Writing to Learn in Science: Effects on Fourth-Grade Students' Understanding of Balance

By

Amy Gillespie

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Special Education

August, 2014

Nashville, Tennessee

Approved:

Donald L. Compton, Ph.D.

Steve Graham, Ed.D.

Lynn S. Fuchs, Ph.D.

Deborah W. Rowe, Ph.D.

Joseph H. Wehby, Ph.D.

To the energetic students and teachers who made this work possible

ACKNOWLEDGEMENTS

I am grateful to my dissertation committee, the individuals who assisted with data collection and fidelity checks for my dissertation study, and all of the professors and staff at Vanderbilt who have assisted me throughout my doctoral program and research endeavors. I would especially like to thank Steve Graham, my initial advisor at Vanderbilt and co-chair of my committee. I have learned so much about research, writing, and life as a professor from your guidance and examples. Additionally, I would like to thank Don Compton, who stepped in as my advisor at Vanderbilt halfway through my doctoral program and served as co-chair of my committee. Not only did you provide the resources I needed to complete my dissertation, you also took time to support me as a person and gave great advice about creating a healthy work-life balance. Lastly, a tremendous thank you to my wonderfully supportive husband, Rob. You were my rock, my comic relief, my personal chef, my running partner, and my unwavering source of inspiration throughout the dissertation journey. Thank you for being there through it all and understanding how to be the kind of partner and friend I needed most.

Additionally, this work was supported with funding from the Institute of Education Sciences, U.S. Department of Education, through the Experimental Education Research Training (ExpERT) program (Grant No. R305B080025). The opinions expressed are those of the author and do not represent the views of the U.S. Department of Education.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure Page

CHAPTER I

INTRODUCTION

Over the last 40 years, there has been an increased emphasis on the value of students learning through writing (Britton, 1970; Gunel, Hand, & Prain, 2007). More recently, the writing to learn movement in education has moved beyond the notion that writing can improve students' reading skills, which has been documented in a considerable number of studies (see Langer & Applebee, 1987; Graham & Hebert, 2010), to examining the impacts of writing on learning across the curriculum (e.g., mathematics, science, and social studies) and out of the English/language arts context (Hand & Prain, 2002).

Specifically, writing to learn in science has received greater attention in the last 10 years, which is a timely topic given recent reforms to writing (Common Core State Standards, 2010) and science standards (National Research Council, 2012). These standards have led to new expectations for the types of writing elementary students should be learning to use and new definitions of what it means to be scientifically literate. Beginning in kindergarten, the Common Core State Standards (CCSS, 2010) require students to compose informational and explanatory texts. Students are expected to do research, analyze and interpret information, and convey their findings appropriately in writing, with evidence to support their claims. Additionally, with the CCSS (2010), there is a strong emphasis on literacy in science and social studies beginning in grade 6. Furthermore, the National Research Council (NRC, 2012) has highlighted the importance of students conducting scientific investigations and explaining their findings in writing beginning in the elementary grades. Science literacy, as defined by the NRC (2012), entails not only being able to read scientific texts, but also being able to record observations, to compose written arguments to support and defend results, and to use writing to convey findings to others.

The use of writing as a tool to support learning is also supported theoretically. Emig (1977) posited that the process and product involved in writing correspond to the higher level cognitive functions involved in learning; stated simply, to be able to write about one's own learning, an individual must be an active and engaged learner. When writing, students make active, personal, and conceptual associations between ideas. Furthermore, writing forces students to analyze and synthesize the new information they have learned. From the written product created, students get feedback and can review and refine their ideas as needed, thus furthering learning (Emig, 1977).

Other researchers have characterized writing as a problem-solving activity during which writers orchestrate a number of mental processes to meet certain goals (Bereiter & Scardamalia, 1987; Hayes & Flower, 1980). To produce a composition, writers draw upon knowledge of the content and process of writing, as well as input from the environment and from the text they are producing (Hayes & Flower, 1980). As writers balance the writing process, goals for writing, and the development of text, they are continually reorganizing information, making connections between ideas, and restructuring prior knowledge of the writing topic (Bangert-Drowns, Hurley, & Wilkerson, 2004; Bereiter & Scardamalia, 1987; Flower & Hayes, 1980; Keys, 1999). In this way, writing contributes to students' learning by helping them reflect and think critically about new information, which permits them to construct new and deeper understandings of subject matter (Bangert-Drowns et al., 2004; Klein, 2000). Additionally, when writing about subject matter, students increase their time on-task and increase the number of times they rehearse new information, both of which have been shown to lead to greater learning. Writing also serves as a self-monitoring strategy, allowing students to monitor their own comprehension of a topic or task, evaluate their own misconceptions, and change their ideas (Bangert-Drowns et al., 2004).

Based on the work of Britton (1970) and other theorists, Klein (1999) described four specific ways writing may contribute to learning. These hypotheses apply to a variety of learners, as they assume differing levels of expertise in writing.

Klein's (1999) first hypothesis about the influence of writing on learning called "shaping at the point of utterance" (p. 211) applies to novice writers who do not plan, review, or revise their work. With shaping at the point of utterance, students produce new knowledge and learning simply by writing down text. As students spontaneously invent content while writing (e.g., a freewriting or expressive writing task), they use language to develop new meanings and understandings. The new content produced by students represents new knowledge. Although this hypothesis does not involve planning, it is assumed that as students write, they may change their writing plan and also change their understanding of the writing topic (Galbraith, 1999).

Klein's (1999) second hypothesis, "forward search" (p. 221), states that because students record ideas in text, they are able to revisit and review them. Thus, as Emig (1977) noted, the actual text produced serves as a form of feedback for generating knowledge. When students revisit their texts, they can build on their ideas, make new connections, make new inferences, and deepen their understanding of new subject matter. This hypothesis applies to students who have greater writing expertise and understand how to make substantive revisions to their writing (Klein, 2000).

Klein's (1999) third and fourth hypotheses both apply to more skilled and experienced writers who understand a variety of writing genres and develop elaborated plans and writing goals. These hypotheses are less likely to be applicable to the younger and more novice writers involved in the current study.

With Klein's third hypothesis called the, "genre hypothesis" (p. 230), he indicated that the genre in which students compose affects their learning. Specifically, expository writing tasks require students to process information on a deeper level and thus result in greater learning than writing in other genres. Additionally, as other researchers have noted, writing tasks prompt students to recall genre schemata, which causes them to search for relevant knowledge and make new connections between ideas (Bereiter & Scardamalia, 1987).

Klein's fourth hypothesis, "backward search" (p. 242), requires the most writing expertise. With this approach, students create new knowledge by setting rhetorical goals for their text. When students create written content to address their rhetorical goals, they retrieve and organize their knowledge, leading to more elaborated understandings and ideas.

Other scholars (Keys, 1999; Holliday, Yore, & Alvermann, 1994) have noted that using writing to communicate ideas to others contributes to learning. They argued that as students write for an audience they are compelled to present their ideas in a logical manner, make explicit connections between concepts, and organize their thoughts coherently. By preparing writing for other readers, students are likely to think more deeply about subject matter so as to serve as an "expert voice" (Keys, 1999, p. 122) on a topic.

In addition, Wallace, Hand, and Prain (2004) asserted that one of the most salient features of writing that may enhance learning is its potential to produce metacognitive awareness. That is, when students write they discover what they know and what they still do not understand. This may prompt students to use other cognitive strategies, such as reviewing what they have learned, searching for new information, or evaluating prior beliefs and ideas. Thus, Wallace and colleagues contended that writing produces metacognition, which in turn contributes to further conceptual growth and understanding.

Studies of Writing to Learn

The initial theories and work of Britton (1970) and Emig (1977) were instrumental in the Writing Across the Curriculum (WAC) movement, which began in the U.S. in the late 1970s and 1980s (Durst & Newell, 1989; Keys, 1999). Most of the studies of the effects of writing on learning during this time were conducted at the college level; many were action research projects undertaken by professors in their own classrooms. Durst and Newell (1989) noted that most of these initial studies examined the impact of writing (e.g., notetaking, short answer questions) on reading comprehension and understanding of subject matter (e.g., Langer & Applebee, 1987). In Ackerman's (1993) review of 35 studies of writing to learn published between 1979 and 1989, he concluded the results of writing to learn intervention studies were inconclusive and the causal relationship between writing and learning was not supported by the available research.

More than a decade later, Bangert-Drowns, Hurley, and Wilkinson (2004) conducted a meta-analytic review of 48 writing to learn studies published between 1926 and 1998. Participants in the studies reviewed ranged from elementary (n = 11) to college (n = 21) students, with six studies conducted with middle school students and 10 conducted with high school students. All but two of the studies (Becker, 1996; Sharp, 1987) examined students writing in expository or informational genres and 36 of the 48 studies reported positive effects for writing to learn interventions. Overall, Bangert-Drowns et al. reported writing to learn interventions had small positive and significant effects (ES = 0.17) on students' academic achievement in a variety of subjects (e.g., math, science, social studies). Later, in their meta-analysis of writing interventions for adolescents, Graham and Perin (2007) also found a small positive and significant effect (*ES* = 0.23) from 26 studies of writing to learn. These studies were conducted with students in grades 4 through 12 in math, science, and social studies. During this time, there was only one study of a writing to learn intervention in science with elementary students with the information needed to calculate an effect size. This study (Mason & Boscolo, 2000), discussed further below, reported large positive effects (*ES* = 0.85) for grade 4 students who completed writing to learn activities during a science unit.

Studies of Writing to Learn in Science

Prior to Bangert Drowns et al.'s meta-analysis, Rivard (1994) published a descriptive review of the literature on writing to learn in science. Similar to reviews of the writing to learn literature, Rivard found the literature he reviewed largely focused on college students who used writing techniques such as notetaking, answering questions, or summarizing during lectures or course readings to help learn new science material (Horton, Frank, & Walton, 1985; Kirkpatrick & Pittendrigh, 1984; Thall & Bays, 1989; VanOrden, 1987). Rivard also noted several studies in which students used journals or expressive writing to respond to lectures and readings in science classes (Malachowski, 1988; Reynolds & Pickett; Willey, 1988). Although more positive than Ackerman's (1993) appraisal of writing to learn, Rivard still suggested more studies were needed to understand the impact writing has on students' conceptual understandings of science material. From his review, Rivard concluded that future researchers should: (a) conduct more studies of writing to learn in science, specifically at the elementary and secondary levels; (b) document and explain contextual factors that may impact the effects of writing to learn interventions in science; and (c) examine moderating variables (e.g., ethnicity, academic ability, gender) that may interact with the use of writing to learn interventions in science.

Despite Rivard's (1994) recommendations, only seven of the studies examined by Bangert-Drowns et al. (2004) in their meta-analysis published ten years later were conducted in science. Of those studies, two were conducted with middle school participants (Ayers, 1993; Rivard, 1996) and two were conducted with high school participants (Nieswandt, 1997; Willey, 1988); the rest of the studies were conducted with college students. None of the studies were published in peer-reviewed journals (i.e., 3 dissertations, 1 conference paper). All of these studies incorporated writing to learn within an entire science unit and involved students composing written responses to readings and lectures (Ayers, 1993; Rivard, 1996; Willey, 1988) or composing written laboratory reports (Nieswandt, 1997), with multiple choice assessments as the primary measure of students' gains in science knowledge and understanding. Both Ayers (1993) and Rivard (1996) reported that students in their control groups outperformed students in their writing to learn treatment groups on posttest assessments. Nieswandt (1997) and Willey (1988) reported positive effects for the writing to learn treatments they investigated. Thus, similar to what other researchers concluded years earlier (Ackerman, 1993; Rivard, 1994), results of writing to learn interventions in science remained inconclusive. Three years later, in Graham and Perin's (2007) review, only one additional study of writing to learn in science was available (Hand, Hohenshell, & Prain, 2004).

More recent research. A notable difference in much of the recent research on writing to learn in science is a shift from writing to learn about texts and lectures to writing to learn while performing scientific investigations (e.g., building models, conducting experiments, reasoning about science problems) and writing for authentic purposes. In these studies, as opposed to the more traditional ways of acquiring science knowledge through vocabulary lessons, textbook readings, and prescribed laboratory experiments using the scientific method, students participate in science practices and use writing to learn from hands-on experiences in inquiry-based formats (Hohenshell & Hand, 2006). Further, students use writing to express their understanding of science subject matter to authentic audiences. Differences in the types of writing to learn activities under investigation likely reflect the recent changes to science standards which emphasized the importance of students being active participants in science practices. By conducting scientific investigations and reporting findings to real audiences, students emulate the practices of true scientists (NRC, 2012).

Studies of writing to learn in science at the secondary level. The research team of Wallace, Hand, Prain, and colleagues conducted a majority of the most recent studies of writing to learn in science with secondary students. It should be noted that this research group and others have undertaken qualitative examinations of writing to learn in science as well. Because the current study was conducted using quantitative methods, only studies of writing to learn in science in which researchers used quantitative measurement of student outcomes are summarized here.

In their book on the topic, *Writing and Learning in the Science Classroom*, Wallace et al. (2004) reported a series of three investigations examining the effects of writing to learn interventions in science with high school students who wrote for authentic purposes and authentic audiences. The first study involved grade 10 students who completed a science unit on genetic engineering (Hand et al., 2004). Four groups of students were examined: (1) students who participated in small group and class discussions to plan before writing an explanation of the science topic (i.e., genetic engineering) to an intended audience of seventh-grade students, (2) students who did not plan before writing an explanation of the science topic to an intended audience of seventh-grade students, (3) students who participated in small group and class discussions to plan before writing an explanation of the science topic to an intended audience of seventh-grade students and to plan before writing a newspaper article on the science topic, and (4) students who did not plan before writing an explanation of the science topic to an intended audience of seventh-grade students and did not plan before writing a newspaper article on the science topic. Students who planned before writing the explanation outperformed students who did not plan before writing the explanation on a posttest exam consisting of 12 recall questions and three conceptual questions about genetic engineering. Students who wrote twice (i.e., explanation and newspaper article) outperformed students who wrote once (i.e., explanation) on three conceptual questions about genetic engineering.

In their second study, Wallace and colleagues (2004) examined grade 11 students who completed a 3-week unit on stoichiometry. At the end of the science unit, students in the treatment group wrote a business letter to seventh-grade students to explain the concepts of stoichiometry. Next, a seventh-grader read and provided feedback on whether he or she could understand the explanation of the science concept. Then, treatment group students wrote a final draft of the letter based on the feedback they received. Students in a comparison condition wrote a chapter summary and answered chapter questions at the end of the science unit. Students in the writing to learn treatment condition outperformed students in the comparison condition on a posttest consisting of short answer recall and conceptual questions about stoichiometry.

In study three (Wallace et al., 2004), grade 10 students participated in a 6-week molecular biology unit. After the unit, students in the treatment group wrote a newspaper article describing the science topic and got feedback on their writing from a local newspaper editor. Students in a comparison condition wrote a chapter summary and answered chapter questions from their textbooks. Students in the writing to learn treatment condition outperformed students in the comparison condition on three conceptual questions about molecular biology.

Hohenshell and Hand (2006) also examined a writing to learn intervention implemented with high school students. In this study, students in grades 9 and 10 completed six laboratory activities during a 7-week cell biology unit. For each laboratory activity, students in the treatment condition completed the Science Writing Heuristic (SWH) which consisted of writing prompts (i.e., What questions do I have? What did I do? What did I see? What inferences can I make? How do I know? How do my ideas compare with others? How have my ideas changed?) designed to scaffold the inquiry process and facilitate students' learning from the laboratory activities. Students in a comparison condition completed traditional laboratory reports (i.e., hypothesis, materials and procedures, results, conclusions). After the six laboratories activities, students in both conditions completed summary reports of the science unit. Students who completed the SWH scored higher on four conceptual questions about cell biology than students in the comparison condition who completed traditional laboratory reports.

Hand, Wallace, and Yang (2004) also used the SWH with grade 7 students who completed an 8-week inquiry-based science unit on cells. The researchers also varied the intended audience (i.e., teacher or peers) for whom students composed. Students in the treatment condition completed the SWH (described above) during all laboratory activities and wrote a research paper to summarize their results at the end of the unit. Students in a comparison condition also completed the SWH during all laboratory activities but they wrote a summary, called a "textbook explanation" (Hand, Wallace, & Yang, 2004, p. 136), intended for their peers to read. Students in a control condition completed traditional laboratory reports and wrote a research paper summarizing their results at the end of the unit. At posttest, students in both SWH conditions (i.e., treatment and comparison conditions), regardless of the audience for whom they wrote, outperformed students in the control conditions on multiple choice questions about cells and on one of three conceptual questions about cells. Students who used the SWH and wrote a textbook explanation of the science topic for their peers (i.e., students in the comparison condition) performed better than students in the treatment condition and students in the control condition on the remaining two conceptual questions about cells.

Similarly, Gunel, Hand, and McDermott (2009) studied students in grades 9 and 10 who were taking a biology course. They used the same writing to learn intervention for both treatment and comparison conditions, while varying the audience for whom students wrote. After finishing a science unit on the human circulatory and respiratory systems, all students wrote an explanation of the coordination of the circulatory and respiratory systems. Depending on the group to which they were assigned, students wrote their explanation for: (a) younger students at a nearby elementary school, (b) their parents, (c) their same-aged peers, or (d) their teacher. Students who wrote explanations for younger students outperformed students who wrote for the teacher on one of three conceptual questions about the circulatory and respiratory systems.

Rivard and Straw (2000) also contributed to the research on writing to learn in science. Grade 8 students were divided into four conditions for five 50-min problem solving sessions during an ecology unit. The unit involved a range of activities, including teacher-led and peer-led discussions, library research, laboratory activities, readings, worksheets, and simulations. During the problem solving sessions, students were given a real-world ecology problem and asked to explain a solution. Students in the talk-only condition discussed the assigned problems with peers during each of the problem solving sessions. Students in the writing-only condition composed written responses to the assigned problems during each of the problem solving sessions. Students in the talk and writing condition discussed assigned problems with peers and then composed written responses to the assigned problems. Students in the control condition completed fill-in-the-blank, true-false, matching, and definition exercises. There were no significant differences between conditions on a multiple choice posttest, but boys in all conditions outperformed girls on this measure. On a delayed multiple choice posttest, given 6 weeks after the ecology unit was completed, students in the talk and writing condition outperformed students in the writing-only condition and students in the control condition; boys outperformed girls on this measure as well.

Rivard (2004) conducted a similar study with grade 8 students who completed an ecology unit. The chief difference between the 2004 study and the previous work by Rivard and Straw (2000) was that Rivard examined students' achievements levels (i.e., final grades in their seventh-grade science course) as a moderator of treatment effectiveness. The treatment, comparison, and control conditions were the same as in the Rivard and Straw study. On a multiple choice posttest, significant differences were found between conditions, but only for students considered low achieving and students considered high achieving based on their previous year's letter grade in science class. Students who were low achieving in science and in the talk-only condition outperformed students who were low achieving and in the control condition at posttest. Students who were high achieving in science and in the writing-only condition outperformed students who were high achieving and in the talk-only condition at posttest. On a delayed multiple choice posttest, given 6 weeks after the ecology unit, students who were low achieving and in the talk-only condition or in the talk and writing condition both outperformed students who were low achieving in the writing-only condition.

Based on the results of recent research, it appears writing to learn interventions in science are effective for improving the recall of science information and conceptual understanding of a variety of science topics for secondary students. Across the relevant studies, students who completed writing to learn tasks during and after they learned new science concepts consistently performed better on recall and conceptual questions than students who completed more traditional science activities (e.g., writing chapter summaries, writing research papers, completing short answer questions). Effective writing to learn interventions included writing for authentic audiences (Gunel et al., 2009; Hand et al., 2004; Wallace et al., 2004) and answering questions designed to scaffold the inquiry process (Hand et al., 2004; Hohenshell & Hand, 2006). Results from one study (Rivard, 2004) revealed writing to learn interventions had different effects for students who were low achieving or high achieving in science. Results from another study revealed writing to learn interventions had greater effects for boys than for girls (Rivard & Straw, 2000). On delayed measures of science knowledge, the effects of writing to learn interventions were less clear (Rivard, 2004; Rivard & Straw, 2000).

Studies of writing to learn in science at the elementary level. Although fewer in number than studies conducted at the secondary level, writing to learn interventions in science have been conducted with elementary-aged students as well. Compared to studies of writing to learn in secondary science, most of these studies involved less rigorous methods (i.e., no randomization of students to conditions, no control or comparison group, less rigorous statistical analyses of student data) and some involved more traditional forms of writing to learn (i.e., notetaking, summarizing) rather than inquiry-based writing and science activities with authentic writing tasks.

Herrenkohl, Palincsar, DeWater, and Kawasaki (1999) conducted a study to examine two groups of students (i.e., one grade 3 and 4 class, one grade 5 class) who completed a science unit on density. During the 10-week unit, students participated in hands-on science investigations, whole group discussions, and individual writing and reflective opportunities. For the pretest and posttest measure of science knowledge, students responded in writing to the following: Write your best explanation and model of why the plastic piece sank in the fresh water but floated in the salt water. Compared to only 4% of students who used density explanations at pretest, 63% of students in the grade 3 and 4 class used density explanations at posttest. For the grade 5 class, no students used density explanations at pretest, while 48% used density explanations at posttest.

Tucknott and Yore (1999) studied Canadian students in grade 4 who completed a 6-week unit on simple machines. During the science unit, teachers posed questions, conducted demonstrations, and facilitated class discussions. Students completed a variety of writing tasks during the unit, including answering short answer questions, notetaking, summarizing, and writing explanations. Students showed significant gains from pretest to posttest on a 40-item assessment (i.e., 32 multiple choice/short answer recall questions, 8 higher level reasoning questions requiring a written explanation).

In Italy, Mason and Boscolo (2000) examined a writing to learn intervention with grade 4 students who participated in a 3-week science unit on photosynthesis. Students in the treatment condition took notes and wrote personal responses to lectures and information learned throughout the unit. Students in the control condition did not use writing to learn information during the science unit. On a 19-item short answer posttest, students in the treatment condition outperformed students in the control condition.

Of particular relevance to the current study, Klein (2000) examined Canadian elementary students who wrote after completing experiments involving density and balance. Students in grades 4, 6, and 8 (i.e., Canadian elementary grades but elementary *and* middle grades in the U.S.) completed one of two experiments: (a) a density experiment, where they tested whether 10 different objects would sink or float in a bucket of water; or (b) a balance experiment, where they tested whether 15 different weight configurations on a beam would balance, tilt left, or tilt right. With both experiments, students: (1) verbally answered a question before beginning the experiment (i.e., What makes objects float or sink? or What makes the beam tilt in one direction or the other, or balance?); (2) completed the experiment and recorded their findings for each trial; (3) verbally answered the same question they were asked before beginning the experiment; (4) responded in writing to the same question they were asked before beginning the experiment; and (5) verbally answered the same question again. Students who stated one level of density or balance explanation after the experiment (step 3 above) and stated a higher level of explanation after writing (step 5 above) were considered to have made explanatory gains during writing. Klein found students made significant explanatory gains after writing about density. He reported that grade level was significantly correlated with level of explanation of density but not significantly correlated with gains made from writing. For the balance experiment, students did not make significant explanatory gains. Level of balance explanation and gains made from writing were not significantly correlated with grade level for balance.

Although limited in number and in rigor, most studies of elementary-aged students reported improvements in students' recall and explanations of a variety of science topics when they used writing to learn activities while learning new science concepts. Effective writing to learn interventions were not as well defined in studies of elementary students as they were in studies of secondary students, and most involved multiple types of writing (Herrenkohl et al., 1999; Mason & Boscolo, 2000; Tucknott & Yore, 1999). Results from the only study comparing a writing to learn intervention to a no writing control group (Mason & Boscolo, 2000) showed students who took notes and wrote personal responses while learning science information performed better on short answer questions than students who did not write while learning the same science information (*ES* = 0.85). Maintenance measures were not reported in studies of writing to learn interventions at the elementary level.

The Present Investigation

Currently, there are a limited number of studies available on writing to learn in science, and few have included elementary-aged participants. Additionally, researchers have not investigated if elementary students' writing skills have an impact on their ability to benefit from writing to learn activities.

In this study, I examined the impact of a writing to learn intervention in elementary science. Grade 4 students completed 30 trials using an equal arm balance beam to explore the concept of balance. Students in the treatment condition wrote during and after the balance trials about what they were learning. To control for time spent writing during the intervention, students in the comparison condition wrote about their favorite parts of the experiment and their favorite parts of science class. Students in the no-treatment control condition participated in business as usual science instruction with their classroom teachers. This investigation extended past research in several important ways and addressed limitations of prior work examining writing to learn in science.

Participants. First, this study was conducted with elementary students in grade 4. Most of the previous work on writing to learn in science has involved secondary and post-secondary students. I purposefully selected fourth grade as the target population for this study because writing becomes increasingly important as students move into the upper elementary grades. Beginning in the upper elementary grades, written products become the primary means by which teachers measure students' knowledge and understanding of subject matter (Graham, 2006). Additionally, the CCSS (2010) for grade 4 English/Language Arts require students to conduct short research projects and produce informative and/or expository texts to convey information clearly, as they did in this study.

Intervention. Second, the writing to learn intervention in this investigation involved balance trials similar to the balance experiment students completed in Klein (2000). Unlike science topics such as biology, balance has been studied in only one investigation of writing to learn in science thus far (Klein, 2000). Additionally, I chose balance as the topic of investigation because: (a) balance is a concept emphasized in elementary science in the state where the study was conducted; (b) the balance trials could be completed in a relatively short session, while still requiring students to think carefully as they analyzed their data (Klein, 2000); and (c) the progression of students' understanding of the concept of balance was established as early as 1958 (Inhelder & Piaget) and has been validated in subsequent research (Hardiman, Pollatsek, & Well, 1986; Klein, 2000; Siegler, 1985).

Although Klein (2000) reported no significant explanatory gains for students completing a similar balance task, this study employed a more rigorous design than Klein's, accounted for students' pre-intervention balance knowledge, and involved students writing both during and after the balance trials. I hypothesized that each of these factors would contribute to more pronounced effects of the writing to learn intervention.

First, the current study involved a more rigorous research design than Klein's (2000). In this study, students were randomly assigned to a writing to learn treatment condition, a comparison condition, or a no-treatment control condition. Additionally, as an extension of Klein's work and all of the other writing to learn interventions in science reviewed here, this study included organized training for research assistants who delivered the study assessments and interventions, videotaping of the treatment and comparison conditions, and calculation of procedural fidelity.

Students' balance knowledge prior to the intervention was also taken into consideration when assigning students to conditions and when calculating treatment effects for this investigation. Klein (2000) did not account for students' initial balance knowledge in his study. One possible explanation of Klein's findings is that many participants already had a solid understanding of balance prior to the writing to learn intervention; thus, they did not make significant gains when explaining the concept of balance. Because initial performance on the balance knowledge pretest was likely a powerful predictor of students' performance on the posttest, I tried to control for initial balance knowledge in two ways. First, I stratified students based on balance knowledge pretest scores before randomly assigning them to conditions; in this way, there was a relatively equal distribution of balance knowledge within each condition. Second, I accounted for pretest balance knowledge in my data analysis of the balance knowledge posttest using a repeated measures ANOVA with time of testing as the within subjects factor.

Furthermore, although students only wrote after completing the balance trials in Klein's (2000) study, students wrote during and after the writing to learn intervention in this study. Treatment effects may be more pronounced using this design, as other researchers (Hand et al., 2004; Hohenshell & Hand, 2006) have reported that students who answered questions designed to facilitate the inquiry process during writing to learn interventions demonstrated greater learning than students in comparison groups.

Student differences. Third, because students' writing skills had the potential to facilitate or hinder their performance on the writing to learn task, I attempted to control for the impact of initial writing skills in two ways. First, I stratified students based on their Test of Written Language-Fourth Edition (TOWL-4; 2009) Story Composition subtest scores prior to randomly assigning them to conditions; thus, there was a relatively equal distribution of writing skills within each condition. Second, I considered writing skills as a potential covariate in my data analysis. As others have argued, a possible reason for the lack of consistent findings about the effects of writing to learn interventions may be that researchers have focused on main effects of writing to learn interventions and ignored students' individual differences (Kieft, Rijlaarsdam, & van den Bergh, 2008; Rivard, 1994). Although researchers have examined gender (Rivard & Straw, 2000) and students' science performance (Rivard, 2004) as potential moderators of the effects of writing to learn interventions in science, students' writing skills have not been taken into account to help explain variations in writing to learn treatment effects. Research has shown that students with weaker writing skills often focus on forming letters properly and spelling words correctly, leaving little room left in working memory to devote to the content, meaning, and coherence of their compositions (Baker, Gersten, & Graham, 2003). Thus, students who were preoccupied with lower level writing skills (e.g., transcription) may have been less likely to reflect on what they were learning about balance and less likely to make connections between new ideas, which may have impeded their science learning during the intervention. As a result, I examined if students' writing performance on the TOWL-4 was related to their performance on balance knowledge measures. When this occurred, I used TOWL-4 scores as a covariate in my analyses.

Pilot study. In 2013, I conducted a pilot study of the writing to learn intervention with 15 grade 4 students. During this pilot study, I served as the primary investigator, meeting with Steve Graham regularly to refine study procedures based on students' responses to the intervention. I videotaped all students completing the writing to learn intervention (i.e., balance trials and writing tasks). During intervention sessions, I recorded anecdotal notes about the balance trials, writing prompts, and researcher prompting. I also watched video of each session to take additional notes about changes that may better the writing to learn intervention. Because the focus of the pilot study was to refine the writing to learn intervention, only variations in the treatment condition were tested. Final decisions about the balance trials, writing tasks, researcher prompts, study procedures, and measures were all based on data from the pilot study.

During the pilot, I tested two variations in the number of balance trials students performed during the writing to learn intervention. Seven students completed 30 balance trials and eight students completed 40 trials. Based on review of video and student performance during the intervention, I decided to use 30 trials in the present study. Forty trials appeared to be overwhelming for some students; many appeared fatigued or less willing to write after completing 40 trials. Furthermore, students' written descriptions and explanation of the concept of balance were as well-developed when completing 30 trials as when completing 40 trials.

I also tested two variations in the amount of researcher prompting during the balance trials. Half of the students had minimal verbal prompting during the balance trials, while the other half had more directed verbal prompts throughout the trials (e.g., What happened to the balance?, Why do you think it balanced/tilted right/tilted left?, Tell me what you think about this.). Although it appeared students better explained their understandings of balance when verbally prompted, this presented a confounding variable and posed threats to implementation fidelity when considering conducting the study on a larger scale with more individuals delivering the intervention. Therefore, for the present study, I replaced the researcher verbal prompts during the trials with four written prompts for students to answer during the balance trials.

Lastly, during the pilot study, I calculated test-retest reliability for the balance knowledge pre- and posttest measure. I administered the assessment to all 15 students, then administered it again three weeks later. Test-retest reliability for the balance knowledge assessment was 0.29 for level 1 questions ($p = 0.29$), 0.85 for level 2 questions ($p < 0.001$), 0.81 for level 3 questions ($p < 0.001$), 0.49 for level 4 questions ($p = 0.06$), and 0.57 for the entire test ($p = 0.03$). Test-retest reliability for level 1 and level 4 questions was lower and correlations at these two levels were not statistically significant. However, with a small sample, outcomes were easily skewed by the variable performance of a small number of students. Exact agreement for questions at level 1 was 0.80 and exact agreement for questions at level 4 was 0.53. Therefore, I decided to keep level 1 and level 4 questions in the balance knowledge pre- and posttest for the current study.

Research questions and hypotheses. Specifically, in the present investigation, I aimed to answer: (1) Do students who write about balance during and after the balance trials perform better on a balance knowledge posttest than students in a comparison condition who do not write about balance during and after the trials and better than students in a no-treatment control condition? and (2) Do students who write about balance during and after the balance trials perform better on a posttest extended writing prompt (i.e., total words written and level of balance understanding) than students in a comparison condition who do not write about balance during and after the trials and better than students in a no-treatment control condition?

I expected that students in the writing to learn treatment condition who wrote about balance during and after the intervention would perform better on the balance knowledge posttest than students in the comparison condition who completed the balance task but wrote about topics other than balance and students in the no-treatment control condition, as the type of writing students in the treatment condition were asked to undertake required them to internalize, integrate, and think further about the new information they were learning about balance. Because a majority of all students (84%) correctly answered the five pretest questions involving the most basic balance principle (i.e., when weights are placed the same distance from the fulcrum on each arm of the balance, the amount of weight on each arm causes the beam to balance or tilt right/left), I did not anticipate significant differences between conditions at the lowest level of balance understanding (i.e., level 1) on the posttest questions. However, I expected students in the treatment condition would perform better than the students in the other two conditions on balance knowledge posttest questions at levels 2, 3, and 4. Additionally, I anticipated students in the comparison condition may outperform students in the control condition on balance knowledge posttest questions at levels 2, 3, and 4, as they may have acquired some new information about balance simply by completing the balance trials.

I also predicted that, on average, the treatment condition would compose longer responses to the posttest extended writing prompt than the comparison or control conditions, as students in the treatment would have learned more about balance and would have more information to include in their written responses. Additionally, I anticipated the writing to learn treatment condition would demonstrate higher levels of balance understanding on responses to the posttest extended writing prompt than the other two conditions, as the writing they completed during intervention would help them reflect on and better understand balance concepts. Although

I anticipated the comparison condition could potentially perform better than the control condition on the balance knowledge posttest, I did not expect differences between comparison and control conditions on the posttest extended writing prompt. I reasoned that even if students in the comparison condition learned some new information about balance via the balance trials, the amount they learned would be too inconsequential to substantially alter their posttest written responses.

CHAPTER II

METHOD

Consent and Assent

Classroom teachers sent home study consent forms to the parents/guardians of all grade 4 students. Parents/guardians were asked to sign and return one copy of the consent form to the classroom teacher and to keep the additional copy of the consent form for their records. The consent form described the study, its purposes and procedures, and the hypothesized impact of the writing to learn intervention. Parents/guardians were asked for permission: (a) for their son/daughter to participate in the study, (b) to videotape their son/daughter while he or she completed the balance trials and writing tasks, and (c) to share the video clips from the treatment or comparison sessions with other professionals (e.g., professional conferences, teacher professional development).

Grade 4 teachers collected signed consent forms in an envelope provided by the primary investigator. One week after the initial consent forms were sent home, classroom teachers sent a second set of forms home with any student who had not returned a signed consent form. No further attempts were made if a parent/guardian did not respond after the second set of consent forms was sent home. Once students received parent/guardian consent to participate in the study, they were assented.

Students who received parent/guardian consent to participate in the study met with the primary investigator or a trained research assistant (RA) who read the student assent form aloud to them. The student assent letter clarified that students who chose to participate in the study would take balance and writing assessments, complete balance trials, and complete writing tasks. Students were asked for their permission to participate in the study and to be videotaped during treatment and comparison sessions (if they were assigned to one of those conditions).

Participants

Seventy-one students received consent and granted assent to participate in the study. Two students were dropped after the TOWL-4 pretest assessments were scored (i.e., one student was eliminated because he did not produce two T-units on this assessment; another student was eliminated because she produced a story that was unrelated to the picture prompt and thus it could not be scored), making the total sample 69 grade 4 students. Fifty-eight percent of the sample was female; 42% was male. The average age was nine years (*SD* = 0.34 years). Twenty percent of the sample received services as English Language Learners (ELL). Two of the participating students received special education and/or speech/language services. A majority of the sample, 86%, received free or reduced lunch.

Setting. Students came from seven grade 4 classrooms in an elementary school in the Southeastern United States. The school served pre-K through grade 4, with an enrollment of 936 students. Of these students, 33% qualified for services as English Language Learners, with 23 languages represented. Forty-two percent of the overall school population was considered Limited English Proficient. Six percent of the students at the school received special education services and 89% qualified for free or reduced lunch.

Each fourth-grade teacher completed a survey about the science instruction she delivered during the timeframe this study was implemented. All teachers indicated they taught an animal studies unit focused on vertebrates and invertebrates, animal habitats and ecosystems during the three weeks this study was carried out.

Measures

All RAs who administered any of the assessments associated with this study were trained prior to administration. The primary investigator reviewed procedures for administering each assessment, demonstrated administration of each assessment, and asked RAs to practice delivering the directions and test items for each assessment. Each RA was required to train to 100% accuracy for administering each assessment before he or she was permitted to administer assessments to participating students.

Pre-intervention. Prior to the implementation of the intervention, the primary investigator administered a writing skills assessment and balance knowledge assessment to all students with consent and assent. Both assessments were group-administered in students' homeroom classes in a single session. To document that all pre-intervention assessments were delivered in the same manner, the primary investigator completed a procedural fidelity checklist as she delivered each assessment.

Writing skills. Students' pre-intervention writing skills were assessed using the TOWL-4 Story Composition subtest (Form B). For the Story Composition subtest, students wrote a story in response to a picture prompt. Students were given 5 min to plan their stories on scratch paper before composing their stories within a 15 min time limit. Test-retest reliability for the Story Composition subtest is 0.80 for grades 4 through 6. Internal consistency reliability is .68 for grade 4. Because the intervention and comparison conditions required considerable writing, an a priori decision was made to eliminate any students who could not write two T-units (i.e., a minimal terminable unit, defined as a simple clause with all of its related subordinate clauses; Hunt, 1965) on the Story Composition subtest; one student was eliminated for this reason.

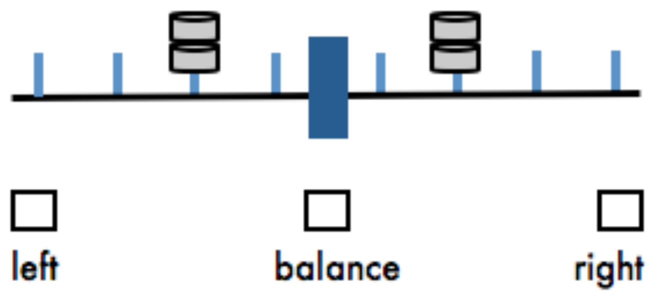
Another student was eliminated because she produced a story that was unrelated to the picture prompt and thus it could not be scored.

Each writing sample produced for the Story Composition subtest was scored for 11 components (i.e., beginning, inclusion of an event before or after the picture prompt, sequence, plot, character emotions, action, ending, writing style, overall impression of writing, inclusion of specific vocabulary words, and overall vocabulary) by the primary investigator. A trained research assistant scored 35% of students' writing samples. Inter-rater reliability, calculated as the correlation between the raters' scores, was 0.86.

Balance knowledge pretest. During the same assessment session, students were also group-administered a 20-item balance knowledge pretest (see Appendix A). In the pilot study, 15 students completed the balance knowledge pretest; students took the test twice, three weeks apart. Test-retest reliability for the balance knowledge assessment was 0.29 for level 1 questions ($p = 0.29$; exact agreement = 0.80), 0.85 for level 2 questions ($p < 0.001$), 0.81 for level 3 questions ($p < 0.001$), 0.49 for level 4 questions ($p = 0.06$; exact agreement = 0.53), and 0.57 for the entire test ($p = 0.03$).

For this assessment, the primary investigator began by reading aloud the directions, which included three diagrams to show what the balance looked like when it balanced, when it tilted left, and when it tilted right. With each diagram, a corresponding box was checked below it, just as students were expected to do on the pretest problems. The primary investigator then asked students to complete sample problem 1. The sample consisted of a diagram of an equal arm balance with two weights, represented by cylinders on the diagram, on the second peg to the left of the fulcrum and two weights on the second peg to the right of the fulcrum (see Figure 1 and Appendix A). After giving appropriate wait time, the primary investigator asked students to raise their hands if they would like to share their answer for sample 1. In all testing sessions, the student who was called on gave the correct answer for sample 1 (i.e., balance) and the primary investigator reminded the rest of the class to make sure they had marked the correct answer. The same process was repeated for sample 2.

Figure 1. Sample Problem 1 From the Balance Knowledge Pretest



Then the primary investigator told students to flip the page and to begin the pretest. Each item on the pretest consisted of a diagram of an equal arm balance with weights placed in a specific configuration, just as students practiced with sample 1 and 2. For each item, students checked a box to indicate if they thought the balance would tilt left, balance, or tilt right with the weights configured as in the diagram. If a student asked for help during any part of the pretest, the primary investigator responded with, "I can't help you with your answers. Please do your best."

The primary investigator scored the balance knowledge pretest as the number of problems correct out of 20 and a trained RA scored 35% of the pretests for reliability purposes. There was 100% reliability of scoring at each level of questioning (i.e., level 1, level 2, level 3, level 4) on the balance knowledge pretest.

Post-intervention. After the implementation of the writing to learn intervention, trained RAs administered a balance knowledge posttest and extended writing prompt to all students in the treatment, comparison, and control conditions. Because the intervention was delivered individually, the post-intervention assessments were delivered individually as well. To make sure there was no bias in administration, RAs who administered the post-intervention assessments had not worked with the same students they tested during treatment or comparison condition sessions. To document that all post-intervention assessments were delivered in the same manner, RAs completed procedural fidelity checklists as they delivered each assessment.

Balance knowledge posttest. One day after students completed the writing to learn intervention, RAs individually administered the balance knowledge posttest (see Appendix A). The balance knowledge posttest was identical to the balance knowledge pretest. A trained RA administered the balance knowledge posttest in the same manner as the pretest was administered (described above), with the exception that the posttest was given individually.

The primary investigator scored the balance knowledge posttest as the number of problems correct out of 20 and a trained RA scored 35% of the posttests for reliability purposes. There was 100% reliability of scoring at each level of questioning (i.e., level 1, level 2, level 3, level 4) on the balance knowledge posttest.

Extended writing prompt. After the student completed the 20 balance knowledge posttest items, the RA read aloud the extended writing prompt questions: What makes the beam balance? What makes it tilt right or tilt left? The RA told the student he or she could have as much time as needed to write a response and two full pages (or more, if needed) on which to write. The RA reminded the student to write in complete sentences and to write neat enough for someone else to be able to read the response but that spelling did not count. If a student asked for help during any part of the posttest extended writing prompt, the RA responded with, “I can’t help you with your answers. Please do your best.” It is important to note that the extended writing prompt questions were the same questions students in the writing to learn treatment condition responded to after completing the balance trials. Because these questions were part of the treatment condition, they were not included on the balance knowledge pretest.

Scoring of written responses on posttest. Before scoring the responses to the posttest extended writing prompt, the primary investigator typed all written responses into Microsoft Word. She corrected written responses for spelling, punctuation, and capitalization to minimize bias that might occur when scoring, as surface level features of text, such as illegible handwriting and spelling mistakes, have been shown to influence examiners’ judgments of writing (Graham, 1999). A trained RA checked 35% of the typed responses for accuracy and correspondence to the written text. Reliability for typed responses was 0.99 and all differences were discussed and resolved by consensus.

Additionally, prior to scoring the written responses, the primary investigator created a rubric for scoring level of balance understanding. This scoring system was developed based on the progression of children’s understanding of balance described by Inhelder and Piaget (1958) and validated in subsequent research (Hardiman et al., 1986; Klein, 2000; Siegler, 1985):

Level 1. Weight: Understanding that when weights are placed the same distance from the fulcrum (same peg) on each arm of the balance, the amount of weight on each arm causes the beam to balance or tilt right/left

Level 2. Distance from fulcrum: Understanding that with equal weight on each arm of the balance, the beam will tilt to the side with the weight placed on the peg furthest from the fulcrum

Level 3. Ratio of weight to distance: Understanding that the ratio of weight to distance compensates and the beam balances (i.e., beam balances with twice as much weight on one arm and twice the distance from the fulcrum on the other arm with half as much weight).

Level 4. Torque: Understanding of how to compute torque (weight times distance from the fulcrum) to determine whether the beam balances or tilts right/left

To create the rubric, the primary investigator completed multiple readings of a random sample of 50% of students' written responses from the posttest. From these written responses, the primary investigator selected examples to demonstrate each level of balance understanding on the rubric (see Appendix B). She also included a category on the rubric for responses which were "incorrect, unintelligible, or unrelated". Prior to scoring, the examples for each level of balance understanding on the rubric were checked and confirmed by an RA unfamiliar with the study intervention.

An RA was trained on all scoring procedures for written responses. Using five random samples obtained in this study of students' responses to the extended writing prompt, the primary investigator demonstrated how to use the word count feature on Microsoft Word to record total words written. Total words written for each response was counted as the number of words (i.e., a letter or group of letters separated by a space) written, regardless of grammar or meaning. The primary investigator also explained how she would score each written response for level of balance understanding using the rubric. Level of balance understanding was scored from 0-4, based on the overall impression gained from reading each written response. Each student received a score for the highest level of balance understanding expressed in his or her response. For example, if a student wrote two sentences expressing understanding of balance at level 2 and one sentence expressing understanding of balance at level 3, then he or she received a balance understanding score of 3. If a student's written response did not contain any of the levels of balance understanding outlined above or expressed an incorrect understanding of balance, it received a score of 0.

After training was completed, the primary investigator scored all responses to the posttest extended writing prompt. The trained RA independently scored 35% of the responses for reliability purposes. Reliability for total words written was 0.99 and reliability for level of balance understanding was 100%. All differences were discussed and resolved.

Conditions

Scores from the TOWL-4 Story Composition subtest and scores from the balance pretest were used for assigning students to treatment (n = 23), comparison (n = 23), and control conditions (n = 23). Participants were stratified based on TOWL-4 and balance knowledge pretest scores. Then, they were randomly assigned to the three conditions using a randomization algorithm (PASS; 2011). Assignment to conditions using this method ensured a relatively equal representation in each condition of the pre-intervention range of writing skills and pre-intervention balance knowledge in the entire sample.

Before the start of the study, the primary investigator conducted a one-day training session for all RAs to learn to administer study assessments as well as learn to implement the writing to learn treatment and the comparison conditions. Each RA received a project notebook with color coded materials, so procedures and materials for treatment and comparison conditions were not easily confused. As RAs practiced carrying out the treatment and comparison conditions, they learned to check off each completed step on a procedural fidelity checklist. After the group training session, each RA completed an individual session with the primary investigator where he or she was required to implement each condition with 100% accuracy. See Table 1 for a description of the differences between the treatment and comparison conditions.

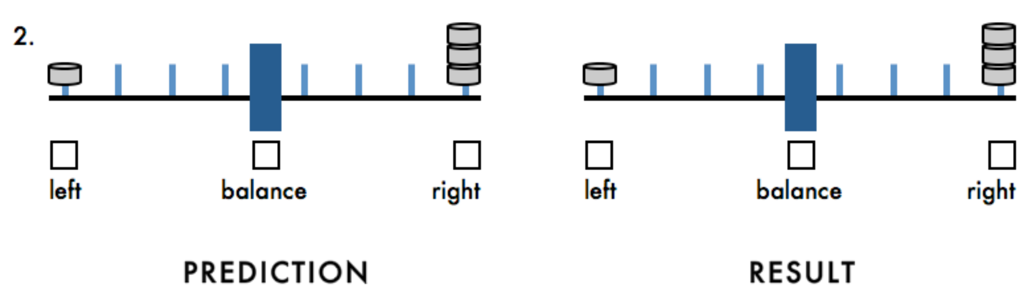
Table 1
Order of Study Activities for Treatment and Comparison Conditions

Writing to learn treatment condition	Comparison condition
Pre-intervention	
TOWL-4 Story Composition subtest	TOWL-4 Story Composition subtest
Balance knowledge pretest	Balance knowledge pretest
Intervention	
30 balance trials	30 balance trials
Four writing prompts: What is happening with the balance and weights? Are you noticing any patterns?	Four writing prompts: What do you like about this experiment? What is your favorite part so far?
Extended response writing prompt: What makes the beam balance? What makes it tilt right or tilt left?	Extended response writing prompt: What do you like most about science class?
Post-intervention	
Balance knowledge posttest	Balance knowledge posttest
Extended writing prompt: What makes the beam balance? What makes it tilt right or tilt left?	Extended writing prompt: What makes the beam balance? What makes it tilt right or tilt left?

Writing to learn treatment condition. The primary investigator and three trained RAs administered the writing to learn intervention individually to each of the 23 students in the treatment condition. The intervention involved two components: 30 balance trials and five written responses. Students required approximately 29 to 52 min (M = 39, SD = 7) to complete the session.

Balance trials. During the treatment, each student completed 30 balance trials using an equal arm balance beam. For each trial, the student was shown diagrams of two identical equal arm balances with weights placed in a particular configuration. The number of weights, represented by cylinder shapes on the diagrams, and the placement of the weights on four different pegs on either side of the fulcrum, represented by lines on the diagram, varied with each balance trial (see Figure 2 for an example and Appendix C for the balance trials sheet). For each balance trial, the student first made a prediction about whether the beam would balance, tilt left, or tilt right with the weights configured as they were in the diagram and checked a box to indicate the prediction; the student recorded his or her answer under the balance diagram labeled *Prediction*. Then the student tested the prediction by actually placing weights, as shown in the diagram, on an equal arm balance. The student recorded what happened when the weights were placed on the balance by checking the appropriate box under the balance diagram labeled *Result*.

Figure 2. Example Problem From the Balance Trials

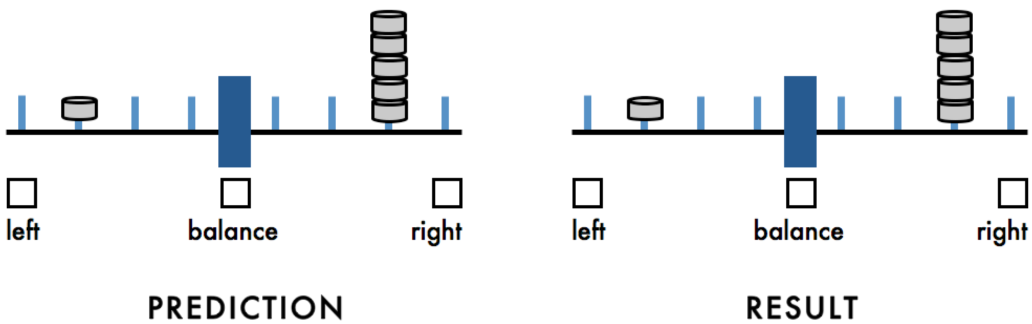


The 30 balance trials were ordered by complexity, based on the same progression of balance understanding used to score students' written responses (Hardiman et al., 1986; Inhelder

& Piaget, 1958; Klein, 2000; Siegler, 1985). Because students in the pilot study demonstrated better understandings of balance knowledge questions at levels 1 (i.e., Student understands that when weights are placed the same distance from the fulcrum [same peg] on each arm of the balance, the amount of weight on each arm causes the beam to balance or tilt right/left) and 2 (i.e., Student understands that with equal weight on each arm of the balance, the beam will tilt to the side with the weight placed on the peg furthest from the fulcrum), the balance trials began with five problems at level 1, followed by five problems at level 2. The next 10 problems were at level 3 (i.e., Student understands that the ratio of weight to distance compensates and the beam balances [the beam balances with twice as much weight on one arm and twice the distance from the fulcrum on the other arm with half as much weight]). The final 10 problems were at level 4 (i.e., Student understands how to compute torques [weight times distance from the fulcrum] on each arm to determine whether the beam balances or tilts right/left).

Prior to beginning the balance trials, the examiner (i.e., the primary investigator or one of three trained RAs) read aloud the directions to the student. The directions began with an explanation and three diagrams to show what the balance looked like when it balanced, when it tilted left, and when it tilted right (i.e., the same directions given on the balance pretest and posttest). With each diagram, a corresponding box was checked below it to indicate the position of the balance (i.e., balance, left, or right). Then the examiner asked the student to complete the sample problem. The sample contained diagrams of two identical equal arm balances, with one weight placed three pegs to the left of the fulcrum and five weights placed three pegs to the right of the fulcrum (see Figure 3). The examiner asked the student to make a prediction about whether the beam would balance, tilt left, or tilt right with the weights configured as they were in the diagram and to check a box to indicate the prediction. Then the student tested the prediction by actually placing weights, as shown in the diagram, on an equal arm balance and recorded the result. After the student completed the sample problem, the examiner explained that the student was the “lead scientist” for the experiment, so although the examiner could help if there was a problem with the balance, he or she was going to let the student do most of the work during the session. Then, the examiner directed the student to flip the page and begin the balance trials. The student completed each of the 30 balance trials, just as he or she practiced in the sample problem.

Figure 3: Sample Problem From the Balance Trials



Written responses. At four designated stopping points, a short answer writing prompt was inserted on the balance trials sheet (see Appendix C). The placement of the writing prompt after problem 5, after problem 10, after problem 20, and after problem 30 corresponded to the completion of each set of trials related to a level of balance understanding (i.e., problems 1-5 were level 1, problems 6-10 were level 2, problems 11-20 were level 3, and problems 21-30 were level 4). The prompt was the same for all four stopping points and was read aloud to the student by the examiner. The prompt read: What is happening with the balance and weights? Are you noticing any patterns? The examiner instructed the student to write his or her response on the five blank lines provided on the balance trials sheet. The examiner told the student he or she could have as much time as needed to write the response. The examiner reminded the student to write neat enough for someone else to be able to read the response later but that spelling did not count. If a student asked for help, the examiner responded, “I can’t help you with your answers. Please do your best.” The same procedure was repeated for each of the four writing prompts on the balance trials sheet.

After completing the balance trials and the fourth short answer response on the balance trials sheet, the examiner gave the student an additional piece of paper with the following writing prompt: What makes the beam balance? What makes it tilt right or tilt left?, followed by two pages of blank lined paper (see Appendix C). The examiner told the student he or she could use the notes recorded on the balance trials sheet to help write an extended response to the questions. The examiner reminded the student that he or she was the “lead scientist” for the day and scientists have the responsibility of sharing their work and explaining their findings to others. The examiner informed the student that he or she should write as if, “telling other children who did not know much about balance.” The examiner told the student he or she could have as much time as needed to write the response and two full pages (or more, if needed) on which to write. The examiner reminded the student to write in complete sentences and write neat enough for someone else to be able to read the response later but that spelling did not count. If a student asked for help, the examiner responded, “I can’t help you with your answers. Please do your best.”

Comparison condition. The primary investigator and three trained RAs administered the comparison condition individually to each of the 23 students in this condition. The comparison condition was identical to the writing to learn treatment condition in most aspects. The only two differences, outlined below, were: the short answer written responses that students completed during the balance trials and the extended written response that students completed after the balance trials. Students required approximately 25 to 52 min ($M = 38$, $SD = 7$) to complete the session. A one-way analysis of variance (ANOVA) showed there was not a statistically significant difference, $F(1, 37) = 0.43$, $p = 0.52$, between the amount of time students spent completing treatment and comparison condition sessions.

Balance trials. In the comparison condition, each student completed 30 balance trials using an equal arm balance beam (see Appendix C), just as described above for the students in the writing to learn treatment condition. The examiner procedures, directions, sample problems, and balance trials were identical to those in the writing to learn treatment condition.

Written responses. To control for the amount of time students spent writing during the balance trials, at the same four designated stopping points as in the writing to learn treatment condition described above, a short answer writing prompt was inserted on the balance trials sheet (see Appendix D). The prompt was the same for all four stopping points and was read aloud to the student by the examiner. The prompt read: What do you like about this experiment? What is your favorite part so far? The examiner instructed the student to write his or her response on the five blank lines provided on the balance trials sheet. The examiner told the student he or she could have as much time as needed to write the response. The examiner reminded the student to write neat enough for someone else to be able to read the response later but that spelling did not count. If a student asked for help, the examiner responded, "I can't help you with your answers. Please do your best." The same procedure was repeated for each of the four writing prompts on the balance trials sheet.

After completing the balance trials and the fourth short answer response on the balance trials sheet, the examiner gave the student an additional piece of paper with the following writing prompt: What do you like most about science class?, followed by two pages of blank lined paper (see Appendix D). The examiner told the student he or she could make notes if needed to help write an extended response to the question. The examiner informed the student that he or she should write as if, “telling other children who do not know much about science class.” The examiner told the student he or she could have as much time as needed to write the response and two full pages (or more, if needed) on which to write. The examiner reminded the student to write in complete sentences and write neat enough for someone else to be able to read the response later but that spelling did not count. If a student asked for help, the examiner responded, “I can’t help you with your answers. Please do your best.”

Control condition. The 23 students in the no-treatment control condition participated in business as usual classroom instruction during the three week study. As noted previously, their teachers did not cover topics related to balance during science instruction while this study was being conducted. Control condition students were only pulled out of class once to take the balance knowledge posttest and to complete the posttest extended writing prompt; these individual testing sessions were conducted during the same timeframe that students in the treatment and comparison conditions were posttested.

Fidelity of Treatment Implementation

To ensure both conditions were delivered as planned, RAs were required to implement the writing to learn treatment and comparison conditions with 100% accuracy before implementing them with participating students. Further, during each session with a student, examiners followed a step-by-step procedure and checked off each item on a procedural fidelity checklist (see Appendix C and Appendix D) as it was completed. These checklists were turned in to the primary investigator weekly, who held individual meetings with RAs who deviated from the procedures.

Additionally, all writing to learn treatment and comparison condition sessions were videotaped. A trained RA, who was not responsible for delivering the treatment or comparison conditions in this study, independently viewed a random selection of 35% of the videos, using the procedural fidelity checklist to record the number of steps completed by each examiner. Fidelity of treatment implementation, calculated as the number of correctly observed procedures divided by the sum of the correctly observed procedures and the incorrectly observed or unobserved procedures, was 0.98 for the treatment condition and 0.99 for the comparison condition.

Data Analysis

To answer research question 1 (i.e., Do students who write about balance during and after the balance trials perform better on a balance knowledge posttest than students in a comparison condition who do not write about balance during and after the trials and better than students in a no-treatment control condition?), I conducted four 2x3 repeated measures ANOVAs. TOWL-4 Story Composition subtest scores were not significantly correlated with any of the balance knowledge posttest outcomes (see Table 2), so this variable was not examined as a covariate. Time of testing (i.e., pre- and posttest) was the within group factor and condition (i.e., treatment, comparison, or control) was the between group factor. Because I ran a separate repeated measures ANOVA for each of the four levels of balance understanding on the balance knowledge pre- and posttest (i.e., levels 1-4), I made a Bonferroni correction, adjusting the level of significance to $p = 0.0125$.

To answer research question 2 (i.e., Do students who write about balance during and after the balance trials perform better on a posttest extended writing prompt than students in a comparison condition who do not write about balance during and after the trials and better than students in a no-treatment control condition?), I used one-way analysis of covariance (ANCOVA). Because TOWL-4 Story Composition subtest scores were significantly correlated with both of the extended writing prompt outcomes at posttest (see Table 2), I examined this variable as a covariate in each of two ANCOVAs. For the first ANCOVA, I examined group differences on the dependent variable total words written. For the second ANCOVA, I examined group differences on the dependent variable level of balance understanding.

For each significant group difference found in the above analyses, I reported the standardized mean difference effect size (*ES*), which was calculated as the difference between group means at posttest divided by the pooled standard deviation of the two groups (Lipsey & Wilson, 2001). For research question 1, because the balance knowledge assessment was given at both pre- and posttest, I first adjusted each posttest mean on this measure by subtracting the pretest mean from it. Then, I calculated the *ES* using the standardized mean difference *ES* formula with the adjusted means.

Table 2
Correlations Among Pre- and Posttest Measures

Measure	1	2	3	4	5	6	7	8	9	10	11
1. TOWL-4	--										
2. Level 1 pre	0.29*	--									
3. Level 2 pre	0.11	0.02	--								
4. Level 3 pre	-0.18	-0.44**	0.11	--							
5. Level 4 pre	-0.04	-0.39**	0.44**	0.20	--						
6. Level 1 post	0.15	0.19	0.11	-0.18	0.07	--					
7. Level 2 post	-0.03	0.14	0.23	0.13	0.08	0.22	--				
8. Level 3 post	0.03	0.10	0.00	-0.04	-0.07	-0.30*	0.04	--			
9. Level 4 post	0.13	-0.07	0.09	0.03	0.33**	-0.16	0.16	0.24*	--		
10. TWW	0.34**	0.18	0.13	-0.18	-0.03	-0.03	-0.06	0.09	0.06	--	
11. LBU	0.44**	0.18	0.14	-0.10	0.13	0.09	0.08	0.22	0.05	0.16	--

Note. TOWL-4 = Tests of Written Language-Fourth Edition, TWW = total words written, LBU = level of balance understanding
* $p < 0.05$, ** $p < 0.01$

CHAPTER III

RESULTS

Chi square analyses and a one-way ANOVA (i.e., for age in years) revealed no statistically significant differences between conditions on four demographic variables: (a) sex: $\chi^2(2, N = 69) = 3.37$, $p = 0.19$; (b) received services as ELL: $\chi^2(2, N = 69) = 0.18$, $p = 0.91$, (c) received free/reduced lunch: $\chi^2(2, N = 69) = 1.64$, $p = 0.44$; and (d) age: $F(2, 66) = 0.87$, $p = 0.43$. A one-way ANOVA revealed no statistically significant differences between conditions prior to intervention on the TOWL-4 Story Composition subtest, $F(2, 66) = 0.02$, $p = 0.98$ ($M = 9.20$, $SD = 2.08$). Four ANOVAs, each corrected using the Bonferroni adjustment ($p = 0.0125$), indicated no statistically significant differences between conditions on balance knowledge pretest scores at each level of balance understanding, level 1: $F(2, 66) = 0.47$, $p = 0.63$; level 2: $F(2, 66) = 1.04$, $p = 0.36$; level 3: $F(2, 66) = .50$, $p = 0.63$; level 4: $F(2, 66) = 1.23$, $p = 0.30$.

Research Question 1

Table 3 displays the means and standard deviations by condition for each level of balance understanding assessed on the balance knowledge pre- and posttest. I conducted a series of four 2x3 repeated measures ANOVAs (significance levels adjusted to $p = 0.0125$) to assess group differences on the balance knowledge posttest at each level.

The repeated measures ANOVA for balance knowledge scores on level 1 questions indicated a significant main effect for time, $F(1, 66) = 23.90$, $p < 0.001$, and a significant main effect for group, $F(2, 66) = 5.99$, $p < 0.01$. The analysis also showed a significant group by time interaction, $F(2, 66) = 5.05$, $p = 0.01$. Post-hoc analysis revealed students in the control condition had significantly higher scores on balance knowledge posttest questions at level 1 than students in the treatment condition ($p < 0.01$), with an *ES* of 0.89. Students in the control condition also had significantly higher scores on the balance knowledge posttest questions at level 1 than students in the comparison condition ($p = 0.01$), with an *ES* of 1.05. The treatment and comparison conditions did not differ statistically on posttest questions at level 1.

Table 3
Pre- and Posttest Balance Knowledge Scores by Level of Balance Understanding and Condition

	Condition					
Balance Understanding	Treatment		Comparison		Control	
	Pre	Post	Pre	Post	Pre	Post
Level 1	4.74 (0.54)	3.96 (1.11)	4.83 (0.49)	4.04 (0.93)	4.87 (0.34)	4.83 (0.39)
Level 2	2.57 (2.09)	3.17 (1.64)	2.22 (1.86)	3.09 (1.73)	3.04 (1.89)	3.96 (1.72)
Level 3	0.13 (0.46)	1.78 (1.81)	0.17 (0.49)	1.91 (1.65)	0.30 (0.82)	0.09 (0.42)
Level 4	0.87 (1.36)	1.52 (1.28)	0.35 (0.78)	1.61 (1.47)	0.65 (1.19)	0.35 (0.65)

The repeated measures ANOVA for balance knowledge scores on level 2 questions revealed a significant main effect for time, $F(1, 66) = 8.18$, $p < 0.01$. There was not a significant main effect for group, $F(2, 66) = 2.23$, $p = 0.12$, or a significant interaction between time of testing and group, $F(2, 66) = 0.12$, $p = 0.89$, so I did not perform follow-up analyses.

For level 3 questions, a repeated measures ANOVA indicated a significant main effect for time, $F(1, 66) = 32.35$, $p < 0.001$; a significant main effect for group, $F(2, 66) = 8.10$, $p < 0.01$; and a significant group by time interaction, $F(2, 66) = 11.77$, $p < 0.001$. Post-hoc analysis revealed students in the treatment condition had significantly higher scores on balance knowledge posttest questions at level 3 than students in the control condition ($p < 0.001$), with an *ES* of 1.42. Additionally, students in the comparison condition had significantly higher scores on posttest questions at level 3 ($p < 0.001$) when compared to students in the control condition (*ES* = 1.62). The treatment condition was not statistically different from the comparison condition on level 3 questions.

The repeated measures ANOVA for balance knowledge posttest scores on level 4 questions indicated a significant main effect for time, $F(1, 66) = 12.07$, $p < 0.01$ but the main effect for group was not significant, $F(2, 66) = 3.12$, $p = 0.05$. However, the analysis showed a significant group by time interaction, $F(2, 66) = 8.71$, $p < 0.001$. Post-hoc analysis indicated students in both the treatment (*ES* = 0.94) and comparison conditions (*ES* = 1.37) significantly outperformed students in the control condition ($p < 0.01$ for each) on posttest questions at level 4. However, the treatment and comparison conditions did not differ statistically on level 4 questions.

Research Question 2

Table 4 displays the means and standard deviations by conditions for total words written and level of balance understanding, both outcomes which were assessed on the posttest extended writing prompt. A preliminary analysis of the homogeneity of regression assumption indicated the relationship between the covariate (i.e., TOWL-4) and the dependent variable (i.e., total words written) did not differ significantly as a function of condition, $F(2, 63) = 0.97$, $p = 0.39$. A one-way ANCOVA, controlling for TOWL-4 Story Composition subtest scores, indicated no statistically significant differences between conditions in total words written on the posttest extended writing prompt, $F(2, 65) = 0.44$, $p = 0.64$. The homogeneity of regression assumption was met for the level of balance understanding outcome as well, $F(2, 63) = 1.97$, $p = 0.15$. There were no statistically significant differences between conditions for level of balance understanding on the posttest extended writing prompt, $F(2, 65) = 1.24$, $p = 0.30$, when controlling for TOWL-4 scores.

Table 4
Posttest Total Words Written and Level of Balance Understanding by Condition

	Condition		
Measure	Treatment	Comparison	Control
Total Words Written	42.57 (25.79)	44.78 (15.50)	39.00 (20.48)
Level of Balance Understanding	0.96 (0.77)	1.17 (0.39)	0.96 (0.56)

CHAPTER IV

DISCUSSION

In this study, 69 grade 4 students participated in a writing to learn intervention in science (treatment, n = 23), a comparison condition (n = 23) or a no-treatment control condition (n = 23). While completing 30 balance trials, students in the treatment condition wrote what they were learning about balance. After the trials, treatment students completed an extended written response about balance. Students in the comparison condition completed the same 30 balance trials. To control for writing time, comparison students wrote about their favorite parts of the balance trials and wrote an extended written response about their favorite parts of science class. Students in the control condition participated in business as usual science instruction in their classrooms; this instruction did not involve balance or any topics related to balance.

Research Question 1: Balance Knowledge

On a 20-item balance knowledge posttest, statistically significant differences were found between conditions on questions at level 1 (i.e., when weights are placed the same distance from the fulcrum [same peg] on each arm of the balance, the amount of weight on each arm causes the beam to balance or tilt right/left), level 3 (i.e., the ratio of weight to distance compensates and the beam balances), and level 4 (i.e., torque [weight times distance from the fulcrum] determines whether the beam balances or tilts right/left).

Contrary to my hypothesis, students in the control condition performed approximately one point better than students in both the treatment (ES = 0.89) and comparison conditions (ES = 1.05) on balance knowledge posttest questions at level 1. The treatment and comparison conditions did not differ statistically on level 1 questions on the balance knowledge posttest. This finding raises the question of whether students in treatment and comparison conditions may have overgeneralized higher level balance understandings to level 1 questions at posttest. Deepening their understanding of balance concepts at more advanced levels may have led some students in the treatment and comparison conditions to abandon their initial correct assumptions (83% of treatment and comparison condition students answered all level 1 questions correctly at pretest) when answering lower level questions on the posttest. Thus, treatment and comparison students may have overextended balance understandings that did not apply to the most basic problems. Indeed, the performance of students in both the treatment and comparison condition was much more variable from pre- to posttest on level 1 questions, with posttest standard deviations at or near double that of pretest standard deviations for both groups. It is possible treatment and comparison group students were "using but confusing", a phrase typically used for literacy skills (Bear, Invernizzi, Templeton, & Johnston, 2011), higher levels of balance understanding such as torque, thus creating their more variable performance on level 1 questions at posttest. Future studies are needed to substantiate these hypotheses about student performance and overgeneralization of higher levels of balance understanding. Studies with a think-aloud component during the balance trials and during the posttest could be used, as students would need to explain their reasoning about balance questions at level 1 in order to fully understand their thought processes when solving and answering these questions.

On balance knowledge posttest questions at the highest levels, students in the treatment condition and students in the comparison condition performed significantly better than students in the control condition (level 3 questions: treatment v. control $ES = 1.42$, comparison v. control $ES = 1.62$; level 4 questions: treatment v. control $ES = 0.94$, comparison v. control $ES = 1.37$), but treatment and comparison conditions did not differ statistically from one another. The findings at levels 3 and 4 were partially aligned with my hypothesis, although I expected students in the treatment condition would outperform students in the comparison condition at these levels of questioning.

It could be argued the lack of significant differences between treatment and comparison conditions on the balance knowledge posttest was because the balance trials, not the writing tasks, were responsible for student gains from pre- to posttest. However, this argument is not supported, as the current study was not designed to provide information about the relative impact of the balance trials compared to the writing tasks students completed during and after the trials. In this study, the science and writing tasks were assessed together, as a combined writing to learn intervention.

A closer examination of the writing students in the treatment condition produced during and after the balance trials may provide an explanation for the lack of significant differences between the treatment and comparison groups on the balance knowledge posttest. First, a majority of students the treatment condition wrote very little during the balance trials; two students wrote only one-word answers (i.e., "no") for each response, and 13 students wrote only one short phrase or one sentence for each response (e.g., "They are the same patterns," "same or different,"). Second, although the short answer questions during the balance trials were intended to facilitate the learning of balance concepts for students in the treatment condition, approximately half of the treatment students ($n = 12$) produced short answer responses with little substance. Four students wrote responses to simply confirm or deny patterns but did not explain what the patterns were or how they were different from other subgroups of questions (e.g., "No," "Yes, I am noticing a pattern"); two students wrote only a phrase or sentence to state that patterns were the same or different, without further explanation of *how* they were alike or different (e.g., “I notice it was the same,” “The patterns are getting different”); and six students recorded only the order of answers for each subgroup of questions (e.g., “The patterns is that it is left, right, balance”). Similarly, although students in the treatment condition were encouraged to write an extended response to the writing prompt they answered after the balance trials, almost half of the students (n = 11) wrote two sentences or less (i.e., four students wrote 2 sentences, seven students wrote one sentence or less). Because many students in the treatment condition produced little substantive writing during and after the balance trials, they may have not benefitted from the potential positive effects of the writing to learn intervention.

A small group of students (n = 5) in the treatment condition demonstrated optimal performance during and after intervention, providing additional support for the hypothesis that the amount and content of writing produced by treatment group students during and after intervention had an influence on their performance at posttest. These students will be referred to as *responders* during this discussion of their performance. It is important to note here that the responders’ incoming balance knowledge (balance knowledge pretest scores) and writing skills (TOWL-4 Story Composition subtest score), were not statistically different from the remainder of the students in the treatment condition and not statistically different from the remainder of the study sample as a whole. Thus, I hypothesize that the responders’ writing during and after intervention impacted their posttest outcomes.

In contrast to the treatment group students described above, the five responders produced more substantive writing during and after the balance trials. Additionally, the responders earned posttest scores from 4 to 11 points higher than their pretest scores. Specifically, the responders all earned four or five (out of five possible) points on level one questions at posttest.

Furthermore, four of the five responders made gains of four to five points from pre- to posttest on level 2 questions; the other responder correctly answered all level two questions but had done so at pretest as well. For level 3 questions, three of the responders improved from 0 points at pretest to four or five correct at posttest. At the highest level of balance understanding, level 4, four responders improved from 0 correct at pretest to 1 (n = 2), 2 (n = 1), or 3 correct (n = 1) at posttest.

The responders' short answer responses during the balance trials ranged from one to four complete sentences (e.g., "If more weights are put on one side, then it goes to that side. If the same amount of weights are put on the balance beam, then it levels itself out"). Additionally, the two responders with the greatest pre- to posttest gains in the treatment condition (i.e., 11 point gain from pre-to posttest for each student) included vocabulary terms (i.e., peg, fulcrum) they were taught at the beginning of the intervention in their short answer responses. For example, one such student wrote, " If the weights are farther from the *fulcrum* it will tilt that way." Similarly, the extended responses written by the responders after completing the balance trials tended to be longer and better developed than most other students in the treatment condition. Four responders wrote three sentences for the extended response, and one responder wrote four sentences. Two of the five responders also included the vocabulary term *fulcrum* in their extended responses written after completing the balance trials.

Because this group of students performed as intended in the treatment condition and subsequently performed 4 to 11 points higher at posttest than pretest, there is potential for maximizing treatment effects. Future studies may need to involve different short answer response formats and extended response questions that will elicit the type of texts written by the responders in this study. Furthermore, as the students who performed optimally often included balance-related vocabulary in their written responses, additional instruction during the writing to learn intervention may be necessary to facilitate students' learning and understanding of new vocabulary terms related to balance. Future research could also examine the use of strategy instruction designed to help students learn to write texts about the balance trials. If students are provided such instruction, their writing during and after the trials may prove more beneficial in helping them internalize the new balance concepts they learn. That is, a task-specific writing strategy might help students produce longer and more substantive responses, which may aid in their reflection on and internalization of new ideas learned during and after the balance trials.

Examining the writing students in the comparison condition produced during and after the balance trials also provides some insight into the lack of significant differences between the treatment and comparison groups on the balance knowledge posttest. To control for time spent writing during the balance trials, students in the comparison condition wrote about their favorite part of the experiment (i.e., answered: What do you like about this experiment? What is your favorite part so far?). To control for writing time after the trials, students in the comparison condition wrote about their favorite part of science class (i.e., answered: What is your favorite part of science class?). However, because these questions were rather open-ended, a small number of students in the comparison condition approximated the type of written responses that were expected of students in the treatment condition; these same students also showed pre- to posttest gains on the balance knowledge assessment.

Specifically, three students in the comparison condition produced short answer responses which more closely aligned with the questions students in the treatment group answered during and after the trials. When noting their favorite parts of the balance trials, these comparison condition students noted: (a) what was happening with the balance and weights (e.g., "One thing

I like about this experiment is discovering that the little hoops [weights] go wherever there is more"); (b) patterns they were noticing (e.g., "It is getting harder and I think it is left or right but it is balance," and "The numbers to balance are getting bigger"); and (c) what made the balance tilt right or left (e.g., "My favorite part is learning which direction it goes according to how much weight you put on it and where you put it"), all topics students wrote about in the treatment condition. These three students also earned balance knowledge posttest scores from 4 to 6 points higher than their pretest scores. Because these students in the comparison condition performed much like what was intended for students in the treatment condition, they potentially learned more from the balance and writing tasks. Future research should involve a comparison condition which completes writing tasks that are more distinct from the treatment condition, so there is not the possibility of spillover of treatment effects.

Another possible explanation for outcomes on the balance knowledge posttest is that students in the treatment condition did not maximally benefit from the writing to learn intervention, as compared to students in the comparison condition, because the task was too cognitively overwhelming for fourth-grade students. Although the CCSS (2010) emphasize teaching students to conduct research, to explain their findings in writing, and to learn to use expository and informative genres in the elementary grades, students in grade 4 may not have the language and writing skills necessary to fully express their learning from a science task. It may be more age- and grade-appropriate to conduct a future iteration of this intervention with students in middle school. The CCSS call for an even greater focus on literacy in science in the middle school grades, so writing to learn studies with this student population would be important.

Research Question 2: Extended Writing Prompt

After controlling for TOWL-4 Story Composition scores, no statistically significant differences were found between conditions on the posttest extended writing prompt for total words written or for level of balance understanding. Across all conditions, students produced relatively few words in their written responses and, on average, students expressed balance understandings at or near level 1.

Although I had hypothesized that the treatment condition, on average, would produce longer responses to the posttest extended writing prompt than the comparison or control conditions, this was not the case. A possible explanation is that students in the treatment condition were less motivated than other students when completing the posttest extended writing prompt, as they had just completed the same task one day earlier during intervention. Although I felt it was important to have an outcome that was proximal to what students in the treatment condition were asked to do during intervention, it may be necessary to use other posttest extended writing prompts in future studies. A prompt similar, but not identical, to what treatment students respond to during intervention could be used at posttest in future work.

The finding that students, on average, produced few words in their extended written responses at posttest was not surprising. Students in this study tended to write very little, as shown in the brevity of the written responses produced by students in the treatment and comparison conditions during and after the balance trials (discussed above). For future studies, the extended writing prompt questions at posttest may need to be reworded to elicit more information from students. Additionally, students may need additional verbal prompts to add more to their extended responses; in a previous study (Graham, 1990), students produced substantially longer texts when given simple prompts to add more information to their writing

(Graham, 1990). Aids for planning and brainstorming (e.g., graphic organizers) may also help students in the treatment condition generate more ideas to write about in their extended responses.

My other hypothesis about the extended writing prompt, that the treatment condition would outperform comparison and control on level of balance understanding, was also not supported by the data. Furthermore, the low level of balance understanding produced by students in the treatment ($M = 0.96$, $SD = 0.77$) and comparison conditions ($M = 1.17$, $SD = 0.39$) on the posttest extended writing prompt did not align with the levels of balance understanding demonstrated by these two groups on the balance knowledge posttest. Only one student in the entire sample, a student in the treatment condition, produced an extended written response at posttest that indicated balance understanding at level 3. Yet, approximately one-fourth ($n = 6$) of students in the treatment condition scored four or five correct (out of five possible) on balance knowledge questions at level 3 on the posttest. Four students in the comparison condition earned four or five correct on posttest questions at level 3, but none demonstrated balance understanding at level 3 in their extended written responses at posttest. Similarly, students in the treatment and comparison conditions earned, on average, three problems correct on balance knowledge posttest questions at level 2 ($M = 3.17$ and $M = 3.09$, respectively). However, only three students in the treatment condition and four students in the comparison condition produced extended written responses demonstrating balance understanding at this level. These data provide further support that the task of explaining one's understanding of balance concepts may have been too difficult for students in grade 4. Adjusting the writing to learn intervention to include specific instruction in how to write about one's understanding of balance concepts, as discussed above, may prove powerful in future studies of this writing to learn intervention with upper elementary-aged students. Additionally, as mentioned previously, a future iteration of this intervention using middle school participants may provide valuable insight into whether producing text to demonstrate one's understanding of complicated science concepts may be better suited for older students.

Limitations

Overall, it is important to interpret the results from this study in light of several limitations. First, with a sample of 69 students, the study lacked power to detect effects. Although effect sizes showed students in both the treatment and comparison condition performed approximately one standard deviation unit (*ES*s ranged from 0.89 to 1.62) higher than students in the control condition on balance knowledge posttest questions at levels 3 and 4, the sample size limitation likely impacted outcomes measured with ANOVAs.

In future studies of this or similar writing to learn interventions, larger student samples should be assessed. Such samples should include more students identified as at-risk for or having learning disabilities with written language deficits; examining the impact of the writing to learn intervention on science learning for these students would provide important information about how to best design writing to learn tasks to meet the specific needs of this student population. Additionally, future studies of the balance writing to learn task need to be conducted with older students, perhaps in the middle school grades. It may be possible, as noted above, that the balance trials and writing tasks created too much of a cognitive demand for fourth-grade students. Students in the middle school grades may have more developed thinking and language skills, which would make them more likely to be able to express what they have learned from the balance trials in written form.

A second limitation of this study was that it was not designed to determine if writing alone facilitated students' learning. That is, the components of the writing to learn intervention (i.e., balance trials and written responses) were not examined separately to assess their impact on student outcomes. Future studies should examine the same treatment condition studied here compared to a condition in which students complete the balance trials but answer short answer and extended response questions verbally. This design would allow for an understanding if writing is the key component which facilitates student learning from the balance task.

Lastly, this study involved a highly controlled intervention with little input or instruction from the examiner. Therefore, the findings lack generalizability to other settings. In the pilot study, when the examiner provided verbal prompts to students they seemed to be better able to explain their understanding of balance concepts in writing. With this in mind and an ultimate goal is to develop effective writing to learn interventions that can be easily transferred into classroom practice, it is important to add prompting and instructional elements to facilitate students' science learning in future studies. Developing short writing to learn units (e.g., two to three weeks) on specific science topics, such as balance, could accomplish this. These units could involve multiple types of writing (e.g., short answer, extended response, journaling), as the cumulative impact of instruction and multiple opportunities to write may have powerful effects on students' learning. Additionally, longer writing to learn units with instruction are better aligned with the type of activities and instruction that would occur in actual classroom settings, increasing the applicability of these types of interventions to practice. Such studies could involve a writing to learn unit that is implemented by a classroom teacher compared to the same unit (i.e., same science content) implemented in similar classroom without the writing tasks. The findings from such a study may prove more generalizable to a variety of classroom settings.

Appendix A

Pre- and Posttest Measures

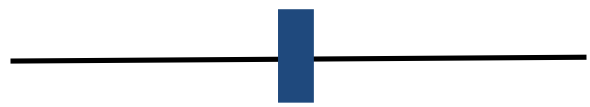
Balance Knowledge Pretest

NAME: ____

Directions:

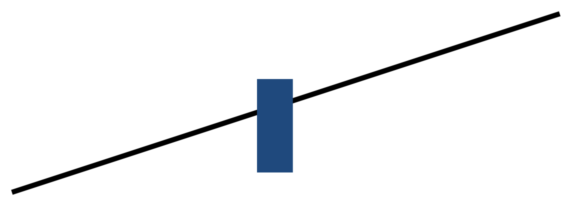
What do you think the balance will do?

If you think it will <u>balance</u>,
check:



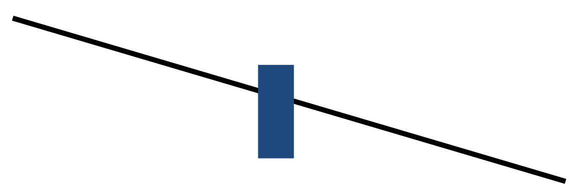
☐ left ☑ balance ☐ right

If you think it will <u>tilt left</u>,
check:



☑ left ☐ balance ☐ right

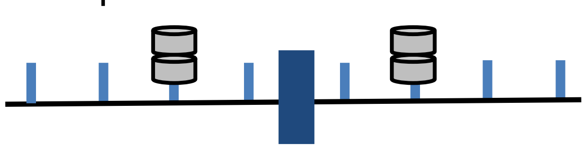
If you think it will <u>tilt right</u>,
check:



☐ left ☐ balance ☑ right

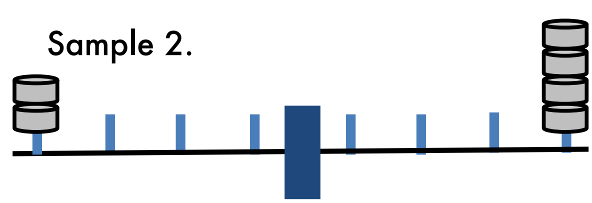
Let's try some together:

Sample 1.



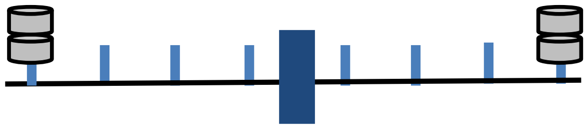
☐ left ☐ balance ☐ right

Sample 2.



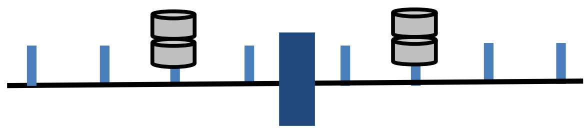
☐ left ☐ balance ☐ right

1.



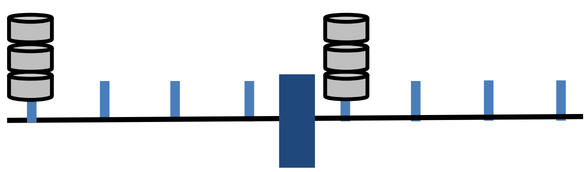
☐ left ☐ balance ☐ right

2.



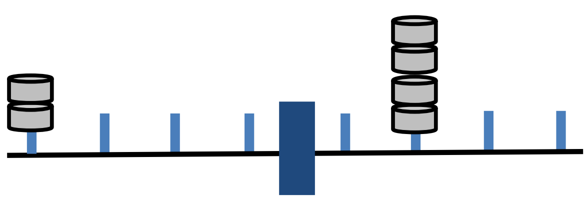
☐ left ☐ balance ☐ right

3.



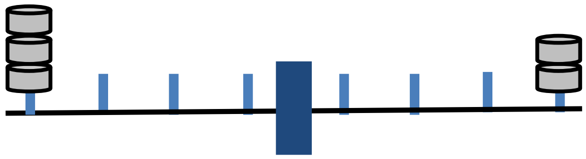
☐ left ☐ balance ☐ right

4.



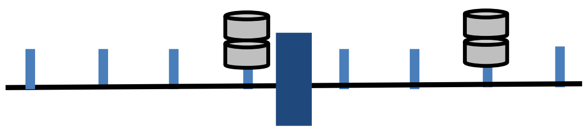
☐ left ☐ balance ☐ right

5.



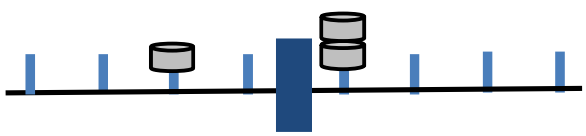
☐ left ☐ balance ☐ right

6.



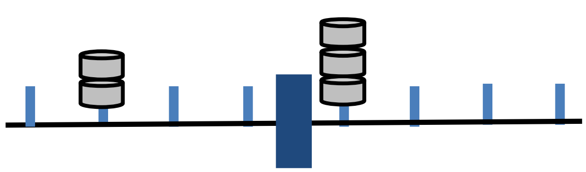
☐ left ☐ balance ☐ right

7.



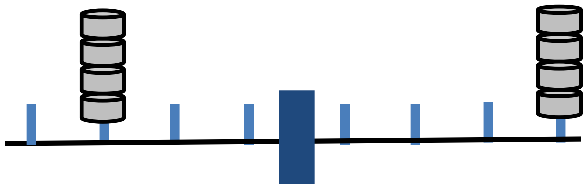
☐ left ☐ balance ☐ right

8.



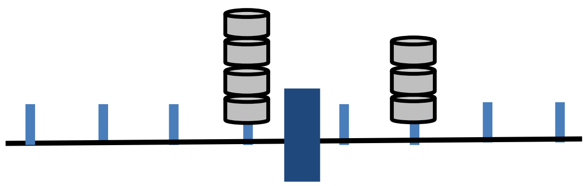
☐ left ☐ balance ☐ right

9.



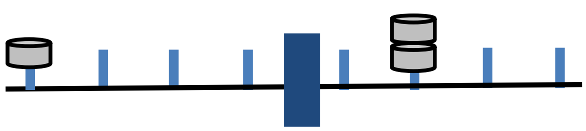
☐ left ☐ balance ☐ right

10.



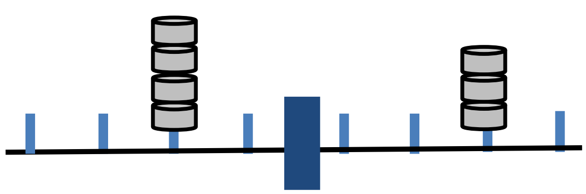
☐ left ☐ balance ☐ right

11.



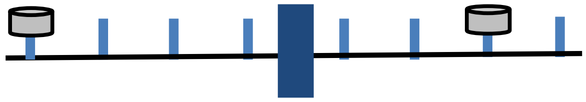
☐ left ☐ balance ☐ right

12.



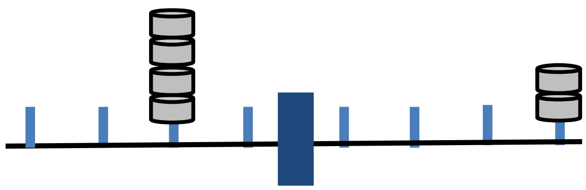
☐ left ☐ balance ☐ right

13.



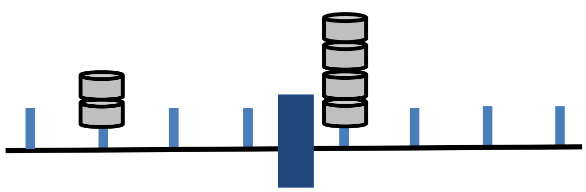
☐ left ☐ balance ☐ right

14.



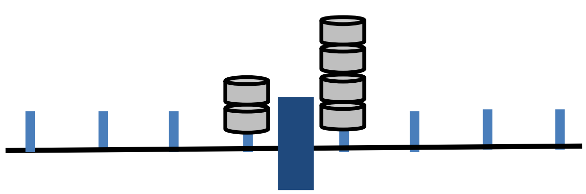
☐ left ☐ balance ☐ right

15.



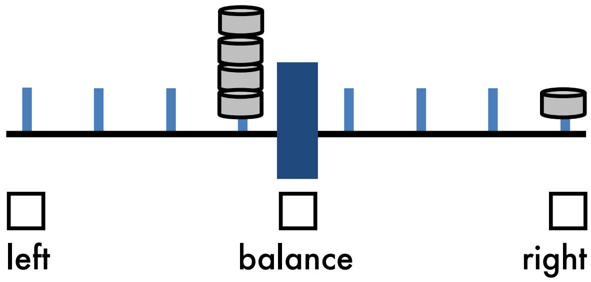
☐ left ☐ balance ☐ right

16.

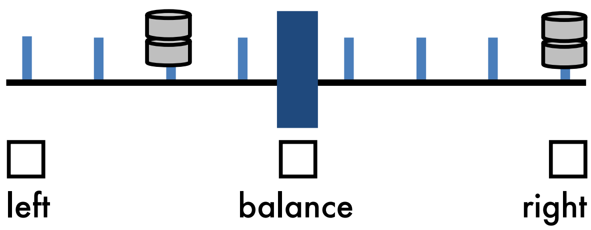


☐ left ☐ balance ☐ right

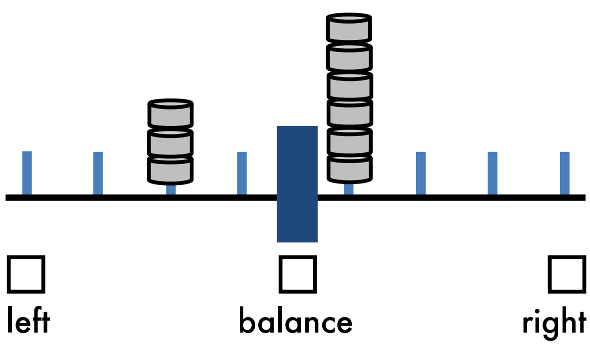
17.


left
balance
right

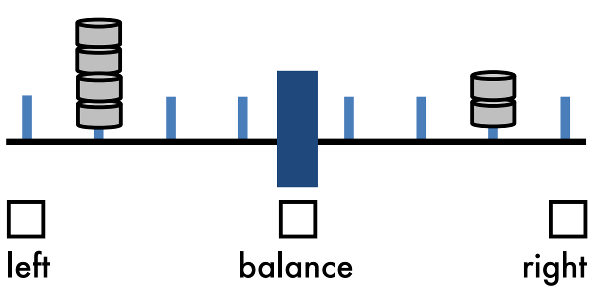
18.


left
balance
right

19.


left
balance
right

20.


left
balance
right

Balance Knowledge Posttest

*This assessment is identical to the balance knowledge pretest, with the addition of the following after problem #20:

What makes the beam balance? What makes it tilt right or tilt left?

Examiner directions for balance knowledge pretest

Directions

- ☐ Hand student packet and pencil.
- ☐ "Today you are going to complete some problems about balance. Please begin by writing your first and last name at the top of the packet."
- ☐ "The drawings on the page in front of you are of a balance beam, like the one I have here (*point to the balance*). You can put weights on the balance like this. (*Put two weights on the left side of the balance and one weight on the right side of the balance*.) See how the balance tipped or tilted left? Depending on the weights you put on it, the balance will tilt left, tilt right or balance."
- ☐ "Let's look at the directions at the top of your page. Read the directions silently as I read them aloud. They say, 'Directions: What do you think the balance will do? If you think it will balance, check balance.' See how the drawing shows a beam that is balanced and the box labeled 'balance' is checked?"
- ☐ "Now, look at the second drawing. It says, 'If you think it will tilt left, check left.' This drawing shows a beam that is tilted left and the checkmark is in the box above the word 'left'."
- ☐ "Now, look at the third drawing. It says, 'If you think it will tilt right, check right.' This drawing shows a beam that is tilted right and the checkmark is in the box above the word 'right'."
- ☐ "Let's try some sample problems together. Mark the correct answer on your paper for sample 1." *Give 5 seconds wait time.* "Which answer did you check?"
    - ☐ "Yes, that's right. The balance beam would balance. You should have put a check above the word balance for sample 1."
    - ☐ "No, look again. The balance beam would balance for sample 1. Erase your answer and put a check above the word balance for sample 1."
- ☐ "Now, mark your answer for sample 2." *Give 5 seconds wait time.* "Which answer did you check?
    - ☐ "Yes, that's right. The balance beam would tilt right in sample 2. You should have put a check above the word right for sample 2."
    - ☐ "No, look again. The balance beam would tilt right in sample 2. Erase your answer and put a check above the word right for sample 2."
- ☐ Do you have any questions about sample 1 or sample 2?"
- ☐ "The remaining questions in your packet are like the sample problems we just practiced. Please work carefully to finish each problem in the packet. There are 20 problems to complete. You may have as much time as you need. Do you have any questions about what you need to do?"

- ☐ "Great! Flip to the next page in your packet. You may begin. When you finish all 20 problems, put your head down on your desk and I will collect your paper."
- ☐ Collect pretest when student finishes.

****If a student asks for help during the test, say, "I can't help you with your answers. Please do your best."****

Examiner directions for balance knowledge posttest

Directions for multiple choice

- ☐ Hand student packet and pencil.
- ☐ "Now that you have practiced using the balance beam and weights, you are going to complete some more problems about balance. These problems look a lot like the ones you have practiced before. Please begin by writing your first and last name at the top of the packet."
- ☐ "Let's look at the directions at the top of your page. Read the directions silently as I read them aloud. They say, 'Directions: What do you think the balance will do? If you think it will balance, check balance.' See how the drawing shows a beam that is balanced and the box labeled 'balance' is checked?"
- ☐ "Now, look at the second drawing. It says, 'If you think it will tilt left, check left.' This drawing shows a beam that is tilted left and the checkmark is in the box above the word 'left'."
- ☐ "Now, look at the third drawing. It says, 'If you think it will tilt right, check right.' This drawing shows a beam that is tilted right and the checkmark is in the box above the word 'right'."
- ☐ "Let's try some sample problems together. Mark the correct answer on your paper for sample 1." *Give 5 seconds wait time.* "Which answer did you check?"
    - ☐ "Yes, that's right. The balance beam would balance. You should have put a check above the word balance for sample 1."
    - ☐ "No, look again. The balance beam would balance for sample 1. Erase your answer and put a check above the word balance for sample 1."
- ☐ "Now, mark your answer for sample 2." *Give 5 seconds wait time.* "Which answer did you check?
    - ☐ "Yes, that's right. The balance beam would tilt right in sample 2. You should have put a check above the word right for sample 2."
    - ☐ "No, look again. The balance beam would tilt right in sample 2. Erase your answer and put a check above the word right for sample 2."
- ☐ Do you have any questions about sample 1 or sample 2?"
- ☐ "The remaining questions in your packet are like the sample problems we just practiced. Please work carefully to finish each problem in the packet. There are 20 problems to complete. You may have as much time as you need. Do you have any questions about what you need to do?"
- ☐ "Great! Flip to the next page in your packet. You may begin. When you finish the first 20 problems, put your head down on your desk and I will tell you what to do next."

Directions for written response

- ☐ "Scientists share their work and explain their findings to others. Because you were the lead scientist during our balance experiment, I have one more question for you to answer about balance. For this question, you will write your answer. Please write your answer as if you were writing it for a boy or girl who doesn't know much about balance."
- ☐ "Read the question to yourself while I read it aloud. It says, 'What makes the beam balance? What makes it tilt right or tilt left?''
- ☐ "Please write your answer on the lines provided. You can have as much time as you need to write and more paper to write on if you need it. Make sure to write in complete sentences and neat enough for someone else to read your writing. Don't worry about spelling; just spell the best you can."
- ☐ If the student produces writing that is illegible, have him or her read the written response aloud to you and write it down on a separate sheet of paper. Staple this response to the back of the posttest.

****If a student asks for help during the test, say, "I can't help you with your answers. Please do your best."****

Appendix B

Balance Understanding Scoring Rubric

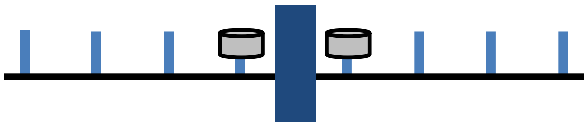
CATEGORY and DEFINITION	EXAMPLES
SCORE OF 1 **Weight** **(Same pegs)** *Understanding that when weights are placed the same distance from the fulcrum (same peg) on each arm of the balance, the amount of weight on each arm of the balance causes it to balance or tilt left/right*	• weight(s) • heavy, heavier • more/less weight • examples with weight (e.g., When you put 2 on the left and 2 on the right it will balance.) • What makes it tilt right is when something on the right side weighs more than what's on the left side • when the weights are equal
SCORE OF 2 **Level 2: Distance from the fulcrum** **(Same weight, different pegs)** *Understanding that with equal weight on each arm of the balance, the balance will tilt to the side with the weight placed on the peg furthest from the fulcrum*	• the places the weight is placed • further/farther • closer to the middle/closer to the end • same/different pegs • same/different places • how many numbers of weights are on each side and where • farther from the fulcrum
SCORE OF 3 **Level 3: Ratio of weight to distance balances** **(Double the weight, half the distance; double the distance, half the weight)** *Understanding that the ratio of weight to distance compensates and the beam balances*	• sometimes the weight is half of the other
SCORE OF 0 **No Score: Incorrect, unintelligible or unrelated**	• there are trick you • If it tilts left, it will go left. • If it goes right, it goes right. • metal • Get 2 spoons and put food on it and whichever has the most food is the answer.

Appendix C

Writing to Learn Treatment Condition Materials

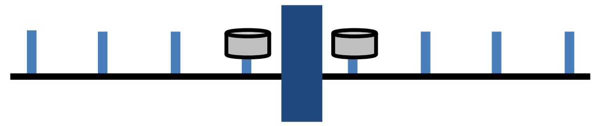
Balance Trials Sheet: Treatment Condition

1.



☐ left ☐ balance ☐ right

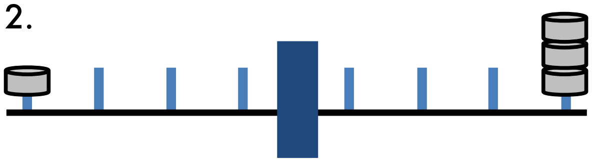
PREDICTION



☐ left ☐ balance ☐ right

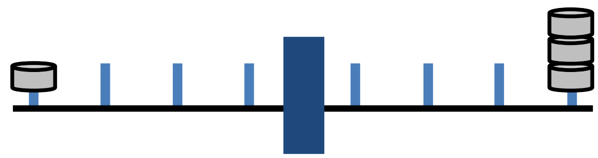
RESULT

2.



☐ left ☐ balance ☐ right

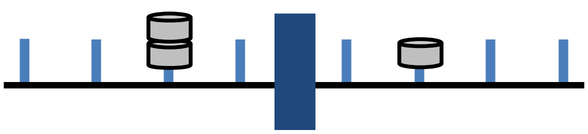
PREDICTION



☐ left ☐ balance ☐ right

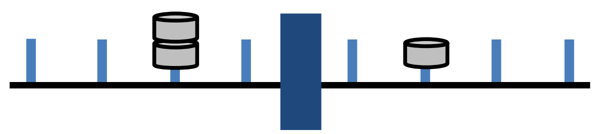
RESULT

3.



☐ left ☐ balance ☐ right

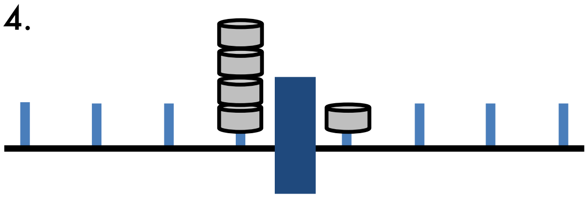
PREDICTION



☐ left ☐ balance ☐ right

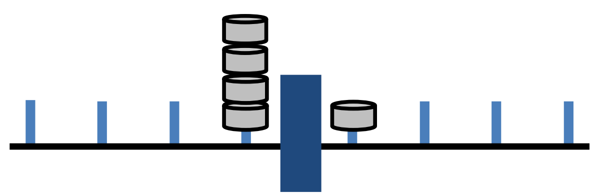
RESULT

4.



☐ left ☐ balance ☐ right

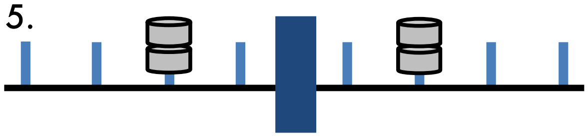
PREDICTION



☐ left ☐ balance ☐ right

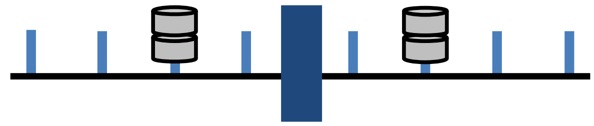
RESULT

5.



☐ left ☐ balance ☐ right

PREDICTION

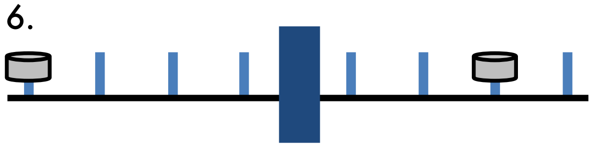


☐ left ☐ balance ☐ right

RESULT

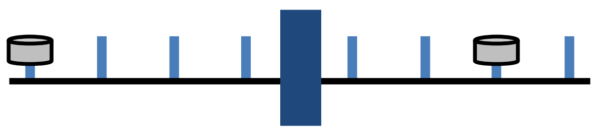
What is happening with the balance and weights? Are you noticing any patterns?

6.



☐ left ☐ balance ☐ right

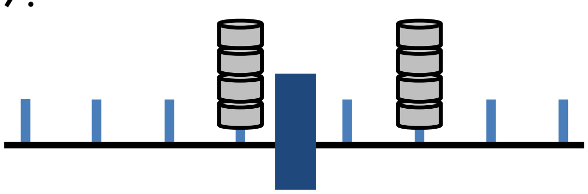
PREDICTION



☐ left ☐ balance ☐ right

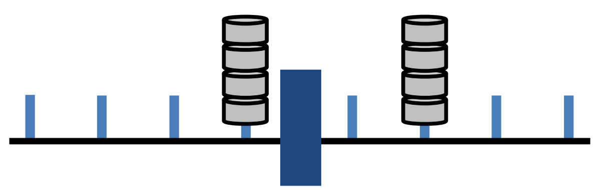
RESULT

7.



☐ left ☐ balance ☐ right

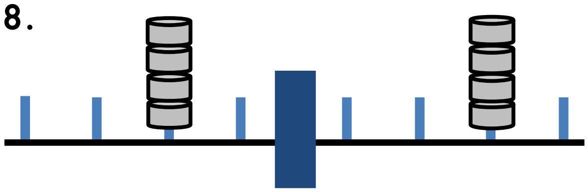
PREDICTION



☐ left ☐ balance ☐ right

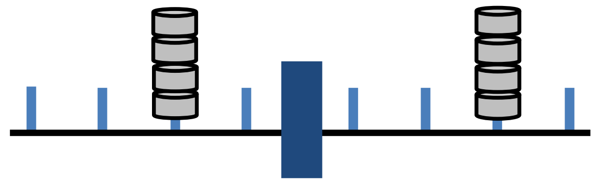
RESULT

8.



☐ left ☐ balance ☐ right

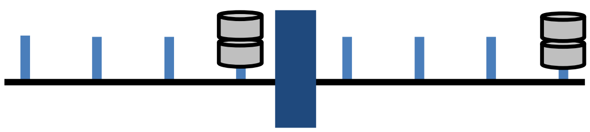
PREDICTION



☐ left ☐ balance ☐ right

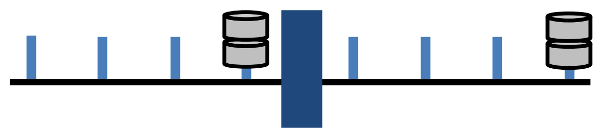
RESULT

9.



☐ left ☐ balance ☐ right

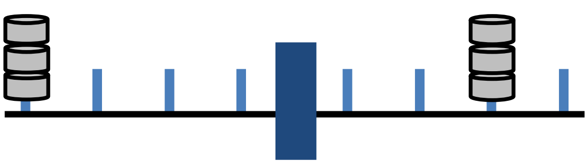
PREDICTION



☐ left ☐ balance ☐ right

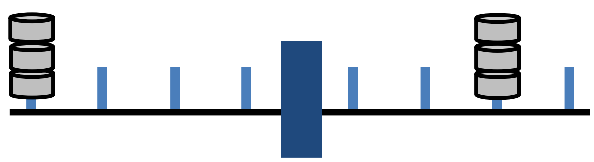
RESULT

10.



☐ left ☐ balance ☐ right

PREDICTION



☐ left ☐ balance ☐ right

RESULT

What is happening with the balance and weights? Are you noticing any patterns?

____

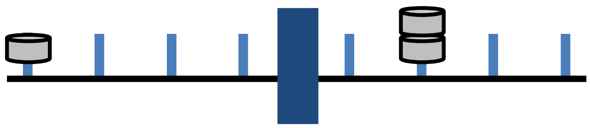
____

____

____

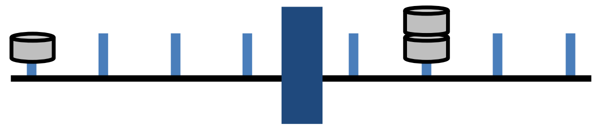
____

11.



☐ left ☐ balance ☐ right

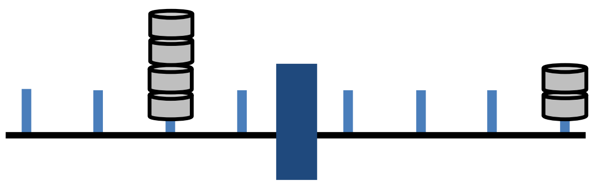
PREDICTION



☐ left ☐ balance ☐ right

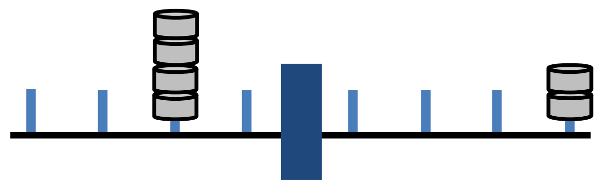
RESULT

12.



☐ left ☐ balance ☐ right

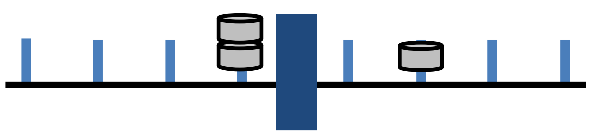
PREDICTION



☐ left ☐ balance ☐ right

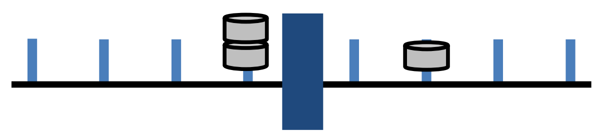
RESULT

13.



☐ left ☐ balance ☐ right

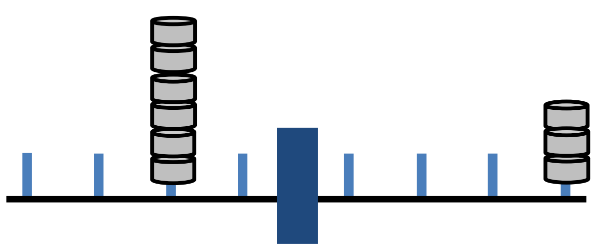
PREDICTION



☐ left ☐ balance ☐ right

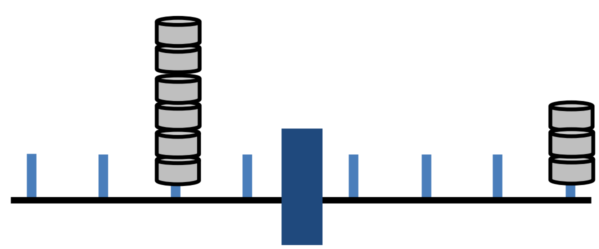
RESULT

14.



☐ left ☐ balance ☐ right

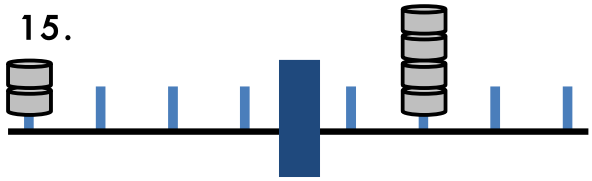
PREDICTION



☐ left ☐ balance ☐ right

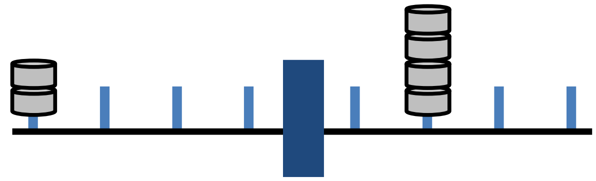
RESULT

15.



☐ left ☐ balance ☐ right

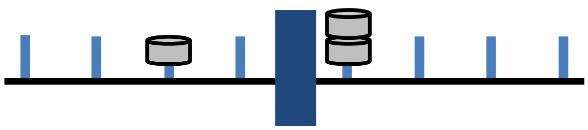
PREDICTION



☐ left ☐ balance ☐ right

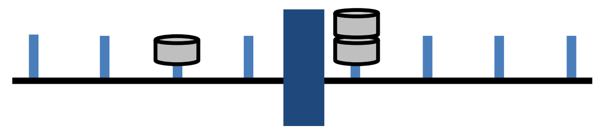
RESULT

16.



☐ left ☐ balance ☐ right

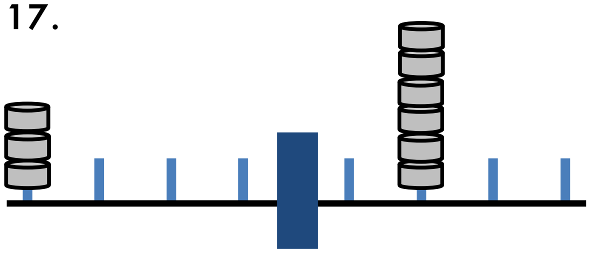
PREDICTION



☐ left ☐ balance ☐ right

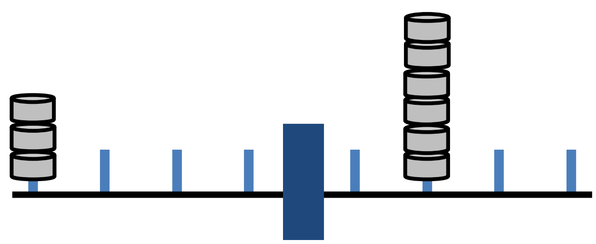
RESULT

17.



☐ left ☐ balance ☐ right

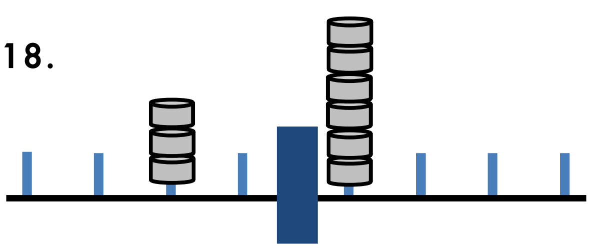
PREDICTION



☐ left ☐ balance ☐ right

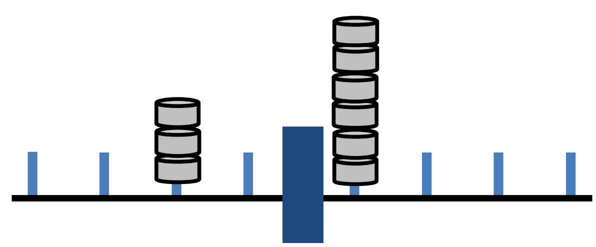
RESULT

18.



☐ left ☐ balance ☐ right

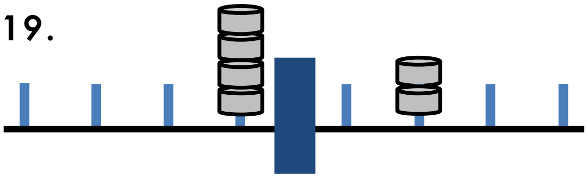
PREDICTION



☐ left ☐ balance ☐ right

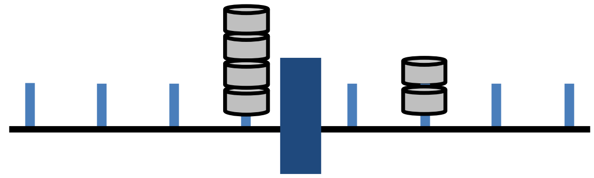
RESULT

19.



☐ left ☐ balance ☐ right

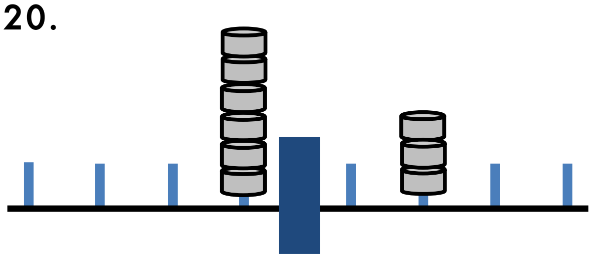
PREDICTION



☐ left ☐ balance ☐ right

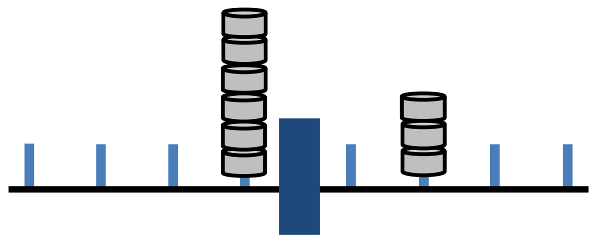
RESULT

20.



☐ left ☐ balance ☐ right

PREDICTION



☐ left ☐ balance ☐ right

RESULT

What is happening with the balance and weights? Are you noticing any patterns?

____

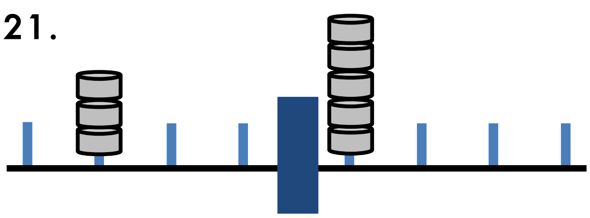
____

____

____

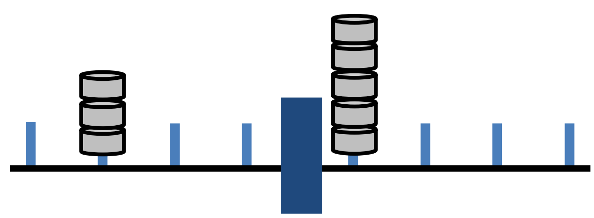
____

21.



☐ left ☐ balance ☐ right

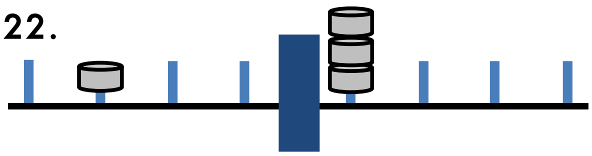
PREDICTION



☐ left ☐ balance ☐ right

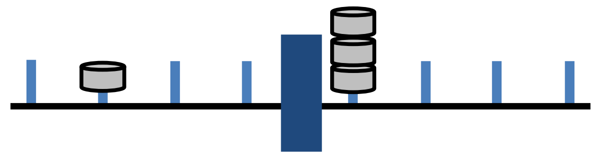
RESULT

22.



☐ left ☐ balance ☐ right

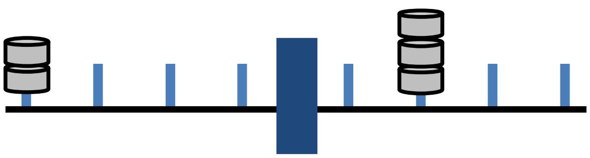
PREDICTION



☐ left ☐ balance ☐ right

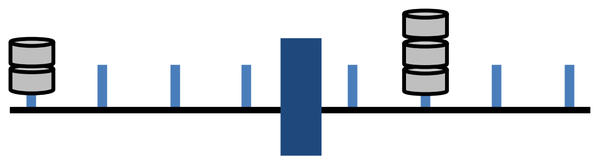
RESULT

23.



☐ left ☐ balance ☐ right

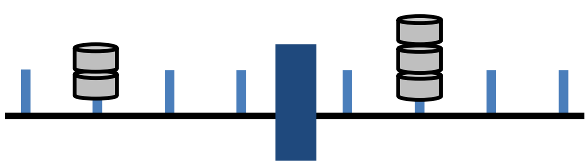
PREDICTION



☐ left ☐ balance ☐ right

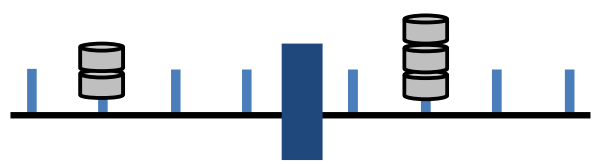
RESULT

24.



☐ left ☐ balance ☐ right

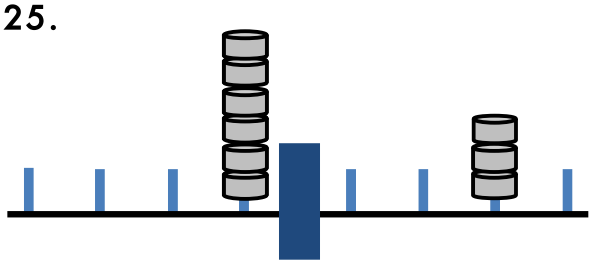
PREDICTION



☐ left ☐ balance ☐ right

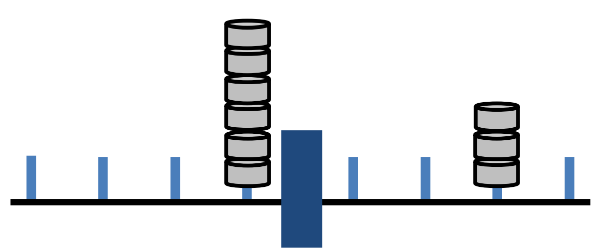
RESULT

25.



☐ left ☐ balance ☐ right

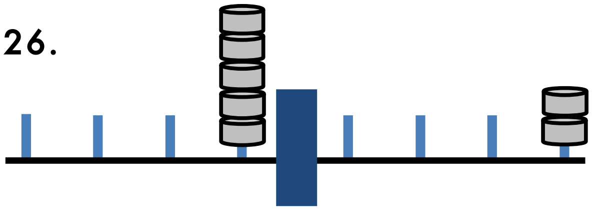
PREDICTION



☐ left ☐ balance ☐ right

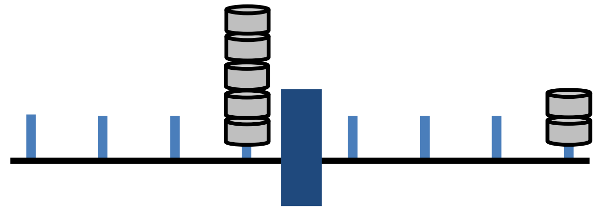
RESULT

26.



☐ left ☐ balance ☐ right

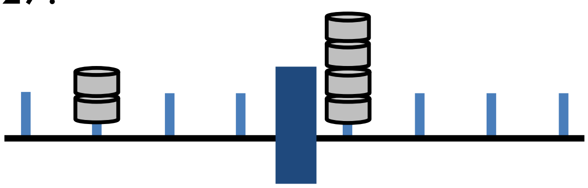
PREDICTION



☐ left ☐ balance ☐ right

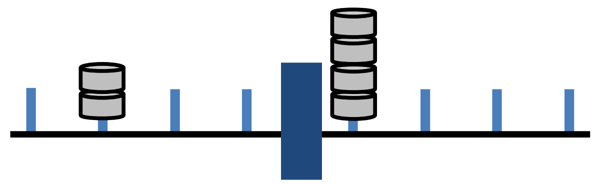
RESULT

27.



☐ left ☐ balance ☐ right

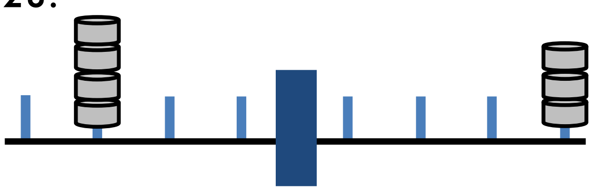
PREDICTION



☐ left ☐ balance ☐ right

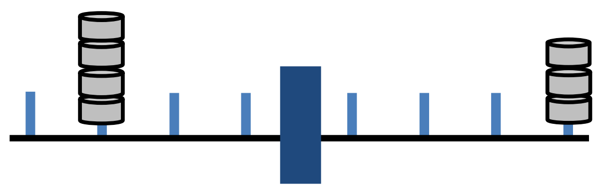
RESULT

28.



☐ left ☐ balance ☐ right

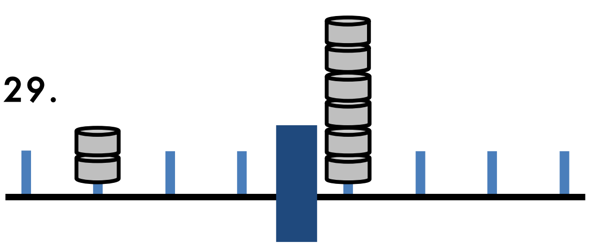
PREDICTION



☐ left ☐ balance ☐ right

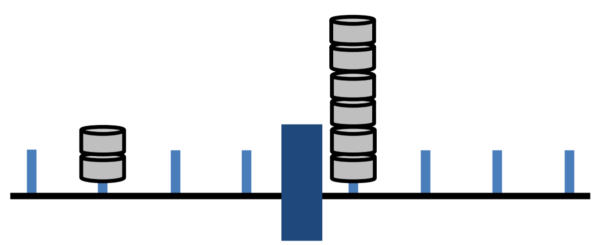
RESULT

29.



☐ left ☐ balance ☐ right

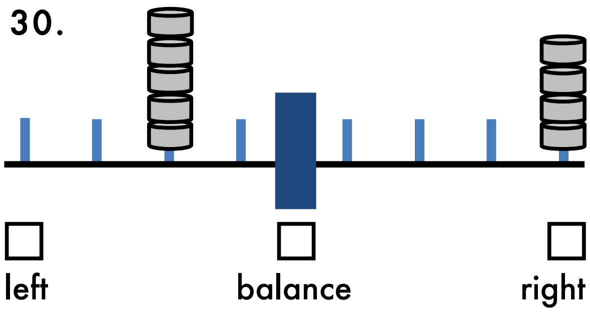
PREDICTION



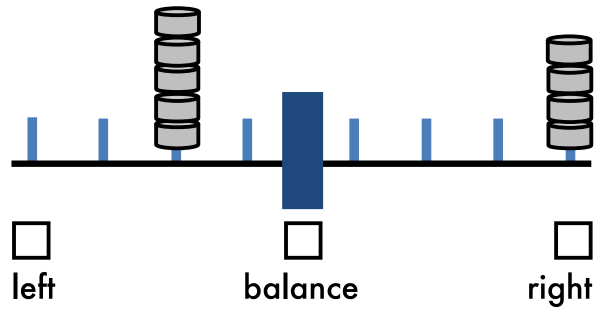
☐ left ☐ balance ☐ right

RESULT

30.



PREDICTION



RESULT

What is happening with the balance and weights? Are you noticing any patterns?

Writing Prompt: Treatment Condition

What makes the beam balance? What makes it tilt right or tilt left?

Procedural Fidelity Checklist: Treatment Condition

☐ <u>****TURN ON CAMERA BEFORE BEGINNING DIRECTIONS****</u>

Directions

- ☐ "Today we are going to do a balance experiment. You are going to make guesses about what you think the balance will do when different weights are placed on it. You'll also get to test each of your guesses by putting these weights on the balance (*Point to the balance and weights*)."
- ☐ "First, let's talk about the balance and each of its parts." Point to and name:
  - ☐ the *balance*
  - ☐ the *left arm* and *right arm*
  - ☐ the *fulcrum*
  - ☐ *pegs* 1, 2, 3, and 4 on the *left arm*
  - ☐ *pegs* 1, 2, 3, and 4 on the *right arm*
  - ☐ the *weights*

☐ Read the directions on page 1 of the trials.

☐ Complete Sample 1 with the student.

Balance trials

☐ "You will continue through the balance experiment problems just as we practiced in Sample 1. You will record your prediction about what the balance beam will do in each problem. Then you will test and record the result of what the balance beam actually does when you put weights on it as shown in the picture."

☐ "I am here to help if you need anything, but I want you to be the lead scientist today. That means I won't help with your predictions or with testing each prediction. You can be in charge of those tasks. I'll just be here in case you have questions or a problem with the balance beam."

☐ Remind student they will have as much time as they need to finish the balance trials.

☐ "At four points during the trials, you will stop to answer some questions in writing. When you get to those points, I will read the questions to you."

☐ Tell student he or she can flip the page and begin working.

☐ Monitor and prompt as necessary if student forgets to make a prediction before using the weights on the beam and/or recording results.

☐ Read aloud the short answer questions (What is happening with the balance and weights? Are you noticing any patterns?) after:

- ☐ problem 5
- ☐ problem 10
- ☐ problem 20
- ☐ problem 30

- ☐ Tell student they have 5 lines to respond for each short answer. Tell student to write neat enough for someone else to read his or her writing and not to worry about spelling.

Written response

- ☐ After student completes short answer question following problem 30 hand them the writing prompt sheet and say, “Scientists share their work and explain their findings to others. Because you were the lead scientist during our balance experiment, I have one more question for you to answer about balance. You can use your notes and findings from the balance experiment to help you write your answer. Please write your answer as if you were writing it for a boy or girl who doesn’t know much about balance. For this question, you have more room to write an extended response to explain your ideas.”
- ☐ Have student write first and last name at top of paper.
- ☐ “Read the question to yourself while I read it aloud. It says, ‘What makes the beam balance? What makes it tilt right or tilt left?’”
- ☐ “Please write your answer on the lines provided. You can have as much time as you need to write and more paper to write on if you need it. Make sure to write in complete sentences and neat enough for someone else to read your writing. Don’t worry about spelling; just spell the best you can.”
- ☐ Collect written response when student finishes writing.

Appendix D

Comparison Condition Materials

Balance Trials Sheet: Comparison Condition

*The balance trials sheet is identical to the balance trials sheet used with the writing to learn treatment condition, except for the short answer writing prompt, which is listed below:

What do you like about this experiment? What is your favorite part so far?

Writing Prompt: Comparison Condition

What do you like most about science class?

Procedural Fidelity Checklist: Comparison Condition

☐ **<u>**TURN ON CAMERA BEFORE BEGINNING DIRECTIONS**</u>**

<u>Directions</u>

- ☐ "Today we are going to do a balance experiment. You are going to make guesses about what you think the balance will do when different weights are placed on it. You'll also get to test each of your guesses by putting these weights on the balance (*Point to the balance and weights*)."
- ☐ "First, let's talk about the balance and each of its parts." Point to and name:
  - ☐ the *balance*
  - ☐ the *left arm* and *right arm*
  - ☐ the *fulcrum*
  - ☐ *pegs* 1, 2, 3, and 4 on the *left arm*
  - ☐ *pegs* 1, 2, 3, and 4 on the *right arm*
  - ☐ the *weights*

☐ Read the directions on page 1 of the trials.

☐ Complete Sample 1 with the student.

<u>Balance trials</u>

☐ "You will continue through the balance experiment problems just as we practiced in Sample 1. You will record your prediction about what the balance beam will do in each problem. Then you will test and record the result of what the balance beam actually does when you put weights on it as shown in the picture."

☐ "I am here to help if you need anything, but I want you to be the lead scientist today. That means I won't help with your predictions or with testing each prediction. You can be in charge of those tasks. I'll just be here in case you have questions or a problem with the balance beam."

☐ Remind student they will have as much time as they need to finish the balance trials.

☐ "At four points during the trials, you will stop to answer some questions in writing. When you get to those points, I will read the questions to you."

☐ Tell student he or she can flip the page and begin working.

☐ Monitor and prompt as necessary if student forgets to make a prediction before using the weights on the beam and/or recording results.

☐ Read aloud the short answer questions (What do you like about this experiment? What is your favorite part so far?) after:

- ☐ problem 5
- ☐ problem 10
- ☐ problem 20
- ☐ problem 30

- ☐ Tell student they have 5 lines to respond for each short answer. Tell student to write neat enough for someone else to read his or her writing and not to worry about spelling.

<u>Written response</u>

- ☐ After student completes short answer question following problem 30 hand them the writing prompt and paper and say, "I have one more question for you to answer about science. If you like, you can make notes first to help you write your answer. Please write your answer as if you were writing it for a boy or girl who doesn't know much about science class. For this question, you have more room to write an extended response to explain your ideas."
- ☐ Have student write first and last name at top of paper.
- ☐ "Read the question to yourself while I read it aloud. It says, 'What do you like most about science class?'"
- ☐ "Please write your answer on the lines provided. You can have as much time as you need to write and more paper to write on if you need it. Make sure to write in complete sentences and neat enough for someone else to read your writing. Don't worry about spelling; just spell the best you can."
- ☐ Collect written response when student finishes writing.

REFERENCES

Ackerman, J. M. (1993). The promise of writing to learn. *Written Communication,* **10**, 334-370.

Ayers, W. E. (1993). *A study of the effectiveness of expressive writing as a learning enhancement in middle school science.* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Full Text. (UMI No. 9408754).

Baker, S., Gersten, R., & Graham, S. (2003). Teaching expressive writing to students with learning disabilities: Research-based applications and examples. *Journal of Learning Disabilities,* **36***,* 109-123. http://dx.doi.org/10.1177/002221940303600204

Bangert-Drowns, R. L., Hurley, M. M., & Wilkerson, B. (2004). The effects of school-based writing-to-learn interventions on academic achievement: A meta-analysis. *Review of Educational Research,* **74**, 29-58. http://dx.doi.org/10.3102/00346543074001029

Bear, D., Invernizzi, M., Templeton, S., & Johnston, F. (2011). *Words their way: Word study for phonics, vocabulary, and spelling instruction, 5th edition.* Upper Saddle River, NJ: Prentice Hall.

Becker, R. J. (1996). *Responses to teacher-directed and student-directed writing prompts and performance on the IEA survey of achievement in literature* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Full Text. (UMI No. 9701474).

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Erlbaum.

Britton, J. (1970). *Language and learning.* London, England: Allen Lane.

Common Core State Standards. (2010). *English language arts standards*. Retrieved from http://www.corestandards.org/ELA-Literacy

Durst, R. K., & Newell, G. E. (1989). The uses of function: James Britton's category system and research on writing. *Review of Educational Research,* **59**, 375-394. http://dx.doi.org/ 10.3102/00346543059004375

Emig, J. (1977). Writing as a mode of learning. *College Composition and Communication,* **28***,* 122-128. http://dx.doi.org/10.2307/356095

Flower, L., & Hayes, J. R. (1980). The cognition of discovery: Defining a rhetorical problem. *College Composition and Communication,* **31**, 21-32. http://dx.doi.org/10.2307/356630

Galbraith, D. (1999). Writing as a knowledge-constituting process. In M. Torrance & D. Galbraith (Eds.), *Studies in writing: Vol. 4. Knowing what to write: Conceptual processes in text production* (pp. 139-160). Amsterdam: Amsterdam University Press.

Graham, S. (1990). The role of production factors in learning disabled students' compositions. *Journal of Educational Psychology*, **82**, 781–791. http://dx.doi.org/10.1037/0022-0663.82.4.781

Graham, S. (1999). Handwriting and spelling instruction for students with learning disabilities: A review. *Learning Disability Quarterly,* **22**, 78-98. http://dx.doi.org/10.2307/1511268

Graham, S. (2006). Writing. In P. Alexander & P. Winne (Eds.), *Handbook of Educational Psychology* (pp. 457-478). Mahwah, NJ: Erlbaum.

Graham, S., & Hebert, M. A. (2010). *Writing to read: Evidence for how writing can improve reading* (A Carnegie Time to Act Report). Washington, DC: Alliance for Excellent Education.

Graham, S., & Perin, D. (2007). *Writing next: Effective strategies to improve writing of adolescents in middle and high schools* (A Carnegie Corporation Report). Washington, DC: Alliance for Excellent Education.

Gunel, M., Hand, B., & McDermott, M. A. (2009). Writing for different audiences: Effects on high-school students' conceptual understanding of biology. *Learning and Instruction,* **19***,* 354-367. http://dx.doi.org/10.1016/j.learninstruc.2008.07.001

Gunel, M., Hand, B., & Prain, V. (2007). Writing for learning in science: A secondary analysis of six studies. *International Journal of Science and Mathematics Education,* **5**, 615-637. http://dx.doi.org/10.1007/s10763-007-9082-y

Hammill, D. D., & Larsen, S. C. (2009). Test of written language (Fourth edition) [Measurement instrument]. Austin, TX: Pro-ed.

Hand, B., Hohenshell, L. M., & Prain, V. (2004). Exploring students' responses to conceptual questions when engaged with planned writing experiences: A study with year ten science students. *Journal of Research in Science Teaching,* **41**, 186–210. http://dx.doi.org/10.1002/tea.10128

Hand, B., & Prain, V. (2002). Teachers implementing writing-to-learn strategies in junior secondary science: A case study. *Science Education,* **86**, 737-755. http://dx.doi.org/10.1002/sce.10016

Hand, B., Wallace, C. W., & Yang, E. (2004). Using a science writing heuristic to enhance learning outcomes from laboratory activities in seventh-grade science: Quantitative and qualitative aspects. *International Journal of Science Education,* **26**, 131-149. http://dx.doi.org/10.1080/0950069032000070252

Hardiman, P. T., Pollatsek, A., & Well, A. D. (1986). Learning to understand the balance beam. *Cognition and Instruction,* **3**, 63-86. http://dx.doi.org/10.1207/s1532690xci0301_3

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

Herrenkohl, L. R., Palincsar, A. M., DeWater, L. S., & Kawasaki, K. (1999). Developing scientific communities in classrooms: A sociocognitive approach. *The Journal of the Learning Sciences,* **8**, 451-493. http://dx.doi.org/10.1080/10508406.1999.9672076

Hohenshell, L. M., & Hand, B. (2006). Writing-to-learn strategies in secondary cell biology: A mixed method study. *International Journal of Science Education,* **28**, 261-289. http://dx.doi.org/10.1080/09500690500336965

Holliday, W. G., Yore, L. D., Alvermann, D. E. (1994). The reading-science learning-writing connection: Breakthroughs, barriers, and promises. *Journal of Research in Science Teaching,* **31**, 877-893. http://dx.doi.org/10.1002/tea.3660310905

Horton, P. B., Frank, R. H., & Walton, R. W. (1985). The effect of writing assignments on achievement in college general chemistry. *Journal of Research in Science Teaching,* **22***,* 535-541. http://dx.doi.org/10.1002/tea.3660220607

Hunt, K. W. (1965). *Grammatical structures written at three grade levels* (NCTE Research Report No. 3). Champaign, IL: National Council of Teachers of English.

Inhelder, B., & Piaget, I. (1958). Equilibrium in the balance. *The growth of logical thinking from childhood to adolescence* (pp. 164-181). New York, NY: Basic Books, Inc.

Keys, C. W. (1999). Revitalizing instruction in scientific genres: Connecting knowledge production with writing to learn in science. *Science Education,* **83**, 115-130. http://dx.doi.org/10.1002/(SICI)1098-237X(199903)83:2<115::AID-SCE2>3.0.CO;2-Q

Kieft, M., Rijlaarsdam, G., & van den Bergh, H. (2008). An aptitude-treatment interaction approach to writing-to-learn. *Learning and Instruction,* **18**, 379-390. http://dx.doi.org/10.1016/j.learninstruc.2007.07.004

Kirkpatrick, L. D., & Pittendrigh, A. S. (1984). A writing teacher in the physics classroom. *The Physics Teacher,* **22***,* 159-164. http://dx.doi.org/10.1119/1.2341502

Klein, P. D. (1999). Reopening inquiry into cognitive processes in writing-to-learn. *Educational Psychology Review,* **11**, 203-270. http://dx.doi.org/10.1023/A:1021913217147

Klein, P. D. (2000). Elementary students' strategies for writing-to-learn in science. *Cognition and Instruction,* **18**, 317-348. http://dx.doi.org/10.1207/S1532690XCI1803_2

Langer, J. A., & Applebee, A. N. (1987). *How writing shapes thinking.* Urbana, IL: National Council of Teachers of English.

Lipsey, M.W., & Wilson, D.B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Malachowski, M. R. (1988). The use of journals to enhance chemical understanding in a liberal arts chemistry class. *Journal of Chemical Education,* **65**, 439-440. http://dx.doi.org/10.1021 /ed065p439

Mason, L., & Boscolo, P. (2000). Writing and conceptual change: What changes? *Instructional Science,* **28***,* 199-226. http://dx.doi.org/10.1023/A:1003854216687

National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: The National Academies Press.

Nieswandt, M. (1997, March). Improving learning in chemistry classes through original writing about chemical facts. Paper presented at the American Educational Research Association, Chicago, IL.

Reynolds, F. E., & Pickett, I. (1989). Read! Think! Write! The reading response journal in the biology classroom. *American Biology Teacher,* **51**, 435-437. http://dx.doi.org/10.2307/4448974

Rivard, L. P. (1994). A review of writing to learn in science: Implications for practice and research. *Journal of Research in Science Teaching,* **31**, 969-983. http://dx.doi.org/10.1002/ tea.3660310910

Rivard, L. P. (1996). The effect of talk and writing, alone and combined, on learning in science: An exploratory study (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Full Text. (UMI No. NN13481).

Rivard, L. P. (2004). Are language-based activities in science effective for all students, including low achievers? *Science Education,* **88**, 420-442. http://dx.doi.org/10.1002/sce.10114

Rivard, L. P., & Straw, S. B. (2000). The effect of talk and writing on learning science: An exploratory study. *Science Education,* **84**, 566-593. http://dx.doi.org/10.1002/1098237X (200009)84:5<566::AID-SCE2>3.0.CO;2-U

Sharp, J. E. (1987). *Expressive summary writing to learn college biology* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Full Text. (UMI No. 8714454).

Siegler, R. S. (1985). Encoding and the development of problem-solving. In S. F. Chipman, J. W. Segal, & R. Glaser (Eds.), *Thinking and learning skills: Vol. 2. Research and open questions* (pp. 161-185). Hillsdale, NJ: Erlbaum.

Thall, E., & Bays, G. (1989). Utilizing ungraded writing in the chemistry classroom. *Journal of Chemical Education,* **66***,* 662-663. http://dx.doi.org/10.1021/ed066p662

Tucknott, J. M., & Yore, L. D. (1999, March). *The effects of writing activities on grade 4 children's understanding of simple machines, inventions, and inventors.* Paper presented at the annual meeting of the National Association for Research in Science Teaching, Boston, MA.

VanOrden, N. (1987). Critical-thinking writing assignments in general chemistry. *Journal of Chemical Education,* **64***,* 506-507. http://dx.doi.org/10.1021/ed064p506

Wallace, C. S., Hand, B., & Prain, V. (2004). *Writing and learning in the science classroom*. Dordrecht: Kluwer Academic Publishers.

Willey, L. H. (1988). *The effects of selected writing to learn approaches on high school students' attitudes and achievement* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Full Text. (UMI No. 8905957).