Hammill Institute on Disabilities

Assessment for Effective Intervention
2017, Vol. 42(2) 97–109

DOI: 10.1177/1534508416670367
aei.sagepub.com
SAGE

Examining the Relations Between Self-Regulation and Achievement in Third-Grade Students

Stephanie L. Day, PhD[1] and Carol McDonald Connor, PhD[1]

Abstract

Children with stronger self-regulation skills generally demonstrate greater overall success in school both academically and socially. However, there are few valid and reliable measures of self-regulation in middle elementary school. Such a measure could help identify whether a child is truly having difficulties. Thus, the Remembering Rules and Regulation Picture Task (RRRP) was developed. The aim of this study was to develop scoring systems for the RRRP and then to examine the associations between RRRP and independent measures of self-regulation and academic achievement in mathematics and reading. Children (N = 282) from 34 third-grade classrooms in Florida participated in this study. Results revealed that the RRRP captured three constructs: working memory, attentional flexibility, and inhibitory control. Hierarchical linear modeling (HLM) demonstrated that the RRRP was significantly and positively associated with other measures of self-regulation. The RRRP was significantly and positively associated with mathematics and reading as well. The RRRP appears to be a promising measure of children's self-regulation skills.

Keywords

self-regulation, vocabulary, academic achievement, reading, executive function, mathematics, language

Research on the relation between self-regulation and academic achievement is becoming increasingly important as we try to learn about all of the potential factors that may influence children's academic achievement (Connor et al., 2010; McClelland & Ponitz, 2011; Zimmerman, 2001). There are a number of child characteristics that can affect literacy achievement and academic success (Duncan et al., 2007; McGee, Prior, Williams, Smart, & Sanson, 2002), and increasingly, in addition to cognitive skills, self-regulation has been noted as a potentially important source of influence. We define self-regulation as being supported by executive functioning, specifically working memory, inhibitory control, and attentional flexibility (Lin, Coburn, & Eisenberg, 2016). Self-regulation is an active process by which people learn to control their own behaviors, cognition, motivation (Pintrich, 2000), and emotions (Lin et al., 2016). Having strong self-regulation skills appears to be crucial for academic success as it allows students to concentrate on instruction, to be more organized, rehearse information that is to be remembered, to use their environmental resources efficiently and effectively so that they may benefit from learning experiences, to hold positive beliefs about their own capabilities, and to set plans and goals for their actions (Schunk, 1989; Schunk & Zimmerman, 1994). Hence, developing an assessment that can reliably and validly assess children's self-regulation would, potentially, allow for better identification and accommodation of students with either weak or strong self-regulation skills during their school career, and was thus the purpose of this study. There have been a number of studies that have examined the associations between self-regulation and academic achievement, but most of these studies have been conducted with younger children, particularly those in preschool and kindergarten (McClelland et al., 2007; Ponitz et al., 2008; Ponitz, McClelland, Matthews, & Morrison, 2009), whereas much less work has focused on middle elementary grades (Del Giudice, 2014).

Research has shown that students who demonstrate more disruptive behaviors and lack strong self-regulation skills in elementary school are more likely to exhibit academic difficulties in the future (Vitaro, Brendgen, Larose, & Tremblay, 2005). Moreover, for children who come from disadvantaged socioeconomic backgrounds, the risk for developing problems with self-regulation may be even greater (McClelland & Ponitz, 2011). We conjecture that strong self-regulation skills may be particularly important

[1]University of California Irvine, USA

Corresponding Author:
Stephanie L. Day, School of Education, University of California Irvine, 3200 Education, Irvine, CA 92697-5500, USA.
Email: sday1@uci.edu

in middle elementary school and beyond, where it is expected that children be able to work more independently and without constant aid from the teacher. The main purpose of this study is to examine how a newly developed measure of self-regulation called the Remembering Rules and Regulation Picture Task (RRRP) measures self-regulation skills in the service of academic achievement in reading and mathematics in third-grade students.

Self-Regulation and Academic Achievement

Over the past decade, studies examining associations between self-regulation and academic achievement have shown that having strong self-regulation skills is associated with school success in both reading and mathematics (Adams & Snowling, 2001; Connor et al., 2010; Day, Connor, & McClelland, 2015; Lin et al., 2016; St. Clair-Thompson & Gathercole, 2006). As reported above, self-regulation is supported by and overlaps with executive functions, specifically attentional flexibility, working memory, and inhibitory control. Our aim with the RRRP was to develop an assessment that relied on the coordinated regulation of all three functions. We discuss each of them below.

Attentional Flexibility

Attentional flexibility is the ability to block out unwanted stimuli to pay attention and focus on tasks while also being able to shift attention away from that task when necessary (Rueda, Posner, & Rothbart, 2004). Having strong attentional flexibility skills appears to be related to being successful in school (Rabiner, Coie, & The Conduct Problems Prevention Research Group, 2000). Children who have poor attentional skills are more likely to have difficulties in reading and mathematics (Holmes, Gathercole, & Dunning, 2009; Rabiner et al., 2000; H. L. Swanson & Sachse-Lee, 2001).

Working Memory

Baddeley (1986) defined working memory as a system of limited capacity for the temporary maintenance and manipulation of information. Working memory is highly correlated with performance on several academic and language-related tasks, such as vocabulary, reading comprehension, mathematics, and problem solving, and is necessary for a wide range of classroom learning situations (Hofmann, Schmeichel, & Baddeley, 2012; H. L. Swanson, 1994). Gathercole and Alloway (2008) found that more than 80% of children who fell into the bottom 10th percentile of those with poor working memory skills struggle quite substantially in reading and mathematics. It has also been found that children who had higher mathematics ability had a higher working memory span in children 7 years of age (Bull & Scerif, 2001).

Inhibitory Control

Inhibitory control (also called task inhibition) is the ability to stop a dominant response in favor of a subdominant, more adaptive response (Ponitz et al., 2009). Children who have poor inhibitory control skills in elementary school are typically more hyperactive and impulsive, and can often be seen aimlessly wandering around the classroom, speaking without raising their hand, playing at their desks, or chatting with other classmates when they are supposed to be doing work (Rimm-Kaufman, La Paro, Downer, & Pianta, 2005). Weak inhibitory control may prevent children from completing their work independently (Day et al., 2015), while creating disruptions for peers and the teacher, which takes away from learning time (McGee et al., 2002; Skibbe, Phillips, Day, Brophy-Herb, & Connor, 2012).

Assessing Self-Regulation

Whereas there continues to be a growing body of research on self-regulation and how it is related to academic achievement, there has not been concurrent advancement in the assessment of self-regulation, particularly in middle to older elementary grades. Currently, there are a number of assessments that rely on self, teacher, parent, or clinician report. The *Strengths and Weaknesses of ADHD Symptoms and Normal Behavior Rating Scale* (SWAN; J. M. Swanson et al., 2006) has been found to be a valid and reliable measure of attentional flexibility and inhibitory control (hyperactivity–impulsivity) problems, and has been found to be an appropriate measure to use in the general population (Sáez, Folsom, Al Otaiba, & Schatschneider, 2012). Other measures include The Child Behavior Questionnaire (CBQ; Rothbart, Ahadi, Hershey, & Fisher, 2001) and the Child Behavior Checklist (CBCL; Achenbach, 2001), both of which capture behavioral inhibition and emotional problems (Polderman et al., 2007). Self-report measures such as the *Impulsivity Scale for Children* (ISC; Duckworth, Quinn, & Tsukayama, 2012) have also been utilized to measure self-regulation skills (Suchodoletz, Larsen, Gunzenhauser, & Fäsche, 2015). However, with the exception of the SWAN, measures such as the CBCL were designed to be given by clinicians and tend to only capture those with more serious attention or behavior problems and when administered to the general population, result in a skewed distribution of scores (Polderman et al., 2007; Whitebread et al., 2009). Furthermore, some studies have reported difficulty finding robust relations between direct measures of self-regulation, classroom behaviors, and parent reports of self-regulation skills (Blair & Razza, 2007; Neuenschwander, Röthlisberger, Cimeli, & Roebers, 2012). In addition, surveys and questionnaires may also be open to observer bias (McClelland & Ponitz, 2011; Ponitz et al., 2008; Whitebread et al., 2009).

Working memory, attentional flexibility, and inhibitory control are also measured in more direct ways, but often as separate constructs. For example, the Knock-Tap, a subtest from a developmental neuropsychological assessment (NEPSY; Korkman, Kirk, & Kemp, 1998), measures task inhibition and working memory (Molfese et al., 2010). However, there appears to be a ceiling effect in 6- to 8-year-olds. The authors note that for typically developing children past the early elementary grades, another measure might be more useful. Another example includes delayed gratification tasks such as the Watch and Wait Task developed by Neubauer, Gawrilow, and Hasselhorn (2012) in which preschool students are instructed to watch an hourglass run out to receive a reward. However, measures of delayed gratification are primarily designed to capture inhibitory control and do not necessarily measure attentional flexibility or working memory. Other commonly used measures of working memory in middle-to-late childhood include backward recall tasks in which students must recall a sequence of numbers or letters in the reverse order as the examiner said (Adams & Snowling, 2001; Holmes et al., 2009; St. Clair-Thompson & Gathercole, 2006). Variations of the Stroop task in which the subject names the colors of a series of congruent and incongruent color words have also been used to measure working memory and inhibition skills (Bull & Scerif, 2001; Neuenschwander et al., 2012; St. Clair-Thompson & Gathercole, 2006). Attentional flexibility is often measured using a visual search task. For example, Adams and Snowling (2001) utilized a task in which children between the ages of 8 through 11 years old search for targets in rows containing target and distracter items. While many of these tasks are done one-on-one, some are completed on a computer, which may not be easy to administer in a school setting.

For younger children (preschool through first grade), the Head–Toes–Knees–Shoulders (HTKS) task was designed to directly measure self-regulation. In this measure, children are asked to inhibit their impulse to touch the body part named and touch the opposite body part instead. For example, when the tester says, "touch your head," the student should then touch their toes and vice versa. While this measure has been shown to be a strong predictor of academic performance in preschool and kindergarten, it is less useful with children past the first grade as these children tend to reach ceiling on the measure (Connor et al., 2010; Day et al., 2015; McClelland & Ponitz, 2011).

The RRRP

The RRRP was designed to be a direct measure of self-regulation skills, including how children managed the coordination of working memory, attentional flexibility, and inhibitory control. The RRRP was also designed to be appropriate to administer to students beyond the first grade, and that was a direct measure as opposed to teacher or parent report. In the RRRP, children are presented with a picture of a park setting and are asked to place different colored blocks on objects in the picture and in a particular order. Children are expected to pay close attention to each question as it as read, to remember which color blocks they should use and which objects (and in which order) they are to place the blocks. The students are also expected to wait for the question to be read in its entirety and for the tester to say "go" before placing the blocks. An extra tax on working memory and inhibitory control is introduced in the second half of the test where students are instructed to switch blue blocks for red blocks and vice versa.

The following research questions guided this study:

> **Research Question 1:** How should the RRRP be scored?

Our aim was to develop a one-dimensional construct but recognize that there may be multiple dimensions (e.g., working memory, inhibitory control, and attentional flexibility).

> **Research Question 2:** What is the association among the RRRP and other measures of self-regulation?

Our aim was to develop an assessment that could be used as either a complement or instead of teacher-report measures frequently used to measure self-regulation, which would suggest adequate construct reliability.

> **Research Question 3:** To what extent does the fall RRRP predict spring academic skills, specifically mathematics, reading, and vocabulary in third-grade classrooms? Furthermore, does fall RRRP predict gains in academic skills from fall to spring?

We hypothesize that the Fall RRRP will be significantly associated with gains in academic skills, which would suggest adequate predictive reliability and would align with other studies showing this association.

Method

Participants

Children from 34 third-grade classrooms in one public school district in Florida participated in this study. Participants were part of a larger cluster-randomized control field trial that was designed to help teachers learn how to individualize their literacy instruction called Individualizing Student Instruction (ISI; Connor et al., 2013). In the larger study, a smaller subset of students was randomly selected to receive an extended battery of tests, which included the RRRP. In each classroom, students were rank-ordered by their fall comprehension scores and then divided into equaled number groups: high, average, and low. Within each group, approximately three students were randomly selected to receive the extended battery of tests. On average, approximately 11 students from each class received the extended battery. In total, the RRRP was administered to 282 students. Forty-five percent of the students qualified for free or reduced price lunch. The majority of the sample was Caucasian (84%; 6% African American, 3% Hispanic, 1% Asian, and 6% Other). Child gender was 57% female and 43% male. The mean age at the start of the study was 8.5 years of age ($SD = 0.415$).

Measures

The RRRP. The RRRP was designed to measure children's self-regulation skills. In the RRRP, children are presented with a picture of a park setting and are asked to place different colored blocks (we used LEGO®s) on objects in the picture and in a particular order. In the second half of the task, students are instructed to switch blue blocks for red blocks and vice versa. To eliminate any possible difficulties with vocabulary and morphosyntactic skills, we selected objects that most elementary students could easily identify by the third grade, and the syntactic forms were generally mastered by children by kindergarten.

Administration. The RRRP consists of 10 items. Students are read the directions by a trained tester followed by three practice trials. They are given blocks in assorted colors that are placed in a pile next to the picture and then given a direction, for example, "Put a blue block on the squirrel by the rock. Go." The child is expected to wait until the examiner says "go" before they place the correct color block (an unexpected task with blocks) on the correct object, in the correct order.

The assessment begins with simpler tasks by including one block and one object, and becomes increasingly more difficult. For example, a more difficult prompt is: "Put a black block on the duck in the lake after you put a red block on any ant. Go." Halfway through the administration, students are told that when the administrator says to use a red block, they should use a blue block, and when the tester says to use a blue block, they should use a red block. Students are also able to ask for each direction to be repeated once. The repetitions were included as a proxy of metacognitive awareness; however examining repetitions was beyond the scope of the present study.

Examiners documented every detail of each student's response for all items on the test. They were to mark whether the question was answered correctly, if the child waited for "go," and if he or she asked for a repetition. If the question was answered incorrectly, the administrator marked what specifically the child did incorrectly (block color, object, and/or order). To establish interrater reliability, we video recorded 10 students taking the pilot test and all assessors were required to watch the video and score the same students. Interrater reliability was then calculated using SPSS. The interrater reliability of the RRRP was .95 (Kappa) in the current study.

In sum, working memory, attentional flexibility, and inhibitory control were tapped through the child's ability to answer each question correctly. For the students to score well on the RRRP, they must pay attention and remember three things while taking the test: the color of the blocks, the objects in the pictures in which they are to place the blocks, and the order in which they are to place the blocks. They must wait for the tester to finish reading the directions and for "go" before they begin placing the blocks. The students must also inhibit the expected response to follow the new color switching rule for the second half of the task.

Other measures of self-regulation and executive functioning. Students' attention and inhibitory control were assessed using the SWAN Rating Scale (J. M. Swanson et al., 2006). The SWAN asks teachers to rate their students' behavior through 18 questions. Teachers rate their students' attention and inhibitory control skills on a 7-point scale, where Far Below Average = 3, Below Average = 2, Somewhat Below Average = 1, Average = 0, Somewhat Above Average = −1, Above Average = −2, and Far Above Average = −3. The first nine items on the test measure attention skills while the last nine measure inhibitory control (hyperactivity–impulsivity). When used with preschool children in a previous validation study, test–retest reliability ranged from .91 to .96 and internal consistency ranged from .71 to .76 (Lakes, Swanson, & Riggs, 2011). The SWAN is designed to capture both strengths and weaknesses of attentional problems, and yields a normal distribution of scores when used in the general population. Thus, it is an appropriate measure to use in a school setting with typically developing children (Polderman et al., 2007).

Memory for Digits Reversed was designed as a measure of working memory. This measure was adapted from the Memory for Digits task from the *Comprehensive Test of Phonological Processing* (CTOPP; Wagner, Torgesen, & Rashotte, 1999). This task was designed by taking the number sequences from the Memory for Digits CTOPP test and instructing the students to repeat back the sequence of numbers in the reverse order in which they are read. The sequence starts with two numbers and then increases to eight numbers. Reliability in the current sample was $\alpha = .77$.

Academic measures. Children's reading and mathematics skills were assessed using subtests from the *Woodcock–Johnson Tests of Achievement–III* (WJ-III; Woodcock, McGrew, & Mather, 2001). The Picture Vocabulary subtest was used to assess vocabulary. The Picture Vocabulary task asks children to identify objects in a picture. The Letter-Word Identification (ID) subtest and the Passage Comprehension subtest of the WJ-III were used to assess children's reading skills. The Letter-Word subtest requires children to simply read a list of words that become increasingly difficult. On the Passage Comprehension subtest, children are asked to orally supply a missing word from a short passage that they read silently to themselves. Mathematics skills were measured using the Applied Problems subtest and the Math Fluency subtest of the WJ-III. On the Applied Problems subtest, word problems are read orally to the child and then the child can use pencil and paper to solve the problem. On the Math Fluency measure, students are asked to solve addition, subtraction, and multiplication problems in 3 min. W scores were used in all analyses, which are a form of a Rasch score providing a common scale that represents both task difficulty and a person's growth. W scores are obtained using the WJ-III Compuscore software program. The W score is centered at 500 for each test, which is set as the approximate average performance of a 10-year-old child. The subtests of the WJ-III have reliabilities ranging from .81 to .94.

Procedure

Students were assessed in the early fall on the academic subtests of the WJ-III, and then again at the end of the school year in the spring. The RRRP was also administered in the fall and spring, which took approximately 7 min per child. The WJ tests, RRRP, and the working memory measure were all administered by trained research assistants outside the classroom in the hallway or in a separate room (when available). Tests were administered in a fixed order for all participants. Each of the WJ tests took approximately 7 to 10 min to administer. Memory for Digits Reversed was given in the spring. Teachers completed the SWAN Rating Scale for each student in their class in early spring of the school year.

Results

Means, standard deviations, ranges, and skewness and kurtosis values were examined for all variables and can be found in Table 1. Normality of the data was evaluated by examining the skew and kurtosis values for all variables. All variables appeared to fall within the normal range for both skew and kurtosis. Bivariate correlations were examined next and can be found in Table 2. Performance on the RRRP was significantly correlated with academic measures of both reading and math, working memory, and teacher's ratings of children's attention and hyperactivity skills from the SWAN. It was also found that children's ability to wait for "go" on the RRRP was correlated with ratings of hyperactivity on the SWAN.

Scoring the RRRP

When the RRRP was administered to participants, examiners were asked to record each child's answer in every detail. In other words, the examiner recorded whether the color, object, and order was correct, whether the child asked for a repetition, and whether they answered each prompt correctly or not. To analyze these results, a number of variables were created using the collected data for children's performance on the RRRP. First, a total score variable was created to represent how many items the child answered correctly out of the 10 questions on the test. For an answer to be considered correct, the child needed to place the correct color blocks on the specific objects, and in the correct order. The question was marked correct if the child answered the prompt correctly on the first or second trial. If the child first answered a question correctly and then asked for the question to be repeated and answered the question incorrectly on the second trial, the answer was marked as incorrect on the score sheet. If a child first answered the question incorrectly and then asked for a repetition and answered the question correctly on the second trial, the answer was marked as correct. A total score was created for both the fall and the spring tests.

In addition, the RRRP could be divided into two parts with Part 1 including the first five questions and Part 2 including Questions 6 through 10. The questions in Part 2 included the color switch for the red and blue blocks, such that the child was required to use the opposite color than what he or she was directed. Total scores for both parts were also created. A variable was also created to represent the total number of times the child waited for "go" on the first trial. A total score of 10 would indicate that the child waited for "go" on all 10 items on the test. Cronbach's alpha was estimated for each scale with values of .83 for Part 1 of the RRRP, .83 for the Part 2 of the RRRP, .97 for the "go" items, and an overall scale reliability of .97. An item response theory (IRT) analysis of fall RRRP scores was also completed to assess the difficulty and discrimination of each item. More information on the IRT analysis can be found in the supplemental materials. We also examined whether students made significant gains on the RRRP from fall to spring by running a repeated-measures ANCOVA. Part 1 scores on the RRRP decreased slightly from Time 1 to Time 2 ($b = -.418$, $p < .001$), but Part 2 (color switch) scores on the RRRP did not change over time ($b = -.024$, $p = .811$), indicating that scores on the RRRP are fairly stable in third grade.

Association of RRRP With Other Measures of Self-Regulation

To examine the associations among measures of attentional flexibility, working memory, inhibitory control, and the RRRP, hierarchical linear modeling (HLM) was used to account for the nested structure of the data (children nested within teacher;

Table 1. Child Descriptives.

Variable	*n*	Minimum	Maximum	*M*	*SD*	Skewness	Kurtosis
School FARL	282	39	59	45.26	7.73	1.00	−0.71
Fall RRRP Part 1 total score	278	1	5	4.20	0.90	−0.98	0.28
Fall RRRP Part 2 total score	278	0	5	3.80	1.13	−0.72	−0.07
Fall RRRP total correct	280	2	10	7.97	1.77	−0.85	0.28
Fall RRRP wait for "Go" total	278	0	13	8.15	3.11	−1.63	1.29
Fall WJ Letter-Word ID W score	278	446	539	492.75	17.35	−0.02	−0.28
Fall WJ Passage Comprehension W score	278	458	515	488.71	10.30	−0.13	−0.14
Fall WJ Picture Vocabulary W score	278	474	535	497.38	10.07	0.35	0.72
Fall WJ Math Fluency W score	278	482	508	493.56	4.84	0.30	−0.24
Fall WJ Applied Problems W score	278	442	531	487.47	16.01	−0.18	−0.21
Spring RRRP Part 1 total score	278	0	5	3.80	1.41	−1.17	0.55
Spring RRRP Part 2 total score	278	1	5	3.77	1.01	−0.42	−0.49
Spring RRRP total correct	279	1	10	7.56	2.09	−0.75	−0.12
Spring RRRP wait for "Go" total	278	0	10	7.79	3.03	−1.30	0.44
Spring WJ Letter-Word ID W score	282	460	556	505.69	16.84	−0.02	0.14
Spring WJ Passage Comprehension W score	282	469	520	495.55	9.49	−0.30	−0.12
Spring WJ Picture Vocabulary W score	282	477	535	502.04	10.46	0.30	−0.01
Spring WJ Math Fluency W score	282	485	516	497.48	5.65	0.58	0.17
Spring WJ Applied Problems W score	282	458	542	503.05	14.55	−0.44	0.43
SWAN Hyperactivity total	267	−3	2.78	−0.33	1.27	−0.55	0.12
SWAN Attention Scale total	266	−3	2.78	−0.17	1.39	−0.28	−0.42
Memory for Digits Reversed total score	278	3	14	7.03	2.22	0.82	0.29

Note. FARL = free and reduced price lunch; RRRP = Remembering Rules and Regulation Picture Task; WJ = Woodcock–Johnson III; ID = Identification; SWAN = *Strengths and Weaknesses of ADHD Symptoms and Normal Behavior Rating Scale*.

Table 2. Correlations Between the RRRP, Self-Regulation Measures, and Child Outcomes.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1.	Fall RRRP	1															
2.	Sp RRRP	.222**	1														
3.	Fall RRRP P1	.405**	.084	1													
4.	Fall RRRP P2	.541**	−.056	.462**	1												
5.	Sp RRRP P1	.277**	.020	.359**	.303**	1											
6.	Sp RRRP P2	.225**	.038	.260**	.276**	.468**	1										
7.	Fall RRRP Go	.886**	.255**	.029	.146*	.117	.095	1									
8.	Sp RRRP Go	.207**	.996**	.063	−.066	−.032	.013	.247**	1								
9.	MfD Reversed	.122*	.038	.198**	.171**	.245**	.169**	.027	.026	1							
10.	SWAN Attention	−.248**	−.186**	−.186**	−.169**	−.202**	−.188**	−.182**	−.174**	−.277**	1						
11.	SWAN Hyperactivity	−.275**	−.242**	−.093	−.110	−.086	−.110	−.261**	−.232**	−.174**	.823**	1					
12.	Sp WJ Letter-Word ID	.158**	−.009	.227**	.266**	.251**	.214**	.033	−.022	.351**	−.349**	−.215**	1				
13.	Sp WJ Passage Comprehension	.182**	−.065	.226**	.245**	.345**	.255**	.062	−.084	.266**	−.375**	−.274**	.623**	1			
14.	Sp WJ Picture Vocabulary	.196**	.009	.213**	.260**	.242**	.186**	.079	−.011	.198**	−.280**	−.176**	.458**	.494**	1		
15.	Sp WJ Math Fluency	−.012	−.089	.106	.082	.122*	.094	−.073	−.098	.243**	−.294**	−.161**	.357**	.266**	.119*	1	
16.	Sp WJ Applied Problems	.139*	.004	.307**	.239**	.346**	.303**	−.005	−.019	.364**	−.418**	−.242**	.492**	.500**	.406**	.450**	1

Note. RRRP = Remembering Rules and Regulation Picture Task; Sp = Spring, P1 = Part 1, P2 = Part 2; MfD = Memory for Digits; SWAN = *Strengths and Weaknesses of ADHD Symptoms and Normal Behavior Rating Scale*; WJ = Woodcock–Johnson III; ID = Identification.

$*p < .05$. $**p < .01$.

Table 3. Associations Between Parts 1 (Top) and 2 (Bottom) of the RRRP and Self-Regulation Measures.

Fixed effect	Coefficient	*SE*	*t*	*df*	*p* value
Fall RRRP Part 1 total score intercept	4.195	0.07	60.00	29	<.001
Memory for Digits Reversed	0.069	0.02	3.31	249	.001
SWAN Attention total score	−0.150	0.07	−2.12	249	.035
SWAN Hyperactivity total score	0.125	0.07	1.81	249	.071
Random effect	*SD*	Variance component	*df*	χ^2	*p* value
Between classroom residual	0.251	0.06	29	52.94	.005
Within classroom residual	0.843	0.71			
Deviance	673.166				
Fixed effect	Coefficient	*SE*	*t*	*df*	*p* value
Fall RRRP Part 2 total score intercept	3.765	0.10	38.78	29	<.001
Memory for Digits Reversed	0.079	0.03	2.55	249	.011
SWAN Attention total score	−0.086	0.11	−0.79	249	.429
SWAN Hyperactivity total score	0.018	0.10	0.19	249	.848
Random effect	*SD*	Variance component	*df*	χ^2	*p* value
Between classroom residual	0.389	0.15	29	61.43	<.001
Within classroom residual	1.050	1.10			
Deviance	790.747				

Note. ICC = 12% (top) ICC = 14% (bottom). RRRP = Remembering Rules and Regulation Picture Task; SWAN = *Strengths and Weaknesses of ADHD Symptoms and Normal Behavior Rating Scale*; ICC = intraclass correlation coefficient.

Table 4. Associations Between "Wait for Go" Total Score of the RRRP and Self-Regulation Measures.

Fixed effect	Coefficient	*SE*	*t*	*df*	*p* value
Fall RRRP "Wait for Go" total score intercept	8.073	0.26	31.04	29	<.001
Memory for Digits Reversed	0.025	0.10	0.24	249	.811
SWAN Attention total score	0.153	0.23	0.66	249	.509
SWAN Hyperactivity total score	−0.846	0.37	−2.30	249	.022
Random effect	*SD*	Variance component	*df*	χ^2	*p* value
Between classroom residual	1.034	1.07	29	62.37	<.001
Within classroom residual	2.831	8.11			
Deviance	1,294.072				

Note. ICC = 1%. RRRP = Remembering Rules and Regulation Picture Task; SWAN = *Strengths and Weaknesses of ADHD Symptoms and Normal Behavior Rating Scale*; ICC = intraclass correlation coefficient.

Raudenbush & Bryk, 2002). Models were run for Parts 1 and 2 of the fall RRRP and for "Wait for Go." On Part 1 of the RRRP, both Memory for Digits and the SWAN Attention subscale were significantly related to performance on the fall RRRP (see Table 3). Ratings of inhibitory control on the SWAN indicated a trend (p = .071). For Part 2 (color switch) of the fall RRRP, only Memory for Digits was significantly and positively associated with performance on Part 2 of the RRRP (see Table 3). Finally, a third model was run for "wait for go," which was designed to capture inhibitory control. In this model, only ratings of inhibitory control were significantly related to children's ability to wait for "go" (see Table 4).

Association With Academic Outcomes

Mathematics. Three models were run for the mathematics measures. The first two-level HLM model was run to examine whether fall scores on the RRRP predicted performance in spring Math Fluency, and a second model was run for the Applied Problems subtest. For Math Fluency, fall RRRP scores (Parts 1 and 2) did not predict spring outcomes in Math Fluency. On the contrary, in the Applied Problems model, Part 1 of the fall RRRP significantly predicted spring Applied Problems outcomes (see Table 5). We then ran a third two-level model that examined whether Fall RRRP predicted growth on the Applied Problems subtest. When controlling for fall Applied Problems scores, Part 1 of the RRRP significantly predicted gains on Applied Problems from fall to spring (see Table 5). Children's ability to wait for "go" was not a significant predictor in any of the academic models and was thus removed from the models. Socioeconomic status (SES) was initially controlled for using school-level free and reduced price lunch (FARL) in the mathematics models as well as the previous models; however, it was not significant ($b = -.010$, $p = .463$), and thus SES was removed from the models.

Reading. To test for the association between fall RRRP and reading performance, we ran separate two-level HLM models for Letter-Word ID, Passage Comprehension, and Picture Vocabulary. In the Letter-Word and Passage Comprehension Models, Part 1 of the RRRP significantly predicted spring outcomes (see Table 5), such that students who performed better on Part 1 of fall RRRP also exhibited higher scores on both the Letter-Word ID and Passage Comprehension subtests in the spring. However, for vocabulary, Part 2 of the fall RRRP predicted spring outcomes on the Picture Vocabulary subtest (see Table 5).

Similar to the mathematics models, we also examined whether RRRP scores predicted gains from fall to spring on each of the three reading measures by running three additional two-level models controlling for fall WJ scores (see Table 5). When controlling for fall Letter-Word ID scores, fall RRRP no longer predicted spring outcomes, which indicates that while fall RRRP predicted spring absolute outcomes on Letter-Word ID, it did not predict gains on the Letter-Word ID tests from fall to spring. Interestingly, in the initial Passage Comprehension model, Part 1 of the RRRP significantly predicted spring scores, however, when controlling for fall Passage Comprehension scores, now Part 2 (color switch) of the fall RRRP predicted gains from fall to spring. Finally, in the Picture Vocabulary model, Part 2 of the fall RRRP significantly predicted gains in vocabulary from fall to spring. These results demonstrate that different aspects of the RRRP are more predicative of certain areas of academic achievement. Furthermore, fall RRRP also predicted growth in achievement in both reading and mathematics in the spring.

Discussion

The primary purpose of this study was to examine the relations of a newly designed measure of self-regulation, the RRRP, to independent measures of self-regulation and achievement. Although research has shown that self-regulation is important for school success (Blair & Razza, 2007; Diamond & Lee, 2011), there has been little advancement in how self-regulation is measured, particularly in middle elementary school (third–fifth grade). Previous studies have utilized measures that were originally designed to be administered by clinicians or in a laboratory setting. These types of measures do not necessarily capture the coordination of self-regulation skills and may not be appropriate for use in a school setting (Polderman et al., 2007; Whitebread et al., 2009). The RRRP is potentially useful in a school setting because it can be administered to the general population of students. In addition, self-regulation measures commonly come in the form of questionnaires that are completed by parents, teachers, or children themselves, which may make them open to observer bias (McClelland & Ponitz, 2011; Ponitz et al., 2008). Measures, such as the RRRP, that capture these skills in more a direct way may be more useful and accurate.

Again, self-regulation is defined as involving attentional flexibility, working memory, and inhibitory control skills (Bronson, 2000; Matthews, Ponitz, & Morrison, 2009; Ponitz et al., 2008); however, these skills have been typically measured independently of one another. It may be that it is the coordination of these three skills that are more important for school success and thus measures that assess self-regulation in this way may be a better predictor of academic achievement. It was for these reasons that the RRRP was developed.

The RRRP was associated with independent measures of working memory, attentional flexibility, and inhibitory control. Upon examining the associations between the RRRP and independent measures of self-regulatory skills, it was found that working memory as measured by Memory for Digits Reversed and scores on the SWAN Attention subscale were significantly related to performance on the Part 1 of the RRRP. Part 1 was designed to tap attention and working memory skills. On Part 2 of the RRRP, only the working memory measure was associated. This finding makes sense as Part 2 of the RRRP included the added difficulty of color switching, which was an extra tax on working memory skills. Furthermore, children's ability to wait for "go" on the RRRP was significantly and positively related to teacher's ratings of inhibitory control, which is to be expected because waiting for the examiner to say "go" on the RRRP is also designed to tap inhibitory control skills. These results suggest that the RRRP is tapping the same skills that teacher-reported symptoms of inhibitory control and attention problems on the SWAN and the Memory for Digits Reversed capture. In addition, scores on the RRRP were also fairly stable over time from fall to spring, which may be an indication that self-regulation skills are generally stable in third grade. This finding supports previous findings that demonstrate the stability of self-regulation skills in middle childhood (Harms, Zayas, Meltzoff, & Carlson, 2014).

Another aim of this study was to examine whether the RRRP predicted spring performance in mathematics and reading. Results revealed that different parts of the RRRP were predicative of spring outcomes, depending on the subject. For

Table 5. Hierarchical Linear Modeling Results for Spring Academic Outcomes as Predicted by Fall RRRP (Top) and Academic Outcomes Controlling for Fall Scores (Bottom).

	WJ Spring Applied Problems			WJ Spring Letter-Word ID			WJ Spring Passage Comprehension			WJ Spring Picture Vocabulary		
Fixed effect	Coefficient	*df*	*t*	Coefficient	*df*	*t*	Coefficient	*df*	*t*	Coefficient	*df*	*t*
Intercept	502.743	29	445.75**	505.347	29	314.64**	495.451	29	664.25**	501.433	29	514.10**
Fall RRRP Part 1 total score	1.009	223	3.97**	2.349	223	2.26*	1.471	223	2.67**	1.063	223	1.80
Fall RRRP Part 2 total score	1.302	223	1.46	1.964	223	1.686	1.105	223	1.635	1.635	223	3.47**
Random effects	Variable	*df*	χ^2	Variable	*df*	χ^2	Variable	*df*	χ^2	Variable	*df*	χ^2
Between classroom residual	18.76	29	57.38*	54.80	29	110.19**	7.92	29	54.58**	18.57	29	87.43**
Within classroom residual	161.32			184.59			74.23			82.93		

	WJ Spring Applied Problems			WJ Spring Passage Comprehension			WJ Spring Picture Vocabulary		
Fixed effect	Coefficient	*df*	*t*	Coefficient	*df*	*t*	Coefficient	*df*	*t*
Intercept	502.942	29	629.37**	495.711	29	1,101.22**	501.827	29	869.99**
Fall WJ W score	0.553	222	10.62**	0.611	222	15.16**	0.685	222	16.82**
Fall RRRP Part 1 total score	2.422	222	2.86**	−0.433	222	−0.86	0.054	222	0.12
Fall RRRP Part 2 total score	0.715	222	0.47	0.854	222	2.11*	1.02	222	3.37**
Random effects	Variable	*df*	χ^2	Variable	*df*	χ^2	Variable	*df*	χ^2
Between classroom residual	7.56	29	46.52*	0.68	29	30.524	4.28	29	49.86**
Within classroom residual	21.02			47.52			47.43		

Note. ICC = 11% (Applied Problems), 26% (Letter-Word ID), 13% (Passage Comprehension), 21% (Vocabulary). RRRP = Remembering Rules and Regulation Picture Task; WJ = Woodcock–Johnson III; ID = Identification; ICC = intraclass correlation coefficient.
*$p < .05$. **$p < .01$.

mathematics, Part 1 of the RRRP was a significant predictor of spring Applied Problems scores, such that higher performance on the RRRP in the fall on the spring Applied Problems. Furthermore, when examining gains from fall to spring on the Applied Problems subtest, fall scores on Part 1 of the RRRP remained significant. Part 1 of the RRRP was designed to tap working memory and attention skills, but did not include the added color switching component that Part 2 of the RRRP includes. As we hypothesized, this may indicate that coordination of attentional flexibility and working memory may be implicated in mathematics learning. To be able to do well in mathematics, students must be able to read and comprehend each question, decide the appropriate mathematical operations to use, and to perform these operations correctly, which may be difficult for students with poor attention and working memory skills.

In reading, Parts 1 and 2 of the RRRP significantly predicted spring reading outcomes. Fall scores on Part 1 of the RRRP significantly predicted spring outcomes in both decoding (Letter-Word ID) and reading comprehension (Passage Comprehension). Finally, for vocabulary, only Part 2 of the RRRP was predicative of spring scores. Keeping in mind that Part 2 of the RRRP included an added difficulty of color switching with red and blue blocks, this component may have placed an extra tax on working memory capacity, and additionally, inhibitory control, as children needed to remember to use the opposite color block that the tester indicated. The Picture Vocabulary measure required children to quickly identify pictures. Providing the appropriate word may be difficult for children who have weaker self-regulation skills to effectively think about their answers before providing a response. In addition, working memory might be related to monitoring the speech stream, making it easier to learn vocabulary in context (Gleason, 1997). When examining gains on the reading measures from fall to spring, we found that fall RRRP scores also significantly and positively predicted growth in both reading comprehension and vocabulary.

Taken together, these findings suggest that with some improvements, the RRRP can be a valid measure of self-regulation that examines taps the coordination of attentional flexibility, working memory, and inhibitory control. The RRRP also predicts academic outcomes in reading outcomes, in regard to absolute status as well as gains from fall to spring and may be a useful measure in a school setting. This study also demonstrates that different aspects of self-regulatory skills may be more closely associated with certain skills. For example, Part 1 of the RRRP, which was most closely associated with attentional flexibility and working memory, predicted spring problem solving, decoding, and comprehension. On the contrary, Part 2 of the RRRP, which was most closely associated with working memory, predicted vocabulary as well as gains in comprehension and vocabulary. These findings extend previous studies that have found that weaker self-regulation skills are related to difficulties in vocabulary, and reading comprehension (Connor et al., 2010; Day et al., 2015; H. L. Swanson, 1994). This study also extends prior research that has shown the importance of having strong attention skills in mathematics (Bull & Scerif, 2001; Holmes et al., 2009; H. L. Swanson & Sachse-Lee, 2001). Mathematics requires a great deal of attention, particularly when multiple steps are involved in the problem-solving process. Students with attention problems may make errors out of carelessness in following rules of multistep problems. During instruction, attention problems may cause students to miss important pieces of information that are needed to understand different concepts.

Study Limitations and Future Directions

There are some limitations to the RRRP and to this study that should be noted. First, although the RRRP appears to measure the anticipated construction of self-regulation and predicted academic outcomes, there was some evidence that the RRRP may be too easy for some children. We plan to add more difficult items by increasing the number of rules to remember, and the length of the directions. Finally, the RRRP must be administered to a larger and more diverse sample of students and in other grades.

In terms of feasibility, assessors reported that the RRRP was generally easy to administer, but required paying close attention to what the students were doing. Testers also reported that most students enjoyed the game-like structure of the RRRP and liked using the blocks. Many of our testers were undergraduate students or graduate students in education, and agreed that the test would be simple enough for a teacher to administer. However, a formal procedure for measuring feasibility was not in place at the time of this study, and thus should be noted as a limitation of this work. Future work should consider having assessors complete more formal feedback forms to accurately measure test feasibility.

It was anticipated that children's ability to wait for "go" would be associated with performance on the academic measures. Upon examination of the correlation table, waiting for "go" was only correlated with the SWAN and Memory for Digits Reversed. This finding may be due to the fact that most children were able to wait for "go" and suggests that some alteration of the assessment in this aspect might be warranted. For example, it might be informative to lengthen the wait time between the instruction and saying "go." Also of interest to examine in future work is the importance of asking for repetitions, which was designed to tap working memory, attentional flexibility, and metacognitive awareness. This variable might be used as a categorical variable, given its relative rareness, as indication of children's attention and working memory skills in the service of comprehension monitoring but this remains to be tested.

The potential ceiling effect was also a limitation in this work. We designed the third-grade version of the RRRP to include two objects per prompt, for example, "Put a blue block on the duck behind the tree before you put a red block on the duck flying last." We anticipated that such prompts would be more challenging for third graders. Future versions of the test should focus on increasing the difficulty and complexity of each item, for example, by adding a third object. It is also important to note that the sample was primarily Caucasian, and 45% of this sample qualified for free or reduced price lunch. Different results may be found in a sample that is more economically and ethnically diverse. While we observed a slight decrease in scores on Part 1 of the RRRP from fall to spring, this may provide evidence of either a ceiling effect or that by third grade, these skills are stable for most children and the decrease represents regression to the mean.

Although we intentionally designed the RRRP to use simple vocabulary that we assumed third-grade students would be familiar with, it is also important to acknowledge that children can vary quite significantly in their language skills, particularly in populations with dual language learners. Future versions of this measure should consider establishing a baseline to ensure that students are familiar with the vocabulary used on the RRRP prior to testing. One final limitation to this study was that the SWAN and Memory for Digits Reversed were only given one time in the spring. It would have been useful to give these tests in the fall as well so that we could measure potential change in self-regulation skills over the school year. One other question that still remains is whether or not a direct measure like the RRRP is a *better* predictor of behavioral and academic outcomes than a survey measure like the SWAN. Future work in this area should take this question into consideration as we continue to test which types of behavioral measures are most appropriate in an academic context. Future work should also consider including more direct measures of self-regulation skills for comparison.

Although the current version of the RRRP may need some revisions, this measure does hold promise. There are a few advantages to using the RRRP in place of other measures. First, the RRRP is designed to measure attentional flexibility, working memory, and inhibitory control skills. While previous studies have measured these skills independently of one other, the RRRP can measure these skills in one test. The RRRP can be administered in approximately 7 min, and this may be a quicker way to assess self-regulation skills as opposed to using multiple measures. The RRRP is also easy to give in a school setting, and may also be a more cost-effective way to measure self-regulation skills, as opposed to tests that must be given on computers or that take longer to administer. Although the current version of the RRRP was designed to be given to children in third grade, the difficulty of the questions can be manipulated to accommodate different age groups.

Overall, this study highlights the important role that self-regulation plays in academic achievement in both reading and mathematics. Developing measures that directly assess self-regulation skills in a reliable, valid, and direct way and that can be used longitudinally may help to identify children who have poor self-regulation skills and may be a useful tool in finding better ways to support students' academic achievement.

Authors' Note

The opinions expressed are ours and do not represent views of the funding agencies.

Acknowledgments

We thank the Individualizing Student Instruction Project team members for their hard work collecting the data. We thank the children, parents, teachers, and school administrators without whom this research would not have been possible.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was funded by U.S. Department of Education, Institute of Education Sciences (Grants R305H04013 and R305B070074), "Child by Instruction Interactions: Effects of Individualizing Instruction," and by Eunice Kennedy Shriver National Institute of Child Health and Human Development (Grants R01HD48539, R21HD062834, and P50 HD052120).

References

Achenbach, T. M. (2001). *Manual for the Child Behavior Checklist/4-18*. Burlington: University of Vermont, Department of Psychiatry. doi:10.1007/978-0-387-79948-3_1529

Adams, J. W., & Snowling, M. J. (2001). Executive function and reading impairments in children reported by their teachers as "hyperactive." *British Journal of Developmental Psychology*, *19*, 293–306. doi:10.1348/026151001166083

Baddeley, A. (1986). *Working memory*. Oxford, UK: Clarendon Press.

Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development*, *78*, 647–663. doi:10.1111/j.1467-8624.2007.01019.x

Bronson, M. B. (2000). *Self-regulation in early childhood: Nature and nurture*. New York, NY: Guilford Press.

Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: Inhibition, switching, and working memory. *Developmental Neuropsychology*, *19*, 273–293. doi:10.1207/S15326942DN1903_3

Connor, C. M., Morrison, F. J., Fishman, B., Crowe, E. C., Al Otaiba, S., & Schatschneider, C. (2013). A longitudinal cluster-randomized control study on the accumulating effects of individualized literacy instruction on students' reading from 1st through 3rd grade. *Psychological Science*. Advance online publication. doi:10.1177/0956797612472204

Connor, C. M., Ponitz, C. C., Phillips, B. M., Travis, Q. M., Glasney, S., & Morrison, F. J. (2010). First graders' literacy and self-regulation gains: The effect of individualizing student instruction. *Journal of School Psychology*, *48*, 433–455. doi:10.1016/j.jsp.2010.06.003

Day, S. L., Connor, C. M., & McClelland, M. M. (2015). Children's behavioral regulation and literacy: The impact of the first grade classroom environment. *Journal of School Psychology*, *53*, 409–428. doi:10.1016/j.jsp.2015.07.004

Del Giudice, M. (2014). Middle childhood: An evolutionary-developmental synthesis. *Child Development Perspectives*, *8*, 193–200. doi:10.1111/cdep.12084

Diamond, A., & Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science*, *333*, 959–964. doi:10.1126/science.1204529

Duckworth, A. L., Quinn, P. D., & Tsukayama, E. (2012). What No Child Left Behind leaves behind: The roles of IQ and self-control in predicting standardized achievement test scores and report card grades. *Journal of Educational Psychology*, *104*, 439–451. doi:10.1037/a0026280

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., . . . Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, *43*, 1428–1446. doi:10.1037/0012-1649.43.6.1428

Gathercole, S. E., & Alloway, T. P. (2008). *Working memory and learning: A teacher's guide*. London, England: SAGE.

Gleason, J. B. (1997). *The development of language*. Boston, MA: Allyn & Bacon.

Harms, M. B., Zayas, V., Meltzoff, A. N., & Carlson, S. M. (2014). Stability of executive function and predictions to adaptive behavior from middle childhood to pre-adolescence. *Frontiers in Psychology*, *5*, 331–342. doi:10.3389/fpsyg.2014.00331

Hofmann, W., Schmeichel, B. J., & Baddeley, A. D. (2012). Executive functions and self-regulation. *Trends in Cognitive Sciences*, *16*, 174–180. doi:10.1016/j.tics.2012.01.006

Holmes, J., Gathercole, S. E., & Dunning, D. L. (2009). Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental Science*, *12*(4), F9–F15. doi:10.1111/j.1467-7687.2009.00848.x

Korkman, M., Kirk, U., & Kemp, S. (1998). *NEPSY: A developmental neuropsychological assessment*. San Antonio, TX: Psychological Corporation.

Lakes, K. D., Swanson, J. M., & Riggs, M. (2011). The reliability and validity of the English and Spanish Strengths and Weaknesses of ADHD and Normal Behavior Rating Scales in a preschool sample: Continuum measures of hyperactivity and inattention. *Journal of Attention Disorders*, *16*, 510–516. doi:10.1177/1087054711413550.

Lin, B., Coburn, S. S., & Eisenberg, N. (2016). Self-regulation and reading achievement. In C. M. Connor (Ed.), *The cognitive development of reading and reading comprehension* (pp. 67-87). New York, NY: Routledge.

Matthews, J. S., Ponitz, C. C., & Morrison, F. J. (2009). Early gender differences in self-regulation and academic achievement.

Journal of Educational Psychology, *101*, 689–704. doi:10.1037/a0014240

McClelland, M. M., Cameron, C. E., Connor, C. M., Farris, C. L., Jewkes, A. M., & Morrison, F. J. (2007). Links between behavioral regulation and preschoolers' literacy, vocabulary, and math skills. *Developmental Psychology*, *43*, 947–959. doi:10.1037/0012-1649.43.4.947

McClelland, M. M., & Ponitz, C. C. (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.), *New Directions for Child and Adolescent Development: Vol. 133. Thriving in childhood and adolescence: The role of self-regulation processes* (pp. 29–44). doi:10.1002/cd.302

McGee, R., Prior, M., Williams, S., Smart, D., & Sanson, A. (2002). The long-term significance of teacher-rated hyperactivity and reading ability in childhood: Findings from two longitudinal studies. *Journal of Child Psychology and Psychiatry*, *43*, 1004–1017. doi:10.1111/1469-7610.00228

Molfese, V. J., Molfese, P. J., Molfese, D. L., Rudasill, K. M., Armstrong, N., & Starkey, G. (2010). Executive function skills of 6–8 year olds: Brain and behavioral evidence and implications for school achievement. *Contemporary Educational Psychology*, *35*, 116–125. doi:10.1016/j.cedpsych.2010.03.004

Neubauer, A., Gawrilow, C., & Hasselhorn, M. (2012). The watch-and-wait task: On the reliability and validity of a new method of assessing self-control in preschool children. *Learning and Individual Differences*, *22*, 770–777. doi:10.1016/j.lindif.2012.05.006

Neuenschwander, R., Röthlisberger, M., Cimeli, P., & Roebers, C. M. (2012). How do different aspects of self-regulation predict successful adaptation to school? *Journal of Experimental Child Psychology*, *113*, 353–371. doi:10.1016/j.jecp.2012.07.004

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. 451–502). San Diego, CA: Academic Press.

Polderman, T. J. C., Derks, E. M., Hudziak, J. J., Verhulst, F. C., Posthuma, D., & Boomsma, D. I. (2007). Across the continuum of attention skills: A twin study of the SWAN ADHD Rating Scale. *Journal of Child Psychology and Psychiatry*, *48*, 1080–1087. doi:10.1111/j.1469-7610.2007.01783.x

Ponitz, C. C., McClelland, M. M., Jewkes, A. M., Connor, C. M., Farris, C. L., & Morrison, F. J. (2008). Touch your toes! Developing a direct measure of behavioral regulation in early childhood. *Early Childhood Research Quarterly*, *23*, 141–158. doi:10.1016/j.ecresq.2007.01.004

Ponitz, C. C., McClelland, M. M., Matthews, J. S., & Morrison, F. J. (2009). A structured observation of behavioral self-regulation and its contribution to kindergarten outcomes. *Developmental Psychology*, *45*, 605–619. doi:10.1037/a0015365

Rabiner, D., & Coie, J. D., & The Conduct Problems Prevention Research Group. (2000). Early attention problems and children's reading achievement: A longitudinal investigation. *Journal of the Academy of Child & Adolescent Psychiatry*, *39*, 859–867. doi:10.1097/00004583-200007000-00014

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Newbury Park, CA: SAGE.

Rimm-Kaufman, S. E., La Paro, K. M., Downer, J. T., & Pianta, R. C. (2005). The contribution of classroom setting and quality of instruction to children's behavior in kindergarten classrooms. *The Elementary School Journal*, *105*, 377–394. doi:10.1086/429948

Rothbart, M. K., Ahadi, S. A., Hershey, K. L., & Fisher, P. (2001). Investigations of temperament at 3–7 years: The Children's Behavior Questionnaire. *Child Development*, *72*, 1394–1408. doi:10.1111/1467-8624.00355

Rueda, M. R., Posner, M. I., & Rothbart, M. K. (2004). Attentional control and self-regulation. In R. F. Baumeister & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory, and applications* (pp. 283–300). New York, NY: Guilford Press.

Sáez, L., Folsom, J. S., Al Otaiba, S., & Schatschneider, C. S. (2012). Relations among student attention behaviors, teacher practices, and beginning word reading skill. *Journal of Learning Disabilities*, *45*, 418–432.

Schunk, D. H. (1989). Self-efficacy and achievement behaviors. *Educational Psychology Review*, *1*, 173–208. doi:10.1007/BF01320134

Schunk, D. H., & Zimmerman, B. J. (1994). *Self-regulation of learning and performance: Issues and educational applications*. Hillsdale, NJ: Lawrence Erlbaum.

Skibbe, L. E., Phillips, B. M., Day, S. L., Brophy-Herb, H. E., & Connor, C. M. (2012). Children's early literacy growth in relation to classmates' self-regulation. *Journal of Educational Psychology*, *104*, 541–553. doi:10.1037/a0029153

St. Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *The Quarterly Journal of Experimental Psychology*, *59*, 745–759. doi:10.1080/17470210500162854

Suchodoletz, A., Larsen, R. A., Gunzenhauser, C., & Fäsche, A. (2015). Reading and spelling skills in German third graders: Examining the role of student and context characteristics. *Educational Psychology*, *85*, 533–550. doi:10.1111/bjep.12090

Swanson, H. L. (1994). The role of working memory and dynamic assessment in the classification of children with learning disabilities. *Learning Disabilities Research & Practice*, *9*, 190–202. Retrieved from http://onlinelibrary.wiley.com/journal/10.1111/%28ISSN%291540-5826

Swanson, H. L., & Sachse-Lee, C. (2001). Mathematical problem solving and working memory in children with learning disabilities: Both executive and phonological processes are important. *Journal of Experimental Child Psychology*, *79*, 294–321. doi:10.1006/jecp.2000.2587

Swanson, J. M., Schuck, S., Mann, M., Carlson, C., Hartman, K., Sergeant, J. A., . . . McCleary, R. (2006). *Categorical and dimensional definitions and evaluations of symptoms of ADHD: The SNAP and SWAN Rating Scales*. Retrieved from http://vrosario.bol.ucla.edu/forms/SWAN.pdf

Vitaro, F., Brendgen, M., Larose, S., & Tremblay, R. E. (2005). Kindergarten disruptive behaviors, protective factors, and educational achievement by early adulthood. *Journal of Educational Psychology*, *97*, 617–629. doi:10.1037/0022-0663.97.4.617

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *Comprehensive test of phonological processes (CTOPP)*. Austin, TX: Pro-Ed.

Whitebread, D., Coltman, P., Pasternak, D. P., Sangster, C., Grau, V., Bingham, S., . . . Demetriou, D. (2009). The development of two observational tools for assessing metacognition and self-regulated learning in young children. *Metacognition and Learning*, *4*, 63–85. doi:10.1007/s11409-008-9033-1

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock–Johnson III tests of achievement*. Itasca, IL: Riverside.

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: Lawrence Erlbaum. doi:10.4324/9781410601032