6-29-2021 2:00 PM

Assessing Self-Regulation in Grade 1 Writers: An Evaluation of the Reliability and Validity of the Interview on Self-Regulation in Early Writing

Ashley A. Bildfell, *The University of Western Ontario*

Supervisor: Klein, Perry, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Education

Abstract

Few studies have considered the reliability and validity of the measures used to assess self-regulation in writing. The purpose of the current study was to evaluate the reliability and validity of the Interview on Self-Regulation in Early Writing, a novel task-specific measure designed to assess self-regulation in writing in Grade 1. This eight-item interview was designed to assess the planning, goal setting, strategy, self-statements, self-monitoring, coping, reviewing, and self-reinforcement aspects of self-regulation on separate transcription and composition dimensions. The data from 117 Grade 1 students were used to evaluate inter-rater reliability and the internal consistency of the scale. A principal components analysis (PCA) was used to examine the internal structure of the scale. Discriminant and predictive validity were also assessed. The Interview on Self-Regulation in Early Writing had excellent inter-rater reliability and good internal consistency once unreliable items were removed. The internal structure and discriminate and predictive validity analyses provided support for the validity of this measure. In sum, the current study provided evidence that supported the reliability and validity of the Interview on Self-Regulation in Early Writing. Further, it provided a description of self-regulatory writing behaviours writing in Grade 1 and filled a gap in the literature between content-neutral self-regulation in early childhood and subject-specific self-regulation in later grades.

Keywords: Self-regulation, Writing, Reliability, Validity, Early Writers, Writing Measure, Assessment

Summary for Lay Audience

Writing is a challenging task that requires balancing a great deal of information in the mind at once, including thinking of ideas, organizing thoughts, and remembering information like spelling and punctuation. Self-regulation is the process that helps people manage the information required to write a story. Many researchers have discovered that teaching self-regulation skills helps older students to write better stories, but little work has been done with early writers. Further, no researchers have ever done any work examining if the tools that are used to measure self-regulation measure self-regulation instead of something else. The purpose of this research was to evaluate a new tool called the Interview on Self-Regulation in Early Writing. This study examined the tool's reliability and validity. Reliability refers to whether the interview measured self-regulation consistently, and validity refers to whether the interview accurately measured self-regulation instead of something else. This interview asked Grade 1 students about aspects of self-regulation including their thoughts about planning stories, setting goals for their writing, the strategies they can use to help them write a good story, reviewing their stories, and the things they can say to motivate themselves to keep writing. This study had three groups of participants. One group was taught the regular curriculum from their teachers, one group was taught writing strategies and one group was taught writing strategies and self-regulation skills. This study found that the interview items measured self-regulation consistently. It also found that the interview predicted writing quality after they had received all their lessons and that the group that had been taught about writing strategies and self-regulation had more knowledge about self-regulation measured by the interview. This confirmed that the Interview on Self-Regulation in Writing was reliable (i.e., consistent), and valid (i.e., measured what it was supposed to). This research helped us

to understand early writing behaviors and how to teach students in Grade 1 to write better stories.

Acknowledgements

To my academic supervisor, Dr. Perry Klein, thank you for your teaching, patience, support, and generosity. From your mentorship and guidance, I grew as a student, writer, and researcher. Thank you for being understanding and supportive as I navigated health issues and for fostering resiliency in me when we had to change projects. I sincerely appreciated the feedback you provided me, the way you inspired growth in me and having the opportunity to learn about mentorship and leadership by observing you.

To my faithful committee member, Dr. Lisa Archibald, thank you for your support and feedback on not one, but two research projects. I appreciate the time you invested, the encouragement you provided and the enthusiasm you demonstrated for my work.

To Dr. Colin King, my committee member and clinical supervisor, thank you for checking in on my research progress and for the confidence you demonstrated in my ability to pivot and persist at critical decision points. Thank you also for consistently entertaining my research and clinical ideas and encouraging me to pursue them. Your feedback and support throughout the process was so helpful.

To the research team, Christine Giese, Jill Dombroski, Serena Thompson, and Kristen Sha, thank you for all the hours you dedicated to coordinating this study, and collecting and analyzing the data. This project was a team effort and would not be possible without your hard work and dedication.

To my cohort, Olivia Faulconbridge, Melissa Coyne-Foresi, and Safi Abbas, thank you for investing in a positive "cohort culture" that enabled us to progress through the program collaboratively and collegially. I appreciate each one of you. Olivia, thank you for your friendship, encouragement, and hospitality. I learned so much from you as a person and

a developing psychologist in the last four years. Thank you to you and Matt for opening your home to me as often as needed. “Wednesday night sleepovers” are certainly one of my favourite memories.

To my writing group, Olivia Faulconbridge, Julia Ranieri, and Armush Salahadin, thank you for your consistency and for holding me accountable to my goals. I admire and appreciate each one of you.

To Leah Kennedy and Dakota Graham, thank you for your friendship, steadfastness, encouragement and “Wine and Feelings”.

To my family, Donna, Bob, and Bobby Hamilton, Joshua Green, and Patty, and Gord Kent, thank you for your support of and patience with me throughout my education. Mom, thank you for helping me keep things in perspective and see the humour in stressful situations. Aunty, thank you for always being there, for any and every reason. Whether I needed a laugh, a walk, to talk or to vent, you are always there, and I am so thankful for that. Bobby, your outlook on life is so inspiring. I appreciate you for always listening and for your fresh take on life’s challenges. Gord, thank you for the best advice I have ever been given. Remembering, “… how to eat an elephant” helped motivate me to get this done. Dad, thank you for the efforts you have made to be there for me. I appreciate all of you.

To my husband Jacob Bildfell, to thank you would not be enough to express how much your support has meant to me for the past 11 years. Your unwavering belief in my potential has inspired me to keep striving for more. I look up to you, Jacob. Your work ethic, standards and drive have shown me what is possible with that type of commitment. I am so thankful we are a team and appreciate everything you have done for me including the sacrifices you have made to make this dream a reality.

Finally, to the children who shared their thoughts about stories, because of you, I now know that pages, capitals, periods, unicorns, dinosaurs, and ninjas are required to make a story really good.

Table of Contents

List of Tables

List of Figures

List of Appendices

Preface

Writing is a cyclical, goal-directed process governed by self-regulation (Hayes & Flower, 1980). Since self-regulation has been included in models of writing, topics such as knowledge of the development of skill-specific self-regulation, the impact of self-regulation on writing quality, and the efficacy of Self-Regulated Strategy Development interventions have been well-researched. However, researchers have only recently begun to consider the reliability and validity of the measures that have been developed to assess self-regulation in writing. This raises the question, "if the reliability and validity for the measures we have used to assess self-regulation in writing are unknown, how can we confidently attribute writing research findings to the construct of self-regulation itself?" For that reason, the purpose of the current study was to examine the reliability and validity of a novel measure designed to assess self-regulation in Grade 1 titled: The Interview on Self-Regulation in Early Writing.

In Chapter 1, relevant theories and empirical research on measurement, self-regulation, writing, and the self-regulation of writing will be presented and reviewed. This will be followed in Chapter 2 by a discussion about the development of self-regulation in writing and a review of relevant research findings on what is known about self-regulation in skilled versus less skilled writers and the impact of self-regulation on writing quality. The most common methodologies for assessing self-regulation in learning, self-report questionnaires and think-aloud protocols, will be presented, followed by a review of the literature on self-report questionnaires used to assess self-regulation in writing. These measures will be critiqued and evaluated according to current theories of self-regulation in writing. Chapters 3. Chapters 4 and 5 will then present the methodology used in the current study to evaluate the reliability and validity of the Interview on Self-Regulation in Early Writing, present the results, and discuss the findings, respectively.

Chapter 1

1 Theoretical Perspectives

The purpose of the current study was to examine the reliability and validity of the Interview on Self-Regulation in Early Writing. To understand the methods used and the justification for the creation of the measure, relevant theories will be presented and reviewed. In this chapter, measurement theory will be presented first followed by cognitive theories of self-regulation, writing, and self-regulation of writing. Finally, the Self-Regulated Strategy Model of instruction will be discussed.

1.1 Measurement Theory

In this section, theories of measurement will be reviewed to provide background information on the approach taken to evaluate the reliability and validity of the Interview on Self-Regulation in Early Writing. This information is provided to give context to the current study.

Measurement is an essential part of science (DeVillis, 2016, p. 2). Measurement or assessment is one of the most difficult challenges facing psychology and the behavioural sciences. Ultimately, assessment is organized observations of different processes including visible behaviours and hidden mental states or skills. For that reason, reliability and validity are paramount. Assessment tools with strong reliability and validity enable psychologists and scientists to conduct scientifically rigorous research, make informed decisions on matters that have a wide social impact and interpret test results in ways that are fair and unbiased (Bornstein, 2011).

Measures serve as proxies for unobservable variables (DeVillis, 2016, ch. 2). The latent variable is both unobservable and able to change with respect to strength or

magnitude. While the latent variable cannot be directly observed, it is assigned a value meant to represent the true score. The inherent assumed relation between the latent variable and a measure means the two should be empirically related (DeVillis, 2016, ch. 2). To better understand how to create a reliable and valid measure, theories of reliability and validity are presented and discussed.

1.2 Reliability

Reliability is the extent to which a measure performs in a consistent and predictable fashion; it represents a core issue in measurement (DeVillis, 2016, ch. 3). One way of considering reliability is to consider the internal consistency of a measure. Internal consistency is, "the homogeneity of the items within a scale" (DeVillis, 2016, p. 42). If the items of a scale are strongly related to each other, this means they should also be strongly related to the latent variable. This is assessed by examining the strength of the correlations between items on a scale. Internal consistency is measured using Cronbach's alpha.

1.3 Inter-Rater Reliability

Inter-rater reliability is the, "extent to which raters generate scores that correspond" (DeVillis, 2016, p. 67). Pearson's correlation coefficient can be used to examine the extent to which, "the frequency of exact agreements between judges exceeds what could be expected by chance" (DeVillis, 2016, p. 67). This type of reliability is important to ensure coding systems are consistent, replicable, and unbiased.

1.4 Validity

Validity is the extent to which the latent variable represents the true score (DeVillis, 2016, ch. 4). Two seminal articles have shaped our traditional understanding

of validity: Cronbach and Meehl's (1955) "Construct Validity in Psychological Tests" and Campbell and Fiske's (1959) "Convergent and Discriminant Validation by the Multitrait Multimethod Matrix." Cronbach and Meehl (1955) operationalized the validity statistic *r*. This statistic represents the magnitude of the relation between a predictor variable and an outcome variable.

Since that time, researchers have proposed that validity should be evaluated by: (1) Content validity or the extent to which items on a measure represent the latent variable that the tool seeks to assess, (2) Criterion-related (also known as predictive) validity or the extent to which a measure predicts a outcome assessed by another measure and (3) Construct Validity or the extent to which the measure can be used to make inferences about the theoretical construct measured (DeVillis, 2016, ch. 4). Convergent and discriminant validity are considered subtypes of criterion-related validity. Convergent validity is the extent to which two measures that should be related are related and discriminant validity is the extent to which two measures that should not be related, are not.

Bornstein (2011) proposed an additional method of assessing validity; the process-focused model. According to this model, validity is defined as "the degree to which respondents can be shown to engage in a predictable set of psychological processes during assessment, with those processes dictated a priori by the nature of the instrument(s) used, and context in which it takes place" (p. 532). This method of assessing validity shifts from correlational to experimental methods which allows researchers to better understand the relations between variables. In sum, the process-focused model of validity emphasizes process over outcome and experimental methods

over correlational methods. The current study uses this methodological approach to examine the validity of the Interview on Self-Regulation in Early Writing in Grade 1.

Hogan and Agnello (2004) reviewed 696 research reports to evaluate the types of validity evidence considered for each measure. Correlations between the measure of interest and another self-report measure were used to assess validity in 87% of cases. Behavioural outcome criteria were only used in 5% of cases and no cases used experimental methods to evaluate validity. Similarly, Cizek et al., (2008) found in a review of 283 studies, only 1.8% assessed response process. Borstein (2011) confirmed these findings in a review of the methods used in leading validity journals between 2006 and 2008. Despite the theoretical shift in thinking on how to assess validity, 91% of studies used correlational methods only to assess validity and only 9% used experimental methods. This highlighted the need for more rigorous methods when considering the validity of a newly constructed measure.

1.5 Best Practices for Assessing the Validity of a Measure

According to the Standards for Educational and Psychological Testing (2014), the best practices for assessing the construct validity of a measure include evidence based on content, response process, internal structure, relations with other variables and the consequences of testing. Evidence based on content examines the extent to which a questionnaire aligns with the most current empirical research on the topic, in this case, self-regulation in writing. Evidence based on response process evaluates if: (1) Students understand the items and types of responses required; (2) Students must access related information from their memories; (3) Students must integrate recalled information into a coherent response; (4) Students must match their recalled knowledge to a response and

(5) Students must select/produce the correct answer. Evidence based on internal structure means that the relations between items or dimensions on an assessment are consistent with current theory. This is generally tested with an exploratory and/or confirmatory factor analyses to evaluate the internal structure of a questionnaire.

Evidence based on relations with other variables is composed of three elements: convergent relations, discriminant relations, and predictive relations. Convergent relations mean that the measure in question shows a strong, positive relation with another established behavioural measure of the same theoretical construct. Discriminant relations mean that no relation is found between the measure in question and a theoretically dissimilar construct. Predictive relations mean that the measure in question can be used to predict theoretically related constructs. Lastly, evidence based on consequences of testing means that any interpretations or consequences that result from the measure are theoretically sound (Wolters et al., 2018).

The theories reviewed guided the approach used to assess the Interview on Self-Regulation in Early Writing in the current study. Now that we have reviewed relevant methodological considerations for measure development, we will turn to a discussion on cognitive theories of self-regulation, writing, and the self-regulation of writing.

1.6 Social Cognitive Theory of Self-Regulation

According to Bandura (1991), self-regulation is a complex and dynamic process that enables individuals to manage personal, environmental, and behavioural factors. In his seminal paper, the author delineated the *structure* and *function* of self-regulatory systems. Self-regulatory systems were structured into three subprocesses: (1) self-observation, (2) judgmental process and (3) self-reaction. Self-observation, also known as

self-monitoring, is an internal evaluation of thought patterns, behaviors, and performances. This subprocess is highly influenced by pre-existing cognitive structures, self-beliefs, and mood. The judgmental subfunction is the process of comparing what one observes about their internal thoughts, behaviours and performances to their personal standards, and self-reaction is the process of how self-judgements and comparison to personal standards are translated into action. This could include tangible outcomes, self-reflection, or self-incentives. Together, the *function* of these self-regulatory mechanisms is to enable personal agency by impacting thoughts, behaviours, and motivation.

The self-efficacy mechanism represents the extent to which an individual believes they are capable and in control of both internal and external factors that influence their functioning. This influences their choices, goals, effort, how they handle difficulties, stress, and coping. This mechanism also influences how individuals perceive successes and failures. Individuals with higher self-efficacy are more likely to set higher goals for themselves and persist to achieve them. While there is a bidirectional relation between the self-regulatory structures and functions, understanding their unique role is an important distinction to make as we transition to discussing assessing self-regulation in writing.

1.7 The Cognitive Model of Writing

Writing is a cyclical, goal-directed activity that requires high levels of self-regulation to manage competing demands. In Hayes and Flower's (1980) influential model of cognitive processes in writing, the authors emphasized that writing cannot be conceptualized as a sequence of stages but as distinct elements that interact and shift over the course of the writing process. In their model, three distinct processes for text

composition, planning, translating (i.e., the transformation of oral language into written form) and reviewing were identified. In addition to the core components of writing, Hayes and Flower (1980) also introduced the concept of self-regulation into their cognitive model of writing. The self-regulatory structures they addressed were goal setting, monitoring, the task environment, and long-term memory.

In their seminal paper, the authors explained that monitoring is the self-regulatory process that alerts the writer when it is time to shift from one phase of the writing process to another (e.g., from planning to translating). Another key element of this model was emphasizing that writing is a goal-directed activity and acknowledging that the act of writing itself may stimulate new writing objectives. In sum, Hayes and Flower (1980) included self-regulation as a key component of their cognitive model of writing and highlighted that self-regulatory processes are essential to facilitate the cyclical and goal-directed nature of the writing process. The structures the authors identified laid the groundwork for future researchers to expand on how self-regulation drives the writing process.

1.8 The Not So Simple Model of Writing

The more current, Not So Simple Model of Writing, added further nuances to our understanding of the writing process to include the interplay between the mental environment where ideas are generated and the external environment where writing is produced. According to this model, writing involves three primary types of processes: *text generation*, *self-regulation*, and *transcription* (Berninger & Chanquoy, 2012). The first process, *text generation*, is rooted in oral language and incudes the production of words, sentences, and discourse. The second process, *self-regulation*, is the ability to be

strategic when writing, stay motivated, navigate problems that may be encountered and actively monitor writing quality. The third process, *transcription*, is the act of translating oral language into written language which includes the production of both handwriting and spelling. These three processes provide an explanation for how ideas from within our mental environments can be expressed in the external environment.

Continuing the development of theories on writing development, Kim et al. (2017) argued that the Simple and Not So Simple Views of Writing lacked specificity with regards to the relation between component skills and particularly, text generation. Their model, the Direct and Indirect Effects Model of Developmental Writing further nuanced our understanding of the writing process by using structural equation modeling to elaborate on previous models in an empirically based way. Based on their findings, discourse-level oral language and transcription skills (handwriting fluency and spelling) were found to be directly related to writing. In contrast, working memory affected writing through the mediation of oral language skills (vocabulary and grammatical knowledge), higher order cognitive skills (inference and theory of mind) and transcription skills (handwriting fluency and spelling). Given the substantial, albeit indirect, effect of working memory on writing, working memory was indicated as a key cognitive ability that contributes to writing.

Taken together, writing is a process that requires text generation, self-regulation and transcription skills which are grounded in working memory. The Self-Regulation Model of Writing will be presented and discussed next.

1.9 The Self-Regulation Model of Writing

While Hayes and Flower (1980) and Berninger and Chanquoy (2012) included the general concept of self-regulation in their model of writing, they did not consider all the ways individual self-regulatory subprocesses are essential to the composing process. In Zimmerman and Risemberg (1997), the authors proposed three classes and ten processes in their Self-Regulation Model of Writing. While it was proposed more than two decades ago, it is still considered an influential model in the present day. Derived from social cognitive and self-regulation theory (Zimmerman, 1989), their model divided self-regulation into three forms: environmental, behavioural and personal/covert. The authors expanded on Flower and Hayes' (1980) theory and posited that these self-regulatory processes interact in a cyclic feedback loop allowing writers to monitor their writing and respond based on the feedback they gathered. This theory influenced the development of the Interview on Self-Regulation in Early Writing.

1.9.1 *Environmental*

The environmental self-regulatory processes include: (1) Environmental structuring and (2) Self-selected models. Environmental structuring consists of choosing, organizing, and generating successful settings and situations to produce written work. Self-selected models include resources writers can use to gain knowledge of writing and the skills required to do so. These may include sample texts or writing styles of people they admire.

1.9.2 *Behavioural*

The behavioural self-regulatory processes include: (3) Self-monitoring, (4) Self-consequenting, and (5) Self-verbalizations. Self-monitoring consists of observing and

being aware of one's progress. This may include tracking the amount of work produced or goals met. Self-consequenting is the process of incentivizing oneself for meeting goals or punishing oneself for not meeting goals and self-verbalizations include things writers say to themselves such as reading their text out loud to help them edit or using positive self-talk to help them overcome barriers to writing.

1.9.3 *Personal/Covert*

The personal/covert self-regulatory processes include: (6) Time planning and management, (7) Goal setting, (8) Self-evaluative standards, (9) Use of cognitive strategies, and (10) Use of mental imagery. Time planning and management involves effectively predicting and accounting for the time required for writing and goal setting requires establishing short- and long-term writing goals with respect to aspects like length and quality. Self-evaluative standards are the personally determined criteria each individual sets for themselves about the structure, content and quality of their writing. Use of cognitive strategies consists of rule-bound methods to organize, generate, and edit writing, and use of mental imagery is the process of recalling or creating a mental picture of a setting, action, or character to facilitate writing about it.

In sum, Zimmerman and Risemberg (1997) integrated Bandura's (1991) social cognitive theory and Hayes and Flower's (1980) model of cognitive processes in writing to specify the self-regulatory processes required to produce a quality composition. This theory clarified our understanding of self-regulation in writing and was used to help develop the Interview on Self-Regulation in Early Writing.

1.10 The Self-Regulated Strategy Development Model

While the focus of the current study is the Interview on Self-Regulation in Early Writing, the interview was administered to students who participated in a larger instructional study. Administering the interview in an instructional setting enabled the researchers to test the sensitivity of the interview to the effects of self-regulation instructions. This was one method used to test the validity of the interview in the current study. For that reason, more information about the Self-Regulated Strategy Development Model (SRSD; Harris & Graham, 1996), the methods they used to teach self-regulation, and the self-regulatory processes taught will now be provided. In the literature, there is a consensus that writing requires self-regulatory processes. However, the question that remains is how to teach self-regulation to developing writers. A strategy instruction approach was initially designed to help children with learning disabilities who have been shown to struggle with maladaptive attributions (e.g., attributing poor performance on a math test to being stupid as opposed to needing more time to understand the concepts) and learned helplessness (i.e., a belief that effort and successful outcomes are unrelated). This approach outlined that if students were struggling with effective strategies to complete academic tasks, then the solution was to *teach* them effective strategies to complete academic tasks (Reid et al., 2013).

A strategy is a willful and effortful tool used to facilitate performance. Practically, it can be defined as, “a series of ordered steps that helps a student perform a task” (Reid et al., 2013, p. 16). Knowledge of information processing, and specifically the role of working memory, is important for strategy instruction. Working memory is where information is, “temporarily stored, processed and manipulated” (Baddeley, 1986, 1996).

It consists of three components: (1) The phonological loop manages spoken and written verbal information, (2) The visuospatial sketchpad manages spatial information and (3) The central executive manages planning, organizing and problem solving. When a student's working memory is overloaded, they are no longer able to receive or process new information. For that reason, considering the amount and timing of the information presented to students in strategy instruction is important for learning.

Student attributional styles, or how they interpret successes and failures and use self-talk to cope, also have a major impact on learning (e.g., Tabassam & Grainger, 2002). Weiner (1979) proposed three attributional categories: (1) Internal versus external, (2) Controllable versus uncontrollable and (3) Stable versus unstable. The first attributional style considers whether students attribute their successes and failures to factors inside or outside of themselves. The second attributional style considers the extent to which the student believes the success or failure was within their control. The third attributional style considers whether a student views their pattern of successes and failures as constant or unchanging over time.

Students with maladaptive attributional styles tend to attribute failures to internal, uncontrollable factors and successes to uncontrollable, external factors (e.g., Tabassam & Grainger, 2002). For this reason, these students are at higher risk of developing a negative attitude towards learning, avoiding academic tasks, and experiencing school-related anxiety. Further, they tend to use negative self-talk, (e.g., "I'm stupid", "I can't do it", "this is impossible"). In contrast, students with a positive attribution style are more likely to attribute successes and failures to internal, controllable factors (e.g., "I'm trying my best", "I haven't had enough practice do this yet", "this is hard, but I can take my

time"). This makes them more likely to persist on challenging tasks. By teaching students to attribute their successes and failures to their own efforts and to use their strategies, teachers can encourage a positive attributional style (e.g., Tabassam & Grainger, 2002).

With respect to instruction, self-regulation is how students learn to manage cognitive (i.e., thinking) and metacognitive (i.e., thinking about thinking) processes (Reid et al., 2013, p. 29). Teaching students the how, when, where and why behind instruction helps to maintain and generalize strategy use. Students also need to be able to self-regulate their efforts to maximize the efficacy of their strategy use. In the SRSD model, the goal is for students to understand where and when to use a strategy, monitor the use of the strategy to evaluate whether it was effective, block negative thoughts that could hinder their performance, internalize the belief that using their strategies makes them better thinkers, and practice using the strategy enough for it to become automatic.

Graham and Harris' (1996) Self-Regulated Strategy Development Model (SRSD) is the most influential model of this type of instruction (Reid et al., 2013). It has extensive research support, considers cognitive, motivational, and academic characteristics of learners, is used in combination with self-regulation strategies to help maintain student focus, effort and motivation and is practical to implement in a classroom environment (Reid et al., 2013). The Self-Regulated Strategy Development Model has six stages: (1) Developing and activating background knowledge, (2) Discussing the strategy, (3) Modeling the strategy, (4) Memorizing the strategy, (5) Supporting the strategy and (6) Independent performance. Each step of the model will now be described in more detail.

1.10.1 *Developing and Activating Background Knowledge*

Stage 1 has two primary goals: (1) Define the skills student need to use a strategy and (2) Assess student knowledge and/or ability of the defined skills (Reid et al., 2013). Defining the skills students need to use a strategy can be done by breaking a task down into *steps* and *skills*. Helpful tools to assess student knowledge in this stage include observation, curriculum-based measurements, and interviewing students about what they are doing by asking how and why. To illustrate, before teaching a long-division strategy, students must be proficient in multiplication and subtraction. If multiplication or subtraction were not mastered by the student, they may not have the working memory resources available to learn a long-division strategy. Therefore, the purpose of this stage of the strategy is to ensure students have the defined skills they require to use the strategy through assessment (Reid et al., 2013).

1.10.2 *Discussing the Strategy*

The purpose of stage 2 of the model is three-fold: (1) Acquire student "buy-in" by "selling" the strategy, (2) Help students develop an awareness of their current level of performance, and (3) Introduce the steps of the strategy (Reid et al., 2013). The purpose of this method of instruction is to help students become self-regulated learners. To do so, they need to understand the logic behind why they are being asked to use a strategy and how it will benefit them (e.g., improved performance, fewer recesses spent inside to correct work etc.). If students do not see the utility in a strategy, they are unlikely to use it. Establishing a baseline level of performance with students and graphing their progress can serve as tangible evidence of the effectiveness of the strategy and serve to reinforce strategy use. Once students understand why they are being asked to use a strategy and

their current ability level, it is time to introduce the steps of the strategy, emphasizing what, how and where each step of the strategy is useful. Key to this stage of the strategy is promoting the message that growth and good outcomes are directly linked to effort and strategy use (Reid et al., 2013).

1.10.3 *Modeling the Strategy*

Modeling the strategy is arguably one of the most important steps of the process (Reid et al., 2013). In this stage of the model, the teacher models the strategy as an "expert" and demonstrates the metacognitive skills required for effective strategy use using "think aloud" procedures. By verbalizing their inner thought processes, the teacher emphasizes the why and how of strategy use to their students. This stage requires the teacher to go beyond just repeating or listing the steps of the strategy. In the literature, modeling the strategy with an emphasis on the metacognitive processes of effective learners is essential. Despite the importance of this step, teachers can find it challenging to identify and vocalize automatic thoughts and actions for their students. To help them identify self-regulatory and metacognitive processes within themselves, teachers can ask themselves these four questions: (1) "Why am I doing this step in the task?" (2) "How did I know to do it?" (3) "What are the important actions, cues or questions?" and (4) "What knowledge do I need?" (Reid et al., 2013, p. 39).

1.10.4 *Memorizing the Strategy*

To use a strategy effectively, students must have the steps committed to memory and be able to recall them automatically (Reid et al., 2013). It is important to note that strategy memorization includes an understanding of how and why each step of the strategy needs to be applied, not just being able to regurgitate them. In contrast to

modeling the strategy, memorizing the strategy is regarded as the easiest step of the process. Strategy memorization can be achieved through classroom games and repetition (Reid et al., 2013).

1.10.5 *Supporting the Strategy*

The goal of this critical step of the process is to have teachers and students work collaboratively to help students be able to perform the strategy effectively and independently (Reid et al., 2013). This is accomplished through scaffolding where the teacher initially completes majority of a task while asking for student input and gradually withdraws their support until the student is performing the more difficult steps of the strategy independently. Essential elements of this stage of the model include ensuring that the release of responsibility from teacher to student is *gradual* and providing students adequate *time* and *support* to achieve mastery of the strategy. This can be accomplished through content, task, and material scaffolding. Content scaffolding involves using easy content to introduce a strategy (e.g., one level below student ability level), using interesting or familiar content and having students perform easier steps of the strategy initially (i.e., while the teacher performs the more difficult steps). Once the students become familiar with the steps, they can progress to completing the more difficult steps of the strategy themselves as well. Task scaffolding consists of strategy performance in a group setting gradually shifting from the teacher to the students (e.g., initially, the teacher elicits the required strategy step from the students, describes it and models using it for them. Then, the teacher elicits the required strategy step from the students, has the students describe the strategy and models using it for them). Finally, the teacher elicits the required strategy step from the students, has the students describe the strategy, and

model how to use it). Material scaffolding involves the gradual fading of prompts or other supporting materials (e.g., a strategy poster that lists all the steps is covered up one step at a time). The purpose of this phase of the model is to ensure students can use the strategy effectively and independently (Reid et al., 2013).

1.10.6 *Self-Regulation Strategies*

Reid et al. (2013, p. 71) described self-regulation strategies that can be effectively taught to students with learning disabilities. These included self-monitoring, self-evaluation, self-instruction, goal-getting and self-reinforcement, each of which will be discussed in turn. These self-regulation strategies are assessed using the Interview on Self-Regulation in Early Writing and are therefore the focus of the current study.

Self-monitoring. Self-monitoring behaviour results in an internal evaluation of the action as successful or unsuccessful (Reid et al., 2013). If the action was successful, the individual provides themselves positive reinforcement for example by saying, “Great work!” which makes it more likely for them to repeat the action in the future. If the action was unsuccessful, the individual provides themselves negative reinforcement for example by saying, “oops” which makes it less likely for them to repeat the action in the future. With respect to writing, self-monitoring can be used to help the student assess whether they have included all the parts of a story (Reid et al., 2013).

Self-evaluation. Self-evaluation is similar to self-monitoring, but there are two key distinctions (Reid et al., 2013). First, if self-monitoring involves an individual observing their behaviours, self-evaluation requires an *accurate* assessment of one’s behaviour to receive reinforcement. Second, self-evaluation often involves a student comparing their self-assessments to an external standard (Reid et al., 2013).

Self-instruction. Self-instruction is the use of self-talk to regulate behaviour and self-talk can be taught to help students learn how to talk themselves through a behaviour (Reid et al., 2013). Self-instruction can be applied at two levels: general and task specific. Self-instruction at the task-specific level would not for example generalize from a writing to a math task. There are six functions of self-instruction: (1) Problem definition (e.g., "OK, what do I need to do now?"), (2) Focusing attention/planning (e.g., "I need to take my time and concentrate"), (3) Strategy related (e.g., "I need to remember to use my strategy"), (4) Self-evaluation (e.g., "I need to check and see how I am doing"), (5) Coping (e.g., "I can do this if I keep at it") and, (6) Self-reinforcement (e.g., "I did it! Great job!"; Reid et al., 2013).

Goal setting. Goal setting is an important self-regulatory behaviour (Reid et al., 2013). Goals serve three primary functions: (1) Goals define specific targets and help determine what effort is required to achieve a goal, (2) Goals help an individual evaluate progress and (3) Goals are both motivating to achieve and reinforcing once achieved. Effective goals meet specificity, proximity, and difficulty parameters. Said otherwise, effective goals are well-defined, able to be accomplished in a timely fashion and are not too easy or too difficult (Reid et al., 2013).

Self-reinforcement. Self-reinforcement is when an individual gives themselves a reward or consequence after evaluating their performance against pre-determined criteria (Reid et al., 2013). For example, this could mean a writer rewards themselves with a special dessert after meeting their weekly writing goals.

1.11 Conclusion

In sum, a measure must be reliable and valid to assess the intended construct, which in this study, is self-regulation in early writing. Self-regulation is an integral part of the writing process that is often overlooked in instruction. However, the Self-Regulated Strategy Development Model provides a framework to incorporate self-regulation instruction into teaching. Now, we turn to a literature review relevant to developing a measure to assess self-regulation in writing.

Chapter 2

2 Literature Review

The purpose of this chapter is to review the literature on the development of self-regulation in writing, the relation between self-regulation and writing quality, assessment of self-regulation in learning and the research on assessing self-regulation in writing.

2.1 Development of Self-Regulation in Writing

While Hayes and Flower (1980) emphasized that writing is a goal-directed activity, McCutchen (1995) found that early writers do not produce goal-directed compositions. With that said, preliminary research supports that self-regulation plays an important role in early writing skill development. To illustrate, Kent et al. (2014) studied the relation between early writing and component skills in kindergarten and writing outcomes concurrently and longitudinally in Grade 1 in a sample of 265 students. The authors examined a writing model that included self-regulation (i.e., attention), reading, spelling, handwriting fluency, and oral language component skills. The attention regulation variable was measured using a teacher-report scale designed to measure symptoms of attention-deficit/hyperactivity disorder. This study measured attention-memory, attention set-shifting, and attention-inhibitory control. While all three aspects of self-regulation are relevant to student writing behaviour, attention during writing was not specifically assessed. The structural equation model was better fitting when self-regulation (i.e., attention) was included as opposed to a model that just included reading and spelling. Self-regulation (i.e., attention) was positively related to writing fluency in kindergarten and predicted writing fluency and quality in Grade 1. Similarly, a study by Gerde et al. (2012) found that controlling for letter knowledge, home literacy

environment, and decoding, self-regulation explained 7.1% of the variance in kindergarten children's name-writing skills. In this case, self-regulation was measured using the Head-Toes-Knees-Shoulders task. This consisted of assessing elements of behaviour regulation including working memory, attention, inhibitory control, and task persistence. Taken together, this evidence suggests that self-regulation may play an important role in writing development (Puranik et al., 2019). However, neither study measured self-regulation during writing behaviours.

2.1.1 *Writers become more self-regulated with age and education*

Comparable to the development of other skills, writing-specific self-regulatory behaviours increase with age, schooling and writing skill development (Berninger et al., 1994; Berninger et al., 1996, Gerde et al., 2012). To illustrate, Bereiter and Scardamalia (1987) found that between Grades 4 and 6, the amount of planning notes doubled, and conceptual planning improved slightly from Grades 4 to 8. An additional study by Boscolo (1990) found that while planning notes for children in Grades 2 to 4 were generally copied sentences with few significant changes, planning notes for children in Grades 6 and 8 were more sophisticated (i.e., reminders, summaries, or synthesized information).

Revising behaviour has also been demonstrated in the literature to increase in frequency and quality with age and experience (Fitzgerald, 1987). To illustrate, in a longitudinal study, Limpo et al. (2014) found that planning and revising abilities grew from Grades 4 to 9 in a sample of 381 students controlling for gender, school achievement, age, handwriting fluency, spelling and text structure. Interestingly, the authors found planning and revising were only related to writing quality in older students.

Given that self-regulation in writing improves with age and schooling, one would expect that self-regulation predicts writing quality- and it does.

2.2 Self-Regulation Predicts Writing Quality

Santangelo et al. (2016) conducted a meta-analysis on Zimmerman and Risemberg's (1997) model of self-regulatory processes. The goal of this meta-analysis was to examine the impact of self-regulation instruction on writing quality. For that reason, effect sizes for writing quality were calculated. The final sample included 78 studies published between 1963 and 2014 (35 studies used a random design and 49 were quasi-experimental; 49 were peer-reviewed journal articles, 25 were dissertations, and seven were either conference papers, book chapters or unpublished manuscripts). Of the ten self-regulatory processes described by Zimmerman and Risemberg, five were studied in at least four papers that met inclusion criteria. The five self-regulatory processes examined in this meta-analysis included: (1) Self-selected models, tutors, or books, (2) Goal setting, (3) Self-evaluative standards, (4) Cognitive strategies and (5) Mental imagery. Cognitive strategies were further divided into cognitive strategies instruction and prewriting.

2.2.1 *Models, Tutors or Books*

Seven studies that included 1,217 students in Grades 4, 5, 6, 8, 9 and 12 with representative literacy skills examined the impact of teaching students to use models, tutors, or books to improve their writing. In the intervention conditions, students were asked to seek and analyze various example texts such as books and teacher or peer compositions. What was emphasized in the teacher's lessons varied from general writing characteristics to specific linguistic techniques. Comparison conditions were taught

writing without models. All seven studies found that the use of models improved the quality of student writing (ES = 0.30).

2.2.2 *Goal Setting*

Eight studies that included 429 students in Grades 4-8 with representative literacy skills and struggling students examined the impact of teaching goal setting on writing quality. In the intervention conditions, students were taught to strive to meet specific goals pertaining to drafting or revising their texts (e.g., include three pieces of supporting evidence). Generally, the goals were determined by the teachers. Comparison conditions were provided with broad goals (e.g., write a good essay). All eight studies found that goal setting improved the quality of student writing (ES = 0.73).

2.2.3 *Self-Evaluation*

Twelve studies that included 1,326 students in Grades 2-8 and 10-12 with representative literacy skills examined the impact of self-evaluation on writing quality. Students in the intervention conditions were taught to self-evaluate using a rubric or guide to follow. Comparison conditions varied but ranged from students being given no instruction to being provided a rubric but without instructions on how to use it. Eleven of the twelve included studies found that self-evaluation improved the quality of student writing (ES = 0.51).

2.2.4 *Cognitive Strategy Instruction*

Thirty-eight studies that included 3,268 students in Grades 2-10 examined the impact of cognitive strategy instruction on writing quality. Students in the intervention conditions were taught to use cognitive strategies in 25 of the 38 studies using the self-regulated strategy development (SRSD) model (e.g., Graham & Harris, 2005).

Comparison conditions were taught writing without cognitive strategies. All 38 studies found that teaching cognitive strategies improved student writing quality (ES = 1.06). It is worth noting that SRSD instruction, a type of cognitive strategy instruction, has been demonstrated to produce effect sizes that are significantly larger than non-SRSD instruction. This suggests that aspects of SRSD programming, such as self-monitoring and self-statements, contribute to writing over and above the strategy steps alone. To illustrate, Graham et al., (2012) found an average weighted ES = 1.14 for SRSD instruction compared to 0.59 for non-SRSD instructional approaches.

2.2.5 *Prewriting*

Thirteen studies that included 1,216 students in Grades 2-12 with representative literacy skills examined the impact of prewriting on writing quality. Students in the intervention condition used drawing, graphic organizers, videos, and relevant materials to help them with the prewriting process. Comparison conditions did not use prewriting at all or used a much less structured prewriting activity. All 12 studies found that prewriting improved student writing quality (ES = 0.55).

2.2.6 *Mental Imagery*

Four studies that included 293 students in Grades 3-6 with representative literacy skills and struggling students examined the impact of mental imagery on writing quality. Students in the intervention conditions were taught to use mental imagery to enhance creativity or character descriptions. Comparison conditions used alternative instruction in three studies and no instruction in one. All four studies found that mental imagery improved student writing quality (ES = 0.76).

Overall, teaching self-regulation was found to improve student writing quality. However, the behaviours the authors considered self-regulatory warrant further discussion. Cognitive processes are not inherently self-regulatory or metacognitive. For example, in prewriting, if a student had practice with graphic organizers, independently chose to use this strategy to help them plan a narrative, and monitored their work, this behaviour would be considered self-regulation. Alternatively, if a teacher provided the student with a graphic organizer as a required part of an assignment, while completing the organizer would facilitate planning, this would not be considered a self-regulatory action on the part of the student. The same argument can be applied to mental imagery. If a student actively and independently chose to use a mental imagery strategy to help with character development, it would be considered self-regulatory or metacognitive. If mental imagery was suggested or required by a teacher, it would not. While this meta-analysis adhered to the most current model of self-regulation in writing, the research included in the paper confounded cognitive processes with metacognitive or self-regulatory processes in the case of prewriting and mental imagery studies.

Further, studies evaluating five of the ten self-regulatory processes outlined in Zimmerman and Risemberg's model of self-regulatory processes were located as part of this meta-analysis. Future research on the following self-regulatory processes is needed: (1) Environmental structuring, (2) Self-monitoring, (3) Self-consequenting, (3) Self-verbalization and (5) Time management and planning. Developmental research on writing development and the development of self-regulatory processes is also badly needed. No studies on self-regulation instruction in Grade 1 were included in this meta-analysis. For this reason, the authors called for research at different grade levels.

2.3 Assessment of Self-Regulation in Learning

Now that we have reviewed the literature on self-regulation and writing, we will examine the issue of the assessment of self-regulation in learning, and then with respect to writing. In the literature, there are two primary ways that self-regulation is assessed: (1) Self-report questionnaires and, (2) Think-aloud protocols. Self-report questionnaires will be discussed followed by think-aloud protocols.

2.3.1 *Self-Report Measures*

Self-report measures used to assess self-regulation can consist of any combination of the following protocols: verbal interviews, surveys, questionnaires, diaries, and stimulated recall (Greene et al., 2018). The context self-regulation is measured in can also vary. These dimensions include subject-specific versus subject-general, activity-focused versus student-focused and online versus offline assessments. Subject-specific measures inquire about self-regulation related to a specific subject such as writing whereas subject-general measures inquire about general self-regulation skills that could be applied to any subject. Activity-focused measures refer to self-regulation of an activity, such as using a graphic organizer to help plan a narrative. In contrast, student-focused measures assess internal aspects of self-regulation such as what a student would say to themselves to stay motivated. Online measures assess self-regulation while a student is engaged in a task and offline measures assess self-regulation when a student is not engaged in a task.

The two most common standardized assessments of self-regulation are the Motivated Strategies for Learning Questionnaire (MSLQ; Pintrich et al., 1993) and the Learning and Study Strategies Inventory (LASSI; Weinstein et al., 1987). The MSLQ measures has two primary domains: (1) Motivation Scales and (2) Learning Strategies

Scales. The Motivation Scale assesses intrinsic goal orientation, extrinsic goal orientation, task value, control of learning beliefs, self-efficacy for learning and performance and text anxiety. The Learning Strategies Scale assesses rehearsal, elaboration, organization, critical thinking, metacognitive self-reflection, time and study environment management, effort regulation, peer learning and help seeking. The LASSI has three primary domains: (1) Affect and Effort Related, (2) Metacognition, and (3) Cognitive and Active Learning. The Affect and Effort Related Scale measures motivation, time management and concentration. The Metacognition Scale measures attitude, test strategies, anxiety and selecting main ideas. The Cognitive and Active Learning Scale measures information processing, study aids and self-testing. Both assessments are general measures of self-regulatory behaviour that can be applied to several learning subjects and contexts. These assessment tools are typically administered to high-school and university-aged students.

MSLQ and LASSI Reliability and Validity. The empirical evidence on the reliability and validity of these measures is mixed. Pintrich et al. (1993) found that the MSLQ had robust scale reliabilities and a confirmatory factor analysis revealed a good factor structure in a sample of 356 university students. The scale also was reasonably predictive of student achievement which demonstrated predictive validity. However, a more recent meta-analysis by Crede and Phillips (2011) found that while their results were largely consistent with theories of self-regulated learning, some specific scales (i.e., rehearsal, elaboration, organization, critical thinking, peer learning and help seeking) were not related to academic performance. The authors posited that a curvilinear relation exists between variables (e.g., low achieving and high achieving students may not

demonstrate the behaviour) and how university grades are earned (i.e., often unrelated to effective use of learning strategies) could serve as possible explanations for their findings.

Given the popularity of the LASSI, there were surprisingly few empirical studies that examined the reliability and validity of the scale. Cano (2006) administered the LASSI to two independent samples of students in their first (N = 527) and final (N = 429) years of university. Data analysis of the first data set demonstrated acceptable psychometrics and revealed a three-factor model (i.e., affective strategies, goal strategies, and comprehension monitoring). The authors used a confirmatory factor analysis to test a three-factor model in the second data set, which supported a three-factor solution. Affective strategies and goal strategies were positively linked to academic performance, which suggests these factors have predictive validity. While there is some empirical evidence to support the reliability and validity of the MSLQ and the LASSI, the evidence is not as strong as expected given how widely these tools are used.

General- Versus Task-Specific Measures of Self-Regulation. Both the MSLQ and the LASSI are general measures of self-regulatory learning behaviours used to assess self-regulation across a variety of subject domains. Based on the literature reviewed with respect to writing, higher levels of self-regulation should be associated with better academic outcomes. However, in a meta-analysis, relations between the MSLQ subscales and grades in individual classes were higher than the scores between the MSLQ and overall GPA for all 15 subscales (Crede & Phillips, 2011). This highlights that self-regulation skills in one subject area do not spontaneously transfer to another subject area.

For this reason, it is imperative to use *subject-specific* measures to accurately assess self-regulatory knowledge.

Advantages and Disadvantages of Self-Report Measures. There are advantages to using self-report measures to assess self-regulation which include being easy to make, administer, score, and convert to data that is ready for analysis. Self-report questionnaires also are easily adapted to different contexts; they can capture a variety of beliefs, attitudes, and strategies in one administration, and they provide insight into unobservable processes (Greene et al., 2018). However, Schneider et al. (2017) identified important methodological issues that need to be considered. First, as previously demonstrated, many standardized self-regulation assessment tools do not find a positive correlation between self-regulation and achievement. This is contrary to expectations and calls for the validity of the current assessment tools to be questioned. The authors posited this finding could be a result of relying on students to judge how often they use a certain strategy. This request more heavily relies on long-term memory than knowledge of strategy use. Second, responding to questions about metacognition requires a high degree of insight and abstract thinking that can be challenging for younger students. Younger students also tend to struggle with social desirability (i.e., providing responses they think their examiner would like to hear) and memory bias (i.e., misremembering past behaviours). Third, the tools used to assess self-regulation are not able to assess metacognitive knowledge separately from strategy use. These two constructs are highly interconnected. Taken together, Schneider et al. (2017) highlighted special considerations when designing, administering, and validating measures intended to assess self-regulation, especially in younger students.

2.3.2 *Think-Aloud Protocols*

Think-aloud protocols are an online assessment of metacognition while engaged in a specific task (Veenman, 2005). For this type of assessment, students are asked to speak aloud their thoughts to give researchers insight into their internal processes. For example, a student may be asked to tell the researcher everything they are thinking while they are writing a paragraph. Generally, these assessments are viewed as more reliable than self-report questionnaires because they are less subject to social desirability bias, memory errors and are not directly interpreted by the researcher. They also do not require any frequency judgements. Numerous empirical studies have found correlations between think-aloud protocols and learning outcomes (e.g., Azevedo et al., 2008; Greene & Azevedo, 2007; Greene et al., 2012) which provides evidence to support their validity. However, most of the research on think-aloud protocols has been in reading, history, science, and mathematics (Greene et al., 2018). While this assessment type has many methodological advantages, think-aloud protocols are time consuming to administer and score.

Unexpectedly, convergent validity between self-report questionnaires and think-aloud protocols, despite supposedly measuring the same constructs, is regularly low (-0.07 to 0.31; Veenman, 2005). There are numerous possible explanations for this finding. First, think-aloud protocols require high levels of insight on the part of the participant and place heavy demands on working memory resources that are already taxed by the nature of the assessment. Second, it is possible that self-report questionnaires and think-aloud protocols measure different aspects of metacognition. However, while general measures of self-regulation reached correlations up to 0.22 with think-aloud protocols, task-specific

measures of self-regulation reached correlations up to 0.42 (Van Hout-Wolters, 2009). Taken together with the research on self-report questionnaires, the available evidence further supports that subject-specific measures should be used to assess self-regulation in writing.

2.4 Research on Self-Regulation in Writing

The purpose of this next section is to examine the available literature on the assessment of self-regulation in writing and relevant empirical findings. This literature review revealed a paucity of empirical research specifically considering the reliability and validity of the measures used to assess self-regulation in writing. While the issues of self-regulation in learning and self-regulation in writing are both well-researched, few studies have considered the validity of the tools they are using to assess this construct. The following section will present the empirical results on self-regulation in writing. A discussion on validity studies will follow.

Self-report interviews and questionnaires have been used to assess self-regulation in writing in the following areas of research: (1) To examine individual differences in writers (Gillespie et al., 2013; Graham et al., 1993; Saddler & Graham, 2007), (2) To examine the effect of elements of self-regulation on writing quality (Graham et al., 2017; Olinghouse & Graham, 2009; WijeKumar et al., 2019) and, (3) To evaluate the effectiveness of a Self-Regulated Strategy Development or other writing interventions (Fidaglo et al., 2008; Graham et al., 2005; Harris et al., 2006; Zumbrunn & Bruning, 2013). Only one study specifically evaluated the reliability and validity of a measure designed to assess self-regulation in writing (Golembek et al., 2019). These measures

have been used with typically achieving writers, less skilled writers, and students with learning disabilities in Grades 2 to 8, high school, and university.

In this body of work, various scales examined numerous elements of self-regulation in writing. It is important to note that most of this literature is based on of the original self-regulation interview designed by Graham et al. (1993). For that reason, the scoring system will be described in more detail before results are presented and discussed.

2.4.1 *Scoring*

There is a lot of variability in how self-report interviews about self-regulation in writing were scored. Most researchers modelled their scoring systems after Graham et al. (1993) which was influenced by Hayes and Flower's (1986) and Zimmerman and Pons (1986) models of writing and self-regulated learning strategies during writing. According to their scoring system, student responses were separated into idea units which were then categorized according to the following dimensions: environmental structuring, production procedures, substantive procedures, seeking assistance, motivational abilities, related other and unrelated other. Environmental structuring reflected student efforts to manage their physical environment to facilitate the writing process (e.g., "find a quiet room") and production procedures referenced the written product itself (e.g., "write it neatly" or "spell the words correctly"). Substantive procedures included statements about the writing process such as planning and revising. This category was further broken down into knowledge and prior knowledge of task (e.g., "I already know what needs to be in this report"), information generation (e.g., "go to the library and get more details"), organizing (e.g., "put my notes in order"), goal setting/planning (e.g., "first, I would think

about how I want to start"), writing and drafting (e.g., "write a first draft"), reviewing, evaluating and revising (e.g., "I would read my paper over to see that needs to be changed") and other.

Motivation, or statements that referred to motivation, rewards, or punishments (e.g., "they give up" or "they keep doing it until they get an A") were also scored. The remaining idea units were categorized into seeking assistance, abilities, no changes, other related and unrelated other. The number of idea units in each category were summed for the purpose of analysis. Most of the scales examined in this literature had adapted at least part of their scoring system from Graham et al. 's (1993) study. Other scoring systems included Likert-type scales. Likert-scales were more commonly used to measure aspects of writing motivation and self-efficacy than cognitive elements of self-regulation in writing.

2.4.2 *Writing Knowledge*

Writing knowledge is commonly assessed in tools that claim to examine self-regulation (Gillespie et al., 2013; Graham et al., 1993; Graham et al., 2005; Olinghouse & Graham, 2009; Saddler & Graham, 2007; WijeKumar et al., 2019; Zumbrunn & Bruning, 2013). Questions about what good writers do were often used to elicit self-regulatory responses. For example, Graham et al. (1993) asked students the following questions about declarative, procedural and conditional knowledge of writing: (1) "Suppose you were asked to be the teacher for one of your classes today and that one of the students asked you, what is good writing? What would you tell that student about good writing?", (2) "When good writers write, what kind of things do they do?" and (3) "Why do you think some kids have trouble writing?"

In the literature, the evidence supported that older and higher-achieving students demonstrated more writing knowledge. Graham et al. (1993) examined differences in self-regulatory abilities in students with and without learning disabilities in Grades 4, 5, 7 and 8. Their sample included 10 Grade 4/5 and 28 Grade 7/8 students with learning disabilities and 11 Grade 4/5 and 18 Grade 7/8 typically achieving students. With respect to writing knowledge, the authors found that typically achieving writers made more substantive, or higher-level comments about what good writers did than students with learning disabilities. Older students were also more likely to make substantive comments than younger students. Students with learning disabilities were more likely to describe production-level activities than typically developing students.

Gillespie et al., (2013) found similar results in a sample of 50 Grade 5 students. When asked what good writers do when they write, 95% of student responses fell into two categories: substantive and production procedures. Most comments on substantive procedures (80%) fell into five categories: (1) Structural elements of writing, (2) Reviewing or revising, (3) Planning and goal setting, (4) Generate or obtain information for writing tasks and (5) Organizing writing content. Comments on production procedures comprised the remaining 15%. Students with more substantive knowledge about how to write knew more about different genres of text (i.e., story, persuasive and informational) controlling for gender, writing achievement and production procedure responses.

Correlational and instructional evidence supports that students with more writing knowledge tend to write higher-quality texts. In a sample of 10 less skilled writers (i.e., scored below the 25th percentile on three subtests of the TOWL-3; Hammill & Larsen,

1996) and 10 average writers (i.e., scored at the 50th percentile or above on three subtests of the TOWL-3; Hammill & Larsen, 1996) in Grade 4, Saddler and Graham (2007) examined at the relation between writing knowledge and story length and quality. Correlations between story length and quality and total number of ideas for substantive procedures and production procedures were calculated separately for skilled and less skilled writers. For more skilled writers, statistically significant correlations were found between length and substantive procedures (r =.73). Significant correlations were also found between substantive procedures and text quality (r =.79) and production procedures and text quality (r =.66). No significant correlations were found between these variables for less skilled writers. This suggested that more skilled writers have more writing knowledge which is associated with story length and quality.

Writing knowledge has also been demonstrated to predict text length and quality. Olinghouse and Graham (2009) tested this question in a sample of 32 Grade 2 and 32 Grade 4 students. Substantive processes, production procedures, motivation, story elements and irrelevant information explained 14% of the variance in story quality, 14% of the variance in story length, and 19% of the variance in vocabulary diversity controlling for grade, gender, basic reading skills, handwriting fluency, spelling, attitude toward writing and advanced written story plan.

The literature also supported that writing interventions are associated with improved writing knowledge and text quality. Graham et al. (2005) examined the effect of a SRSD intervention with (N = 24) and without a peer component (N = 25) on writing knowledge in comparison to a control condition (N = 25). At pretest, there were no significant differences between conditions in writing knowledge, 43% of comments about

what made good and poor writing were about production procedures and 39% focused on substantive processes. At posttest, students in the SRSD conditions described good and poor writing using more substantive procedures than the control group, but no difference was found between the SRSD conditions. This means SRSD instruction improved student substantive writing knowledge. Similarly, in a sample of 179 Grade 5 students, WijeKumar et al. (2019) examined the impact of a persuasive writing intervention on writing knowledge, text quality and number of words written. At pretest, writing knowledge uniquely predicted writing knowledge and the number of words in a persuasive text. After a persuasive writing intervention, writing knowledge, strategic behaviours and skills explained unique variance in writing quality and writing knowledge and strategic behaviours predicted number of words written. Taken together, writing interventions have been demonstrated to improve writing knowledge and text quality.

Only one study used this methodology to examine writing knowledge in Grade 1 (Zumbrunn & Bruning, 2013). Six Grade 1 students participated in a qualitative interview on writing knowledge designed to elicit comments about their writing knowledge. Four of six students mentioned production procedures like the need for neat printing and punctuation. All six students also included substantive procedures like details and exciting words in their responses. In sum, empirical evidence supported that higher-achieving and older students have more writing knowledge, that writing knowledge is linked to writing quality and that instruction that targets writing knowledge improves both writing knowledge and text quality.

2.4.3 *Writing Strategy Knowledge*

Writing strategy knowledge, such as knowledge of planning, editing, revising, and the elements to include in a text was also often evaluated in self-report interviews about self-regulation (Fidaglo et al., 2008; Graham et al., 1993; Graham et al., 2005; Graham et al., 2017; Harris et al., 2006). Sample questions used to assess this construct included: "Teachers often ask students to write a short paper outside of class on a famous person such as Abraham Lincoln; when you are given an assignment like this, what kinds of things do you do to help you plan and write the paper?" and "Teachers often ask students to change their papers to make them better; if you were asked to change your paper to make it better or improve it, what kinds of changes would you make?" (Graham et al., 1993).

Similar to general writing knowledge, research evidence supported that higher achieving students demonstrated more writing strategy knowledge. In support, Graham et al., (1993) found that when asked about planning and revising behaviours, students with learning disabilities were more likely to describe mechanical revisions. Writing strategy knowledge has also been demonstrated to predict text quality. In a sample of 227 Grade 4 students, Graham et al. (2017) found that strategic writing behaviour explained an additional 5.4% of variability in personal narrative writing quality controlling for gender and motivation. Strategic writing behaviour did not predict number of words written.

The literature also supports that writing interventions are associated with improved writing strategy knowledge and text quality. In Graham et al. (2005), students asked about planning provided mostly substantive process responses (82%; e.g., made a list, wrote ideas down, organized notes etc.) Students in the SRSD plus peer support

condition provided significantly more substantive processes for planning compared to the SRSD-only and comparison conditions. No difference was found between the SRSD-only and control conditions. When students were asked to recall what needed to be included in a story, 91% of their ideas described the parts of a story (e.g., beginning, middle, end, characters, feelings, setting etc.) Students in the two SRSD conditions provided significantly more attributes and elements of a persuasive paper than students in the control condition. No difference was found between the SRSD conditions. This suggested that the intervention increased student strategy knowledge, but that the peer component did not result in additional knowledge as expected.

Harris et al. (2006) examined the impact of an SRSD intervention on writing strategy knowledge in a sample of 66 Grade 2 students randomly assigned to three conditions: SRSD-only (N = 22), SRSD plus peer support (N = 22) and comparison (N = 22). In their study, there was a statistically significant difference at posttest between the number of substantive processes included between the control and SRSD groups responses to questions about planning. Students in the intervention conditions also included more specific story elements when asked about the parts of a story. Fidaglo et al. (2008) studied 56 Grade 8 students who had previously received a planning and revising intervention and 21 similarly achieving students who had not received the intervention. Students provided written responses to eight questions on writing metaknowledge. Metaknowledge was treated as a dichotomous variable for analysis and explained an additional 25% of the variance in writing quality. Mentioning structuring content, monitoring their texts for errors, mentioning spelling, and not mentioning grammar each contributed individually to this finding.

Taken together, empirical evidence supports that higher-achieving students have more writing strategy knowledge, that writing strategy knowledge is linked to writing quality and that instruction that targets writing strategy knowledge improves both writing strategy knowledge and text quality.

2.4.4 *Transcription-Level Writing Processes*

Self-regulation of transcription-level writing processes were only assessed in one study (Fidaglo et al., 2008). In their study, students responded to eight open-ended questions. A coding category labelled, "low-level processes," was included in their coding system. Any self-regulatory student responses that referred to neatness/appearance, spelling or grammar were coded at the transcription-level. While other coding systems included these types of responses in other categories (e.g., Graham et al., 1993; coded comments about spelling or grammar under mechanical revisions), Fidaglo et al., (2008) were the only authors to include transcription-level responses in their primary coding system. They found in a single, stepwise multiple regression with all metacognitive (including low-level processes), self-efficacy and process variables, identifying spelling as a concern was the fifth best predictor of writing quality and mentioning the importance of neatness and appearance was the ninth best predictor of writing quality. This suggested that while other variables were better predictors of writing quality, students were including "low-level" ideas in their responses.

2.4.5 *Motivation*

Motivation, or attitude towards writing, was assessed throughout this body of literature (e.g., Graham et al., 1993; WijeKumar et al., 2019). To better understand student motivation for writing, the following types of questions were rated on a Likert

scale: (1) “I enjoy writing”, (2) “Writing is fun”, (3) “I like to write at school”, (4) “I like to write at home”, (5) “Writing is a good way to spend my time”. Graham et al. (1993) found that typically developing students had a more positive attitude towards writing than students with learning disabilities and older students had a more positive attitude than younger students. In Saddler and Graham (2007), correlations between story length, quality and motivation were calculated separately for skilled and less skilled writers. For more skilled writers, a statistically significant correlation was found between text length and motivation (r=.65). There was no relation between quality and motivation. No significant correlations were found between these variables for less skilled writers. This suggests that motivation is not related to writing skill or text quality. WijeKumar et al. (2019) found similar results in a sample of 179 Grade 5 students. In their study, writing motivation did not predict writing quality or number of words written at pretest or posttest. The literature did not support that motivation is linked to text quality.

2.4.6 *Self-Efficacy*

Scales were used to assess self-efficacy in writing (Fidaglo et al., 2008; Graham et al., 1993; Graham et al., 2005; Graham et al., 2017; WijeKumar et al., 2019). Self-efficacy is the degree of belief one has in their own ability to complete challenging tasks such as writing. Examples of items used to assess self-efficacy in writing included, “it is easy for me to get ideas/get started/make all the changes I need to make” and “it is hard for me to organize my ideas/keep the paper going/correct my mistakes”. In the literature, the findings on the impact of self-efficacy on text quality and the number of words written were mixed. Graham et al. (1993) found that both students with learning disabilities and typically developing students viewed their writing abilities favorably. In

WijeKumar et al. (2019), the authors found at pretest that self-efficacy was related to the number of words written whereas at posttest, self-efficacy was related to writing quality and the number of words written. Similarly, Graham et al., (2017), found that attitudes toward writing and self-efficacy explained an additional 5.2% of the variability in writing quality controlling for gender and strategic writing behaviour. Attitude toward writing and self-efficacy predicted number of words written after controlling for gender and strategic writing behaviours. In contrast, Graham et al. (2005) found no instructional effect on struggling writer's self-efficacy before and after the intervention and Fidaglo et al. (2008) did not find any individual self-efficacy subscale scores that were related to overall writing quality. Taken together, the findings on the impact of self-efficacy on writing quality were unclear.

2.5 Reliability

Most studies that used an interview to assess self-regulation in writing examined the questionnaire's inter-rater reliability. In the studies reviewed, inter-rater reliability was generally acceptable (i.e., > .80), but ranged from 0.70 to 0.99.

2.6 Validity of Self-Report Questionnaires of Self-Regulation in Writing

Validity refers to if a scale is measuring the desired construct. Before evaluating the validity of the self-report questionnaires on self-regulation in writing that have been used in the literature, the questionnaires and scoring systems themselves will be reviewed according to The Self-Regulation Model of Writing (Zimmerman & Risemberg, 1997).

2.6.1 *Environmental*

The environmental self-regulatory processes include: (1) Environmental structuring and (2) Self-selected models. Environmental structuring was assessed in

numerous studies (Fidaglo et al., 2008; Gillespie et al., 2013; Golembek et al., 2019; Graham et al., 1993; Graham et al., 2005; Olinghouse & Graham, 2009; Saddler & Graham, 2007; WijeKumar et al., 2019; Zumbrunn & Bruning, 2013). Self-selected models were not included as a unique code in any of the assessment tools included in this literature review.

2.6.2 *Behavioural*

The behavioural self-regulatory processes include: (3) Self-monitoring, (4) Self-consequenting, and (5) Self-verbalizations. Self-monitoring was examined in two studies (Fidaglo et al., 2008; Golembek et al., 2019). While the idea of rewards and punishments were included in some coding systems, self-consequating and self-verbalizations were not explicitly measured in any of the measures reviewed.

2.6.3 *Personal/Covert*

The personal/covert self-regulatory processes were the most consistently measured elements of self-regulation in writing: (6) Time planning and management (Gillespie et al., 2013; Golembek et al., 2019; Graham et al., 1993; Graham et al., 2005; Graham et al., 2017; Harris et al., 2006; Olinghouse & Graham, 2009; Saddler & Graham, 2007; WijeKumar et al., 2019; Zumbrunn & Bruning, 2013), (7) Goal setting (Gillespie et al., 2013; Golembek et al., 2019; Graham et al., 1993; Graham et al., 2005; Olinghouse & Graham, 2009; Saddler & Graham, 2007) (8) Self-evaluative standards (Fidaglo et al., 2008; Gillespie et al., 2013; Golembek et al., 2019; Graham et al., 1993; Graham et al., 2005; Saddler & Graham, 2007), and (9) Use of cognitive strategies (Gillespie et al., 2013; Graham et al., 1993; Graham et al., 2005; Harris et al., 2006;

Olinghouse & Graham, 2009; Saddler & Graham, 2007; Zumbrunn & Bruning, 2013). No coding systems or items measured use of mental imagery.

2.6.4 *Validity Studies*

Factor analyses are one way to assess the validity of a measure by examining if related questions group together statistically. In this body of work, three studies completed factor analyses to examine the structure of the self-report questionnaire designed to measure self-regulation (Graham et al., 2005; Graham et al., 2017; WijeKumar et al., 2019) and one study specifically considered the validity of their assessment tool (Golembek et al., 2019). Each of these studies will be discussed in further detail.

Graham et al. (2005) measured writing knowledge using an interview (i.e., declarative knowledge, knowledge about planning and genre-specific knowledge) in their study designed to evaluate a Self-Regulated Strategy Development (SRSD) intervention with a peer support element. Their study included 73 Grade 3 students in three conditions (i.e., SRSD only: N = 24; SRSD plus peer support: N = 24; and comparison: N = 25). The authors did not complete a factor analysis for any writing knowledge measures. However, they did conduct a factor analysis to see if the self-efficacy scale was unidimensional. The items included in the five-item scale were: (1) “When my class is asked to write, my paper is one of the best,” (2) “When writing a paper, it is hard for me to decide what goes first, second, third and so on,” (3) “When writing a paper, I have trouble finding the right words for what I want to say,” (4) “When I plan a paper, my plan is one of the best in the class” and, (5) “When writing a paper, it is easy for me to keep thinking of things to say.”

A factor analysis of student responses at pretest revealed a two-factor solution: (1) Self-efficacy for planning/writing and (2) Self-efficacy for generation/organization.

Graham et al. (2017) looked at whether strategic writing behaviour, attitudes towards writing and self-efficacy predicted writing quality and number of words written in a sample of 227 Grade 4 students. The strategic writing behaviour scale included 10 items that were rated on a 5-point Likert-type scale that ranged from strongly disagree to strongly agree. The 10 items inquired about detail included when writing, organization, reading the assignment instructions, using a detailed outline, giving examples or definitions, finding good words, using a topic to guide writing, time management and considering the writer's audience. The 5-item attitude toward writing scale was also rated on a Likert-scale that ranged from strongly disagree to strongly agree which included questions about if a student enjoyed writing, if writing was fun, if they liked writing at school, if they liked writing at home and whether writing was a good way to spend time. Lastly, the self-efficacy scale for writing included 13 items scored on a 100-point scale that ranged from 0 = no chance to 100 = completely certain. This scale asked students about if they could spell words correctly, if they could write complete sentences, if they could use punctuation and grammar, if they could generate and organize their ideas in a text, if they could focus/avoid distractions while writing and if they knew when and where to use writing strategies.

The authors conducted a confirmatory factor analysis to see if the three scales represented different constructs. The first three-factor model did not adequately fit the data. This necessitated an examination of the data that revealed that three items on the self-efficacy scale related to sentence writing (i.e., "I can write complete sentences", "I

can punctuate sentences correctly", and "I can write grammatically correct sentences") needed to be allowed to covary to fit the model. Once completed, the three-factor solution adequately fit the data. The authors justified and accepted this modification because all three items were related to sentence writing. It is important to note that because these questions assessed self-efficacy, their relation to self-regulation is indirect.

WijeKumar et al. (2019) assessed attitudes towards writing, self-efficacy, and strategic writing behaviours in 148 Grade 8 students. All three constructs were assessed using the same measures as in Graham et al., (2017). In contrast to their previous study, the researchers conducted separate factor analyses for each scale. An exploratory factor analysis was used to determine if the strategic writing behaviours scale was unidimensional as was reported in Graham et al., (2017). The initial two-factor solution revealed that three items double-loaded on both factors (i.e., "I use a lot of examples and definitions to make things clear in my writing", "I easily find good words for what I want to say when writing and I think about my readers when I write"). The remaining seven items loaded on a single factor.

On the attitudes towards writing scale, factor analysis produced a single factor. However, two factors emerged from the self-efficacy for writing scale: (1) Self-efficacy for writing mechanics and (2) Self-efficacy writing regulation. The items on the first factor asked about spelling, writing complete sentences, punctuation, grammar, paragraph structure, word choice, idea generation, and idea placement. The second factor's items inquired about ability to focus, avoid distractions, task initiation, and when and where to use writing strategies. This was a similar finding to Graham et al., (2005).

In all three studies, the initial hypothesized factor solutions did not fit the data and required follow-up analyses for at least one scale. The only study designed to test the validity of self-regulation measures will be presented and discussed next.

Golombek et al. (2019) designed and validated the Self-Efficacy for Self-Regulation of Academic Writing (SSAW) scale. Their scale was developed in accord with current theory and evaluated in three studies. The first study aimed to examine the items and psychometric properties, the second study examined the scale's factorial validity, and the third study examined the scale's convergent validity. The scale items were selected in accord with Bandura's (2006) guidelines for creating self-efficacy scales and based on Zimmerman and Kitsantas' (2007) cyclical model of self-regulation by writers. The initial scale included 73-items that captured self-regulation in the forethought, performance, and reflection phases of writing. The rating scale included 11 categories that ranged from zero (= no chance) to 100 (= completely certain). A cognitive pretest to assess all items and instructions were clear resulted in the elimination of three items and minor revisions to the wording of some items.

The first study had 121 German-speaking university students complete the 70-item SSAW, the 10-item self-report self-regulation scale (Schwarzer & Renner, 2000) and the 10-item general self-efficacy self-report (GSE; Schwarzer & Jerusalem, 1995). Both the self-regulation scale and GSE were Likert-type scales that ranged from (1 = not at all true to 4 = certainly true). Principal axis analyses using promax rotation resulted in the scale being reduced to 29 items (i.e., all items with item-total correlations $r < .30$ were excluded; Bortz & Doring, 2006). A principal axis analysis completed with the new 29-item scale resulted in a two-factor solution that reduced the scale to 22 items. The final

principal axis analysis with the 22-item scale revealed one underlying factor. The scale's internal validity was excellent $\alpha = .95$ and the item-total correlations ranged from .47 to .79. Content analysis indicated that six items were consistent with the forethought phase (e.g., "I can set myself specific writing goals), nine items were consistent with the performance phase (e.g., "I can concentrate on writing"), and seven items were consistent with the self-reflection phase (e.g., "I can judge what I have to do differently next time"). The authors also found preliminary support for convergent validity of the scale. Correlations between the SSAW scale and the general self-efficacy and self-regulation scales ranged from .32 to .62. Overall, the SSAW scale had good psychometric properties and preliminary evidence for convergent validity. Even though the analyses revealed one factor, based on a content analysis, the authors proposed three subscales (with good internal consistencies).

The second study examined the SSAW scale's factorial validity in a sample of 660 German-speaking university students. The authors used a structural equation modeling analysis to examine three competing models: a one-factor model and two models that tested the three proposed scales. The analysis revealed that the one-factor model was not acceptable but a three-factor model that delineated the three subscales (i.e., forethought, performance, and self-reflection) was the best-fitting model.

The third study examined the SSAW scale's convergent validity. The authors hypothesized that the SSAW scale would positively correlate with general self-regulation measures, general self-efficacy measures, instruments that assessed self-efficacy for self-regulated learning and academic efficacy and SSAW. One-hundred and eighty-eight German-speaking university students completed the same general self-efficacy and self-

regulatory measures as the first study, an 11-item self-regulated learning scale (Zimmerman, 1992) and the scale for academic self-efficacy (Schyns & Collani, 2002). The results for study three supported all four of the author's hypotheses. That is, significant positive correlations were found between the SSAW scale and general self-regulatory skills, general self-efficacy, self-efficacy for self-regulated learning and the scale assessing academic self-efficacy. Taken together, these results provided support for convergent validity.

In sum, the authors found support for the psychometric properties, content validity, factorial validity, and convergent validity of the SSAW. Now, the methodological and validity issues in this body of work will be discussed further.

Self-Report Interview Adherence to The Self-Regulation Model of Writing. Graham et al. 's (1993) original interview was based on Hayes and Flower's (1986) and Zimmerman and Pons (1986) models, which were current at the time of publication. This raises two issues: (1) Graham et al. 's (1993) interview is not consistent with the most current Self-Regulation Model of Writing and (2) Graham et al. 's (1993) interview and scoring system has been adapted numerous times to answer specific research questions which has resulted in somewhat of a "theoretical shift" since the time of publication. While each of the self-report questionnaires reviewed measured aspects of self-regulation in writing, none did so comprehensively or in a way that was consistent with the most current model. Environmental structuring, self-monitoring, time management and planning, goal setting, self-evaluative standards and use of cognitive strategies were commonly assessed using measures of self-regulation in writing. However, self-selected models, self-consequating, self-verbalizations or use of mental imagery were rarely

considered. This highlights a gap in the measurement literature that needs to be considered and addressed.

Attitudes Towards Writing and Self-Efficacy. Many of the self-report questionnaires included items or scales that measured attitudes towards writing, self-efficacy, or motivation. However, we would argue that only by measuring the underlying self-regulatory mechanisms can we attribute gains in writing attitudes, self-efficacy or motivation to self-regulation itself and the evidence supports this. As we saw in Graham et al., (2017), strategic writing behaviour, attitudes towards writing and self-efficacy measured distinct constructs. This suggests that more work examining the *processes* of self-regulation in writing is badly needed, and that these processes need to be conceptualized as separate from the *function* of self-regulation, or attitudes towards writing and self-efficacy.

Transcription-Level Versus Composition-Level Self-Regulation. One issue that has not been formally explored in the literature is the issue of transcription-level versus composition-level self-regulation in writing. In Berninger and Chanquoy (2012), the authors explained that transcription-level (i.e., spelling and handwriting) skills are foundational to self-regulation. However, if we conceptualize self-regulation as a complex and dynamic process that enables individuals to manage personal, environmental, and behavioural factors (Bandura, 1991), self-regulation should be possible at the transcription- and composition-level. To illustrate, a student may demonstrate transcription-level self-regulatory behaviors by checking their spelling and composition-level self-regulatory behaviours by checking if they included all the parts of a story. Arguably, checking to see if all parts of a story were included is more

sophisticated than checking for correct spelling. However, both behaviours are self-regulatory in nature.

This query is supported by the finding in Graham et al., (2017) that items related to sentence writing (i.e., I can write complete sentences, I can punctuate sentences correctly, and I can write grammatically correct sentences) needed to be allowed to covary to fit the statistical model for a self-efficacy scale. In further support, factor analyses of self-efficacy scales in both Graham et al. (2005) and WijeKumar et al. (2019) found two-factor solutions that separated self-efficacy for writing mechanics and self-efficacy for writing regulation. These findings could also reflect the skill level of the participants or the accuracy of their self-reported judgements of their abilities.

The Questions. In the questionnaires reviewed, most of the questions asked students about what they themselves would do in hypothetical scenarios. Given that the composing process is largely an internal experience, it seems logical to ask about a student's experience. However, children especially have more difficulty responding to questions about themselves than others. Asking questions about the student's personal experience or asking them to apply a hypothetical scenario to themselves may make them more likely to respond in a way that is socially acceptable or provide an answer that they think the researcher would like to hear. Knowing that social desirability bias is a methodological issue in assessment, it is worth examining the structure of the questions used to measure self-regulation in writing.

Validity of the Measures used to Assess Self-Regulation in Writing. At the beginning of this section, the question was posed, "Are the measures we have used to measure self-regulation in writing valid". Correlational and instructional research

evidence supports that self-regulation in writing is related to writing quality. However, the only study to formally examine the validity of a novel measure to assess self-regulation was designed to assess self-efficacy in academic writing in post-secondary students. This literature review revealed issues with outdated theories informing self-regulation assessments, that self-efficacy and attitudes towards writing are often assessed as a proxy for the structural elements of self-regulation, that transcription-level and composition-level self-regulation are often confounded, and that the internal nature of the questions may make this literature especially subject to social desirability bias.

2.7 The Current Study

The Interview on Self-Regulation in Early Writing was designed to address numerous gaps found in the literature. First and foremost, this interview was designed to assess whether the processes in self-regulation instruction mediate the gains in writing quality that have been previously demonstrated in the literature. Second, the interview was developed to be consistent with internal aspects of the Self-Regulation Model of Writing. Third, the interview's scoring system was specifically designed to capture self-regulation at the transcription- and composition-levels. Fourth, the interview questions were written to minimize social desirability bias. Finally, the developmental appropriateness of the interview specifically considered given that the measure was designed for 6- and 7-year-old students in Grade 1.

The goal of the current study was to evaluate the reliability and validity of a novel measure designed to measure self-regulation in writing in Grade 1, specifically for the genre of personal narrative because it is a common genre of writing for these students. The purpose of this work was to apply rigorous methodological standards to ascertain

whether this newly constructed tool was measuring the intended construct. It is important to note that the current research was not for the purposes of measure standardization.

To do so, the Interview on Self-Regulation in Early Writing was administered to students who participated in a SRSD writing intervention. The study had three conditions: (1) The SRSD group received instruction about a writing strategy and the aforementioned elements of self-regulation, (2) The strategy-only (SDO) group received instruction about a writing strategy only and (3) The waitlist control group received instruction as usual. The Interview on Self-Regulation in Early Writing was administered to students in all three condition at posttest to evaluate their knowledge of self-regulation in writing. Writing samples were collected from participants in all three conditions at pretest and posttest. These writing samples were used to obtain writing quality and spelling scores. A subsample of low, medium, and high-achieving writers also completed a think-aloud writing activity.

Important qualities of an assessment include construct validity, convergent validity, discriminate validity, and predictive validity (Gravetter & Forzano, 2012). Each of these were tested statistically using correlational methods (Gravetter et al., 2020). This leads to the following hypotheses:

2.8 Hypotheses

2.8.1 *Reliability*

1. The Interview on Self-Regulation in Early Writing will have inter-rater reliability of scoring the interview on total score, the transcription and composition subscales and items.

2. The Interview on Self-Regulation in Early Writing will have internal consistency as operationalized by Cronbach's alpha.
3. All items included in the Interview on Self-Regulation in Early Writing will be able to be retained after examining mean item-total, inter-item correlations and Cronbach's alpha if item deleted statistics.

2.8.2 *Validity*

1. The Interview on Self-Regulation in Early Writing will have evidence based on internal structure. Based on the current literature review, a principal components analysis will result in a two-component solution: A transcription-level component and a composition-level component.
2. The Interview on Self-Regulation in Early Writing will have evidence based on relations with other variables.
    1. Convergent Relations. Convergent validity is when measuring the same concept in different ways produces similar results. In this study, self-regulatory knowledge was measured in two different ways: (1) The Interview on Self-Regulation in Early Writing and (2) A think-aloud writing sample. Based on the literature reviewed, we expected to find a significant correlation between these two measures. This would demonstrate convergent validity.
    2. Discriminate Relations. Self-regulation ability is expected to be related to writing ability, but independent from spelling ability. This will be demonstrated by no significant correlation between self-regulation and

spelling ability measured by the proportion of words spelled correctly in the writing samples collected at posttest.

3. Predictive Relations.
    1. Given the hypothesised relation between self-regulation and writing quality, the Interview on Self-Regulation in Early Writing will predict writing quality at posttest. Further, the Interview on Self-Regulation in Early Writing will mediate the effects of instruction on text quality. This would demonstrate hypothesis-testing validity.
    2. Known-group validity looks at a known attribute of two groups and assesses whether a measured construct is different across groups. In this study, one group of students received instruction on self-regulatory functions (SRSD), one group of students received instruction on strategy (SDO), and the waitlist control group received instruction as usual. For that reason, we would expect students in the SRSD group to have more knowledge about self-regulation than the SDO and waitlist control groups. As such, students in the SRSD instructional group should score higher on the Interview on Self-Regulation in Early Writing than students in the SDO and waitlist control groups. This would demonstrate known-group validity.

Chapter 3

3 Method

The design of the Interview on Self-Regulation in Early Writing, item structure, administration, scoring and rater training will be discussed followed by the data gathering methods.

3.1 The Interview on Self-Regulation in Early Writing

3.1.1 *Scale Construction*

The Interview on Self-Regulation in Early Writing was designed to be consistent with Zimmerman and Risemberg's (1997) Self-Regulation Model of Writing and measure the self-regulation processes outlined in Reid et al. (2013). It specifically assessed self-verbalizations, self-monitoring, self-consequating, time management and planning, goal setting, use of cognitive strategies, and self-evaluative standards.

3.1.2 *Items*

The Interview on Self-Regulation in Early Writing items are presented below:

1. Ravi is in Grade 1. He is going to write a story about himself. It is about something fun that he did **[planning]**
    a. What is the first thing he should do?
    b. Should he make a plan before writing? How could he make a plan?
2. Ravi is writing a story about himself. **[criteria / goal setting]**
    a. What makes a story really good?
    b. Is there anything else that makes a story really good?

3. Jessica is a girl in Grade 1. She is writing a story about herself. **[strategy/ text structure]**
    a. What parts should her story have?
    b. Are there any other parts she should put in her story?
4. Jessica is writing the story about herself. [**self-statements: open-ended**]
    a. Is there anything she could say to herself to help write a good story? [If yes] What could she say?
    b. Could she say anything else to herself? [If yes] what?
5. Mark is a boy in Grade 1. He is writing a story about himself. He is in the middle of writing [**self-monitoring**]
    a. What could he do to make sure that his story is good?
    b. Should he check his story? [If yes] what should he check for?
6. Mark is writing his story. But he is tired of writing. He is bored. He is worried that he might not finish. [**coping**]
    a. What could Mark do?
    b. Is there anything Mark should say to himself?
7. Fatima is in Grade 1. She has just finished writing a story about herself. **[reviewing/ self-assessment**]
    a. What should she do now?
    b. After Fatima finishes her story, should she check it? [If yes] What should she check for?
8. Now Fatima has read her story and checked it. It is really good [**self-reinforcement**]

a. What could she do now?

b. Is there anything she should say to herself? [If yes] what?

3.1.3 *Item Structure*

Each item presented the student with a hypothetical scenario about another student in Grade 1 who was in various phases of writing a personal narrative. Each item was designed to elicit student knowledge of self-regulation. The scale was designed to limit social desirability bias by externalizing the questions. Each item had two prompts: one general and one more specific. The purpose of this item structure was to capture a students' ability to apply self-regulatory strategies spontaneously.

3.1.4 *Administration*

The interview was introduced to students with the following statement, "We are studying how children write stories. I would like to ask you some questions. If you do not want to answer a question, you do not have to." During the interview, all answers were accepted, and researchers were trained to make periodic encouraging comments (e.g., "I like to hear your ideas about writing stories.") Interviews took between 5- and 10-minutes. After the scenario was presented to the student, if they did not provide a concrete, self-regulatory response to prompt A, students were asked a more directive, follow-up question. Each interview was audiotaped.

3.1.5 *Scoring*

Given that this scale was designed for use in Grade 1, the scoring system was designed to capture self-regulatory functions at the transcription- and composition-level. Transcription-level responses referred to any response that focused on the mechanics of writing. Compositional-level responses referred to any response about generating ideas or

discourse. Each item was scored on both subscales for a maximum of eight-points per item and 64-points total. Students earned one point for each of the following criteria: (1) Any self-regulatory response; (2) A concrete self-regulatory response; (3) An elaborated self-regulatory answer that contains two or more ideas, (4) Any self-regulatory response to prompt A. They were awarded a point for each attribute of the response they provided, up to four points for composition, and four points for transcription self-regulation to a maximum of eight points per item. Students were not awarded any points if they did not respond, provided an answer they had previously given, an answer that applied to the same phase of writing or an unrelated response.

3.1.6 *Inter-Rater Reliability Training*

Coding tables with sample items were constructed (see Appendix A). A second rater participated in a two-hour training session. Six interviews were coded (representative of low, medium, and high writers at pretest). Three were modeled for the rater and three were scored together. Discrepancies were resolved between raters. A representative sample of thirty interviews (including low, medium, and high writers at pretest) were coded by two raters and inter-rater reliability was calculated. Raters were blind to condition.

3.2 Data-Gathering Methods

3.2.1 *Participants*

One hundred and twenty students in nine Grade 1 classrooms were recruited through a public school board in southwestern Ontario. All students in participating classrooms received the study's program and participated in writing activities. However, only consenting students completed the assessments and had their writing samples

collected for analysis. Table 1 contains the demographic information for the students in each condition.

Table 1

Student demographic information by condition

		Control	SDO	SRSD	Overall Scores
		n = 36	n = 45	n = 39	n = 120
Mean Age in Months		79.92(3.57)	76.78(3.49)	76.79(3.32)	76.83(3.43)
		%	%	%	%
Student Gender					
	Male	63.9	51.1	53.8	55.8
	Female	36.1	48.9	46.2	44.2
Primary Language at Home					
	English	83.3	88.9	92.3	88.3
	Other	16.7	11.1	7.7	11.7

3.2.2 *Design*

The purpose of the study was to evaluate if the self-regulation aspect of a Self-Regulated Strategy Development (SRSD; Harris & Graham, 1996) instructional program improved student writing quality over and above a strategy-only instructional program. This study was a quasi-experimental, pretest-posttest design with three groups: (1) Self-Regulated Strategy Development (SRSD), (2) Strategy Development Only (SDO), and

(3) Waiting list control. Each group included three participating classrooms. Before the start of the lessons, all nine classrooms completed a pretest personal writing sample. The SRSD and SDO groups participated in 10 lessons unique to each condition while the waiting list control group received instruction as usual from their classroom teacher. At posttest, all participating classrooms completed another personal narrative writing sample and consenting students completed the Interview on Self-Regulation in Writing. A subsample of students completed a think-aloud writing sample. Students in all three conditions also completed a posttest personal writing sample.

3.2.3 *Professional Development*

Teachers in all three conditions participated in two half-days of professional development. The professional development sessions were conducted by the research team and consisted of discussing the strategy development, reviewing the content of the teacher guides, lesson modeling by the research team, teachers completing student activities to understand them, and teachers practicing teaching the lessons.

3.2.4 *Instruction*

Teachers in all three conditions were provided with condition-specific teacher guides and all required materials. For all three groups of students, lesson plans were designed in accordance with expectations from the Ontario Ministry of Education Language curriculum which requires instruction on strategy instruction in writing and learning multiple forms of text such as personal narratives. In all three groups, students learned a personal narrative writing strategy. Globally, the strategy taught the students to include the following six elements in their texts: (1) Topic, (2) Setting, (3) Beginning, (4) Middle, (5) End and (6) Feeling. To learn the strategy, the groups participated in the

following activity types: reading picture books as a class, teacher-led discussions, teacher modelling, teacher support, partner writing, independent practice, and feedback. All lessons took between 30 and 40 minutes.

Strategy Development Only (SDO). The purpose of the SDO unit was to teach students a strategy they could use to help them improve their personal narrative writing abilities. Students participated in ten lessons with the following structure: Set the Context for Student Learning (e.g., review); Reading Aloud; Develop the Strategy; Student Activity; and Wrap-Up (sharing, feedback). Students were taught to use a strategy to include the following elements in their compositions: (1) Setting, (2) Beginning, (3) Middle, (4) End and, (5) Feeling. While the instruction focused on the goal and the steps of the strategy, it **did not** include any other aspects of self-regulation such as self-monitoring or self-reinforcement etc.

Teachers were instructed to scaffold the lessons; gradually transitioning from modeling the strategy for the students to the students using the strategy independently. Encouraged teaching approaches included teacher-led discussions, modelling of the strategies, shared writing with the teacher, students writing individually with teacher support, and independent writing. Teachers were instructed to encourage their students to try to keep writing if they did not know how to spell a word and write on every second line to accommodate editing.

Self-Regulated Strategy Development (SRSD). Students in the SRSD condition were taught the same strategy as the students in the SDO condition. However, they were also taught additional aspects of self-regulation during writing. The number (10) and structure of the lessons in the SRSD condition was the same as the SDO condition, with

one additional element: develop the strategy **and** self-regulation. The following elements of self-regulation were taught in the lessons: (1) Goal setting, (2) Self-instruction, (3) Self-monitoring, (4) Coping, (5) Self-reinforcement. The first lessons in the unit focused on teaching the strategy for personal narrative and the later lessons focused on teaching self-regulatory functions. Just as in the SDO condition, teachers were instructed to encourage their students to try to keep writing if they did not know how to spell a word and write on every second line to accommodate editing.

3.2.5 *Assessment*

Personal Narrative Writing Sample. To measure Grade 1 personal narrative writing development, the quality of their texts was evaluated using three dimensions of text quality: (1) Holistic test quality, (2) Word count, and (3) Personal Narrative Structure (Kim et al., 2014). Students wrote responses to different prompts at pretest and posttest. The prompts (listed below) were designed to be relatable experiences for all Grade 1 students:

- A Time I Went to a Party
- A Time I Played Outside

Students had 20 minutes to complete their writing activity. However, most students finished in less than 15 minutes. Upon completion, the researcher reviewed each student text, clarified any illegible words, and wrote them in for the student. All writing samples were transcribed with corrected spelling before they were analyzed for quality. Student texts were analyzed by two coders blind to treatment condition and time of testing (i.e., pretest or posttest).

Holistic Text Quality. Holistic text quality has been demonstrated in the literature to be a valid measure of writing development (Graham et al., 2005; Kim et al., 2014). This variable was included to test predictive validity. Writing samples were given a score from one (i.e., lowest quality) to nine (i.e., highest quality). Coders were instructed to consider "ideation, organization, grammar, sentence structure, and aptness of word choice" and that all factors should be given equal weight when judging writing quality (Graham et al., 2005). The first coder chose three writing samples to serve as examples of scores of 3, 5 and 7. Using these anchor texts, an additional coder scored each writing sample to judge holistic quality. Using this method, inter-rater reliability was expected to be approximately r = .80.

Think-Aloud Writing Sample. The purpose of this assessment was to obtain a behavioural measure of posttest self-regulation while writing to evaluate convergent validity. Generally, these assessments are viewed as more reliable than self-report interviews because they are less subject to social desirability bias, memory errors and are not interpreted by the researcher. They also do not require any frequency judgements. Numerous empirical studies have found correlations between think-aloud protocols and learning outcomes which provides evidence for their validity (e.g., Azevedo et al., 2008; Greene & Azevedo, 2007; Greene et al., 2012).

In the current study, students voiced their thoughts aloud while writing with prompting. This assessment was audio recorded. Given that this activity was time intensive, two students from each class in the first, second and third quartiles in writing quality at pretest were selected to participate. Each participating student completed a training activity to allow them to practice thinking out loud. The researcher asked

students to order four pictures to form a story. As they worked, they were reminded to speak their thoughts out loud (e.g., please tell me about what you are thinking, tell me about what you are doing, I see that you were quiet for a moment, what were you thinking?).

> "Here are four pictures. Please put them in order to tell a story. Please talk about what you are doing. Just say whatever you think of, even if it does not seem important. Sometimes I might remind you to tell me about what you are thinking."

The interviewer encouraged students to continue expressing their thoughts about sequencing the pictures by prompting them to "Please tell me about what you are thinking about." Once the practice activity was completed, students wrote a narrative in response to the prompt, "Write about a time you went shopping". The researcher read from the following script:

> "Have you ever gone shopping? We are going to write a story about a time that you went shopping. Please close your eyes with me. Remember a time that you went shopping. [Pause]. Now open your eyes. Please write a story. Tell me about a time you went shopping."
>
> "As you write, please talk aloud about what you are doing. This is just like before, when you sorted the pictures. Just say whatever comes to mind, even if it does not seem important. Sometimes, I might remind you to tell me about what you are thinking." Pauses in writing are often indicative of planning, so if the participant is silent, the researcher asked, "Tell me about what you are thinking."

Scoring. For analysis, each student's speech was transcribed into tables where each row was an utterance, and each column represented the student speech and text written

together with that utterance. Beginning writers are expected to transcribe (or speak out loud) each word they write whereas more developed writers have been found to make more self-regulatory comments about writing (e.g., Bereiter & Scardamalia, 1987).

Spelling. Spelling was measured by calculating the percent of the words spelled correctly in written compositions at posttest. This variable was included to evaluate discriminate validity. This method of measuring spelling is a valid, widely used curriculum-based measure in young writers (e.g., McMaster et al., 2009).

Chapter 4

4 Results

A series of analyses were conducted to evaluate the reliability and validity of the Interview on Self-Regulation in Early Writing. The reliability analyses are discussed first, followed by the validity analyses.

4.1 Missing Data

The Interview on Self-Regulation in Early Writing was administered to 118 of the 120 students that participated in this study. Two students were absent at the time of data collection and alternate administration dates were not able to be arranged. Item 4 was not administered to one participant. Because all missing data points were determined to be missing completely at random, and less than 5% of data were missing from analysis, the three participants with missing data from the Interview on Self-Regulation in Early Writing were removed from the dataset (Allison, 2001). One of the removed participants was in the control group, one was in the SDO group and one was in the SRSD group. The following analyses were conducted with complete data from 117 students.

4.2 Scale Descriptives

4.2.1 *Transcription Subscale*

On the transcription subscale, students did not frequently provide transcription-level responses to the planning (item 1), goal setting (item 2), strategy (item 3), self-statements (item 4), coping (item 6) or self-reinforcement (item 8) items. For these items, 93%, 71%, 91%, 89%, 94% and 97% of students respectively did not mention self-regulation of transcription. For both the self-monitoring (item 5) and the reviewing (item 7) items, 56% of students did not mention self-regulation of transcription. This suggests

that students were most likely to provide transcription-level responses to items that asked about self-monitoring (i.e., what should Mark do to make sure his story is good/check it?) and reviewing (i.e., after Fatima finishes her story should she check it/what should she check for?). When students were asked what they needed to review or check in their stories, nearly half of the students responded with transcription-level answers like periods, finger spaces, or spelling. Item means by condition and overall are provided in Table 2.

4.2.2 *Composition Subscale*

On the composition subscale, the overall item means for the planning (item 1), goal setting (item 2), self-statements (item 4), self-monitoring (item 5), reviewing (item 7) and self-reinforcement (item 8) ranged from $M = 1.51$ (1.04) to $M = 1.75$ (1.22). This suggests a similar level of difficulty for these items. On the strategy (item 3) and coping (item 6) items, the overall means were $M = 2.12$ (1.36) and $M = 2.63$ (1.22) respectively. This suggests that these items were somewhat easier for the students. Item means by condition and overall are provided in Table 2.

Table 2

Item difficulty by condition

	Control	SDO	SRSD	Overall Scores
	n = 34	n = 45	n = 38	n = 117
Transcription Subscale	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Planning	.03 (.17)	.08 (.52)	.37 (.91)	.16 (.63)
Goal setting	.81 (1.26)	.42 (.90)	.71 (1.19)	.63 (1.11)
Strategy	.47 (1.12)	.23 (.89)	.21 (.74)	.29 (.94)
Self-statements	.47 (1.13)	.14 (.65)	.32 (.98)	.29 (.92)
Self-monitoring	1.42 (1.44)	1.30 (1.54)	.67 (1.22)	1.13 (1.44)
Coping	0.06 (.34)	.07 (.45)	.35 (.96)	.16 (.65)
Reviewing	.93 (1.23)	1.20 (1.30)	.76 (1.12)	.98 (1.25)
Self-reinforcement	.00 (0.00)	.04 (.18)	.03 (.16)	0.02 (.14)
Composition Subscale				
Planning	1.31 (.92)	1.54 (.93)	1.64 (1.26)	1.51 (1.04)
Goal setting	1.28 (1.34)	1.72 (1.23)	1.74 (1.52)	1.60 (1.36)
Strategy	1.63 (1.21)	2.10 (1.36)	2.59 (1.35)	2.12 (1.36)
Self-statements	.88 (1.12)	1.34 (1.28)	2.37 (1.58)	1.54 (1.46)
Self-monitoring	1.69 (1.17)	1.44 (1.24)	2.07 (1.44)	1.72 (1.31)
Coping	2.44 (1.23)	2.61 (1.20)	2.81 (1.25)	2.63 (1.22)
Reviewing	1.63 (1.29)	1.53 (1.17)	2.11 (1.15)	1.75 (1.22)
Self-reinforcement	1.09 (1.13)	1.34 (1.09)	1.71 (.94)	1.39 (1.08)

Note. SDO is strategy-development only and SRSD is self-regulated strategy development.

4.3 Inter-Rater Reliability

Pearson's product-moment correlations are a common method of calculating inter-rater reliability in the literature that considers the shared variance between two continuous variables (Cohen, 1988). A statistically significant strong relation was found between raters on the transcription subscale $r(118) = .92$, $p < .001$, composition subscale $r(118) = .89$, $p < .001$ and scale total $r(118) = .91$, $p < .001$. This suggested that the inter-rater reliability for the Interview on Self-Regulation in Early Writing was excellent. For that reason, the average of the two raters scores was taken and used for the remainder of the analyses.

Table 3

Inter-rater reliability by item, subscale and overall score measured by Cohen's Kappa

	κ	Strength of agreement
Transcription Subscale		
Transcription total	.38	Fair
Planning	.63	Good
Goal setting	.58	Moderate
Strategy	.85	Very Good
Self-statements	.60	Moderate
Self-monitoring	.78	Good
Coping	.68	Good
Reviewing	.64	Good
Self-reinforcement	N/A	N/A
Composition Subscale		
Composition total	.17	Poor
Planning	.43	Moderate
Goal setting	.49	Moderate
Strategy	.76	Good
Self-statements	.69	Good
Self-monitoring	.53	Moderate
Coping	.36	Fair
Reviewing	.49	Moderate
Self-reinforcement	.60	Moderate
Scale Total	.15	Poor

Note. For all reported Cohen's Kappa scores, $p < 0.05$.

4.4 Inter-Item Reliability

Inter-item reliability statistics were examined to assess the internal consistency of the Interview on Self-Regulation in Early Writing. First, the process used to examine inter-item reliability is described followed by the evaluation of the scale. SPSS statistical software was used to obtain inter-item reliability statistics. This procedure produces the following statistics: Cronbach's alpha, item means, an inter-item correlation matrix, scale mean if item deleted, corrected item-total correlation and Cronbach's alpha if item deleted.

Cronbach's alpha coefficient is an estimate of internal consistency for the overall agreement between items. Generally, scores above .7 indicate high internal consistency (Lared Statistics, 2015). If the items on a scale each measure a similar construct, the item means should be similar. If any means are dissimilar from the rest, it could be an indication that the item(s) may need to be removed. For this reason, item means are considered. If items were measuring similar constructs, one would expect the items on the scale to correlate. For that reason, inter-item correlations are examined. If any items do not have a correlation of at least $r = .3$ with another item, it may indicate that the item needs to be removed. Scale means if item deleted are examined to see if removing any item would make the scale means more similar to each other. The corrected item-total correlation represents how much each item correlates with the overall questionnaire. Correlations less than $r = .3$ suggest that an item should be removed from the questionnaire. One way to examine internal consistency is to look at how the overall consistency of the measure would change if you removed each item. If this score goes up if any item were deleted, the item could possibly need to be removed to make the scale

more reliable. These statistics are presented for the sixteen-item scale, each eight-item subscale (i.e., transcription and composition) and for the revised scale after items were removed based on the analysis.

The Cronbach's alpha scores for the scale total, transcription subscale and composition subscale are presented in Table 4. Cronbach's alpha scores above .7 indicate high internal consistency. With all 16 items, internal consistency was $\alpha = .63$. Given that Cronbach's alpha for the full scale was below .7, further examination of the individual items was warranted.

Table 4

Inter-item reliability measured by Cronbach's Alpha

	All 16 Items (α)	Items Removed (α)
Total scale	.63	.68
Transcription subscale	.46	.61
Composition subscale	.75	N/A

Note. In the reduced items column, items 1, 2, 3, 4, 6 and 8 were removed from the transcription subscale.

Item means were examined (see Table 2). On the transcription subscale, the planning (item 1), goal setting (item 2), strategy (item 3), self-statements (item 4), coping (item 6) and self-reinforcement (item 8) item means were substantially lower than the rest of the items on the scale. On the composition subscale, the strategy (item 3) and coping (item 6) item means were higher than the rest of the items on the scale. This meant that other statistics needed to be consulted to determine if removing any items would make the scale more reliable.

Next, inter-item correlations were examined. Items without any correlations above .3 with other items indicated that they may be measuring different constructs and should be removed to improve scale reliability. None of the planning (item 1), goal setting (item 2), strategy (item 3), self-statements (item 4), coping (item 6) and self-reinforcement (item 8) correlated with any other items at or above the r = .3 level. The self-monitoring (item 5) and reviewing (item 7) items of the transcription scale and all items on the composition scale had at least one correlation with one other item equal to or greater than *r* = .3 (see Table 5). Taken together with the item means, the lack of correlation with other scale items provided initial evidence that these items could be removed to improve scale reliability.

Table 5

Inter-item correlation matrix

	t.q1	t.q2	t.q3	t.q4	t.q5	t.q6	t.q7	t.q8	c.q1	c.q2	c.q3	c.q4	c.q5	c.q6	c.q7	c.q8
t.q1	1	.07	.01	.12	.09	-.06	-.03	-.05	.06	-.04	-.11	.23	.02	.04	.12	.08
t.q2	.07	1	.09	.06	0.06	.13	.16	.17	-.02	-.27	.11	.11	-.08	.11	.15	.06
t.q3	.01	.09	1	.07	.09	-.08	.09	-.06	.01	-.00	-.21	.03	.09	.13	-.07	-.01
t.q4	.12	.06	.07	1	.22	.12	.19	.23	-.00	-.01	-.09	-.23	-.10	.04	.01	-.13
t.q5	.09	.06	.09	.22	1	.01	.44	-.03	-.11	.07	-.01	-.12	-.22	.06	-.13	.07
t.q6	-.06	.13	-.08	.12	.01	1	.17	-.04	-.09	.05	.15	.04	-.01	-.12	-.04	.03
t.q7	-.03	.16	.09	.19	.44	.17	1	.03	.12	.23	.15	.02	-.02	.06	-.21	.14
t.q8	-.05	.17	-.06	.24	-.03	-.04	.03	1	.11	-.11	.07	-.15	-.19	.07	.11	-.18
c.q1	.06	-.02	.01	-.01	-.11	-.09	.12	.11	1	.39	.34	.30	.42	.40	.11	.14
c.q2	-.04	-.27	-.00	-.01	.07	.05	.23	-.11	.39	1	.26	.18	.41	.29	.09	.15
c.q3	-.11	.11	-.21	-.09	-.01	.15	.15	.07	.34	.26	1	.26	0.30	.31	.30	.23
c.q4	.23	.11	.03	-.23	-.12	.04	.02	-.15	.30	.18	.26	1	.45	.26	.17	.40
c.q5	.02	-.08	.09	-.10	-.22	-.01	-.02	-.19	.42	.41	.30	.45	1	.33	.30	.23
c.q6	.04	.11	.13	.04	.06	-.12	.06	0.07	.40	.29	.31	.26	.33	1	.14	.38
c.q7	.12	.15	-.07	.01	-.13	-.04	-.21	.11	.11	.09	.30	.17	.30	.14	1	.17
c.q8	.08	.06	-.01	-.13	.07	.03	.14	-.18	.14	.15	.23	0.40	.23	.38	.17	1

Note. t= transcription dimension and c= composition dimension.

Table 6 contains the corrected item-total correlations and Cronbach's alpha if item deleted scores for all 16 items. The corrected item-total correlations represent the degree to which the individual items correlate with the whole scale. Correlations less than $r = .3$ could indicate that the item should be removed. None of the transcription and all the composition items except the reviewing (item 7) item had correlations with the scale that were equal to or greater than $r = .3$. The internal consistency of the scale with all 16 items was $\alpha = .63$. Examining how the internal consistency of the scale would change if each item was removed from the scale indicated that the strategy (item 3), self-statement (item 4), self-monitoring (item 5) and self-reinforcement (item 8) items possibly needed to be removed from the transcription scale because they decreased the scale's overall internal consistency and that reviewing (item 7) should be retained because it did not decrease the scale's internal consistency. Given that the scale's internal consistency would decrease if you deleted any of the composition items, the analysis suggested that these items should be retained (see Table 6).

Table 6

Item total statistics all items

	Scale Mean if Item Deleted *M*	Corrected Item-Total Correlation *r*	Cronbach's Alpha if Item Deleted *α*
Transcription Subscale			
Planning	17.77	.11	.63
Goal setting	17.30	.11	.63
Strategy	17.64	.04	.64
Self-statements	17.64	.03	.64
Self-monitoring	16.80	.07	.65
Coping	17.77	.06	.63
Reviewing	16.95	.27	.61
Self-reinforcement	17.91	-.01	.63
Composition Subscale			
Planning	16.42	.42	.59
Goal setting	16.33	.35	.60
Strategy	15.81	.40	.59
Self-statements	16.39	.38	.59
Self-monitoring	16.21	.41	.59
Coping	15.30	.48	.58
Reviewing	16.18	.20	.62
Self-reinforcement	16.54	.38	.60

Scale reliability statistics for each subscale (transcription and composition) were also examined to determine if any items should be removed. Internal consistency on the transcription subscale was low $\alpha = .46$. The low internal consistency for the transcription subscale indicated that item(s) needed to be removed to make the scale more reliable. As seen in Table 5, only the self-monitoring (item 5) and reviewing (item 7) items had correlations above the $r = .3$ level. Similarly, only the self-monitoring (item 5) and reviewing (item 7) items had correlations above the $r = .3$ level on the corrected item-total correlation with the whole scale and resulted in a decrease in internal consistency. Taken together, all the evidence suggested that the self-monitoring (item 5) and reviewing (item 7) items on the transcription subscale should be retained, and that the planning (item 1), goal setting (item 2) strategy (item 3), self-statements (item 4), coping (item 6) and self-reinforcement (item 8) items should be removed to make the scale more reliable. However, the transcription subscale did not perform well as an independent subscale. This could suggest that the transcription and composition subscales of the Interview on Self-Regulation in Early Writing measured different constructs.

Table 7

Item total statistics transcription subscale

Transcription Subscale	Scale Mean if Item Deleted *M*	Corrected Item-Total Correlation *r*	Cronbach's Alpha if Item Deleted *α*
Planning	3.51	.07	.47
Goal setting	3.04	.18	.44
Strategy	3.38	.11	.47
Self-statements	3.38	.28	.40
Self-monitoring	2.54	.35	.35
Coping	3.51	.11	.46
Reviewing	2.69	.41	.31
Self-reinforcement	3.65	.08	.47

In contrast to the transcription subscale, internal consistency on the composition subscale was high $\alpha = .75$. All items had at least one correlation with another item that was greater than $r = .3$, the corrected item-total correlations were all greater than $r = .3$, and the level of internal consistency was constant on the Cronbach's alpha if item deleted statistic. Taken together, this suggested that the composition subscale of the Interview on Self-Regulation in Early Writing was reliable.

Table 8

Item total statistics composition subscale

Composition Subscale	Scale Mean if Item Deleted *M*	Corrected Item-Total Correlation *r*	Cronbach's Alpha if Item Deleted *α*
Planning	12.75	.50	.72
Goal setting	12.66	.40	.73
Strategy	12.14	.46	.72
Self-statements	12.72	.46	.72
Self-monitoring	12.54	.59	.70
Coping	11.63	.49	.72
Reviewing	12.51	.29	.75
Self-reinforcement	12.87	.40	.73

Based on the reliability analyses, the planning (item 1), goal setting (item 2), strategy (item 3), self-statements (item 4), coping (item 6) and self-reinforcement (item 8) items on the transcription subscale on the Interview on Self-Regulation in Early Writing were removed from the full scale. This decision was supported by reliability analyses and low response rates from students to these items.

The reliability of the ten-item scale was reassessed. Internal consistency for the ten-item full scale was high $\alpha = .68$ (see Table 4). The corrected item-total correlations

were less than $r = .3$ for the transcription items and the reviewing (item 7) item on the composition subscale. Internal consistency was stable for all items. For that reason, all items were retained. For the final scale descriptives, see Table 9.

Table 9

Final scale descriptives

	M	*SD*
Transcription Subscale		
Self-monitoring	1.13	1.44
Reviewing	.98	1.25
Composition Subscale		
Planning	1.51	1.04
Goal setting	1.60	1.36
Strategy	2.12	1.36
Self-statements	1.54	1.46
Self-monitoring	1.72	1.31
Coping	2.63	1.22
Reviewing	1.75	1.22
Self-reinforcement	1.39	1.08
Total	16.37	6.51

4.5 Principal Components Analysis

A principal components analysis (PCA) was conducted on the Interview on Self-Regulation in Early Writing to evaluate the scale's internal structure. To conduct this analysis, four assumptions needed to be met (Lared Statistics, 2015). The first assumption was that there are multiple continuous variables. These variables can be ordinal. The second assumption was that there should be a linear relationship between all variables. This was assessed using a matrix scatterplot. The third assumption was that there should be no outliers more than three standard deviations from the mean. The fourth assumption was that there should be a large sample size. A minimum of 5 to 10 cases per variable are required.

The Interview on Self-Regulation in Early Writing satisfied all four assumptions. There were multiple ordinal variables, there was a linear relationship between all variables (for the final solution, see Figure 1) and there were no outliers more than three standard deviations from the mean. With all 16 items retained in the scale, the interview met the large sample size assumption with five cases per variable. However, with the reduced ten-item scale, the interview met the large sample size assumption with ten cases per variable.

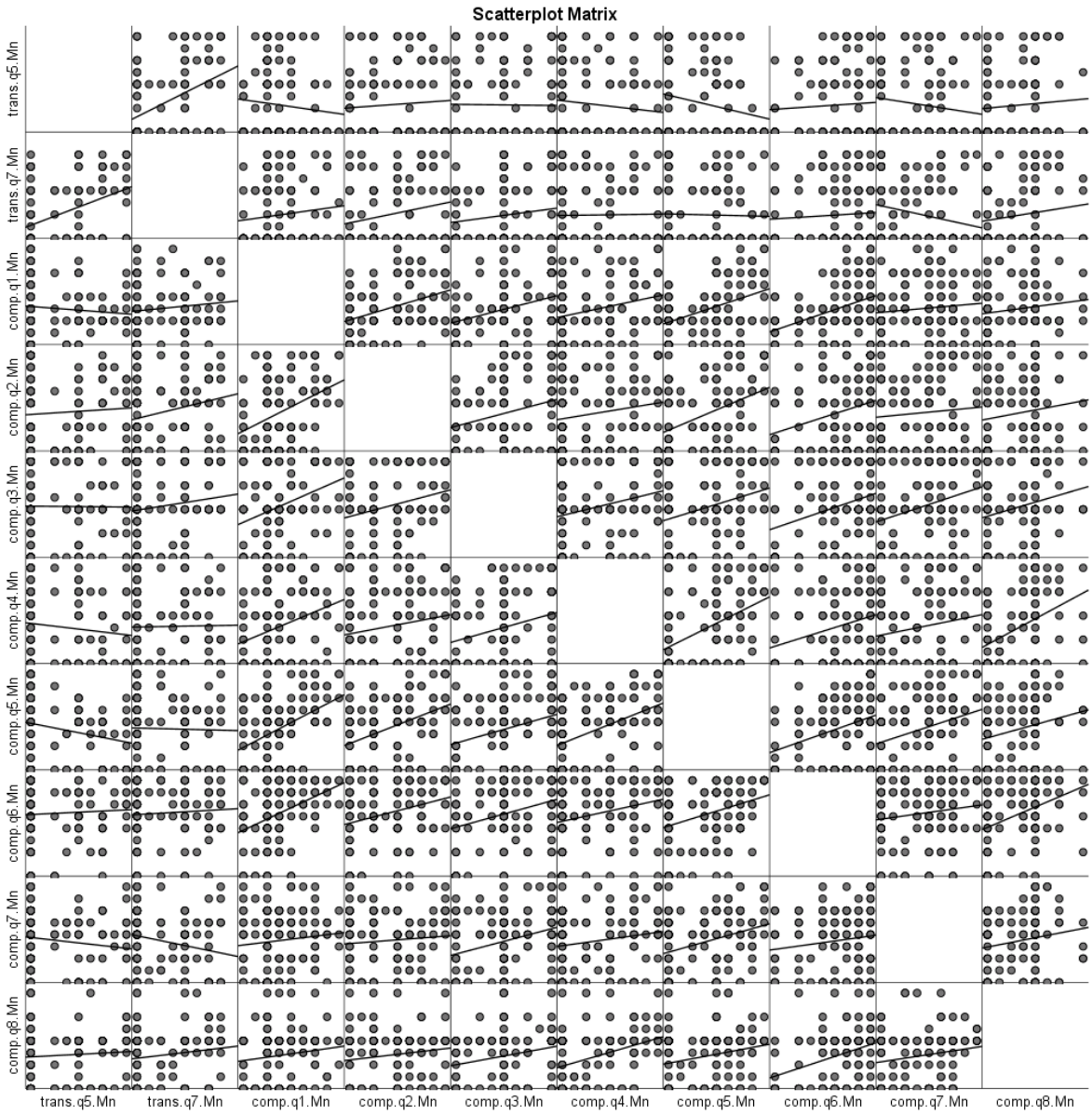

Figure 1. Scatterplot Matrix to examine linear relations between items for PCA

Two PCA's were run on the Interview on Self-Regulation in Early Writing; one with all sixteen items and one with ten items. It was determined that a PCA with all sixteen items was not appropriate (this is described in detail below). For that reason, the planning (item 1), goal setting (item 2), strategy (item 3), self-statements (item 4), coping

(item 6) and self-reinforcement (item 8) items on the transcription subscale on the Interview on Self-Regulation in Early Writing were removed. Once these items were removed, it was determined that a PCA was appropriate.

To run a PCA, each item must have at least one correlation equal to or greater than $r = .3$ with another item. As seen in Table 5, the planning (item 1), goal setting (item 2), strategy (item 3), self-statements (item 4), coping (item 6) and self-reinforcement (item 8) items on the transcription subscale did not have any correlations equal to or greater than $r = .3$ with any other items. This was the first indication that a PCA may not be appropriate for the Interview on Self-Regulation in Early Writing with all 16 items included. The overall Kaiser-Meyer-Olkin (KMO) measure was .62 with unacceptable individual KMO measures (see Table 10). To proceed with a PCA, all individual KMO measure must be equal to or greater than .5. Bartlett's Test of Sphericity was statistically significant ($p < .0005$) which suggested that the data was likely factorizable. Given that the data for the 16-item Interview on Self-Regulation in Early Writing was not appropriate for a PCA, the planning (item 1), goal setting (item 2), strategy (item 3), self-statements (item 4), coping (item 6) and self-reinforcement (item 8) items on the transcription subscale were removed and the assumptions for the ten-item scale were reassessed.

Table 10

Individual KMO measures all 16 items

Variable	KMO Measure
Transcription planning	.39
Transcription goal setting	.43
Transcription strategy	.44
Transcription self-statements	.54
Transcription self-monitoring	.57
Transcription coping	.42
Transcription reviewing	.50
Transcription self-reinforcement	.50
Composition planning	.75
Composition goal setting	.68
Composition strategy	.70
Composition self-statements	.71
Composition monitoring	.73
Composition coping	.67
Composition reviewing	.52
Composition self-reinforcement	.66

Whether or not these data were suitable for a PCA was assessed again. As seen in Table 5, each item had at least one correlation with another item that was equal to or greater than $r = .3$. The overall Kaiser-Meyer-Olkin (KMO) measure was .72 with acceptable individual KMO measures (i.e., greater than .5, see Table 11). Bartlett's test of

sphericity was statistically significant ($p < .0005$). This suggested that the data were likely factorizable.

The PCA revealed three components that had eigenvalues greater than 1.0 which explained 29.9%, 16.5% and 10.7% of the total variance, respectively. Visual inspection of the scree plot (see Figure 2) however indicated that two components should be retained. A two-component solution met the interpretability criterion. For that reason, two components were retained.

The two-component solution explained 46.5% of the total variance. An Oblimin with Kaiser Normalization Rotation was used for two reasons: (1) It helps with interpretability; and (2) It allows the variables to correlate. The rotated solution produced a simple factor structure (Thurstone, 1947). The interpretation of the data was consistent with the aspects of self-regulation the interview was designed to measure with strong loadings of self-regulation of composition on Component 1 and strong loadings of self-regulation of transcription on Component 2. Component loadings and communalities of the rotated solution are presented in Table 12.

Table 11

Individual KMO measures 10 items

Variable	KMO Measure
Transcription self-monitoring	.54
Transcription reviewing	.52
Composition planning	.78
Composition goal setting	.78
Composition strategy	.80
Composition self-statements	.75
Composition self-monitoring	.76
Composition coping	.76
Composition reviewing	.65
Composition self-reinforcement	.69

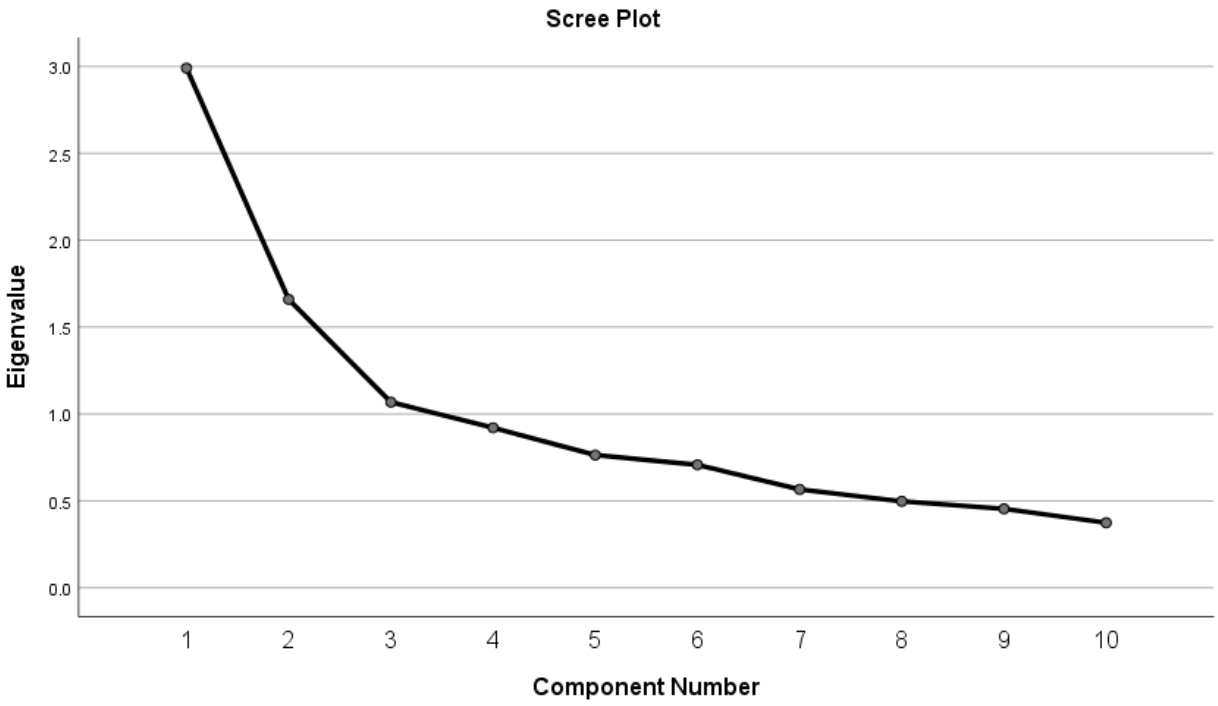

Figure 2. Scree Plot to Evaluate Component Extractions

Table 12

Rotated pattern matrix for PCA with Oblimin with Kaiser Normalization Rotation

Items	Rotated Component Coefficients		
	Component 1	Component 2	Communalities
Composition self-monitoring	**.70**	-.29	.59
Composition planning	**.67**	-.01	.45
Composition coping	**.66**	.09	.44
Composition strategy	**.62**	.01	.38
Composition goal setting	**.61**	.22	.41
Composition self-statements	**.61**	-.18	.41
Composition self-reinforcement	**.55**	.11	.31
Transcription reviewing	.24	**.81**	.69
Transcription self-monitoring	.01	**.79**	.62
Composition reviewing	**.34**	**-.46**	.35

Note. Major loadings for each item are bolded.

4.6 Outliers

For the following analyses, outliers were winsorized to the next non-outlying value for both pretest (2% of the data) and posttest holistic writing quality (5% of data).

4.7 Concurrent Validity

To assess concurrent validity, a subsample of 51 low-, medium- and high-achieving students completed a think-aloud writing activity to get a measure of their "online" self-regulation behaviours. Contrary to expectations, only five of the 51 students made any self-regulatory comments while writing their compositions. All five students

that made self-regulatory comments were in the self-regulated strategy development group. The types of self-regulatory comments students made most often referred to the use of a strategy (e.g., “So title, setting, beginning, middle, ending, actually, this, I’m doing another ending…”) or reviewing behaviour (e.g., “I’m just reading to see what, like if I went to like the middle or end or beginning or something”). Only one student made a metacognitive comment, (i.e., “I am thinking about what I was doing”). For these reasons, concurrent validity was unable to be assessed statistically.

4.8 Discriminant Validity

As previously discussed, Pearson’s product-moment correlations are a common method of calculating inter-rater reliability in the literature that considers the shared variance between two continuous variables (Cohen, 1988). A statistically significant strong relation was found between raters on the curriculum-based measure of spelling $r(117) = .89$, $p < .001$. This suggested that the inter-rater reliability for the curriculum-based measure was excellent. For that reason, the average of the scores of the two raters was taken and used for this analysis.

Based on a review of the literature, one would not expect that a relation would be found between self-regulation of composition and student spelling knowledge (e.g., Kent et al., 2014). Spelling was measured by calculating the percent of the words spelled correctly in student written compositions at posttest. This method of measuring spelling is a valid, widely used curriculum-based measure in young writers (e.g., McMaster et al., 2009). A Pearson’s product-moment correlation was run to examine the relation between the composition dimension of the Interview on Self-Regulation in Early Writing and student spelling knowledge (Lared Statistics, 2018). There was no statistically significant

relation between the composition dimension of the Interview on Self-Regulation in Early Writing and the curriculum-based measure of spelling $r(117) = .10$, $p = .27$. Given that a relationship between the variables was not expected, these results provided evidence for discriminate validity.

4.9 Predictive Validity

4.9.1 *Correlations*

Based on a review of the literature, one would not expect that transcription-level self-regulation would predict writing quality, but that composition-level self-regulation would. A series of Pearson's product-moment correlations were run to examine the relations between the transcription and composition dimensions of the Interview on Self-Regulation in Early Writing and posttest writing quality (Lared Statistics, 2018). There was no statistically significant relation between the transcription dimension of the Interview on Self-Regulation in Early Writing and posttest writing quality $r(117) = .12$, $p = .19$. In contrast, a statistically significant moderate relation was found between the composition dimension of the Interview on Self-Regulation in Early Writing and posttest writing quality $r(117) = .49$, $p < .001$. Therefore, these results provided preliminary evidence for predictive validity.

4.9.2 *Linear Regression*

To assess whether the Interview on Self-Regulation in Early Writing had predictive validity, a linear regression was conducted to evaluate if composition of self-regulation predicted posttest holistic writing quality. To run a linear regression, seven assumptions were evaluated (Lared Statistics, 2015). The assumptions were: (1) There was one dependent continuous variable, (2) There was one independent continuous

variable, (3) There was a linear relation between the dependent and independent variables, (4) There was independence of observations, (5) There were no significant outliers, (6) There was homoscedasticity and (7) The regression residuals were approximately normally distributed.

In this analysis, there was one dependent and one independent continuous variable. This satisfied the first two assumptions. A scatterplot of self-regulation of composition against posttest holistic writing quality with a superimposed regression line was used to assess linearity (see Figure 3). Visual inspection indicated a linear relation between these two variables. There was independence of residuals, as assessed by a Durbin-Watson statistic of 1.61. No outliers were detected. Based on a visual inspection of standardized residuals versus standardized predicted values, there was homoscedasticity (see Figure 4). Residuals were normally distributed as assessed by visual inspection of a normal probability plot (see Figure 5). Given that all seven assumptions were met, a linear regression was run to determine if self-regulation of composition predicted posttest holistic writing quality.

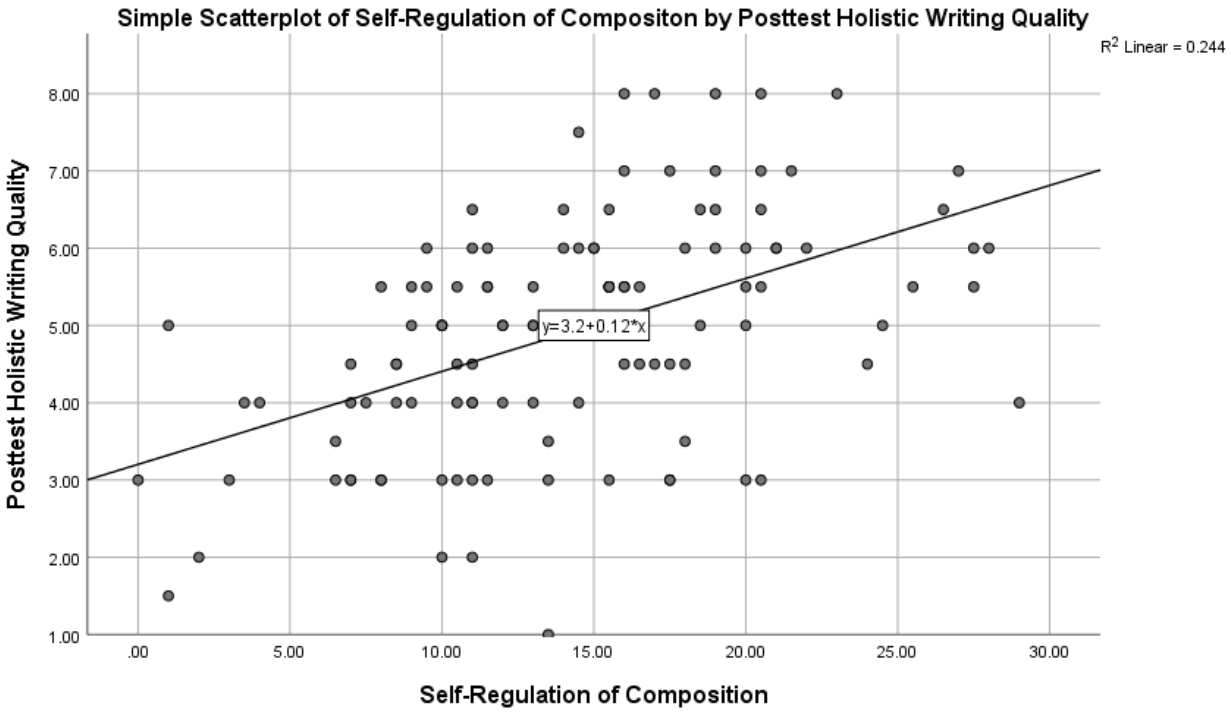

Figure 3. Scatterplot of Posttest Holistic Writing Quality by Self-Regulation of Composition

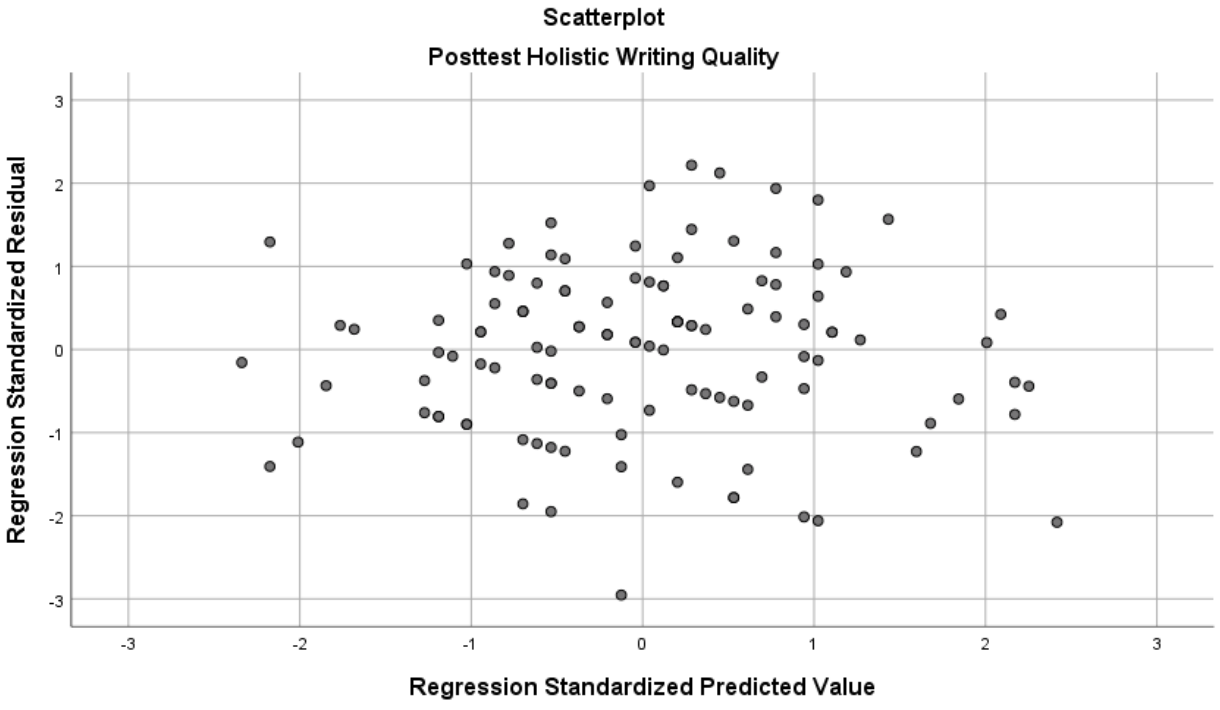

Figure 4. Scatterplot to Evaluate Homoscedasticity for a Simple Linear Regression

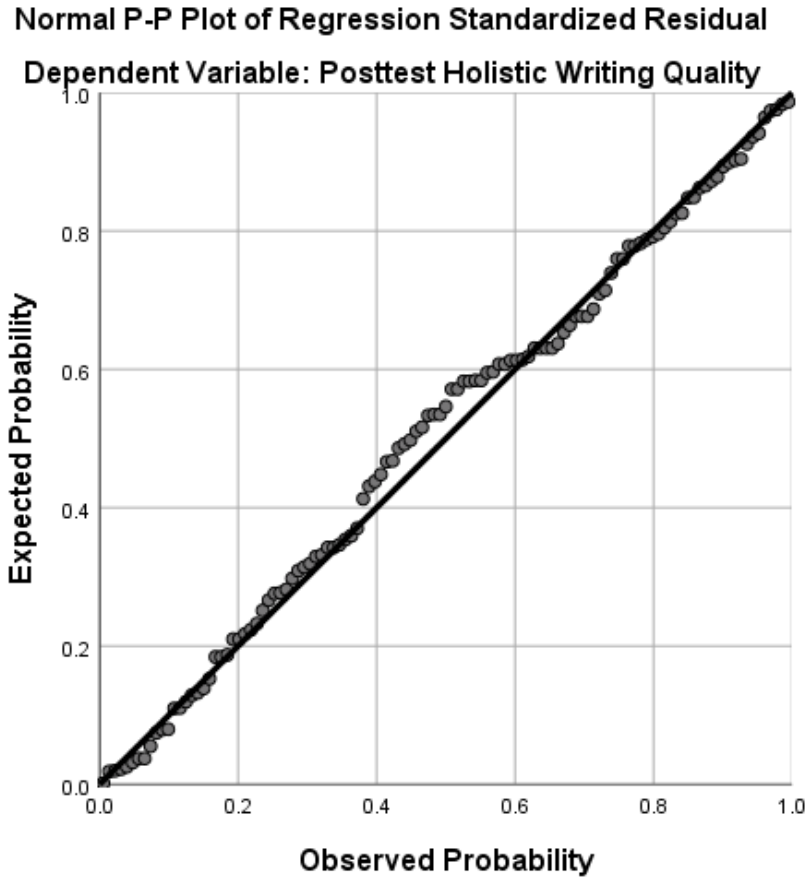

Figure 5. P-P Plot of Regression Standardized Residual for Linear Regression

A linear regression established that self-regulation of composition could statistically significantly predict posttest holistic writing quality, $F(1, 115) = 37.2$, $p < .0005$. Self-regulation of composition accounted for 24.4% of the explained variability in posttest holistic writing quality. The regression equation was: posttest holistic writing quality = 3.20 + .12x (self-regulation of composition).

4.9.3 *Mediation Analysis*

A mediation analysis was conducted to test if the effect of the intervention condition on writing quality from pretest to posttest was mediated by the composition scale on the Interview on Self-Regulation in Early Writing. Accordingly, the assumptions

for two standard multiple regression models were tested (Lared Statistics, 2015). Two models are required to test a mediation analysis to first assess the effect of the independent variable(s) on the dependent variable and then to observe changes to that direct model when a new, or mediator, variable is introduced that is hypothesized to impact or explain some of the relation between the independent and dependent variables. In this analysis, condition and pretest holistic writing quality were the independent variables and posttest holistic writing quality was the dependent variable. Self-regulation of composition was the hypothesized mediator variable. Therefore, model 1 examined the impact of condition and pretest holistic writing quality on posttest holistic writing quality and model 2 examined how these relations changed when self-regulation of composition was added into the analysis.

To run multiple regression analyses, eight assumptions needed to be met. The assumptions were: (1) There was one dependent continuous variable, (2) There were two or more independent continuous or nominal variables, (3) There was independence of errors (residuals), (4) There was a linear relation between the predictor variables (and composite) and the dependent variable, (5) There was homoscedasticity of residuals (equal error variances), (6) There was no multicollinearity, (7) There were no significant outliers, high leverage points or highly influential points and, (8) the errors (residuals) were approximately normally distributed. Assumptions were checked for each model and are reported below.

In the first model, condition and pretest holistic writing quality were the predictor variables and posttest holistic writing quality was the dependent variable. Given that there was one continuous dependent variable, one ordinal independent variable (condition

with three levels of self-regulation training) and one continuous independent variable (pretest holistic writing quality), the first two assumptions were met. There was independence of residuals, as assessed by a Durbin-Watson statistic of 1.98. Two methods were employed to test if there was a linear relation between the predictor variables and the dependent variable: (1) A scatterplot of the studentized residuals against the unstandardized predicted values (see Figure 6) and (2) Partial regression plots were examined between the continuous independent and dependent variables (see Figure 7; nominal variables were excluded). Visual inspection of both charts indicated a linear relation between variables. There was homoscedasticity based on a visual inspection of a scatterplot of studentized residuals versus unstandardized predicted values (see Figure 6). There was no evidence of multicollinearity as assessed by tolerance levels greater than 0.1. There were no studentized deleted residuals greater than three standard deviations above or below the mean, no leverage values greater than 0.2, and values for Cook's distance above 1. The assumption of normality was met as assessed by a Q-Q plot (see Figure 8).

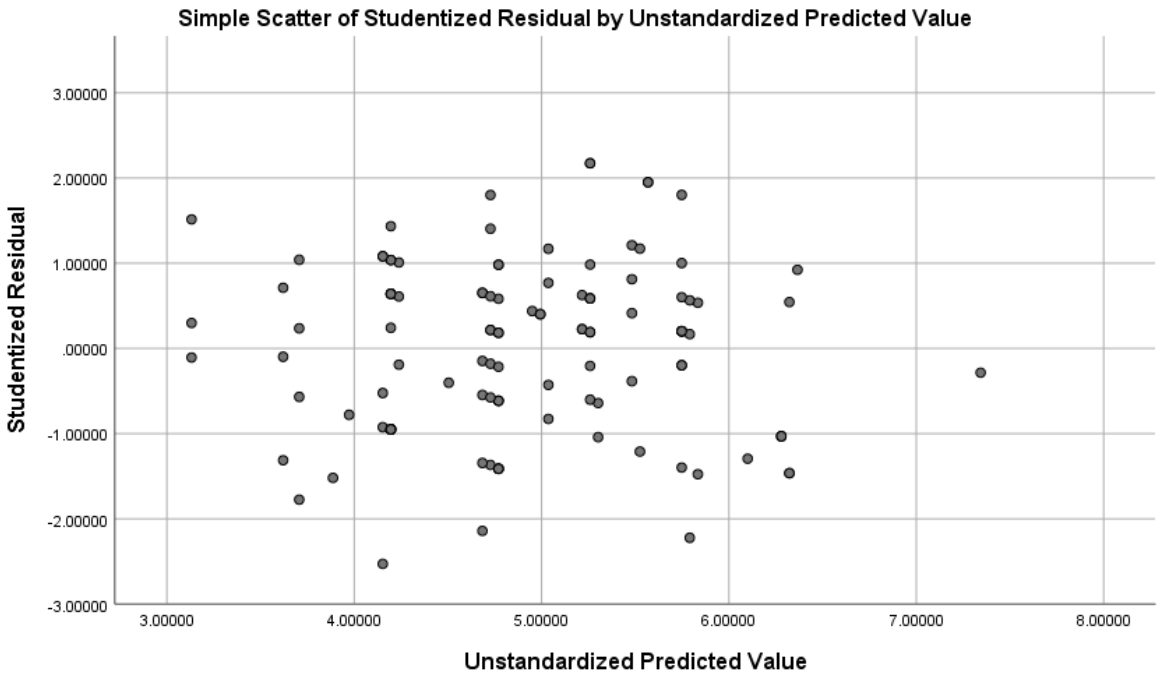

Figure 6. A Scatterplot of the Studentized Residuals against the Unstandardized Predicted Values Model 1 Mediation Analysis

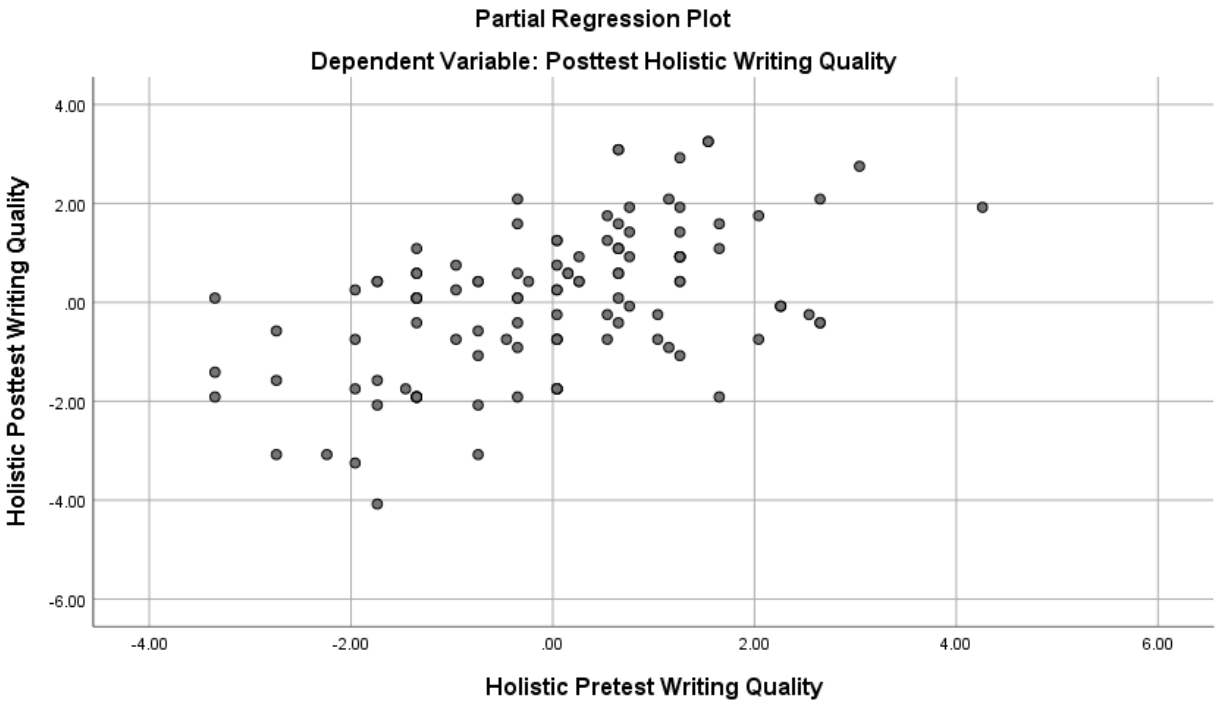

Figure 7. Partial Regression Plot between Pretest and Posttest Holistic Writing Quality Model 1 Mediation Analysis

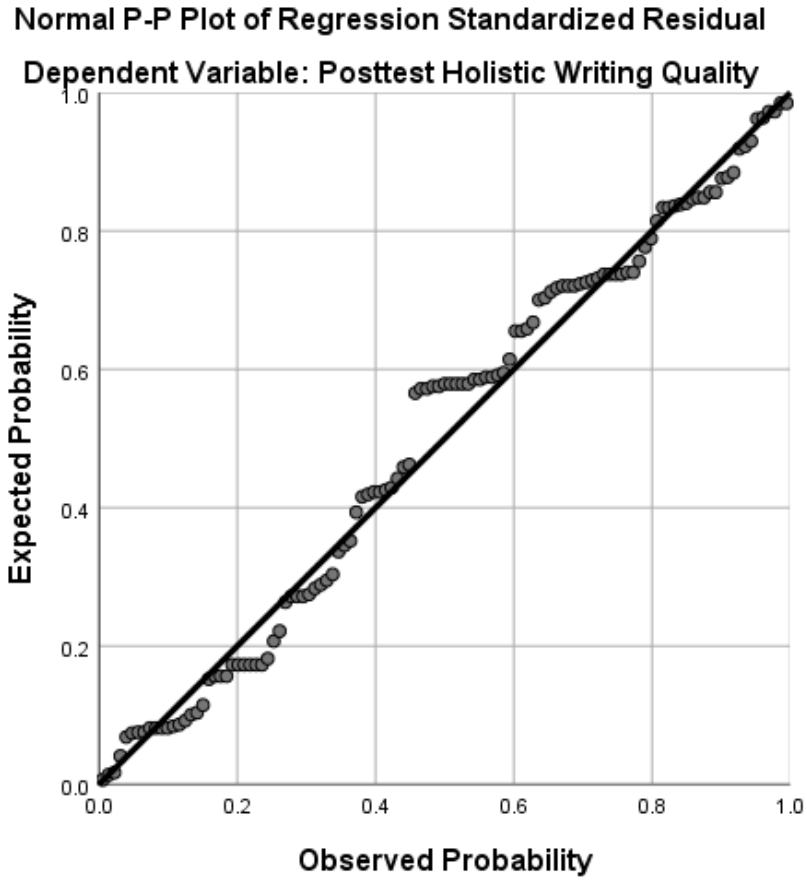

Figure 8. P-P Plot of Regression Standardized Residual for Mediation Analysis

Model 1 Mediation Analysis

In the second model, condition, pretest holistic writing quality and self-regulation of composition were the predictor variables and posttest holistic writing quality was the dependent variable. Given that there was one continuous dependent variable, one nominal independent variable (condition) and two continuous independent variables (pretest holistic writing quality and self-regulation of composition), the first two assumptions were met. There was independence of residuals, as assessed by a Durbin-Watson statistic of 1.87. Two methods were employed to test if there was a linear relation between the predictor variables and the dependent variable: (1) A scatterplot of the studentized

residuals against the unstandardized predicted values (see Figure 9) and (2) Partial regression plots were examined between the continuous independent and dependent variables (see Figures 10 and 11; nominal variables were excluded). Visual inspection of all charts indicated a linear relation between variables. There was homoscedasticity based on a visual inspection of a scatterplot of studentized residuals versus unstandardized predicted values (see Figure 9). There was no multicollinearity as assessed by tolerance levels greater than 0.1. There were no studentized deleted residuals greater than three standard deviations above or below the mean, no leverage values greater than 0.2, and values for Cook's distance above 1. The assumption of normality was met as assessed by a Q-Q plot (see Figure 12). Given that all eight assumptions were met for both models, a mediation analysis was conducted.

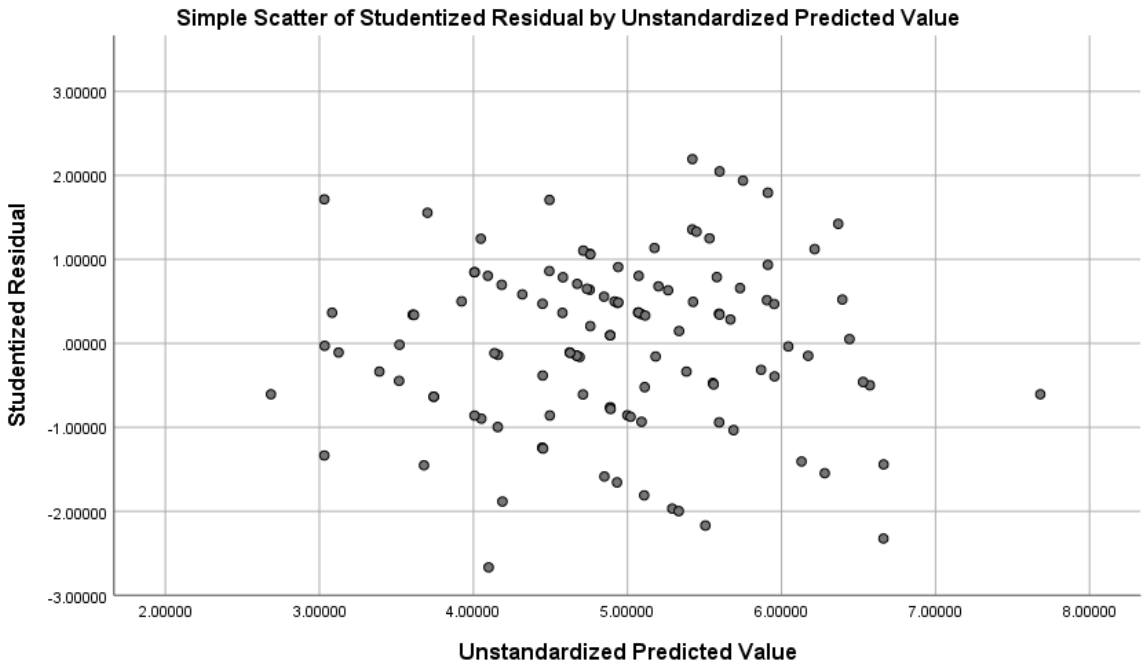

Figure 9. A Scatterplot of the Studentized Residuals against the Unstandardized Predicted Values Model 2 Mediation Analysis

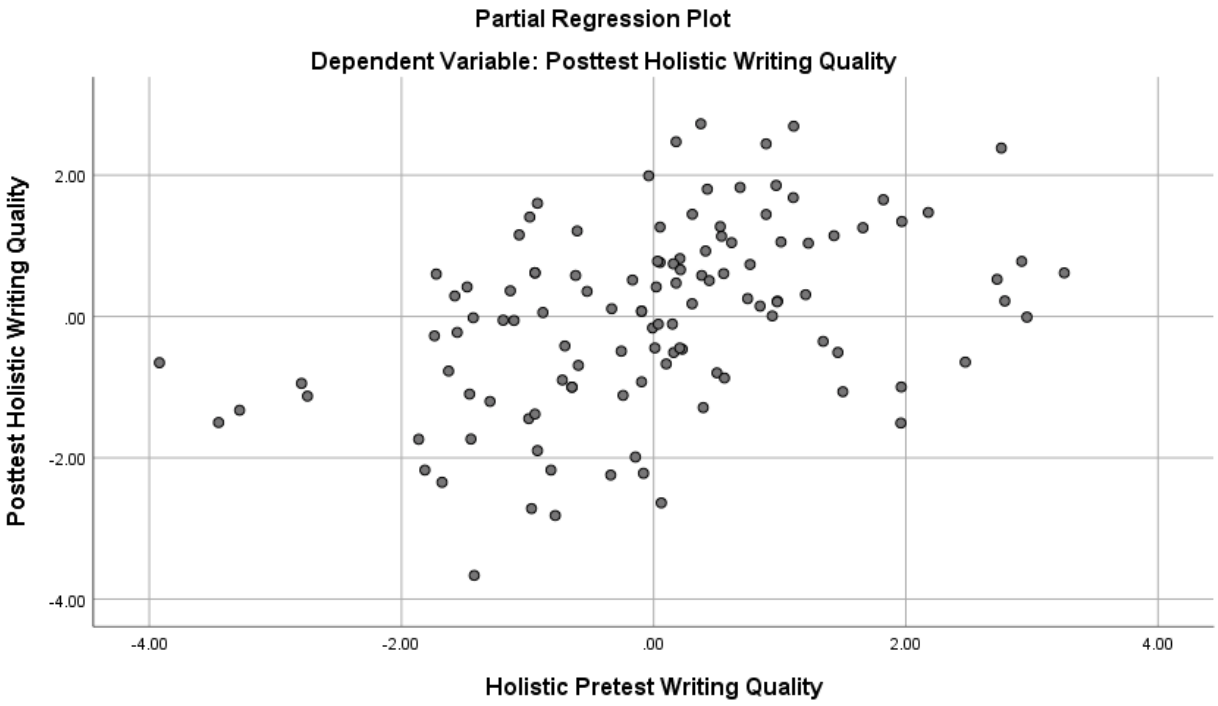

Figure 10. Partial Regression Plot between Pretest and Posttest Holistic Writing Quality Model 2 Mediation Analysis

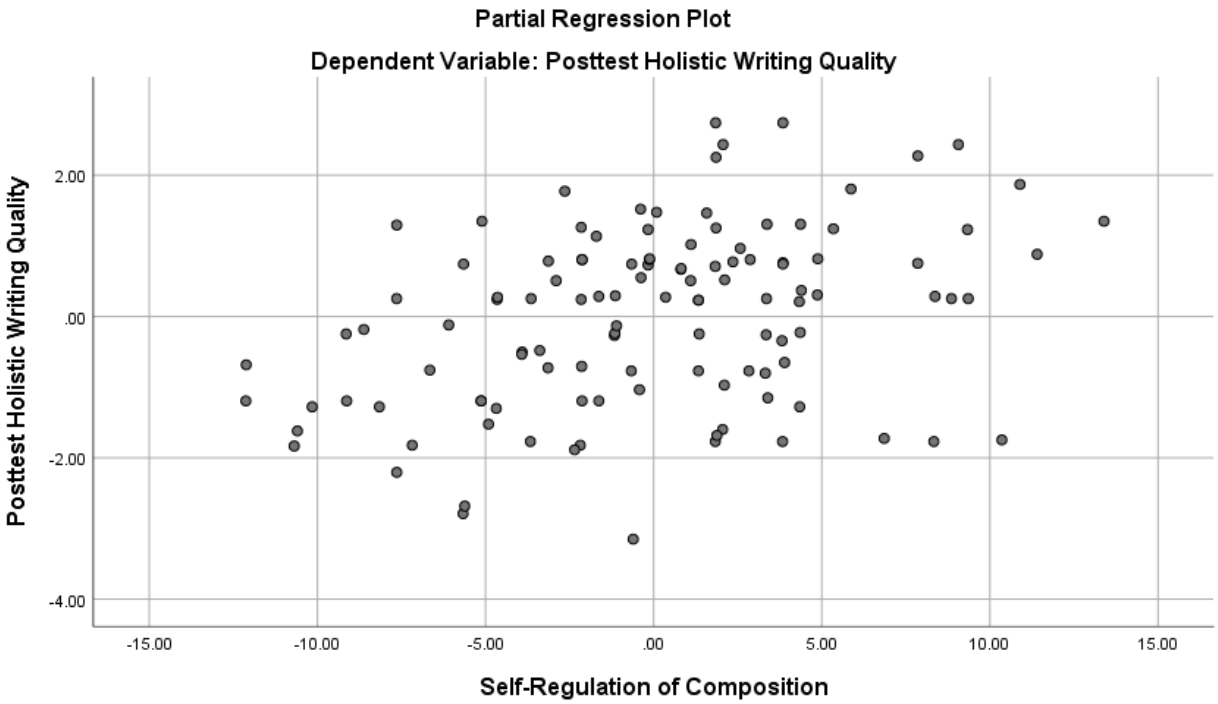

Figure 11. Partial Regression Plot Between Self-Regulation of Composition and Posttest Holistic Writing Quality Model 2 Mediation Analysis

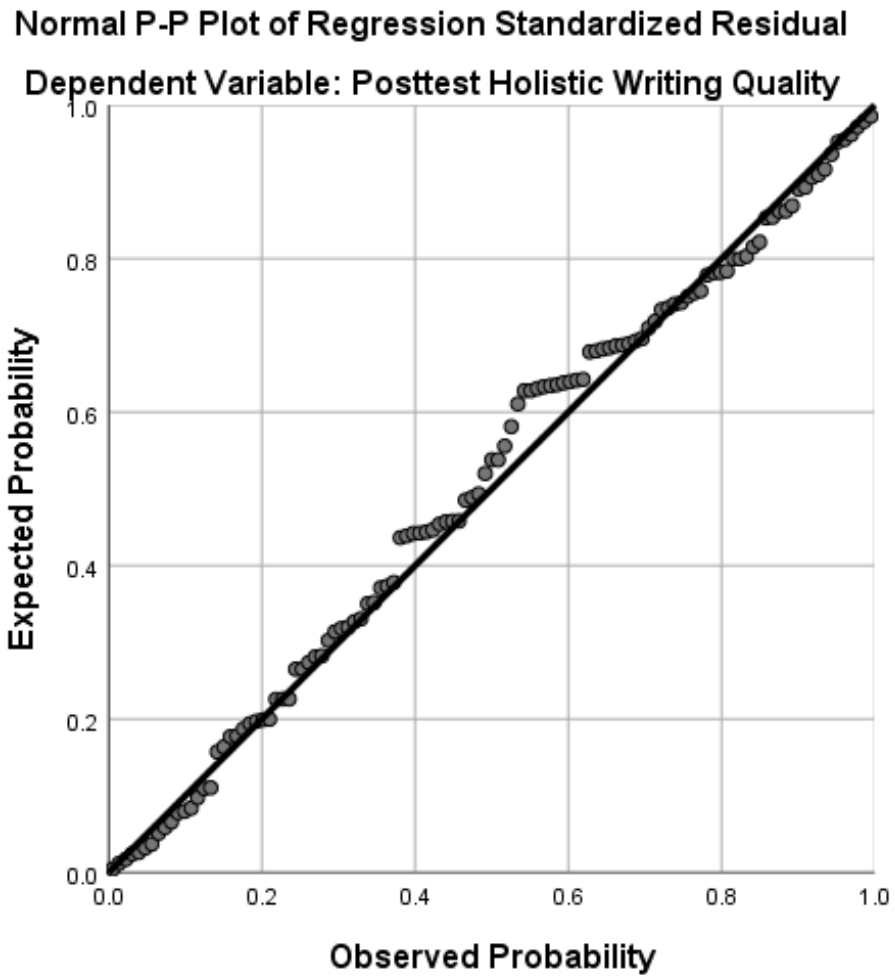

Figure 12. P-P Plot of Regression Standardized Residual for Mediation Analysis

Model 2 Mediation Analysis

With all assumptions tested and met, a mediation analysis was conducted to test if the effect of the intervention condition on writing quality from pretest to posttest was mediated by the composition scale on the Interview on Self-Regulation in Early Writing. Accordingly, two standard multiple regressions models were conducted. In the first model, a multiple regression was run to predict posttest holistic writing quality from condition (0 = control, 1 = SDO, and 2 = SRSD) and pretest holistic writing quality. The model statistically significantly predicted posttest holistic writing quality, $F(2, 114) =$

22.45, $p < .0005$, adj $R^2 = .27$. Both predictor variables added statistically significantly to the prediction, $p < .005$. For regression coefficients and standard errors, see Table 13. To test the mediated model, a multiple regression analysis was conducted to predict posttest holistic writing quality from condition (0 = control, 1 = SDO, and 2 = SRSD), pretest holistic writing quality and self-regulation of composition. The model significantly predicted posttest holistic writing quality, $F(3, 113) = 23.32$, $p < .0005$, adj $R^2 = .37$. In this model, holistic pretest writing quality and self-regulation of composition added statistically significantly to the prediction $p < .005$. With self-regulation of composition included in the model, the effect of the treatment condition was no longer significant. This suggests that the effect of condition on posttest holistic writing quality was largely mediated by self-regulation of composition. For regression coefficients and standard errors, see Table 14.

Table 13

Multiple regression results for posttest writing quality model 1

Posttest writing quality	*B*	95% CI for B		*SE B*	β	R^2	ΔR^2
		LL	*UL*				
Model							
Constant	2.11***	1.23	2.99	.44***		.28	.27***
Condition	.49**	.18	.80	.16**	.26		
Pretest writing quality	.53***	.37	.69	.08***	.55		

Note. Model = "Enter" method in SPSS Statistics; *B* = unstandardized regression coefficient; CI = confidence interval; *LL* = lower limit; *UL* = upper limit, *SE B* = standard error; β = standardized coefficient; R^2 = coefficient of determination; ΔR^2 = adjusted R^2.

*p < .05. **p < .01. *** p < .001.

Table 14

Multiple regression results for posttest writing quality model 2

Posttest writing quality	*B*	95% CI for B		*SE B*	β	R^2	ΔR^2
		LL	*UL*				
Model							
Constant	1.75***	.91	2.59	.42***		.38	***
Condition	.18	-.14	.51	.10	.10		
Pretest writing quality	.40***	.24	.56	.41***	.41		
Self-regulation	.09***	.05	.13	.02***	.36		

Note. Model = "Enter" method in SPSS Statistics; *B* = unstandardized regression coefficient; CI = confidence interval; *LL* = lower limit; *UL* = upper limit, *SE B* = standard error; β = standardized coefficient; *R2* = coefficient of determination; Δ*R2* = adjusted *R2*.

*p < .05. **p < .01. *** p < .001.

4.10 One-Way Analysis of Variances (ANOVA)

A one-way analysis of variances (ANOVA) was conducted with one between-subjects factor (group: SRSD, SDO, waitlist control) to assess whether the SRSD group demonstrated more knowledge of self-regulation of composition than the control and strategy-only groups. Six assumptions were required to run a one-way analysis of variance (ANOVA). The six assumptions were: (1) There was one dependent variable that was measured at the continuous level; (2) There was one independent variable that consisted of two or more categorical, independent groups, (3) There was independence of observations, (4) There were no significant outliers, (5) The dependent variable was approximately normally distributed for each group of the independent variable and (6)

There was homogeneity of variances (Lared Statistics, 2017). The first three assumptions were met by the study's design. There were no outliers, and the data were normally distributed for each group, as assessed by boxplot (see Figure 13) and a Shapiro-Wilk test ($p > .05$), respectively. Homogeneity of variances was violated, as assessed by Levene's Test of Homogeneity of Variance ($p = .02$).

Because homogeneity of variances was violated, a one-way Welch ANOVA was conducted to determine if students in the three instructional conditions performed differently on the composition dimension of the Interview on Self-Regulation in Early Writing. Scores on the composition dimension of the interview were statistically significantly different between instructional groups, Welch's $F(2, 71.37) = 6.\ 11$, $p < .004$, $\eta^2 = .12$. Scores on the composition dimension of the interview increased from the control group ($M = 11.96$, $SD = 5.05$), to the strategy-only group ($M = 13.64$, $SD = 4.93$), to the self-regulated strategy development group ($M = 17.05$, $SD = 7.18$; see Figure 14). Games-Howell post hoc analysis revealed that the mean increases from the control group to the SRSD group (mean difference = 5.10, $SE = 1.37$, $p = .001$) and the strategy-only group to the SRSD group (mean difference = 3.41, $SE = 1.28$, $p = .02$) were statistically significantly different. There was no significant difference between the control and strategy-only groups on self-regulation of composition (mean difference = 1.69, $SE = 1.32$, $p = .41$). In sum, students in the SRSD condition demonstrated more knowledge of self-regulation of composition than both the control and strategy-only groups but the difference in means between the control and strategy-only groups was small and not statistically significant.

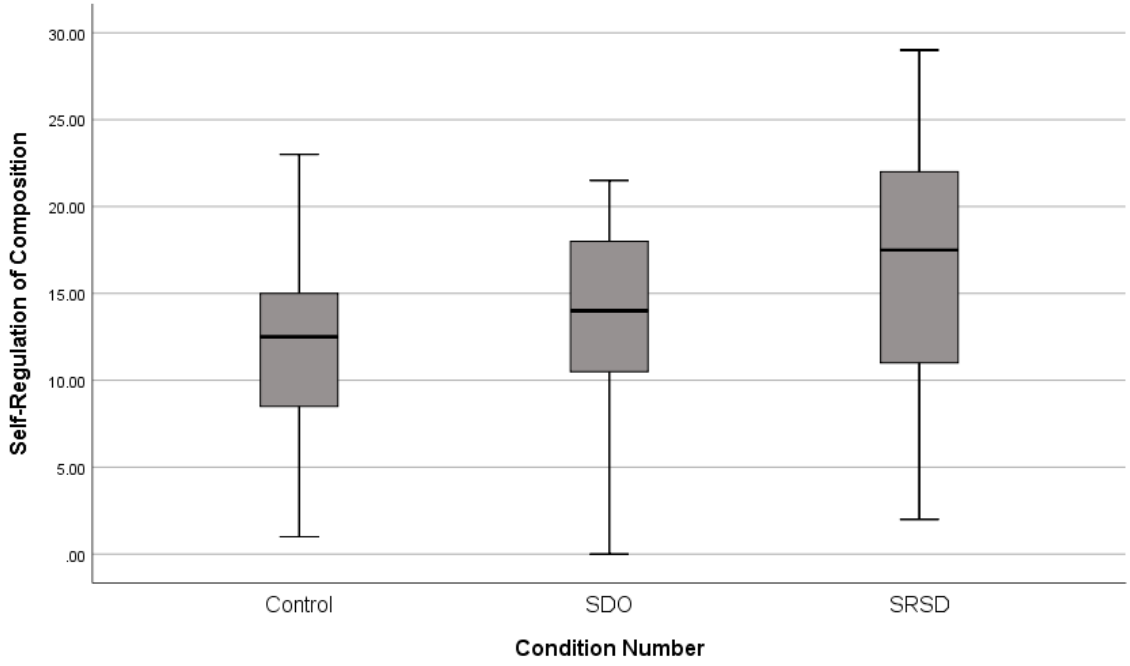

Figure 13. Boxplots of Self-Regulation of Composition by Condition to Check for Outliers

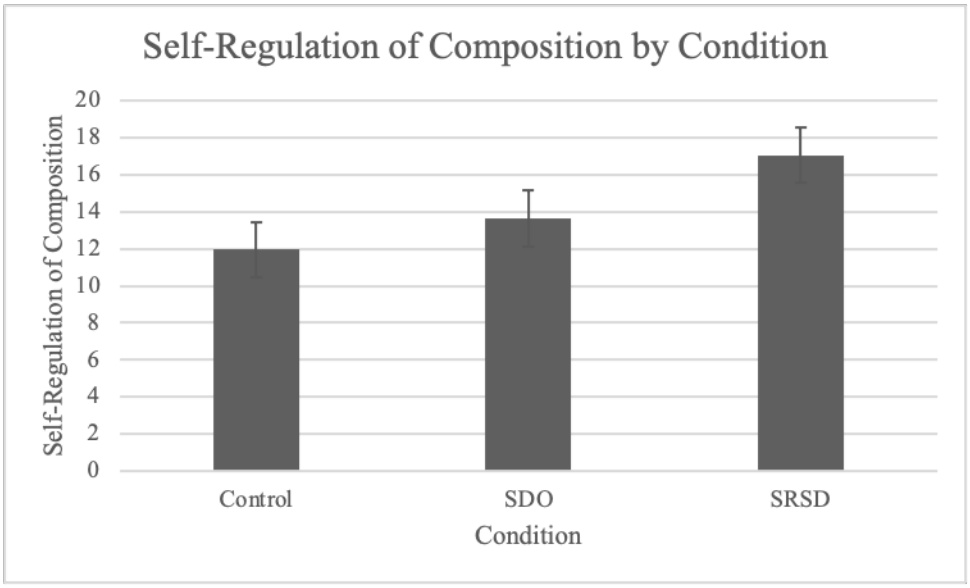

Figure 14. Bar Graph of Self-Regulation of Composition by Condition

Chapter 5

5 Discussion

Self-regulation is a complex and dynamic process that is central to the writing process (Bandura, 1991; Berninger & Chanquoy, 2012; Hayes & Flower, 1980). Writing is a goal-directed and cyclical activity that requires high levels of self-regulation to execute successfully (Hayes & Flower, 1980). According to Berninger and Chanquoy (2012), self-regulation in the context of writing is the ability to be strategic when writing, stay motivated, navigate problems that may be encountered, and actively monitor writing quality. Zimmerman and Risemberg (1997) extended our understanding of the role of self-regulation in writing by specifying the environmental, behavioural, and personal/covert processes required to produce a quality composition. While there has been research examining the impact of self-regulation on writing quality across different ages and grades, a meta-analysis by Santangelo et al. (2016) called for research on self-regulation in writing in Grade 1. To study the role of self-regulation in writing development, we needed a reliable and valid tool to assess self-regulation in early writing. For that reason, we developed the Interview on Self-Regulation in Early Writing. The purpose of this study was to evaluate the reliability and validity of this novel measure.

Measurement is one of the most difficult challenges facing psychology and the behavioural sciences (DeVillis, 2016, p. 2). To have confidence that an assessment is measuring the intended constructs, the measure needs to be evaluated to determine whether it provides consistent results internally, whether it can be used consistently by

different raters, and whether it is sensitive to differences in writing skill and instruction. The results of the reliability and validity analyses are discussed, followed by the unique contributions of this work to the literature, implications, limitations, and future directions for research.

5.1 Reliability

Reliability is the extent to which a measure performs in a consistent and predictable fashion (DeVillis, 2016, ch. 3). To evaluate whether the Interview on Self-Regulation in Early Writing produced consistent internal results and was able to be used consistently by separate raters, the reliability of the interview was assessed in three ways: (1) Inter-rater reliability, (2) Internal Consistency, and (3) Scale Reliability.

5.1.1 *Inter-Rater Reliability*

Inter-rater reliability is the, "extent to which raters generate scores that correspond," (DeVillis, 2016, p. 67). This type of reliability is important to ensure coding systems are consistent, replicable, and unbiased. Consistent with the study's hypothesis, the Interview on Self-Regulation in Early Writing had excellent inter-rater reliability. According to Pearson's product-moment correlations, inter-rater reliability was high for the transcription dimension, composition dimension and scale total (Cohen, 1988). This means that separate raters were able to reliably code self-regulatory behaviours at the transcription- and composition-levels in Grade 1 students using the Interview on Self-Regulation in Early Writing's coding scheme.

Internal Consistency

Internal consistency reliability is, "the homogeneity of the items within a scale" (DeVillis, 2016, p. 42). If the items of a scale are strongly related to each other, this

means they should also be strongly related to the latent variable. This was assessed by examining the strength of the correlations between items on a scale.

Full Scale. Contrary to the study's hypothesis, the Interview on Self-Regulation in Early Writing with all 16 items did not have high internal consistency and all items were not able to be retained. Based on item means, low correlations with other items and the scale overall, it was decided that the transcription-level planning, goal setting, strategy, self-statements, coping and self-reinforcement items would be removed to improve the reliability of the full scale. This was consistent with the low response rate to these items. By removing these six items, the internal reliability of the full scale increased from $\alpha = .63$ to $\alpha = .68$ which is considered an acceptable level of internal consistency. The final scale consisted of two transcription-level items (i.e., self-monitoring and reviewing) and eight composition-level items (i.e., planning, goal setting, strategy, self-statements, self-monitoring, coping, reviewing and self-reinforcement). While these analyses did not support the study's hypothesis, reducing the number of items improved the reliability of the full scale.

Transcription Subscale. When analyzed separately, the two-item transcription subscale did not have high internal consistency $\alpha = .61$. This suggested that these items formed a somewhat independent subscale.

Composition Subscale. When analyzed separately, the eight-item transcription subscale had high internal consistency $\alpha = .75$. This suggested that these items formed an independent subscale.

5.2 Validity

Validity is the extent to which the latent variable represents the true score (DeVillis, 2016, ch. 4). The current study examined validity by evaluating the following constructs: internal structure, convergent validity, divergent validity, and predictive validity.

5.2.1 *Internal Structure*

One way to consider validity is by examining the internal structure of a measure. Using data reduction techniques to better understand the component structure of a scale can help determine if items measure the hypothesized constructs. A PCA of the ten-item version of the Interview on Self-Regulation in Early Writing resulted in a two-component solution; these were interpreted as a transcription-level component and a composition level component. This supported the study's hypothesis that a PCA would result in a two-component solution. The two-component solution explained 46. 5% of the total variance in the scale and produced a simple factor structure. This provided strong evidence that self-regulation at the transcription- and composition-levels are separate aspects of self-regulation in early writing. Because this finding was consistent with the study's hypothesis, this provides support for the internal structure of the Interview on Self-Regulation in Early Writing. It also provides evidence for one aspect of validity.

5.2.2 *Convergent Validity*

Convergent validity is the extent to which two measures that should be related, are (DeVillis, 2016, ch. 4). Think-aloud protocols are an online assessment of metacognition while engaged in a specific task (Veenman, 2005). For this assessment, students were asked to speak aloud their thoughts to give researchers insight into their

internal processes while completing a personal narrative writing task. The goal of this assessment was to obtain another measure of self-regulation in writing to assess convergent validity. However, students rarely provided self-regulatory responses in the think-aloud measure. For that reason, convergent validity could not be formally assessed.

With that said, the poor performance on this measure was revealing. Writing is a demanding task that taxes working memory, especially in young writers that are still developing the prerequisite transcription skills to express their ideas in writing (e.g., Baddeley, 1986, 1996). In a think-aloud writing activity, writers are asked to speak aloud their thoughts. This activity requires high levels of working memory and metacognitive awareness to be able to identify a thought and express it while completing a writing task. For young writers that are still devoting working memory resources to spelling and printing, asking them to express their thoughts while also generating ideas for a story appeared too difficult. Only five of the 51 students that completed the activity provided self-regulatory thoughts.

5.2.3 *Divergent Validity*

Divergent validity is the extent to which two measures that should not be related, are not (DeVillis, 2016, ch. 4). Based on a review of the literature, one would not expect that a relation would be found between self-regulation of composition and student spelling knowledge (e.g., Kent et al., 2014). Indeed, no relation was found between these variables in the current study. This provided further evidence for the validity of the Interview on Self-Regulation in Early Writing by explaining not only how self-regulation of composition relates to other variables, but also describing what it does not relate to. By completing this analysis, we obtained a more in-depth understanding of the role of self-

regulation in early writing development that was consistent with past literature (e.g., Gerde et al., 2012).

5.2.4 *Predictive Validity*

Predictive validity is the extent to which a measure predicts an outcome assessed by another measure (DeVillis, 2016, ch. 4). Bornstein (2011) proposed an additional method of assessing validity: the process-focused model. Overall, the process-focused model of validity emphasizes process over outcome and experimental methods over correlational methods. A process-focused approach to validity testing was employed in the current study to evaluate the Interview on Self-Regulation in Early Writing. Experimental methods, namely, a waitlist control design and two instructional conditions to isolate unique effects of self-regulation on writing quality, facilitated a process-focused approach to validity testing in the current study. By using experimental over correlational methods, we were able to evaluate predictive validity by assessing self-regulation from pretest to posttest, across instructional groups. All planned analyses provided support for the predictive validity of the Interview on Self-Regulation in Early Writing by illustrating that self-regulation instruction resulted in significantly higher scores on the novel tool designed to measure self-regulation.

While self-regulation of composition correlated with posttest holistic writing quality, self-regulation of transcription did not. Based on the review of the literature on the relation between self-regulation and writing quality, it was expected that self-regulation of composition would predict writing quality, which it did (e.g., Santangelo et al., 2016). In a regression analysis, composition of self-regulation accounted for 24. 4% of the variance in posttest holistic writing quality. A mediation analysis was conducted to

better understand the effect of condition and pretest holistic writing quality on posttest holistic writing quality once self-regulation of composition was added into the model. In the initial model, condition and pretest holistic writing quality explained 27% of the variance in posttest writing quality. Both predictors added statistical significance to the prediction. Once self-regulation of composition was added into the model, pretest holistic writing quality and self-regulation of composition explained 37% of the variance in posttest holistic writing quality. The previously statistically significant effect of condition was no longer significant. This suggested that the effect of instruction on posttest holistic writing quality was mediated by knowledge of self-regulation of composition. This finding supported the overall purpose of the current study by suggesting that SRSD instruction is effective *because* it increases knowledge of self-regulation. It also reinforced that self-regulation is important for learning. For these reasons, it is necessary to be able to assess self-regulation for both research and practice.

In a one-way analysis of variance, students in the SRSD instructional condition demonstrated significantly more knowledge of self-regulation of composition than both the control and strategy-only groups. Taken together, these findings provided experimental evidence for the predictive validity of the Interview on Self-Regulation in Early Writing. By taking a process-focused approach to validity testing, we have strong evidence that the Interview on Self-Regulation in Early Writing measured the intended constructs and demonstrated predictive validity.

5.2.5 *Transcription-Level Versus Composition-Level Self-Regulation*

In Berninger and Chanquoy (2012), the authors explained that transcription-level (i.e., spelling and handwriting) skills are foundational to, and distinct from, self-

regulation. The discriminant validity findings from the current study support this notion. If we conceptualize self-regulation as a complex and dynamic process that enables individuals to manage personal, environmental, and behavioural factors (Bandura, 1991), then self-regulation should be possible at the transcription- and composition-levels. The current study was the first to design and test a measure that distinguished between levels of self-regulation in this way. The reason it was important to conceptualize self-regulation in a developmentally sensitive way was because of the young population this measure was designed for.

Generally, young writers seem to need to consciously think about transcription (e.g., Kim et al., 2017). For that reason, it was expected that self-regulation of transcription could contribute to writing quality. In the current study, some early writers demonstrated that they used self-regulation while engaged in transcription-level processes, especially when reviewing their spelling. The current findings suggested that transcription itself is subject to self-regulation to some extent, at least during learning. However, the results of the current study did not fully support the idea that self-regulation of transcription contributes to writing quality. While previous research has demonstrated that transcription is important, self-regulation of transcription itself does not appear to be an important skill for early writers.

The notion that early writers can demonstrate self-regulatory behaviours at the transcription- and composition-levels was confirmed in a PCA. The transcription and composition items on the interview formed different components. Nearly half of the students (44%) provided transcription-level responses to these items which provided support that early writers execute self-regulatory behaviours at the transcription-level

(i.e., spelling and punctuation), especially when asked about reviewing behaviours. However, contrary to expectations, students only provided self-regulatory responses at the transcription-level to the self-monitoring and reviewing items that asked students about what they needed to check their work for. That is, students regulated transcription specifically by "checking" their writing, and not by, for example, planning etc.

5.3 Standards for Educational and Psychological Testing (2014)

According to the Standards for Educational and Psychological Testing (2014), the best practices for assessing the construct validity of a measure include evidence based on content, response process, internal structure, relations with other variables and the consequences of testing. The Interview on Self-Regulation in Early Writing was evaluated according to these standards which is discussed below.

5.3.1 Does the Interview on Self-Regulation in Writing have Evidence Based on Content?

Evidence based on content examines the extent to which a questionnaire aligns with the most current empirical research on the topic, in this case, self-regulation in writing. The Interview on Self-Regulation in Early Writing was designed to be consistent with Zimmerman and Risemberg's (1997) Self-Regulation Model of Writing and measure the self-regulation processes outlined in Reid et al. (2013). It specifically assessed self-verbalizations, self-monitoring, self-consequating, planning, goal setting, use of cognitive strategies, and self-evaluative standards. While self-monitoring, planning, goal setting, use of cognitive strategies, and self-evaluative standards have been formerly assessed in the literature, no studies have specifically assessed self-verbalizations or self-consequating (Santangelo et al., 2016). Not only was the Interview

on Self-Regulation in Early Writing developed to have validity evidence based on content, it made a unique contribution to the literature by exploring these concepts with younger children and in the context of writing and learning a new writing strategy and new self-regulatory processes.

5.3.2 Does the Interview on Self-Regulation in Writing have Evidence Based on Response Process?

Evidence based on response process evaluates if: (1) Students understand the items and types of responses required, (2) Students must access related information from their memories, (3) Students must integrate recalled information into a coherent response, (4) Students must match their recalled knowledge to a response and (5) Students must select/produce the correct answer.

Validity evidence based on response process was considered in two ways in this study: (1) The way the questions were developed and (2) Item reliability analyses. In the questionnaires examined in the literature review, most of the questions asked students about what they themselves would do in hypothetical scenarios. To minimize social desirability bias, the hypothetical scenarios asked about a fictional student in Grade 1. The goal was to reduce the likelihood that students would feel pressured to respond in a way that was socially acceptable or provide an answer to please the researcher. By minimizing social desirability bias, students should have been better able to understand the items and types of responses required, access related information from their memories, integrate recalled information into a coherent response, match their recalled knowledge to a response and select/produce the correct answer. While data was not

collected on the processes students used to generate their answers, the design of the questions intended to minimize this potential issue.

Item reliability analyses also explored validity evidence based on response process. By examining item means and reliability statistics, it was decided that six items needed to be removed from the transcription-dimension of the scale. For the removed items, student response rates were low. By removing those items, the Interview on Self-Regulation in Early Writing improved its adherence to this standard.

5.3.3 Does the Interview on Self-Regulation in Writing have Evidence Based on Internal Structure?

Evidence based on internal structure means that the relations between items or dimensions on an assessment are consistent with current theory. Based on the results of a PCA, the Interview on Self-Regulation in Early Writing demonstrated validity evidence based on internal structure. Consistent with Berninger and Chanquoy's (2012) cognitive models of writing development, the current study found discriminant validity between a curriculum-based measure of spelling and the measure of self-regulation. Therefore, the Interview on Self-Regulation in Early Writing demonstrated validity evidence based on internal structure. The current study also extended our understanding of writing development by highlighting that young writers seem to self-regulate elements of the transcription process, such as spelling, when they are learning.

5.3.4 Does the Interview on Self-Regulation in Writing have Evidence Based on Relations with Other Variables?

Evidence based on relations with other variables is composed of three elements: convergent relations, discriminant relations, and predictive relations. Convergent

relations demonstrate that the measure being assessed shows a strong, positive relation with another established behavioural measure of the same theoretical construct. Unfortunately, it was not possible to assess convergent validity in this study because students tended to generate and transcribe text during the think-aloud measure instead of overt self-regulation. This could be a function of their age, or because they had internalized the text structure.

Discriminant relations means no relation is found between the measure being assessed and a theoretically dissimilar construct. Based on a review of the literature, one would not expect that a relation would be found between self-regulation of composition and student spelling knowledge (e.g., Kent et al., 2014). In support of discriminant relations, there was no significant correlation between the composition dimension of the Interview on Self-Regulation in Early Writing and the curriculum-based measure of spelling. This finding provided further validity evidence for the Interview on Self-Regulation in Early Writing.

Predictive relations demonstrate the measure being assessed can be used to predict theoretically related constructs. The Interview on Self-Regulation in Early Writing demonstrated predictive validity in numerous ways. First, self-regulation of composition explained 24.4% of the variability in posttest holistic writing quality. This finding was consistent with the extant research on self-regulation and writing quality (Fidaglo et al., 2008). Second, in a mediation analysis, once self-regulation of composition was added into a multiple regression model with condition and pretest holistic writing quality as predictors, the previously significant effect of instructional condition on posttest holistic writing quality was nullified. Third, students in the SRSD

instructional condition demonstrated significantly more self-regulation of composition knowledge than both the strategy-only and control groups which suggested that self-regulation is sensitive to instruction. Because all three analyses supported the study's hypotheses, they provided support for validity evidence based on relations with other variables.

In sum, the current study provided evidence for validity based on predictive and discriminate relations with other variables.

5.3.5 Does the Interview on Self-Regulation in Writing have Evidence Based on Consequences of Testing?

Lastly, evidence based on consequences of testing means that any interpretations or consequences that result from the measure are theoretically sound (Wolters & Won, 2018). The Interview on Self-Regulation in Writing was developed to be consistent with current theories self-regulation in writing and SRSD instruction (e.g., Reid et al., 2013; Zimmerman & Risemberg, 1997). Given the novelty of this tool, the theoretical soundness of the interpretations or consequences of testing are unknown at this time.

Based on the recommendations for the best practices for evaluating validity according to the Standards for Educational and Psychological Testing (2014), the Interview on Self-Regulation in Early Writing demonstrated some validity evidence based on content, response process, internal structure, and predictive relations with other variables. The consequences of testing are currently unknown because of the novelty of the tool.

5.4 Research Implications

In the literature, both task-specific and task-general measures were used to assess self-regulation. Based on a literature review, higher levels of self-regulation should be associated with better academic outcomes. However, Crede and Phillips (2011) found in a meta-analysis that a general measure of self-regulation did not correlate strongly with overall GPA. This suggested that it may be important to use subject-specific measures to assess self-regulation and that self-regulation skills in one subject area do not spontaneously transfer to another subject area. In the current study, the Interview on Self-Regulation in Early Writing was not only specific to writing, but specific to the genre of personal narrative. This enabled researchers to obtain a task-specific measurement of the student's knowledge of self-regulation in writing, specific to the genre of personal narrative. The results of the current study described domain-specific self-regulation in young children which was a step towards bridging this gap in the literature.

The Interview on Self-Regulation in Early Writing was designed to address previously identified methodological issues with using self-reports in younger students such as reliance on long-term memory, and underdeveloped abstract thinking and/or insight (e.g., Schneider et al., 2017). Firstly, because the Interview on Self-Regulation in Early Writing was a task-specific measure that assessed student knowledge of self-regulation in the areas of planning, criteria/goal setting, strategy/text structure, self-statements, self-monitoring, coping, reviewing and self-reinforcement, there was no need for students to rely on their long-term memories to make frequency judgements. Secondly, the questions were structured to minimize the amount of insight and abstract thinking required on the part of the student by providing concrete, age-appropriate

hypothetical scenarios and providing more specific follow-up questions to support student understanding when required. However, students who were able to provide self-regulatory responses to interview questions without additional prompts were rewarded in the coding scheme because being able to generate self-regulatory responses without prompting is itself reflective of their level of self-regulatory knowledge. Schneider et al. (2017) also highlighted that there are special considerations needed when designing, administering and validating measures intended to assess self-regulation, especially in younger students. Given the developmental considerations that were used to design, administer and validate the Interview on Self-Regulation in Early Writing, we would argue that the current study addressed Schneider et al. 's (2017) concerns and made significant contributions to the literature on assessing self-regulation in writing and assessing self-regulation in younger students.

In contrast to previous studies that used general and behavioural measures to measure the role of self-regulation in writing development, the current study was the first to develop an interview tool specifically designed to measure task-specific self-regulation of writing in early writers and illustrate good self-regulation behaviours in early writing (Kent et al., 2014; Gerde et al., 2012). Using the Interview on Self-Regulation in Early Writing in studies that examine the development of self-regulation in early literacy could add new information to the literature.

In their paper, they called for research at different grade levels. The current study addressed an identified gap in Santangelo et al.'s (2016) meta-analysis and added to the literature by developing a novel task-specific and developmentally appropriate tool to measure self-regulation in writing in Grade 1.

Nuanced analyses demonstrated that self-regulation may explain more of the variance in writing quality than previously thought. In the literature on self-regulation and writing quality, there was variability in the findings about the extent to which elements of self-regulation explained variance in writing quality (e.g., Fidaglo et al., 2008; Olinghouse & Graham, 2009; Graham et al., 2017). The current study challenged and extended these previous findings. In a mediation analysis, once self-regulation of composition was added into the model, pretest holistic writing quality and self-regulation of composition explained 37% of the variance in posttest holistic writing quality. The previously statistically significant effect of condition was no longer significant. It is possible that measuring self-regulation with a developmentally appropriate and task-specific tool illuminated this finding. The current study began to bridge the gap in the literature between content-neutral self-regulation in early childhood and subject-specific self-regulation in later grades.

Another way the current study contributed to the literature was by considering the reliability and validity of the interview used to assess self-regulation in writing. In the literature review, it was unexpected that these issues of measurement were rarely explicitly considered. While Graham et al., (2005), Graham et al., (2017) and WijeKumar et al., (2019) used factor analyses to examine the structure of the self-report questionnaires they used to measure self-regulation in writing, only one study found by Golembek et al., (2019) explicitly examined the reliability and validity of a novel measure designed to measure self-regulation in writing. Therefore, the current study was the first to examine the reliability and validity of a novel, task-specific measure designed to assess self-regulation of writing knowledge in Grade 1.

5.5 Clinical Implications

While there is ample research on reading disorders, assessment, and intervention, far less is known about writing disorders (Zoccolotti & Friedmann, 2010). There is a lack of established tools that examine self-regulation in writing for both typically developing and clinical populations. This study examined how to assess self-regulation in typically developing writers. While the current study found preliminary support for its reliability and validity, it is not a standardized measure (i.e., it is not age-normed). This measure is also specific to the genre of personal narrative. While this is one of the most common genres for beginning writers, it is not the only writing genre. For these reasons, we are not suggesting that the Interview on Self-Regulation in Early Writing be used formally in clinical assessments. However, school and child clinical psychologists would benefit from an awareness of the role of self-regulation in writing and an understanding of the developmental trajectory of learning to write.

In the literature, there is ample support (including the results of the current study) that self-regulation is associated with text quality (e.g., Santangelo et al., 2016). Clinically, the questions from the Interview on Self-Regulation in Early Writing could serve as the basis for an informal interview for struggling writers. By asking students about hypothetical writing scenarios, clinicians would be able to get a sense of their knowledge of their planning, goal setting, strategy, self-statement, self-monitoring, coping, reviewing and self-reinforcement skills. With further development, the Interview on Self-Regulation in Early Writing could be used to assess whether students have knowledge about self-regulation in writing and make recommendations to teach to their learning deficits. Students with a variety of disorders including learning disabilities and

attention-deficit/hyperactivity disorder have been demonstrated to have poor self-regulation with respect to academic tasks (e.g., Baird et al., 2009; Re & Cornoldi, 2010). Assessing and addressing gaps in self-regulation knowledge in relation to strategy instruction in writing could benefit these students.

5.6 Study Limitations

The current study had numerous limitations that are important to discuss. First and most importantly, this instrument was only tested once, rather than iteratively. This limitation of the current study resulted from time constraints and the influence of the onset of the COVID-19 pandemic on the ability to collect data in schools. Second, given that this was an instructional study, one could argue that this study "tested to the instruction the students received". That is, the measure we used to assess self-regulation was designed to test elements of self-regulation that are taught in the SRSD instructional units. While this is certainly a potential limitation of the current study, the coding system was intentionally designed to help mitigate this bias. In the Reference for Raters (Appendix A), for each score category for each item, two examples were provided: one that was consistent with the instruction provided in the study and one that was not derived from the instruction. This helped ensure that students who were not in the instructional group had an equal opportunity to gain points in each score category for each item. Third, the author of this dissertation who helped develop the coding system for the interview was one of two raters that completed the coding of the Interview on Self-Regulation in Early Writing. While this is a potential source of bias, the high inter-rater reliability suggested that the author was coding the interviews similarly to the other rater who was not involved in the development of the coding system. Fourth, with respect to

reliability, a common way to evaluate scale reliability is test-retest reliability to evaluate the stability of the measure over time. This was not included as a part of the current study and would be beneficial in future studies using this tool. Fifth, because of the poor response rate to the think-aloud assessment, concurrent validity was unable to be assessed. To ensure that the Interview on Self-Regulation in Early Writing truly captures knowledge of self-regulation, future research would benefit from comparing student performance on the current study's interview with a measure of self-regulation that is not related to instruction. This was a significant limitation of the current study. Lastly, the Interview on Self-Regulation in Early Writing is specific to the genre of personal narrative. This would limit the generalizability of this writing assessment to other genres.

5.7 Future Directions for Research

Based on the results of the current study, there are numerous future directions for research. First and most importantly, the instrument needs to be administered to another sample of Grade 1 writers to evaluate the reliability and validity of the tool with unreliable items removed. This process is an important aspect of measure development that is missing from the current study. Further, to address the potential bias that students in the strategy-only and SRSD conditions would have an advantage on this measure, the study's methodology could be reconceptualized. Administering the interview to Grade 1 students who had not received instruction in self-regulation and examining the differences between low- and high-achieving writers could examine whether it was sensitive to differences in knowledge of self-regulation, independent of instruction. If higher-achieving students demonstrated more knowledge of self-regulation than lower-achieving students, this would provide additional support for the validity of the Interview

on Self-Regulation in Early Writing. Alternatively, students could be administered both the Interview on Self-Regulation in Early Writing and a behavioural measure of self-regulation to evaluate the relation between self-regulation of writing and self-regulation of behaviour. To illustrate, student behaviours that could be observed such as self-monitoring with a checklist, consulting lists of self-statements and rereading their own texts are examples of self-regulatory behaviours that could be measured observationally and then correlated with the interview.

Writing is a demanding task for most students, but it is particularly challenging for students with exceptionalities such as learning disabilities and attention-deficit/hyperactivity disorder who tend to struggle with negative self-talk and executive functioning skills like working memory (e.g., Baird et al., 2009; Re & Cornoldi, 2010). Given that these students struggle inherently with self-regulation as a function of their exceptionalities and that we know writing difficulties tend to emerge early and persist across the lifespan, more research on the ways to assess and intervene with respect to writing with these populations is badly needed (e.g., Abbott et al., 2010; Costa et al., 2015; Costa et al., 2018; Scarborough, 1998). Administering the Interview on Self-Regulation in Early Writing to these populations may help researchers and clinicians to understand how knowledge of self-regulation in writing contributes to overall writing difficulties which in turn could help with the development of targeted assessment and intervention strategies for exceptional populations.

5.8 Conclusion

The Interview on Self-Regulation in Early Writing is a 10-item measure designed to measure knowledge of transcription- and composition-level self-regulation. Multiple

aspects of reliability and validity were examined. The data from one-hundred and seventeen Grade 1 students were used to evaluate inter-rater reliability and the internal consistency of the scale. A principal components analysis (PCA) evaluated the internal structure of the scale. Discriminant and predictive validity were also evaluated. The Interview on Self-Regulation in Early Writing had acceptable inter-rater reliability and internal consistency (once unreliable items were removed). The Interview on Self-Regulation in Early Writing also demonstrated validity evidence based on content, response process, internal structure, and predictive relations with other variables. In conclusion, the results of the current study provided empirical support for the reliability and validity of the Interview on Self-Regulation in Early Writing and contributed to the literature by describing self-regulatory behaviours in Grade 1 and beginning to bridge the gap in the research literature between content-neutral self-regulation in early elementary and subject-specific self-regulation in later grades.

References

Abbott, R. D., Berninger, V. W., & Fayol, M. (2010). Longitudinal relationships of levels of language in writing and between writing and reading in grades 1 to 7. *Journal of Educational Psychology, 102*(2), 281-298. http://dx.doi.org.proxy1.lib.uwo.ca/10.1037/a0019318

Allison, P. D. (2009). Missing data. In R. E. Millsap, & A. Maydeu-Olivares (Eds.), *The Sage handbook of quantitative methods in psychology; The Sage handbook of quantitative methods in psychology* (pp. 72-89, Chapter xxvi, 759 Pages). Sage Publications Ltd, Thousand Oaks, CA. http://dx.doi.org.proxy1.lib.uwo.ca/10.4135/9780857020994.n4

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Publisher.

Azevedo, R., Moos, D. C., Greene, J. A., Winters, F. I., & Cromley, J. G. (2008). Why is externally-facilitated regulated learning more effective than self-regulated learning with hypermedia? *Educational Technology Research and Development, 56*(1), 45-72. http://dx.doi.org.proxy1.lib.uwo.ca/10.1007/s11423-007-9067-0

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

Baddeley, A. D. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 49A*, 5-28.

Baird, G. L., Scott, W. D., Dearing, E., & Hamill, S. K. (2009). Cognitive self-regulation in youth with and without learning disabilities: Academic self-efficacy, theories of intelligence, learning vs. performance goal preferences, and effort attributions. *Journal of Social and Clinical Psychology, 28*(7), 881-908. http://dx.doi.org.proxy1.lib.uwo.ca/10.1521/jscp.2009.28.7.881

Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes, 50*(2), 248-287. http://dx.doi.org.proxy1.lib.uwo.ca/10.1016/0749-5978(91)90022-L.

Bandura, A. (2006). Guide for constructing self-efficacy scales. *Self-efficacy beliefs of adolescents*, *5*(1), 307-337.

Scardamalia, M., & Bereiter, C. (1987). Knowledge telling and knowledge transforming in written composition. *Advances in applied psycholinguistics*, *2*, 142-175.

Berninger, V. W., & Chanquoy, L. (2012). What writing is and how it changes across early and middle childhood development: a multidisciplinary perspective. In E. L. Grigorenko, E. Mambrino & D. D. Preiss (Eds.), *Writing: A Mosaic of New Perspectives* (65-84). New York; NY: Psychology Press.

Berninger, V. W., Cartwright, A. C., Yates, C. M., Swanson, H. L., & Abbott, R. D. (1994). Developmental skills related to writing and reading acquisition in the intermediate grades: Shared and unique functional systems. *Reading and Writing: An Interdisciplinary Journal, 6*(2), 161-196. http://dx.doi.org.proxy1.lib.uwo.ca/10.1007/BF01026911

Berninger, V., Whitaker, D., Feng, Y., Swanson, H. L., & Abbott, R. D. (1996). Assessment of planning, translating, and revising in junior high writers. *Journal of*

School Psychology, 34(1), 23-52. http://dx.doi.org.proxy1.lib.uwo.ca/10.1016/0022-4405(95)00024-0

Bornstein, R. F. (2011). Toward a process-focused model of test score validity: Improving psychological assessment in science and practice. *Psychological Assessment, 23*(2), 532-544. http://dx.doi.org.proxy1.lib.uwo.ca/10.1037/a0022402

Bortz, J., & Döring, N. (2006). Quantitative methods of data collection. In *research methods and evaluation* (pp. 137-293). Springer, Berlin, Heidelberg.

Boscolo, P. (1990). The construction of expository text. *First Language, 10*(3), 217-230. http://dx.doi.org.proxy1.lib.uwo.ca/10.1177/014272379001003003

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*(2), 81-105. http://dx.doi.org.proxy1.lib.uwo.ca/10.1037/h0046016

Cano, F. (2006). An In-Depth Analysis of the Learning and Study Strategies Inventory (LASSI). *Educational and Psychological Measurement, 66*(6), 1023-1038. http://dx.doi.org.proxy1.lib.uwo.ca/10.1177/0013164406288167

Cizek, G. J., Bowen, D., & Church, K. (2010). Sources of validity evidence for educational and psychological tests: A follow-up study. *Educational and Psychological Measurement, 70*(5), 732-743. http://dx.doi.org.proxy1.lib.uwo.ca/10.1177/0013164410379323

Cohen, J. (1988). Set correlation and contingency tables. *Applied Psychological Measurement, 12*(4), 425-434. http://dx.doi.org.proxy1.lib.uwo.ca/10.1177/014662168801200410

Costa, L. C., Edwards, C. N., & Hooper, S. R. (2016). Writing disabilities and reading disabilities in elementary school students: Rates of co-occurrence and cognitive burden. *Learning Disability Quarterly, 39*(1), 17-30. http://dx.doi.org.proxy1.lib.uwo.ca/10.1177/0731948714565461

Costa, L., Green, M., Sideris, J., & Hooper, S. R. (2018). First-grade cognitive predictors of writing disabilities in grades 2 through 4 elementary school students. *Journal of Learning Disabilities, 51*(4), 351-362. http://dx.doi.org.proxy1.lib.uwo.ca/10.1177/0022219417721182

Credé, M., & Phillips, L. A. (2011). A meta-analytic review of the Motivated Strategies for Learning Questionnaire. *Learning and Individual Differences, 21*(4), 337-346. http://dx.doi.org.proxy1.lib.uwo.ca/10.1016/j.lindif.2011.03.002

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*(4), 281-302. http://dx.doi.org.proxy1.lib.uwo.ca/10.1037/h0040957

DeVellis, R. F. (2003). *Scale development: Theory and applications*. Sage Publications.

DeVellis, R. F. (2016). *Scale development: Theory and applications*. Sage Publications.

Fidalgo, R., Torrance, M., & García, J. (2008). The long-term effects of strategy-focussed writing instruction for grade six students. *Contemporary Educational Psychology, 33*(4), 672-693. http://dx.doi.org.proxy1.lib.uwo.ca/10.1016/j.cedpsych.2007.09.001

Fitzgerald, J. (1987). Research on revision in writing. *Review of Educational Research*, *57*(4), 481-506.

Gerde, H. K., Skibbe, L. E., Bowles, R. P., & Martoccio, T. L. (2012). Child and home predictors of children's name writing. *Chid Development Research*, *2012*. https://doi.org/10.1155/2012/748532

Gillespie, A., Olinghouse, N. G., & Graham, S. (2013). Fifth-grade students' knowledge about writing process and writing genres. *The Elementary School Journal, 113*(4), 565-588. http://dx.doi.org.proxy1.lib.uwo.ca/10.1086/669938

Golombek, C., Klingsieck, K. B., & Scharlau, I. (2019). Assessing self-efficacy for self-regulation of academic writing: Development and validation of a scale. *European Journal of Psychological Assessment, 35*(5), 751-761. http://dx.doi.org.proxy1.lib.uwo.ca/10.1027/1015-5759/a000452

Graham, S. (2006). Writing. In P. Alexander & P. Winne (Eds.), *Handbook of educational psychology* (pp. 457– 478). Routledge.

Graham, S., & Harris, K. R. (1996). Self-regulation and strategy instruction for students who find writing and learning challenging. In C. M. Levy, & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 347-360). Erlbaum.

Graham, S., Harris, K. R., & Mason, L. (2005). Improving the writing performance, knowledge, and self-efficacy of struggling young writers: The effects of self-regulated strategy development. *Contemporary Educational Psychology, 30*(2), 207-241. http://dx.doi.org.proxy1.lib.uwo.ca/10.1016/j.cedpsych.2004.08.001

Graham, S., Kiuhara, S. A., Harris, K. R., & Fishman, E. J. (2017). The relationship among strategic writing behavior, writing motivation, and writing performance with

young, developing writers. *The Elementary School Journal, 118*(1), 82-104. http://dx.doi.org.proxy1.lib.uwo.ca/10.1086/693009

Graham, S., Schwartz, S. S., & MacArthur, C. A. (1993). Knowledge of writing and the composing process, attitude toward writing, and self-efficacy for students with and without learning disabilities. *Journal of Learning Disabilities, 26*(4), 237-249. http://dx.doi.org.proxy1.lib.uwo.ca/10.1177/002221949302600404

Gravetter, F. J., & Forzano, L. A. B. (2012). *Research methods for the behavioral sciences* (4th ed., pp. 71-106). Cengage Learning.

Gravetter, F. J., Wallnau, L. B., Forzano, L. A. B., & Witnauer, J. E. (2020). *Essentials of statistics for the behavioral sciences*. Cengage Learning.

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. http://dx.doi.org.proxy1.lib.uwo.ca/10.2190/G7M1-2734-3JRR-8033

Greene, J. A., Deekens, V. M., Copeland, D. Z., & Yu, S. (2018). Capturing and modeling self-regulated learning using think-aloud protocols. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (pp. 323-337). Routledge.

Greene, J. A., Hutchison, L. A., Costa, L., & Crompton, H. (2012). Investigating how college students' task definitions and plans relate to self-regulated learning processing and understanding of a complex science topic. *Contemporary Educational Psychology, 37*(4), 307-320. http://dx.doi.org.proxy1.lib.uwo.ca/10.1016/j.cedpsych.2012.02.002

Hammill, D., & Larsen, S. (1996). Test of Written Language- Third Edition. PRO-ED Inc.

Harris, K. R., Graham, S., & Mason, L. H. (2006). Improving the writing, knowledge, and motivation of struggling young writers: Effects of self-regulated strategy development with and without peer support. *American Educational Research Journal*, *43*(2), 295-340. http://dx.doi.org.proxy1.lib.uwo.ca/10.3102/00028312043002295

Hayes, J. R., & Flower, L. S. (1986). Writing research and the writer. *American Psychologist, 41*(10), 1106-1113. http://dx.doi.org.proxy1.lib.uwo.ca/10.1037/0003-066X.41.10.1106

Hayes, L., & Flower, L. (1980). Identifying the organization of writing processes. In L. Gregg & E. Steinberg (Eds.), *Cognitive processes in writing* (pp. 3-30). Erlbaum.

Hogan, T. P., & Agnello, J. (2004). An empirical study of reporting practices concerning measurement validity. *Educational and Psychological Measurement*, *64*(5), 802-812. http://dx.doi.org.proxy1.lib.uwo.ca/10.1177/0013164404264120

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika, 39*(1), 31-36. http://dx.doi.org.proxy1.lib.uwo.ca/10.1007/BF02291575

Kent, S., Wanzek, J., Petscher, Y., Al Otaiba, S., & Kim, Y. S. (2014). Writing fluency and quality in kindergarten and first grade: The role of attention, reading, transcription, and oral language. *Reading and Writing*, *27*(7), 1163-1188. http://dx.doi.org.proxy1.lib.uwo.ca/10.1007/s11145-013-9480-1

Kim, Y. G., & Schatschneider, C. (2017). Expanding the developmental models of writing: A direct and indirect effects model of developmental writing

(DIEW). *Journal of Educational Psychology, 109*(1), 35-50. http://dx.doi.org.proxy1.lib.uwo.ca/10.1037/edu0000129

Kim, Y., Al Otaiba, S., Folsom, J. S., Greulich, L., & Puranik, C. (2014). Evaluating the dimensionality of first-grade written composition. *Journal of Speech, Language, and Hearing Research, 57*(1), 199-211. http://dx.doi.org.proxy1.lib.uwo.ca/10.1044/1092-4388(2013/12-0152)

Laerd Statistics (2015). Cohen's kappa using SPSS Statistics. *Statistical tutorials and software guides*. Retrieved from https://statistics.laerd.com/

Laerd Statistics (2015). Cronbach's alpha using SPSS Statistics. *Statistical tutorials and software guides*. Retrieved from https://statistics.laerd.com/

Laerd Statistics (2015). Multiple regression using SPSS Statistics. *Statistical tutorials and software guides*. Retrieved from https://statistics.laerd.com/

Laerd Statistics (2015). Principal components analysis (PCA) using SPSS Statistics. *Statistical tutorials and software guides*. Retrieved from https://statistics.laerd.com/

Laerd Statistics (2015). Simple linear regression using SPSS Statistics. *Statistical tutorials and software guides*. Retrieved from https://statistics.laerd.com/

Laerd Statistics (2017). One-way ANOVA using SPSS Statistics. *Statistical tutorials and software guides*. Retrieved from https://statistics.laerd.com/

Laerd Statistics (2018). Pearson's product-moment correlation using SPSS Statistics. *Statistical tutorials and software guides*. Retrieved from https://statistics.laerd.com/

Limpo, T., Alves, R. A., & Fidalgo, R. (2014). Children's high-level writing skills: Development of planning and revising and their contribution to writing quality. *British Journal of Educational Psychology, 84* (2), 177-193. http://dx.doi.org.proxy1.lib.uwo.ca/10.1111/bjep.12020

McCutchen, D. (1995). Cognitive processes in children's writing: Developmental and individual differences. *Issues in Education: Contributions from Educational Psychology*, *1*(2), 123-160.

McMaster, K. L., Du, X., & Pétursdóttir, A. (2009). Technical features of curriculum-based measures for beginning writers. *Journal of Learning Disabilities, 42*(1), 41-60. http://dx.doi.org.proxy1.lib.uwo.ca/10.1177/0022219408326212

Olinghouse, N. G., & Graham, S. (2009). The relationship between the discourse knowledge and the writing performance of elementary-grade students. *Journal of Educational Psychology, 101*(1), 37-50. http://dx.doi.org.proxy1.lib.uwo.ca/10.1037/a0013462

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. http://dx.doi.org.proxy1.lib.uwo.ca/10.1177/0013164493053003024

Puranik, C. S., Boss, E., & Wanless, S. (2019). Relations between self-regulation and early writing: Domain specific or task dependent? *Early Childhood Research Quarterly, 46*, 228-239. http://dx.doi.org.proxy1.lib.uwo.ca/10.1016/j.ecresq.2018.02.006

Re, A. M., & Cornoldi, C. (2010). ADHD expressive writing difficulties of ADHD children: When good declarative knowledge is not sufficient. *European Journal of Psychology of Education, 25*(3), 315-323. http://dx.doi.org.proxy1.lib.uwo.ca/10.1007/s10212-010-0018-5

Reid, R., Lienemann, T. O., Hagaman, J. L. (2013). *Strategy instruction for students with learning disabilities* (2nd edition). The Guilford Press.

Saddler, B., & Graham, S. (2007). The relationship between writing knowledge and writing performance among more and less skilled writers. *Reading & Writing Quarterly*, *23*(3), 231-247. http://dx.doi.org.proxy1.lib.uwo.ca/10.1080/10573560701277575

Santangelo, T., Harris, K., & Graham, S. (2016). Self-regulation and writing: Meta-analysis of the self-regulation processes in Zimmerman and Risemberg's model. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 174-193). The Guilford Press.

Scarborough, H. S. (1998). Early identification of children at risk for reading disabilities: Phonological awareness and some other promising predictors. In B.K. Shapiro, P.J. Accardo & A.J. Capute (Eds.) *Specific reading disability: A view of the spectrum* (75-119). York Press, Inc.

Schneider, W., Lingel, K., Artelt, C., & Neuenhaus, N. (2017). Metacognitive knowledge in secondary school students: Assessment, structure, and developmental change. In *Competence Assessment in Education* (pp. 285-302). Springer.

Schwarzer, R., & Jerusalem, M. (1995). Generalized self-efficacy scale. *Measures in health psychology: A user's portfolio. Causal and control beliefs*, *1*(1), 35-37.

Schwarzer, R., & Renner, B. (2000). Social-cognitive predictors of health behavior: action self-efficacy and coping self-efficacy. *Health psychology*, *19*(5), 487.

Schyns, B., & Von Collani, G. (2002). A new occupational self-efficacy scale and its relation to personality constructs and organizational variables. *European Journal of Work and Organizational Psychology*, *11*(2), 219-241. http://dx.doi.org.proxy1.lib.uwo.ca/10.1080/13594320244000148

Tabassam, W., & Grainger, J. (2002). Self-concept, attributional style and self-efficacy beliefs of students with learning disabilities with and without attention deficit hyperactivity disorder. *Learning Disability Quarterly, 25*(2), 141-151. http://dx.doi.org.proxy1.lib.uwo.ca/10.2307/1511280

Thurstone, L. L. (1947). *Multiple factor analysis*. University of Chicago Press.

Van Hout-Wolters, B. H. A. M. (2009). Leerstrategieën meten. Soorten meetmethoden en hun bruikbaarheid in onderwijs en onderzoek. [Measuring Learning strategies. Different kinds of assessment methods and their usefulness in education and research]. Pedagogische Studiën, 86, 103–110.

Veenman, M. V. J. (2005). The assessment of metacognitive skills: What can be learned from multimethod designs? In C. Artelt & B. Moschner (Eds.), *Lernstrategien und Metakognition: Implikationen für Forschung und Praxis* (pp. 75–97). Waxmann.

Weiner, B. (1979). A theory of motivation for some classroom experiences. *Journal of Educational Psychology*, *71*(1), 3-25. http://dx.doi.org.proxy1.lib.uwo.ca/10.1037/0022-0663.71.1.3

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

Wijekumar, K., Graham, S., Harris, K. R., Lei, P., Barkel, A., Aitken, A., Ray, A., & Houston, J. (2019). The roles of writing knowledge, motivation, strategic behaviors, and skills in predicting elementary students' persuasive writing from source material. *Reading and Writing: An Interdisciplinary Journal, 32*(6), 1431-1457. http://dx.doi.org.proxy1.lib.uwo.ca/10.1007/s11145-018-9836-7

Wolters, C. A., Won, S., Schunk, D. H., & Greene, J. A. (2018). Validity and the use of self-report questionnaires to assess self-regulated learning. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (pp. 307-322). Routledge.

Zimmerman, B. J. (1989). A social cognitive view of self-regulated academic learning. *Journal of Educational Psychology, 81*(3), 329-339. http://dx.doi.org.proxy1.lib.uwo.ca/10.1037/0022-0663.81.3.329

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal, 45*(1), 166-183. http://dx.doi.org.proxy1.lib.uwo.ca/10.3102/0002831207312909

Zimmerman, B. J., & Pons, M. M. (1986). Development of a structured interview for assessing student use of self-regulated learning strategies. *American Educational Research Journal, 23*(4), 614-628. http://dx.doi.org.proxy1.lib.uwo.ca/10.2307/1163093

Zimmerman, B. J., & Risemberg, R. (1997). Becoming a self-regulated writer: A social cognitive perspective. *Contemporary Educational Psychology, 22*(1), 73-101. http://dx.doi.org.proxy1.lib.uwo.ca/10.1006/ceps.1997.0919

Zimmerman, B., & Kitsantas, A. (2007). Reliability and validity of self-efficacy for learning form (SELF) scores of college students. *Zeitschrift für Psychologie/Journal of Psychology*, *215*(3), 157-163. http://dx.doi.org.proxy1.lib.uwo.ca/10.1027/0044-3409.215.3.157

Zoccolotti, P., & Friedmann, N. (2010). From dyslexia to dyslexias, from dysgraphia to dysgraphias, from a cause to causes: A look at current research on developmental dyslexia and dysgraphia. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior, 46*(10), 1211-1215. http://dx.doi.org.proxy1.lib.uwo.ca/10.1016/j.cortex.2010.09.003

Zumbrunn, S., & Bruning, R. (2013). Improving the writing and knowledge of emergent writers: The effects of self-regulated strategy development. *Reading and Writing: An Interdisciplinary Journal, 26*(1), 91-110. http://dx.doi.org.proxy1.lib.uwo.ca/10.1007/s11145-012-9384-5

Appendix A: Coding Composition—Reference for Raters

Question	0 for no response / no self-reg / double dipping	+1 for any self-reg response	+1 for concrete how or what of self-reg	+1 for any elaborated answer (**or** 2 or more examples or ideas)	+1 for any self-reg response to prompt "a"
1. Ravi--Planning		-Play with his friends [i.e., gave an example of a topic] - Thinking about it first -Like kind of planning stuff out before he makes the move - Think about the things that he's going to write -Think really hard	-You could take a paper and write about your plan you make - He should write the title if there is a title	- He should make a plan by remembering what he did a long time ago -[By thinking" of any of the following]: topic, beginning, middle, end, feeling	
2. Ravi, Goal setting / Criteria.		-If people like it -Your best writing -Doing your best -A good story has stuff about your friends	-Adding good details -Putting a setting - You could say I felt happy at the end of it -A happy ending	- Detail and lots of pictures - Setting, topic, beginning, middle, end and feelings	

3. Jessica: Self-instruction / Strategy / Text structure		-She should put in things like doing something new in Grade 1 -Parts about her friends -That she had fun	-Where the story happened - First it should have the topic -Feeling	- When it happened, where it was and if there was anyone with you - Topic, Beginning, Middle, End, Feeling	
4. Self-statements: open-ended		-Do good writing -Do a strategy	-I'm going to keep on trying -I'm doing a good job -Wow! I did a great job	-I'm not giving up, I know I can do this - Keep on trying, you're doing a good job	
5. Self-monitoring: Checking		-Check it -Have a happy ending	- To see if it makes sense -He could read it all again -For the topic	- If someone's trying to talk to him, he shouldn't answer because he's too busy. -For the beginning, middle, end and feeling	
6. Mark: Coping		-He could stop writing for a minute -Finish it another day	- Ask the teacher for help - Say it's my choice -Say keep on trying	-He could just try to keep going, work hard you never know sometimes it takes me a little longer because I have	

			-He could say take your time	to think about what I'm writing about - Take a break and get back to it later -He could ask a friend to help him or get it done quick	
7. Fatima: Self-assessment / Reviewing		-Check it - Read it over again -Make sure everything's right - Everything that you need on a story	- She should check for what she missed - She should read it to her teacher -Check for the feeling	- She should check for what parts of the story that she missed -For the beginning, middle, end and feelings	
8. Fatima: Self-Reinforcement		- Take a big long break…	- That she did good -You did a great job!	-Say that she did a good job and that she's a good writer -Say, "you used your strategies" -Give herself a rollercoaster cheer!	

*Plus an additional point if response is offered to first, general prompt, rather than specific prompt.

Coding Transcription—Reference for Raters

0 for no response / no self-reg / double dipping	+1 for <u>any</u> self-reg response	+1 for <u>concrete how or what</u> of self-reg	+1 for any <u>elaborated answer</u> (**or** 2 or more examples or ideas)	+1 for any self-reg response <u>to prompt "a"</u>
	-Write faster -Write better	-Doing capitals and periods/make sure there's periods/check for periods -No messy writing -Do finger spaces -If he missed some letters or something -Check for strait writing -She should check if the words are right -Good commas -To make sure it's neat	- Write in the lines, with lots of detail -Sound out the words. Do the beginning sound, and then a line and then the end sound -He could say you should always have periods when you're done your sentence -Periods at each sentence and capital letters at the start of each sentence -To add a period, question mark or exclamation mark. Check for capitals and lower cases	

*Plus an additional point if response is offered to first, general prompt, rather than specific prompt.

Curriculum Vitae

Name:	Ashley Bildfell
Post-secondary Education and Degrees:	Western University London, Ontario, Canada 2017-2021 Ph.D. School and Applied Child Psychology
	Carleton University Ottawa, Ontario, Canada 2014-2016 M.A., Psychology
	Carleton University Ottawa, Ontario, Canada 2010-2014 B.A. Honours, With High Distinction, Psychology
Honours and Awards:	Province of Ontario Graduate Scholarship 2019-2020, 2020-2021
	Joseph Armand Bombardier Canada Graduate Scholarship Recipient funded by the Social Sciences and Humanities Research Council of Canada 2014-2015
	Carleton University Senate Medal for Outstanding Academic Achievement 2014
Related Work Experience	Clinical Psychology Resident London Family Court Clinic 2020-2021
	Advanced Practicum Student Child and Youth Development Clinic 2019-2020
	Psychoeducational Consultant London District Catholic School Board 2019
	Practicum Student London District Catholic School Board 2018-2019

Practicum Student
Maitland Valley and Clinton Family Health Teams
2018

Practicum Student
Mary J. Wright Center for Research and Education
2017-2018

Teaching Assistant
Western University
2019-2020

Research Assistant
Western University
2017-2020

Teaching Assistant
Carleton University
2014-2016

Research Assistant
Carleton University
2014-2016

Publications

Klein, P. D., Bildfell, A., Dombroski, J. D., Giese, C., Sha, K., & Thompson, S. C. (accepted). Self-regulation in early writing strategy instruction. *Reading and writing quarterly.*

Klein, P. D., Haug, K. N., & Bildfell, A. (2018). Best practices in writing to learn. In S. Graham, C. A. MacArthur, & M. Hebert (Eds.), *Best practices in writing instruction* (3rd ed.; pp. 162-184). New York, NY: The Guilford Press.

Sénéchal, M., Whissell, J., & Bildfell, A. (2017). Starting from Home: Home Literacy Practices That Make a Difference. In K. Cain, D. Compton & R. Parrila, (Eds.). *Theories of Reading Development*. Philadelphia, PA: John Benjamins Publishing Company.

Poster Presentations

Malette, M., Bildfell, A., Whissell, J., & Sénéchal, M. (2016, July*). Learning to read in opaque languages: The rate of word and non-word reading acquisition in French and English.* Poster session presented at the meeting of Society for the Scientific Study of Reading, Porto, Portugal.

Bildfell, A., Sénéchal, M. (2015, June). *French immersion education: the role of a child's emergent reading skills and a language's orthography on language transfer effects.* Poster session presented at the 45th annual meeting of the Jean Piaget Society, Toronto, ON.

Hamilton*, A., Sénéchal, M. (2014, April). *The influence of orthography on bilingual automatic reading acquisition.* Poster session presented at the 3rd annual Psychology Undergraduate Research Event, Carleton University, Ottawa, ON. (*Maiden name)

Conference Talks

Bildfell, A., Klein, P. D., Dombroski, J. D., Giese, C., Sha, K., & Thompson, S. C. (2021, August). *Assessing Self-Regulation in Grade 1 Writers: A Reliability and Validity Study.* Single paper to be presented at the annual meeting of the European Association for Research on Learning and Instruction, Gothenburg, Sweden.

Mussar, R., Bildfell, A., Celebi, H., & Sénéchal, M. (2019, July). Shared writing: Learning to mark vowels, consonants, and silent letters. In S. Martin-Chang (Chair), *Tell me something good: Concurrent and ongoing effects of positive literacy environments*. Symposium conducted at the annual meeting of Society for the Scientific Study of Reading, Toronto, Canada.

Klein, P. D., Bildfell, A., Archibald, L. M. D. & Auckland, K. (2019, April). *Early intervention in writing: Recent research on what works.* Presentation at Faculty Research Partners Day. Faculty of Education. London, Ontario, Canada.

Mussar, R., Celebi, H., Bildfell, A., & Sénéchal, M. (2018, May). Young Children's Construction of Vocabulary Knowledge. In, *On interactions that enhance language development: Examining the role of parents, teachers, and technology*. Symposium conducted at the annual meeting of the Jean Piaget Society, Amsterdam, Netherlands.

Bildfell, A., Sénéchal, M. (2016). *Is .82 too good to be true? The role of maternal sensitivity and responsiveness in a child's early literacy development.* Conference talk presented at the first annual Psychology Graduate Student Conference, Carleton University, Ottawa, ON.