296-309.

Pressley, M., & Harris, K. R. (2008). Cognitive strategy instruction: From basic research to classroom application. *Journal of Education, 189*, 77–94.

Puustinen, M., & Pulkkinen, L. (2001). Models of Self-regulated Learning: A Review. *Scandinavian Journal of Educational Research, 45*(3), 269-286. doi:10.1080/00313830120074206

Ruban, L., & Reis, S. M. (2006). Patterns of self-regulatory strategy use among low-achieving and high-achieving university students. *Roeper Review, 28*(3), 148-156. doi:10.1080/02783190609554354

Rudolph, K. D., Lambert, S. F., Clark, A. G., & Kurlakowsky, K. D. (2001). Negotiating the transition to middle school: The role of self-regulatory processes. *Child Development*, *72*(3), 929-946.

Ryan, A. M., Patrick, H., & Shim, S. O. (2005). Differential profiles of students identified by their teacher as having avoidant, appropriate, or dependent help-seeking tendencies in the classroom. *Journal of Educational Psychology, 97(*2), 275-285.

Ryan, A. M., & Pintrich, P. R. (1997). "Should I ask for help?" The role of motivation and attitudes in adolescents' help seeking in math class. *Journal of Educational Psychology, 89*, 329-341.

Ryan, A. M., Pintrich, P. R., & Midgley, C. (2001). Avoiding help seeking in the classroom: Who and why? *Educational Psychology Review, 13*, 93-114.

Sattler, J. M. (2008). *Assessment of children: Cognitive foundations.* La Mesa, CA: Jerome M. Sattler Publisher.

Schlomer, G. L., Bauman, S., & Card, N. A. (2010). Best practices for missing data management in counseling psychology. *Journal of Counseling Psychology, 57*(1), 1-10.

Schmitz, B., Klug, J., & Schmidt, M. (2011). Assessing self-regulated learning using diary measures with university students. In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 251-266). New York: Routledge.

Schunk, D. H. (2001). Social cognitive theory and self-regulated learning. In B. J. Zimmerman & D. H. Schunk (Eds.), *Self-regulated learning and academic achievement: Theoretical perspectives* (pp. 125–152). Mahwah, NJ: Erlbaum.

Schunk, D. H. (2008). Attributions as motivators of self-regulated learning. In B. J. Zimmerman & D. H. Schunk (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 245-266). New York: Erlbaum.

Schunk, D. H., & Ertmer, P.A. (2000). Self-regulation and academic learning: Self-efficacy enhancing interventions. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 631-650). San Diego, CA: Academic Press.

Schunk, D. H., & Usher, E. L. (2013). Barry J. Zimmerman's theory of self-regulated learning. In H. Bembenutty, T. J. Cleary, & A. Kitsantas (Eds.), *Applications of self-regulated learning across diverse disciplines: A tribute to Barry J. Zimmerman* (pp. 1-28). Charlotte, NC: Information Age.

Sitzmann, T., & Ely, K. (2011). A meta-analysis of self-regulated learning in work-related training and educational attainment: What we know and where we need to go. *Psychological Bulletin*, *137*(3), 421-442. doi:10.1037/a0022777

Stolarova, M., Wolf, C., Rinker, T., & Brielmann, A. (2014). How to assess and compare inter-rater reliability, agreement and correlation of ratings: An exemplary analysis of mother-father and parent-teacher expressive vocabulary rating pairs. *Frontiers in Psychology, 5*, 509. http://doi.org/10.3389/fpsyg.2014.00509

Thomas, J. S. (2017). Interrater reliability/agreement. In S. G. Rogelberg (Ed.), The SAGE encyclopedia of industrial and organizational psychology (2nd ed.) (pp. 748-750). Thousand Oaks, CA: SAGE Publications.

Tonks, S. M., & Taboada, A. (2011). Developing self-regulated readers through instruction for reading engagement. In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 173-186). New York: Routledge.

Vansteenkiske, M., Sierens, E., Goosens, L., Soenens, B., Dochy, F., Mouratidis, A., … Beyers, W. (2012). Identifying configurations of perceived teacher autonomy support and structure: Associations with self-regulated learning, motivation and problem behavior. *Learning and Instruction, 22*, 431-439.

Veenman, M. V. J. (2011). Alternative assessment of strategy use with self-report instruments: A discussion. *Metacognition and Learning, 6*(2), 205-211. doi:10.1007/s11409-011-9080-x

Weinstein, C. E., & Acee, T. W. (2013). Helping college students become more strategic and self-regulated learners. In H. Bembenutty, T. J. Cleary, & A. Kitsantas (Eds.), *Applications of self-regulated learning across diverse disciplines: A tribute to Barry J. Zimmerman* (pp. 197-236). Charlotte, NC: Information Age.

Weinstein, C. E., Husman, J., & Dierking, D. R. (2000). Self-regulation interventions with a focus on learning strategies. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 728–744). San Diego, California: Academic.

Weinstein, C. E., Palmer, D. R., & Schulte, A. C. (1987). *Learning and study strategies inventory*. Clearwater, FL: H & H Publishing.

Whitcomb, S. A., & Merrell, K. W. (2013). *Behavioral, social, and emotional assessment of children and adolescents* (4th ed.). New York: Routledge/Taylor & Francis Group.

Wigfield, A., Klauda, S.L., & Cambria, J. (2011). Influences on the development of academic self-regulatory processes. In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 33-48). New York: Routledge.

Winne, P. H. (2010). Improving measurements of self-regulated learning. *Educational Psychologist, 45*(4), 267-276. doi:10.1080/00461520.2010.51715.

Winne, P. H. (2011). A cognitive and metacognitive analysis of self-regulated learning. In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 15-32). New York: Routledge.

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker & J. Dunlosky (Eds.), *Metacognition in Educational Theory and Practice.* Mahwah, NJ: Erlbaum.

Winne, P. H., & Jamieson-Noel, D. (2002). Exploring students' calibration of self reports about study tactics and achievement. *Contemporary Educational Psychology, 27*(4), 551–572. http://dx.doi.org/10.1016/S0361-476X(02)00006-1

Winne, P. H., Nesbit, J. C., Kumar, V., Hadwin, A. F., Lajoie, S. P., & Azevedo, R. A., & Perry, N. E. (2006). Supporting self-regulated learning with gStudy software: The Learning Kit project. *Technology, Instruction, Cognition and Learning, 3*(1), 105–113.

Winne, P. H., & Perry, N. E. (2000). Measuring self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 531-566). San Diego, CA: Academic Press.

Wolters, C. A., & Rosenthal, H. (2000). The relation between students' motivational beliefs and their use of motivational regulation strategies. *International Journal of Educational Research, 33*, 801–820.

Zimmerman, B. J. (1986). Becoming a self-regulated learner: Which are the key subprocesses? *Contemporary Educational Psychology, 11*(4), 307–313.

Zimmerman, B.J. (1989). A social cognitive view of self-regulated academic learning. *Journal of Educational Psychology*, 81, 329-339.

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist, 25*, 3-17.

Zimmerman, B. J. (1998). Developing self-fulfilling cycles of academic regulation: An analysis of exemplary instructional models. In D. H. Schunk, & B. J. Zimmerman (Eds.) *Self-regulated learning: From teaching to self-reflective practice.* New York: Guilford Publications.

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–35). San Diego, California: Academic.

Zimmerman, B. J. (2001). Theories of self-regulated learning and academic achievement: An overview and analysis. In B. J. Zimmerman, & D. H. Schunk (Eds.), *Self-regulated learning and academic achievement: Theoretical perspectives* (2nd ed.) (pp. 1–38). Mahwah, NJ: Erlbaum.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice*, *41*(2), 64-70.

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal*, *45*(1), 166-183.

Zimmerman, B. J. (2011). Motivational sources and outcomes of self-regulated learning and performance. In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 49-64). New York: Routledge.

Zimmerman, B. J., & Kitsantas, A. (2014). Comparing students' self-discipline and self-regulation measures and their prediction of academic achievement. *Contemporary Educational Psychology, 39*(2), 145-155. doi:10.1016/j.cedpsych.2014.03.004

Zimmerman, B. J. (2013). From cognitive modeling to self-regulation: A social cognitive career path. *Educational Psychologist, 48*(3), 135-147. doi: 10.1080/00461520.2013.794676

Zimmerman, B. J., & Martinez-Pons, M. (1986). Development of a structured interview for assessing student use of self-regulated learning strategies. *American Educational Research Journal, 23*, 614-628.

Zimmerman, B. J., & Martinez-Pons, M. (1988). Construct validation of a strategy model of student self-regulated learning. *Journal of Educational Psychology, 80*(3), 284–290.

Zimmerman, B. J., & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology, 82*, 51-59.

Zimmerman, B. J., & Schunk, D. H. (Eds.). (1989). *Self-regulated learning and academic achievement: Theory, research, and practice.* New York: Springer.

Zimmerman, B. J., & Schunk, D. H. (Eds.). (2011). *Handbook of self-regulation of learning and performance*. New York: Routledge.

Zusho, A. (2017). Toward an integrated model of student learning in the college classroom. *Educational Psychology Review, 29*(2), 301-324. doi: 10.1007/s10648-017-9408-4

Appendix A: Literature Review

The current study examined the reliability and validity of the SRSI-TRS, a rating scale designed to measure teacher perceptions of students' SRL in the classroom context (Callan & Cleary, 2014). The purpose of the literature review is to provide background information on relevant constructs, including self-regulated learning (SRL), SRL assessment, and teacher ratings of student behavior. The discussion of the literature is intended to support the rationale for the need to develop psychometrically strong teacher rating scales to measure student SRL. The existing research, particularly studies including the SRSI-TRS, is elaborated in order to highlight gaps in the literature that are addressed by the current study.

Theoretical Overview of SRL

Historical background. Research on academic SRL emerged in the 1970s and 1980s as a new approach to explain academic achievement (Andrzejewski, Davis, Bruening, & Poirier, 2016; Zimmerman & Schunk, 2011). SRL theories shifted focus from how students' learning is influenced by their innate abilities, environments, and the quality of teaching they receive to student behaviors and skills (Zimmerman, 1986). Within an SRL framework, students are active contributors to the learning process, and they have control over the attainment of their goals (Schunk, 2001; Winne, 2010). Zimmerman (2001) states that, "Neither a mental ability nor an academic performance skill, self-regulation refers instead to the self-directive *process* through which learners transform their mental abilities into task-related academic skills" (p. 1).

Early research related to SRL focused on metacognitive and cognitive strategy use, motivational processes, and behavioral control (Zimmerman & Schunk, 2011). For example, researchers examined how using strategies like elaboration would impact students' memory (Pressley, 1982), or how training impulsive children to use self-talk could enhance their self-control (Meichenbaum & Goodman, 1971). From studies like these emerged a realization that implementing strategies successfully required other factors, such as metacognition and motivation. The different strands of research were eventually integrated into an inclusive construct of SRL, which encompasses metacognition, strategy use, and motivational processes (Zimmerman, 2008). Researchers drew from many paradigms, including behaviorism, constructivism, and social cognitive theory, to form a theoretical understanding of how students self-regulate their learning (Boekaerts et al., 2000).

There are several theories of SRL that have given rise to different models and frameworks (e.g., Boekaerts & Niemvirta, 2000; Pintrich, 2000; Winne & Hadwin, 1998; Zimmerman, 2000). While the theories vary with regard to the specific processes, strategies, or behaviors that they highlight, they do share several common features and assumptions. All theories tend to emphasize SRL as involving self-initiated thoughts, feelings, and behaviors that students use to attain their learning goals (Puustinen & Pulkkinen, 2001). The different theories also stress the importance of students being aware that they are using processes for the purpose of improving their academic performance (Zimmerman, 2001). Many of the models characterize SRL as a three-phase process that includes a preparatory/planning phase, a performance phase, and a reflection phase (Alexander, Dinsmore, Parkinson, & Winters; 2011; Puustinen & Pulkkinen, 2001). Lastly, most theories of SRL include a motivational component that explains why students choose to regulate in certain contexts under certain conditions (Karabenick & Zusho, 2015; Panadero, 2017; Zimmerman, 2001).

Contemporary research has emphasized other important themes underlying the theoretical construct of SRL, such as contextual factors, contingencies, and dynamic relations (Ben-Eliyahu & Bernacki, 2015). Contextual factors refer to how the environment interacts with the learner and influences their efforts to self-regulate (Winne, 2010), and is a significant variable affecting SRL (Panadero, 2017). Contingencies capture the cyclical nature of SRL; a learner's regulatory behavior is based on what happened previously or during preceding phases. Dynamic relations underscores the idea that processes (i.e., goal setting, strategy use, motivation, affect, beliefs) occurring within an SRL cycle mutually influence each other (Karabenick & Zusho, 2015).

Zimmerman's social-cognitive model of SRL. The current study is grounded in Zimmerman's (2000) social-cognitive model of SRL. Social cognitive theory emphasizes the framework of triadic reciprocality, in which an individual's personal factors, their behaviors, and the environment all interact reciprocally to influence and determine future behavior (Bandura, 1986). SRL occurs when a student exerts control over their personal characteristics, behaviors, and the environment in order to achieve specific learning goals. The outcomes of the student's behavior as well as environmental factors in turn influence the student's personal characteristics, thus demonstrating the notion of triadic reciprocality (Schunk & Usher, 2013; Zimmerman, 1989).

An early social-cognitive model of self-regulation incorporated three subprocesses: self-observation, self-judgement, and self-reaction (Bandura, 1986). Self-observation refers to students' deliberate monitoring of their performance and behavior during learning. Self-judgement occurs when students use their observations to compare their performance against goals and standards. Students' self-judgements lead to positive or negative self-reactions that then influence future behaviors (Bandura, 1991; Schunk & Usher, 2013).

Zimmerman expanded on this early theory of self-regulation by elaborating a model that includes forethought, performance, and reflection phases. These phases correspond to regulatory processes that occur before, during, and after task engagement (Schunk & Usher, 2013;

Zimmerman, 2000). Zimmerman's model depicts self-regulation as a cyclical process that entails coordinated attempts by a learner to manage and control personal, environmental, and behavioral components as they purse goals (Zimmerman, 1998). All three phases of Zimmerman's model are important components of the learning process, and can be points of instruction for teachers seeking to enhance students' regulatory skills (Dorrenbacher & Perels, 2016). Indeed, many SRL intervention programs use this model as a framework (Cleary et al., 2017; Perels, Dignath, & Schmitz, 2009).

The first phase, forethought, includes task analysis processes of *goal setting* and *strategic planning*. Before beginning a task, learners set goals for specific outcomes, and they plan the strategies they will use to help them achieve those goals. Possessing task analysis skills is not enough; a learner also needs to be motivated to use them. Self-motivation beliefs, such as self-efficacy, outcome expectations, intrinsic interest, and goal orientation play an important role in the forethought phase (Zimmerman, 2002). For example, students with a greater sense of self-efficacy who perceive themselves to be more capable at a task will set challenging goals and be firmly committed to achieve those goals (Pajares, 2008).

The performance/volitional control phase includes two main processes, *self-control* and *self-observation*. Self-control processes, including self-instruction, focusing, and applying strategies, are key processes to keep a learner engaged in the task, while self-observation is the process of monitoring different aspects of one's performance while learning (Schunk & Usher, 2013; Zimmerman, 2002). Motivation beliefs are important to the performance/volitional control phase as well. For example, students with a high level of self-efficacy use more cognitive and metacognitive strategies while learning and are more likely to persist in the face of difficulty (Pajares, 2008). Performance control strategies also impact motivation beliefs; for example, students' self-monitoring and perceptions of progress can cause an increase in their self-efficacy (Schunk & Ertmer, 2000).

The final phase, self-reflection, encompasses two processes, *self-judgment* and *self-reactions*. Self-judgement is the evaluation of one's performance and subsequent attribution of the outcomes to perceived causes. Learners use information from self-observations to judge their performance based on different types of criteria, and then make causal attributions about the outcomes; specifically, whether the quality of their performance is due to inherent ability or effort (Cleary & Labuhn, 2013). Self-judgement is linked to self-reactions, such as perceptions of satisfaction or dissatisfaction and the associated affect. Adaptive or defensive inferences (i.e.; decisions about how to adapt future learning efforts) also result from students' self-reactions (Zimmerman, 2013). Again, motivation beliefs both influence and are influenced by reflection phase processes. For example, students with low self-efficacy may attribute success to uncontrollable factors such as luck, and then be unmotivated to continue putting forth effort. Attribution of success to use of learning strategies will enhance a learner's motivation and self-efficacy (Schunk, 2008).

Strategy use in SRL. Within Zimmerman's model of SRL, a key characteristic of self-regulated learners is their use of strategies before, during, and after task engagement. Strategies are processes or operations that facilitate learning and performance (Malmberg, Jarvenoja, & Jarvela, 2010; Weinstein & Acee, 2013). While there are different frameworks for classifying strategies, researchers agree that SRL encompasses metacognitive and cognitive strategies as well as motivation and volitional strategies (Boekaerts & Cascallar, 2006; Wolters & Rosenthal, 2000; Weinstein, Husman, & Dierking, 2000). Cognitive strategies include organization, elaboration, and rehearsal of information and task-specific tactics such as drawing a picture to solve a mathematics problem (Weinstein & Acee, 2013; Dent & Koenka, 2016). Metacognitive strategies include monitoring comprehension and self-evaluating (Broadbent, 2017; Pintrich, Smith, Garcia & McKeachie, 1991). Examples of motivational strategies are self-talk and self-consequences, while volitional strategies include paying attention in the classroom and refocusing when distracted while working (Boekaerts & Cascallar, 2006, Corno, 2001).

Beyond knowledge of strategies, SRL requires thoughtful application of the strategies at appropriate times and in relevant contexts. Ineffective self-regulators often do not use strategies in a proactive manner. Rather, they use them reactively; for example, skipping meals when they see that they weigh too much instead of initiating a healthy diet (Cleary, in press; Zimmerman, 2001). Weinstein and Acee (2013) differentiated three forms of strategy knowledge. Learners need to know about a variety of strategies (declarative knowledge), they need to know how to use those strategies effectively and efficiently (procedural knowledge), and they needs to know in which contexts and situations particular strategies would be useful (conditional knowledge.) Strategy knowledge and self-regulation go hand in hand; the knowledge informs strategic planning and strategy use, while self-monitoring and reflection generate new and updated knowledge about strategy use and effectiveness (Weintein & Acee, 2013). Thus, a key to effective strategy use can be captured by Zimmerman's (2000) quote, "No self-regulatory strategy will work equally well for all persons, and few if any, strategies will work optimally for a person on all tasks or occasions" (p. 17).

Callan, Marchant, Finch, & German (2016) conducted a global study with a sample of 475,460 students from 63 countries to examine the relationship between strategy use, achievement, and demographics (i.e., SES and gender). Data was taken from the Programme for International Student Assessment (PISA, 2009), an achievement test designed to assess students' ability to apply reading, science, and mathematics content to real-life situations. PISA includes metacognitive and learning strategies indexes, on which students are given scenarios and asked to rate the quality and usefulness of strategies to reach a certain learning goal. Students' ratings are compared to a set of ratings developed by experts. Callan and colleagues found that students' scores on the learning strategies index, which included memorization, elaboration, and control strategies, was not strongly associated with achievement. The use of metacognitive strategies (e.g., checking for understanding, summarization), on the other hand, demonstrated a strong correlation with achievement across all subject areas ($r = .50$ for reading, $r = .46$ for math, and $r = .48$ for science; all $p < .001$). Metacognitive strategy use was also a significant predictor of achievement after controlling for SES, while learning strategy use was not. The authors caution that this research does not indicate that learning strategies are not important; rather, because learning strategies tend to be more task-specific, their utility might not be captured adequately through PISA, whereas metacognitive strategies that are not as contextually sensitive may relate better to global achievement. This research does provide strong support that metacognitive strategies are important for achievement, and underscores the idea that not all strategies are equally effective in facilitating learning.

Other studies have shown variability in effectiveness of different strategies. Some researchers have differentiated between deep processing strategies, such as organization of information, as opposed to surface processing strategies, such as routine memorization. In a study with 180 undergraduates, Ruban & Reis (2006) found that high achieving students reported using deep strategies such as condensing and reorganizing notes and using mnemonic devices, while low achievers reported surface strategies such as creating flashcards, memorizing information, and reviewing notes. In another study with undergraduate students, Broadbent

(2017) found that strategies of time management and elaboration were significant predictors of achievement for online and blended learning college students. Rehearsal, on the other hand, was predictive of a lower grade. These studies and others point to the key role of strategy use in promoting academic achievement, but also support the importance of metacognitive and motivational factors to implement and adapt strategy use effectively.

Help seeking is an important performance phase strategy that is particularly relevant to classroom contexts (Karabenick, 2011). The SRSI-TRS includes several items pertaining to help-seeking, such as, "The student asks questions about errors he or she makes on tests or assignments." As students enter middle school and high school, the decision to seek help may be difficult because peer acceptance becomes a more prominent concern (Ryan, Pintrich, & Midgley, 2001). Advances in technology have also changed how learners engage in help seeking, because there are many more resources and anonymous sources of help (Karabenick & Berger, 2013). As with other strategies, the way students use help seeking impacts its utility and value for their learning. For instance, students can have different help seeking goals. *Instrumental help seeking* refers to students seeking help in order to develop skills or increase understanding. This type of help seeking provides long-term benefit, because it decreases the need for future help. *Executive help seeking*, on the other hand, is done in order to get answers or assistance to complete a task quickly. While providing a benefit of lessening students' work load, this type of help seeking is maladaptive because it does not reduce the need for continued help (Ryan, Patrich, & Shim, 2005). The research on help seeking supports the notion that how and when students apply strategies is an important factor in the effectiveness of their strategy use (Karabenick & Berger, 2013; Ryan & Pintrich, 1997).

Furthermore, the strategy of help seeking is unique in that it encapsulates Zimmerman's full cyclical model of SRL. Within the help seeking process, students need to determine if there is a problem and whether they need help, and decide whether to seek help, for what reason, and whom to ask for help. These decision points pertain to the forethought phase in Zimmerman's model. Then, students need to solicit and obtain help, which corresponds to the performance phase. Finally, after receiving help, students need to evaluate whether their goals were met and judge how satisfied they are with the help they received, which encompasses the self-reflection phase (Karabenick & Berger, 2013). According to this model, help seeking is a dynamic and cyclical process, and can be viewed as a microcosm of SRL.

In a study on help-seeking behaviors of sixth grade students, Ryan and colleagues (2005) asked teachers to identify students' help-seeking tendencies as avoidant, appropriate, or dependents. They found that teachers and student reports of help-avoidance were well-aligned, suggesting that teachers can accurately observe students' avoidant behavior with regard to help-seeking. They also found that students with an avoidant approach to help-seeking were more likely to have performance goal orientations, while students with appropriate help-seeking behaviors were more likely to have mastery goal orientations. Similarly, Du and colleagues (2016) used multilevel analysis techniques to examine the factors that influence eighth graders' help seeking in the context of math homework using a sample of 796 students from 46 classes. They found that help-seeking was positively related to mastery orientation at both the class and individual levels. The authors suggested that teachers can promote students' help-seeking behaviors (even in the home context) by placing a greater emphasis on fostering a mastery goal orientation to learning.

SRL and Academic Achievement

SRL skills have been shown to improve performance and functioning in a wide range of fields, including sports, business, music, and health (Zimmerman & Schunk, 2001). Most importantly, an extensive literature base documents the prominent role that SRL skills play in academic achievement (Greene, Moos, & Azevedo, 2011; Dent & Koenka, 2016; Zusho, 2017). SRL has been shown to differentiate between high and low achievers (DiFrancesca, Nietfeld, & Cao, 2016; Zimmerman & Martinez-Pons, 1990). In addition, SRL skills have been shown to mediate how individual characteristics and context influence achievement (Dent & Koenka, 2016; Pintrich 2000). Promoting students' SRL skills can ensure equity in education by eliminating impact of student characteristics such as socioeconomic status and ethnicity (Peeters et al., 2016). SRL skills are also important for lifelong learning (Puustinen & Pulkkinen, 2001; Dignath van Ewijk, 2011; Zimmerman, 2002), which is a major goal of education.

SRL skills are important for students in all grades across different academic subjects (Dorrenbacher & Perels, 2016; Hofer et al., 1998; Perry et al., 2002; Zusho, 2017), and are applicable to varied learning contexts, with recent research showing an increased focus on SRL in computer based learning environments (Broadbent, 2017; Greene et al., 2011). The current study focuses on a sample of middle school students, and the measures given were contextualized to the students' mathematics class. SRL skills are particularly important for middle school students, who often experience a decline in motivation and academic performance (Cleary & Chen, 2009; Dembo & Eaton, 2000; Eccles et al., 1993). In middle school settings, students are expected to take more initiative and responsibility for their learning and balance numerous demands and expectations from multiple teachers (Bell & Pape, 2014; Rudolph et al., 2001), thus intensifying the need for strong SRL skills. SRL skills are also particularly important for mathematics, which requires students to use knowledge and skills flexibly in order to solve problems (De Corte, Mason, Depaepe, & Verschaffel, 2011). Research has shown that effective SRL is linked to higher achievement in mathematics (Bell & Pape, 2014; Montague, 2007; Ozcan, 2016).

Some early work on SRL and academic achievement was done by Zimmerman and Martinez-Pons (1986). They developed a structured interview, the Self-Regulated Learning Interview Schedule (SRLIS) for assessing students' use of self-regulation during class, while doing homework, and while studying. The interview asked students to identify the methods and strategies they use in six typical learning situations, such as studying for final exams, completing writing assignments outside of class, and taking tests. They interviewed 80 tenth grade students from a middle class suburban high school, 40 of whom were from a high achieving track and 40 from a low achieving track. Students responses were coded and were given measures for strategy use, strategy frequency, and strategy consistency. The coding framework encompassed 14 strategies derived from SRL research: goal-setting, environmental structuring, self-consequences, self-evaluating, organizing and transforming information, seeking and selecting information, rehearsal and mnemonic strategies, seeking social assistance, and reviewing textbooks, class notes, and previous tests. The high achieving group reported significantly greater use of 13 of the 14 strategies as compared to the low achieving group (only self-evaluation was not significant). The researcher found that 93% of the students could be correctly classified into their achievement group based on the measure of strategy use. The self-regulation score was also predictive of students' standardized achievements scores in math and English.

More recently, Zimmerman and Kitsantas (2014) and Cleary and Kitsantas (2017) used structural equation modeling to explore the role of SRL behaviors in predicting academic performance. Zimmerman and Kitsantas (2014) examined the difference between self-discipline (SD) and self-regulation (SR) in predicting students' GPA and standardized achievement test scores. Both SR and SD were measured using several student self-reports as well as teacher ratings, which were combined to form composite scores for each construct. Using a sample of 507 high school students, they found that SR was as significant predictor of both achievement measures, while SD was not. Cleary and Kitsantas (2017) explored the role of SRL behaviors and self-efficacy beliefs as predictors of mathematics performance within a complex model that included background variables (prior achievement, socioeconomic status) as well as other motivation beliefs (task interest, school connectedness). While prior achievement was the strongest predictor of mathematics grades, both SRL behaviors and self-efficacy emerged as significant predictors as well. In addition, self-efficacy was a significant mediator between prior achievement and SRL behavior and between task interest and SRL behaviors. SRL behaviors were a significant mediator between self-efficacy and mathematics performance and between SES and mathematics performance. It should be noted that both of these studies included teacher ratings as measures of SRL and stressed the importance of a multisource approach to SRL assessment due to limitation of self-reports.

Dent & Koenka (2015) conducted a meta-analysis to explore the relation between SRL and academic achievement. They looked at two components of SRL: metacognitive processes and students' use of cognitive strategies, and how each of these components relates to achievement. Using a total of 118 studies for the meta-analyses, the authors found that both metacognitive and cognitive components had small but significant correlations with academic achievement (r = .20 and r = .11, respectively). The correlation for metacognitive processes (i.e., goal setting, planning, self-monitoring, self-control, and self-evaluation) was significantly stronger. The authors presented a large number of findings across the different processes and strategies examined as well as other moderating variables (academic subject, grade, type of achievement measure, type of academic performance measure). A few notable findings are relevant to this project. Firstly, the association between metacognitive processes and academic performance was significantly stronger after the transition to middle school. Also, studies that used online (i.e., contextualized) measures of metacognition and cognitive strategies had a stronger correlation with academic achievement than studies that used offline (i.e., decontextualized) measures ($r = .30$ for cognitive and $r = .39$ for metacognitive for online measures; $r = .10$ for cognitive and $r = .15$ for metacognitive for offline measures). A meta-analysis is descriptive and does not give causal information, and Dent and Koenka's study did not control for other factors and also did not consider the full scope of SRL. However, the results they present provide a compelling overview of the wide literature base documenting the relation between cognitive strategy use, metacognition, and academic achievement.

The importance of SRL for academic achievement pushes the question of how to teach SRL skills to the forefront. SRL is inextricably intertwined with social and contextual factors, and therefore teachers can create classroom environments that promote SRL (Perry & Rahim, 2011; Schunk, 2001). Perry and colleagues focused on both identifying classroom characteristics that foster SRL and training teachers to incorporate these factors into their classrooms (Perry et al., 2002; Perry, Hutchinson, & Thauberger, 2008; Perry & VandeKamp, 2000). They found that students most often use SRL when they have opportunities to engage in complex activities, exercise choice over learning processes, control the difficulty level of their tasks, and are involved in evaluating and reflecting on their learning. Vansteenkiske et al. (2012) found that students who perceived high levels of autonomy support and clear expectations from their teachers used more regulatory strategies. Boekaerts and colleagues found that characteristics such as clarity and pace of instruction, structure and autonomy, and teacher expectations about students' capacity had an effect on the way students self-regulated their learning (Boekaerts & Cascallar, 2006; Boekaerts, de Kroning, & Vedder, 2006). In addition, teachers can infuse core SRL processes (e.g., metacognition, reflection) into their classroom practices in order to promote students' regulatory behaviors and skills (Moos & Ringdal, 2013).

In addition to studying classroom characteristics that promote or foster SRL, over the past couple of decades researchers have focused on developing interventions to promote students' SRL (Dignath & Buttner, 2008; Sitzmann & Ely, 2011). Specific programs that enhance students' regulatory skills and academic performance have been developed and applied in a variety of settings and subjects (Butler et al., 2005; Cleary et al., 2017; Graham & Harris, 2009). Studies have shown that interventions that combine strategy and motivation training and are situated within a specific content area the most effective (Boekaerts & Cascallar, 2006; Cleary et al., 2017). In a meta-analysis, Dignath and Buttner (2008) found that SRL interventions were effective at improving academic performance for both primary and secondary school students, with some differences between the types of interventions that were most effective for each group.

SRL interventions may be particularly beneficial for closing the achievement gap faced by minority or disadvantaged students (Andrzejewski et al., 2016; Cleary et al., 2017). For example, Andrzejewski and colleagues (2016) developed an intervention to promote the use of goal setting, monitoring, and reflection skills for ninth grade students in Earth Science class. While they found that the intervention had no significant overall effects on students' achievement, when the data was disaggregated based on race, minority students benefitted differentially and exhibited significantly higher achievement than minority students who had not received the intervention. Similarly, Cleary et al. (2017) used the Self-Regulation Empowerment Program (SREP; Cleary & Zimmerman, 2004) as an intervention for academically at-risk middle school students. Students in the intervention group showed a more positive trend in mathematics achievement scores as compared to a control group, and showed significantly higher SRL strategy use on microanalytic and structured interview measures at post-test.

Assessment of SRL

The discussion thus far has explored the theoretical basis of SRL, its importance for academic achievement, and how SRL principles and instruction can be applied within school settings. However, as noted by Boekaerts and colleagues (2000), "It is evident that a sine qua non for the development of a sound knowledge base for furthering theory and applications in this area is the use of reliable and valid measures" (p. 757). Thus, SRL assessment occupies an integral place in both basic and applied research, and examining the psychometric properties of SRL measures remains an important focus (Panadero, Klug, & Jarvela, 2016).

The role of assessment for SRL is multifaceted. First and foremost, measurement is used to inform theory and explore relationships between various components of SRL (Boekaerts et al., 2000); thus, the measures used by researchers have mirrored their conceptualizations of SRL (Ben-Eliyahu & Bernacki, 2015). Assessment is also important for identifying areas of deficit in order to apply interventions, and is used to evaluate the effectiveness of those interventions (Boekaerts & Corno, 2005; Greene et al., 2011). Teachers likewise need accurate knowledge of students' regulatory skills in order to foster and promote SRL in their classrooms (Peeters et al., 2016). Additionally, sometimes measures themselves can be used an interventions; for example, learning diaries are both an SRL measure and an intervention to promote students' self-monitoring and strategy use (Panadero et al., 2016).

Historically, much SRL research relied heavily on self-report measures (Dinsmore et al., 2008, Winne & Perry, 2000). In recent years researchers have begun to shift their attention to other measures, with a particular focus on fine-grained, contextualized measures that capture the process of SRL. Winne and Perry (2000) classified SRL measures into two broad categories: event measures and aptitude measures. Both categories encompass a wide variety of measures, and many have been shown to predict academic achievement.

Event measures. Event measures are based on the conceptualization of SRL as a "dynamic series of cognitive, metacognitive, motivational, and behaviors events that students consciously enact, monitor, and control over the course of a learning task" (Greene et al., 2011; p. 313). These types of measures seek to assess fine-grained behaviors or processes as they are occurring in real time in authentic contexts (Cleary, 2011; Schmitz, Klug, and Schmidt, 2011). Event measures emphasize the fact that SRL is contextual; that is, the regulatory behavior and processes used by a learner will be shaped by the context (i.e., task demands, presence of other people) in which the learning task is occurring (Winne, 2010). Some examples of events measures are traces, think-aloud protocols, and microanalysis.

Traces (or logfiles) are time-stamped data about tactics and strategies that students use while learning or studying, such as annotating or highlighting sections of a text. These data are valuable because they do not rely on students' perceptions and can be used for examining calibration between students' self-reports and what they actually do while learning (Hadwin et al., 2007). Measuring SRL using trace data is particularly suitable for computer-based learning environments (Winne, 2010). Another method, think-aloud protocols, involves asking students to verbalize their thinking as they are engaged in a learning task. Think-aloud protocols have the advantage of capturing regulatory processes as they are happening without relying on learners' memories. In addition, they are open ended so that students can fully express the processes they use (Greene et al., 2011). Microanalysis, on the other hand, uses structured probes to assess a variety of processes within the cyclical phases of SRL. Microanalytic protocols typically use an open ended format to ask learners about strategic processes before, during, and after a task, as well as metric or quantitative questions about motivation beliefs, affect, and self-evaluation (Cleary & Callan, 2017).

Data from event measures are important because they give information about what learners do at specific moments in time and the particular contexts in which the behaviors happen (Greene & Azevedo, 2010). However, there are disadvantages to these methods. Some of the methods, such as think-aloud protocols, can be intrusive to learners and influence processing (Greene & Azevedo, 2010) and may not fully capture the full range of regulatory processes if students don't verbalize all of their thoughts (Veenman, 2011). Traces only capture overt behavior and give no information about the motivation or intent behind them (Veenman, 2011). Microanalysis and think-aloud protocols are also both forms of self-report, and may thus be susceptible to biases (Winne, 2010). Another disadvantage of event measures is that many of them are time intensive and laborious to administer, score, and analyze (Hadwin et al., 2007; Winne & Perry, 2000; Veenman, 2011). The limitations notwithstanding, event measures can be particularly helpful for identifying maladaptive SRL beliefs or processes that can be targeted for intervention (Cleary, 2011; Zimmerman, 2008; DiBendetto & Zimmerman, 2013).

Aptitude measures. Aptitudes represent stable traits that are considered to be relatively enduring over time and across contexts (Winne & Perry, 2000). Aptitude measures of SRL assess the frequency or likelihood with which students use regulatory strategies and behaviors in a typical situation. Students are often asked to respond to items that depict particular behaviors, perceptions, or learning events and situations using a Likert type scale. In many cases, these assessment tools require students to estimate the frequency with which they use specific processes or strategies or their capability in using them (Winne, 2010). Responses are then aggregated to give an overall score (McCardle & Hadwin, 2015). While aptitude measures do not capture the contextualized aspect of SRL as it applies to specific learning tasks, they are useful for evaluating students' overall use of SRL (Cleary et al., 2015). Aptitude measures include structured interviews, self-report questionnaires, and teacher and parent ratings scales. Because the focus of this study is aptitude measures, they are explored in greater detail below.

Structured interviews. Structured interviews include a script that queries students to describe their use of SRL based on memories of what their typical behavior would look like in certain learning situations, such as preparing for a big exam or completing a difficult homework assignment (Winne & Perry, 2000; Zimmerman & Martinez-Pons, 1986). Like self-report questionnaires, structured interviews rely on students' report of their strategy use. Unlike questionnaires, however, structured interviews follow a free response style to avoid suggesting certain strategies to students, and they are contextualized because they ask about specific, albeit fictitious, learning situations (Zimmerman & Martinez-Pons, 1986). Student responses are typically scored by identifying categories of SRL that students mentioned and assigning qualitative values based on the coding scheme used (Winne & Perry, 2000). Zimmerman & Martinez-Pons (1986) developed a structured interview, the SRLIS, to assess students' use of SRL, particularly in non-classroom scenarios (e.g., studying at home.) It was shown to be predictive of students' achievement and to differentiate between high and low achievers

(Zimmerman & Martinez-Pons, 1986). Cleary and colleagues (2017) administered a hypothetical test preparation scenario, which was a variation of one of the SRLIS questions. They found that students who received an SRL intervention were able to generate a more comprehensive strategic plan for test preparation than students in a control group.

Self-report questionnaires. Self-report questionnaires are widely used to measure SRL, and they are helpful for assessing students' memories and interpretations of their behaviors as well as their cognitive, metacognitive, or motivational processes that cannot be observed (Patrick & Middleton, 2002; Zusho, 2017). Questionnaires are easy to design, administer, and score, and they often boast strong psychometric properties (Winne & Perry, 2000). They are also feasible to administer to large samples (Callan et al., 2016; Zusho, 2017). Historically, SRL research has relied heavily on self-report questionnaires (Matthews, Schwean, Campbell, Saklofske, & Mohamed, 2000). In the Handbook of Self-Regulation (2000), Boekaerts et al. emphasized the importance of using performance and observational measures because of the limitations of self-reports. More recently, Dinsmore et al. (2008) found that 59% of studies on SRL between 2003 and 2007 used self-report questionnaires. The limitations of questionnaires are well documented in the literature, and researchers have shown that students' reports of their behavior do not always calibrate with traces or other objective measures (Winne & Jamieson-Noel, 2002). Further, self-reports are susceptible to response bias and memory distortions (Veenman, 2011; Winne & Jamieson-Noel, 2002). If students are not primed properly, they may not recall the full range of strategies they use in different contexts (Karabenick & Zusho, 2015); on the other hand, retrospective questions and prompts may cause students to report strategies that they didn't use due to social desirability effects (Veenman, 2011).

Ratings scales. Parent and teacher ratings of student SRL represent another form of aptitude measures. Parent ratings scales assess students' use of SRL at home while studying and doing homework across different tasks, information that would be difficult to obtain otherwise (Cleary et al., 2015). Teacher rating scales provide information about students' use of SRL in the classroom (Cleary & Callan, 2014; Zimmerman & Martinez-Pons, 1988). Rating scales share some advantages with self-report questionnaires; namely, they are easy to design, administer and score and have strong psychometric properties. Like rating scales, they can also be susceptible to response biases and subjective interpretation of items (Merrell, 2008).

The use of rating scales to measure SRL is particularly important because most other measures, both event and aptitude (e.g., questionnaires, think-alouds, microanalysis, diaries) rely primarily on the student as the source of data. As was noted above, self-report measures have come under criticism in the literature because they have been shown not to calibrate with other, more objective measures (Veenman, 2011). McCardle and Hadwin (2015) defend the use of self-report measures, stating that learners' perceptions are central to SRL because learners regulate themselves based on their self-monitoring and self-judgement, whether it is accurate or not. Using a sample of 263 undergraduate students, they studied alignment across two self-report measures (a questionnaire and diary), and found that while some students had high overlap between the qualitative data from the diary and the quantitative scores on the self-report measure, other students had little overlap. The authors conclude that while self-report measures are valuable, there is a need for complementary methods as well. Teacher and parent ratings scales are a good method for corroborating student self-report data gathered from other measures (Loeber et al., 1989), and are a cost-effective and efficient way to do so.

Convergence between SRL measures. The literature on SRL measures provides evidence of the strengths and limitations of each method. Many early studies on SRL used a single measure (i.e., either aptitude or event), but recent focus has been on a multisource approach to collect information about motivation, strategy use, cognition, metacognition, and affective components of SRL (Ben-Eliyahu & Bernacki, 2015, Butler, 2011). SRL is a multifaceted construct, and as is the case in assessment of many domains, it cannot be fully captured by one method or measure (Karabenick & Zusho, 2015; Meyer et al., 2001).

Along with increased interest in multisource, multimethod approaches has come a focus on triangulation across measures (Winne & Perry, 2000; Greene et al., 2011). Butler (2011) states that different measures may provide “unique, overlapping, or complementary information” (p. 351.). For example, surveys and think-aloud protocols can complement trace data by giving more information about student’s interpretation and memories of the events (Winne, 2010). Self-reports may be inaccurate measures of actual behavior, but they are good for assessing learner’s knowledge, beliefs, and theories (McCardle & Hadwin, 2015). Likewise, observations are not so useful for assessing covert processes but are good for measuring actual behavior. Because different measures yield different perspectives of SRL, it is important to triangulate data in order to understand how data from different measures represents different aspects of SRL (Winne & Perry, 2000). The issue of convergence has received some attention in the literature, with emphasis on comparing event and aptitude measures.

To date, studies on convergence between measures have yielded mixed results (Zusho, 2017). DiBenedetto and Zimmerman (2013) examined the convergence between microanalysis and the RSSRL, a teacher rating scale. They found that teacher ratings had medium, significant correlations with students’ responses to microanalytic questions for performance phase metacognitive monitoring and self-reflection, but not for strategic planning or task strategies. Cleary and Callan (2014) found that the SRSI-TRS had medium, significant negative correlations with two student self-reports measuring maladaptive regulatory behavior. Zimmerman and Kitsantas (2014) found a significant correlation of .43 between the MLSQ (a self-report measure) and the RSSRL. Cleary, Callan, Malatesta, and Adams (2015) found nonsignificant correlations between the MLSQ and microanalytic measures.

In a recent study, Callan and Cleary (2017) examined convergent validity of four different SRL measures. The authors collected data from two aptitude measures (SRSI-SR and SRSI-TRS) and two event measures (micronanalysis and traces). The microanalysis was administered during a mathematics problem solving practice session, and trace data was taken from the same task. The microanalysis was given for both an easy and a difficult mathematics problem in order to examine the convergence between measures across task difficulty. The sample for the study included 100 eighth grade students from an urban school district. The authors found that the two aptitude measures showed significant, medium correlations (r = .30), as did the two event measures (r = .43 for easy mathematics problems, r = .35 for difficult mathematics problems). The correlations between measurement classes were not significant, regardless of task difficulty. The authors noted that even within measurement class, there is a fair amount of unshared variance between the different types of measures. Thus, different measures may be capturing different aspects of students' SRL.

The studies reviewed emphasize the disparate findings of convergence across methods, both within and between groups of measurements. Clearly, the level of convergence varies based on the method, the measures used, and other factors, such as the task or sample. This continues to be an area of interest in the research, and as understanding of SRL expands and evolves, measurement will continue to be used to consider different facets of the construct and how they relate to or complement each other.

Teacher Ratings of Student SRL

An extensive literature base documents the value of teacher ratings for assessing students' academic, social, emotional, and behavioral functioning (Achenbach et al., 1987; Gerber & Semmel, 1984; Frick et al., 2010; Merrell, 2008). Gerber and Semmel (1984) argued in support of "teacher-as-tests," stating that teachers observe thousands of behaviors over the course of the days and months that they spend with students in the classroom, and therefore have a rich sample upon which to base their judgements. Data from teachers can also be used to corroborate information gathered from student self-reports (Achenbach et al., 1987; Loeber et al., 1989).

With regard to SRL assessment, an additional point in favor of using teacher ratings is that they are well aligned with school psychology assessment practices. Best practice in school psychology assessment mandates a multimethod, multisource approach, and teachers are an important group of informants who can provide information about students' functioning in the classroom (Sattler, 2007; Whitcomb & Merrell, 2013). While assessment has always been a fundamental cornerstone of school psychology practice, as educators and researchers develop and deepen their understanding of the factors that influence students' achievement, the need to expand the scope of assessments to include SRL and other academic enablers becomes more apparent (Cleary, 2006; DiPerna, Volpe, & Elliot, 2002; Linnenbrink & Pintrich, 2002). Studies have shown that both teachers and school psychologists perceive assessment of SRL to be important (Cleary & Zimmerman, 2006; Cleary et al., 2010; Peeters et al., 2016). Despite these perceptions, however, Cleary and colleagues found that school psychologists do not routinely conduct evaluations of students' SRL and motivation skills, and they report little to no familiarity with published, available self-report and ratings scales designed to measure SRL and motivation (Cleary, 2009; Cleary et al., 2010). One compelling reason for their lack of familiarity is that very few parent and teacher rating scales targeting student SRL skills exist. Therefore, developing valid and reliable measures that are feasible and acceptable for school psychology practice may be an important step in promoting SRL assessment in the schools. School psychologists often rely on teacher ratings to assess various areas of student functioning (Callan & Cleary, 2017), so teacher ratings of SRL may be a particularly useful tool to add to their repertoire.

While there is a wide research base exploring the use of teacher ratings across many areas of student functioning, teacher ratings of SRL have received minimal attention in the literature. Four main studies that examined teacher ratings of SRL are discussed below. These studies use the RSSRL (Zimmerman & Martinez-Pons, 1988) and the SRSI-TRS, which is the focus of this dissertation.

Zimmerman and Martinez-Pons (1988) developed the RSSRL as a method for validating their strategy model of student SRL. On this measure, teachers rate overt self-regulation strategies that are observable in the classroom (e.g., seeking information) and outcomes for more covert strategies (e.g., preparedness for class as an outcome of planning). The items on the RSSRL were developed based on the 14 previously identified categories of self-regulation strategies (Zimmerman and Martinez-Pons, 1986). Several items about students' motivation and task interest are included on the RSSRL as well. A sample of 80 tenth grade students from a middle class suburban high school were administered a structured interview (SRLIS; Zimmerman & Martinez-Pons, 1986), and their teachers rated them using the RSSRL. Results found a high canonical correlation (R = .70) between teacher ratings and students' reports of SRL. Further, factor analysis of the RSSRL and student Mathematics and English scores revealed a single underlying construct of student SRL that was separate from achievement. Most of the 14 SRL strategies measured by the student interview were significantly correlated with the canonical root of the teacher ratings. (Only self-evaluation, environmental structuring, goal setting and planning, and reviewing notes were not significantly correlated.) The results of this study validated the strategy model as a theoretical basis of SRL. Regarding the RSSRL, the authors concluded that while teacher ratings are a valuable source of information about students' SRL, they are somewhat limited because teachers cannot necessarily observe all strategies included. Teacher rating data, therefore, should be corroborated with other forms of assessment, such as outcome data, observations, and ratings by other informants.

DiBenedetto and Zimmerman (2013) used the RSSRL in a study that aimed to determine construct and predictive validity of a microanalytic approach to assessing SRL subprocesses. In this study, the RSSRL was shown to be predictive of students' achievement on a science task; it accounted for 26% of the variance in achievement scores, and it also shared significant variance with the microanalytic measures. (It should be noted that the micronalytic measures did in fact account for more variance than the RSSRL.) In addition, the RSSRL was significantly correlated with microanalytic measures for metacognitive monitoring and self-evaluation, but not for measures of strategy use. The authors used multiple regression analyses to assess the effects of prior achievement on the SRL outcome measures, and found a linear effect in that higher levels of achievement predicted greater use of regulatory strategies. However, they did not control for prior achievement in their prediction models for post-test achievement scores, so the results should be interpreted with caution.

Cleary and Callan (2014) developed a different teacher rating scale, SRSI-TRS, which is the first teacher rating scale to parallel a student self-report questionnaire (SRSI-SR; Cleary, 2006). This measure was created to address the paucity of available teacher rating scales for SRL, and primarily targets students' use of help-seeking, self-motivation, and organization behaviors. The authors initially drew from the SRSI-SR to create a pool of 20 items pertaining to SRL behaviors that are observable in the classroom. The items were then given to teachers to evaluate in terms of readability and applicability to the classroom context. Following a feedback session with the teachers, seven items were removed from the scale because they were related to covert behaviors or behaviors not done within the school context. The predictive and convergent validity of the SRSI-TRS was then examined in a study with 87 ninth grade students. Results of the study indicated that the SRSI-TRS accounted for a medium amount of variance (9.4%) in students' mathematics test scores, after controlling for prior achievement and self-reported motivation beliefs and self-regulation behaviors. The SRSI-TRS also had a positive correlation with students' self-reported mathematics task interest, and a negative correlation with students' self-reported maladaptive regulatory behaviors.

In a more recent study, Callan and Cleary (2017) examined convergent and predictive validity of four measures of SRL: two aptitude measures (SRSI-TRS, SRSI-SR) and two event measures (microanalysis, traces). Convergence was examined both between and within measurement classes. Predictive validity was examined for three achievement outcomes (practice session math problems, math posttest, and standardized mathematics assessment). A sample of 100 eighth grade students from an urban school district was used for the study. As was noted previously, the authors found that the two aptitude measures showed significant, medium correlations (r = .30), as did the two event measures (r = .43 for easy mathematics problems, r =

.35 for difficult mathematics problems). The correlations between measurement classes were not significant. Three regression analyses were conducted to predict each of the outcomes using the four measures. Students' classroom was controlled in the analyses because of significant differences in mathematics achievement between classrooms. Results of regression analyses showed that the SRSI-TRS uniquely accounted for a medium amount of variance in students' posttest scores and standardized test scores (11% and 12%, respectively). It was not predictive of students' performance on the practice session math problems. The micronalytic measures of metacognitive monitoring were predictive of all three outcomes, while the self-report questionnaire and behavioral traces did not explain a significant amount of unique variance for any outcomes. Overall, teacher ratings emerged as the strongest predictor of students' standardized mathematics scores. This study underscored the value of teacher ratings of SRL as an important predictor of academic achievement, particularly with regard to global measures of achievement such as standardized test scores.

These four aforementioned studies support the premise that teacher ratings can provide important information about student SRL. The utility of the information generated by teacher ratings, however, is dependent on the psychometric properties of the measures used. Studying the reliability and validity of teacher ratings is an essential step to developing measures that can provide accurate and useful information to inform further practice and intervention.

There are several aspects of reliability and validity that are important to consider with regard to the SRSI-TRS. No research to date has examined the factor structure of this measure. Callan and Cleary (2017) noted that the high alpha level they found for the SRSI-TRS may indicate redundancy in the construct measures, so further exploration of construct validity is an important goal for future research. Another new thread of research is interrater reliability, which has been examined for teacher ratings in general but not for teacher ratings of SRL. In their meta-analysis on cross-informant agreement for child behavioral and emotional problems, Achenbach et al. (1987) found large correlations between ratings given by pairs of teachers (r = .64), while other studies have found lower correlations (Nickerson & Nagle, 2001). One factor influencing the interrater reliability is whether the two teachers work in the same classroom and therefore observe students in the same situations. This latter point is relevant to SRL, as students' use of regulatory strategies and behaviors can vary from classroom to classroom (Perry & Rahim, 2011). Thus, examining interrater reliability for teachers in the same classroom can give preliminary evidence as to whether a measure will yield stable results across respondents in the same context.

Another area of interest is convergence of teacher ratings with student reports. In general, studies that examine the accuracy of teacher ratings of academic performance often use students' performance on standardized achievement tests as their criterion measure for calculating correlations (Demaray & Elliot, 1998; Hoge & Butcher, 1984; Hoge & Coladarci, 1989). For student behavior, there are no standardized criterion measures; therefore, teacher-student agreement is often used as a benchmark with which to measure accuracy. It is important to examine cross-informant agreement for the SRSI-TRS and compare the findings to previous studies as well as the general literature on teacher ratings.

Typically, research on teacher-student agreement for ratings of behavioral or emotional problems have yielded low rates of agreement. In their meta-analyis, Achenbach and colleagues (1987) reviewed 17 studies that examined cross-informant agreement between teachers and children. They found a significant but small mean correlation (r = .20), and similar correlation have been found in other studies as well (De Los Reyes et al., 2015; Phares et al, 1989). The level of agreement between teacher and students ratings appears to vary based on the type of behaviors being rated. Achenbach et al. (1987) found a significant difference between the mean correlation for studies examining externalizing behaviors and the mean for internalizing behaviors (r = .34 and r = .16, respectively.) Other studies have similarly found higher correlations for externalizing behaviors than internalizing behaviors (De Los Reyes et al., 2015). This is relevant with regard to SRL, since SRL encompasses covert behaviors (e.g., use of strategies such as help-seeking) as well as more covert processes (e.g., self-monitoring, motivation).

Disagreement between teacher and student self-report need not suggest a low level of accuracy on the part of either informant, but may reflect on the complexity of the construct being measured and may indicate that each measure is providing unique information (De Los Reyes et al., 2015; Meyer et al., 2001). While different data sources can be viewed as complementary, it should be underscored that the quality of teacher ratings matters, because teacher use their judgements of student functioning in order to inform instruction and intervention (Dicke, Ludtke, Trautwein, Nagy, & Nagy, 2012; Gerber & Semmel, 1984). With regard to SRL, if teachers cannot reliably assess students' regulatory skills and behaviors, they cannot effectively foster and promote SRL in their classrooms (Peeters et al., 2016). Thus, continued focus on the accuracy, reliability, and validity of teacher ratings is an important goal for research.

Conclusion

The research reviewed thus far provides robust support for the importance of SRL and its strong relation with academic achievement. In order to promote students' SRL skills through the use of effective classroom practices and targeted interventions, those skills must be assessed accurately and comprehensively. While SRL assessment encompasses a wide range of useful methods and measures, teacher ratings are well aligned with school psychology practice, are cost-effective and efficient to use, and can be linked to intervention strategies. In addition, teacher ratings have been shown to be a valuable measure of many areas of student functioning in school and clinical settings. Teacher ratings for SRL have been utilized in several studies, but additional validity and reliability information is needed to bolster support for their use in research and practice. The current study seeks address this gap in the literature by comprehensively examining the reliability and validity of the SRSI-TRS, an existing teacher rating scale, in order to strengthen the evidence for this measure as a psychometrically-sound and beneficial method of assessing SRL.