THE EFFECTS OF FEEDBACK DURING EXPLORATORY MATHEMATICS PROBLEM SOLVING: PRIOR KNOWLEDGE MATTERS

By

Emily R. Fyfe

Thesis

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

MASTER OF SCIENCE

in

Psychology

May, 2012

Nashville, Tennessee

Approved:

Professor Bethany Rittle-Johnson

Professor Bruce McCandliss

ACKNOWLEDGEMENTS

This research was supported by National Science Foundation grant DRL-0746565 to Bethany Rittle-Johnson and Institute of Education Sciences, U.S. Department of Education, training grant R305B080025. A special thanks to Laura McLean, Abbey Loehr, Maryphyllis Crean, Lucy Rice, Rachel Ross, and Polly Colgan for their assistance with data collection and coding as well as the staff, teachers, and children at Westmeade Elementary School, Jones Paideia Elementary Magnet School, and St. Pius X Classical Academy for their participation in this research project. Most importantly, I would like to express my heartfelt gratitude to my advisor Bethany Rittle-Johnson for her ever-present guidance and support, to my family, especially my parents Mike and Denise and my husband Matthew, for their unwavering confidence in my ability to accomplish this work, and to my friends for their loving companionship through it all.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Contemporary learning theorists often endorse guided discovery learning, as opposed to discovery or instruction alone, as the best method to facilitate understanding (e.g., Mayer, 2004; Schwartz et al., 2011; Wise & O'Neill, 2009). Providing an exploratory activity with subsequent instruction is one form of guided discovery that has been shown to aid learning (e.g., Schwartz & Bransford, 1998). However, the level of guidance provided during the exploratory activity has largely gone unstudied. I propose feedback as one form of guidance that could potentially boost the efficacy of exploration by guiding the learner's search for relevant information. In two experiments, I examined how and for whom feedback might enhance learning during exploration prior to direct instruction. I investigated these questions in the context of children exploring mathematical equivalence, a fundamental concept in arithmetic and algebra.

Guided Discovery Learning

An emerging consensus is that people learn best through some form of *guided discovery*, which combines exploration and instruction (e.g., Aflieri et al., 2011; Hmelo-Silver, Duncan, & Chinn, 2007; Lorch, et al., 2010; Mayer, 2004). As Mayer (2004) notes, "students need enough freedom to become cognitively active in the process of sense making, and…enough guidance so that their cognitive activity results in the construction of useful knowledge" (p. 16). There is currently not a precise definition of guided discovery, largely because the term captures such a broad range of activities including problem-based learning (Barrows & Tamblyn, 1980), inquiry learning (Rutherford, 1964), and constructivist learning (Steffe & Gale, 1995). Here, I adopt the general framework outlined by Alfieri and colleagues (2011) and define guided discovery as exploratory learning tasks that are supplemented with some form of instructional guidance. Learning tasks are exploratory if learners have not received instruction on how to complete them and instructional guidance encompasses a variety of tools, from in-depth instruction manuals to minimal feedback or coaching. Alfieri et al.'s (2011) recent meta-analysis revealed the superiority of guided discovery over both direct instruction and unguided discovery learning.

Providing exploratory activities prior to direct instruction is one form of guided discovery that has been recommended by researchers in education and psychology alike (e.g., Hiebert & Grouws, 2007; Schwartz, Lindgren, & Lewis, 2009), and it is the form that I focus on in this paper. For example, several mathematics education researchers promote the belief that "each person must struggle with a situation or problem first in order to make sense of the information he or she hears later" (Stigler & Hiebert, 1998, p. 3). Similarly, Schwartz and colleagues suggest that exploratory activities facilitate the development of differentiated knowledge of the target domain, which prepares people to learn more deeply from future instruction than would be possible otherwise (e.g., Schwartz & Bransford, 1998; Schwartz, Lindgren, & Lewis, 2009).

There is a growing body of evidence to support the claim that exploration prior to instruction is beneficial (e.g., DeCaro & Rittle-Johnson, 2011; Schwartz & Bransford, 1998; Schwartz & Martin, 2004; Schwartz, Chase, Oppezzo, & Chin, 2011). For example, college students who explored novel examples learned more from a subsequent lecture on psychology principles than students who merely summarized a relevant text prior to the lecture (Schwartz & Bransford, 1998). Further, the timing of instruction matters. For example, elementary school children learned new math concepts better if they solved unfamiliar problems *before* receiving instruction, rather than vice versa (DeCaro & Rittle-Johnson, 2011). Similarly, middle school students who explored a set of novel density problems prior to receiving instruction exhibited better transfer than students who heard the lecture first and practiced the problems afterward (Schwartz et al., 2011). However, further research is needed to optimize this form of guided discovery. For example, the level of guidance provided during the exploratory phase has largely gone unstudied. In this study, I examine the effects of providing guidance versus not providing guidance during exploration prior to instruction.

Using Feedback as One Form of Guidance

Feedback is touted as one form of guidance that may be particularly effective. *Feedback* is any information about performance or understanding that the learner can use to confirm, reject, or modify prior knowledge (Mory, 2004). Based on their meta-analysis, Alfieri and colleagues (2011) specifically recommend "providing timely feedback" as an optimal form of guidance (p. 13). Additionally, Mayer's review (2004) indicates that guided discovery methods that provide feedback or scaffolding during problem solving enable deeper learning than unguided discovery methods. For example, kindergarteners generally struggle with conservation tasks; however, Brainerd (1972) showed that providing outcome feedback to children solving novel conservation tasks improved their explanations on subsequent problems. Although not directly related to exploration provided *prior* to direct instruction, these reviews do suggest that guidance *during exploratory problem solving* can be beneficial.

In addition to these endorsements, there are several reasons to suggest feedback is beneficial. First, decades of research have demonstrated powerful effects of feedback for student achievement in the classroom (see Hattie & Timperley, 2007). Indeed, one meta-analysis comparing feedback interventions to control groups with no feedback included 470 effects and revealed an average positive effect size of .38 for feedback on student performance measures (e.g., reading errors, arithmetic computations; Kluger & DeNisi, 1996). The effects spanned a variety of feedback types, task types, and means of presentation.

Second, past research indicates that feedback's primary function is to identify errors and encourage the adoption of correct alternatives (e.g., Anderson, Kulhavy, & Andre, 1972; Birenbaum & Tatsuoka, 1987; Kulhavy, 1977). In these studies, learners who receive feedback have a higher chance of correcting their errors than learners who do not. For example, Phye and Bender (1989) examined the role of feedback on a memory task and found that when feedback was not available, perseveration (i.e., making the same error multiple times) was the most frequent error type—a pattern not demonstrated in the feedback conditions. Also, though feedback may not be necessary for new strategy generation, feedback has been shown to facilitate strategy generation relative to no-feedback (Alibali, 1999). Together, these studies suggest feedback helps learners reject erroneous ideas and search for more plausible alternatives.

Given these positive effects of feedback, it seems likely that it would improve the efficacy of exploration prior to direct instruction. However, all reviews of feedback research note the wide variability of effects on learning (e.g., Hattie & Timperley, 2007; Kluger & DeNisi, 1998; Mory, 2004). In fact, even though Kluger and DeNisi (1996) found an average positive effect size for feedback, over one third of the effects were negative, indicating that feedback in those cases decreased performance. The authors noted the need for further research to determine the conditions under which feedback is effective. In this study, I explored the possibility that feedback may enhance the benefits of exploration for *only* a certain subset of learners.

The Role of Prior Knowledge

The possibility that feedback during exploration is only beneficial for some learners is consistent with past work on aptitude by treatment interactions. These occur when environments that have positive effects for one kind of person, have neutral or even negative effects for another (Cronbach & Snow, 1977; Snow, 1978; Snow & Lohman, 1984). Importantly, these interactions often occur when instructional approaches differ in the amount of *guidance* provided. For example, Snow and Swanson (1992) suggest tutors "should provide more scaffolding for less able learners and less scaffolding for more able learners" (p. 610). Feedback represents one form of scaffolding whose benefits may depend on certain learner characteristics.

Particularly relevant to the current study is the expertise reversal effect, a specific example of an aptitude by treatment interaction. It "occurs when an instructional procedure that is relatively effective for novices becomes ineffective for more knowledgeable learners" (Kalyuga & Sweller, 2004, p. 558). For example, novices learn more from studying worked examples than from solving problems unaided. But, as knowledge increases, independent problem solving becomes the superior learning activity (e.g., Renkl & Atkinson, 2003). To be clear, the expertise reversal effect involves interactions between instructional methods and learners' knowledge, and need not involve experts (e.g., Ayres, 2006; Homer & Plaas, 2009; Kalyuga et al., 2003). Importantly, novices seem to benefit from more external guidance, while learners with more knowledge seem to benefit from less external guidance (Kirschner, Ayres, & Chandler, 2011). Indeed, Kalyuga and colleagues (2003) posit, "instructional guidance, which may be essential for novices, may have negative consequences for more experienced learners" (p. 24).

One explanation for the expertise reversal effect stems from cognitive load theory (Sweller, van Merrienboer, & Paas, 1998). For novices who lack schemas in the domain, novel tasks can easily overload working memory; thus, they often need external guidance. In contrast, higher-knowledge learners have relevant schemas that help them complete tasks without overloading working memory; thus, they often do not need external guidance. In the present study, providing feedback, a source of guidance, may help learners with low prior knowledge; but higher-knowledge learners may not need feedback and may even perform better without it.

Current Study

I examined the effects of feedback during exploration prior to direct instruction for children with varying levels of prior knowledge. I focused on mathematical equivalence problems (problems with operations on both sides of the equal sign, such as 3 + 4 + 5 = 3 + __) because children are rarely exposed to them and often solve them incorrectly. Mathematical equivalence problems are relatively novel for second- and third-grade children as they are not typically included in elementary mathematics curricula (Perry, 1988; Seo & Ginsburg, 2003). Indeed, a recent analysis revealed that, of all instances of the equal sign in a textbook series for grades 1 through 6, equations with operations on both sides of the equal sign accounted for just 4% of instances (Rittle-Johnson, Matthews, Taylor, & McEldoon, 2011). Further, decades of research have shown that elementary school children exhibit poor performance on mathematical equivalence problems (e.g., McNeil, 2008; Rittle-Johnson & Alibali, 1999), which often stem from misinterpretations of the equal sign as an operator symbol meaning “get the answer,” as opposed to a relational symbol meaning “the same amount” (Baroody & Ginsburg, 1983; Kieran, 1981; McNeil & Alibali, 2005). Thus, mathematical equivalence problems are unfamiliar and difficult for elementary school children, providing an apt domain to investigate exploratory problem solving.

In the context of exploratory mathematics problem solving, two types of feedback seem particularly relevant: *outcome feedback* provides a judgment about the accuracy of the learner’s response, whereas *strategy feedback* provides a judgment about how the learner obtained that response. Outcome feedback has been studied extensively and is generally related to positive learning outcomes (e.g., Kluger & DeNisi, 1996). In contrast, few empirical studies have examined the effects of strategy feedback (Ahmad, 1988; Luwel, et al., 2011). The limited evidence suggests that strategy feedback can improve strategy selection relative to outcome feedback; however, more research is needed to examine its benefits across tasks and outcome measures. My primary goal was to compare the effects of providing any feedback to providing no feedback during exploration prior to direct instruction. However, I included two types of feedback to explore whether different types differentially impact learning and also to bolster the empirical evaluation of strategy feedback as a learning tool.

In the current study, children received a tutoring session that included exploratory problem solving followed by brief instruction. During problem solving, children received (a) no-feedback, (b) outcome-feedback, or (c) strategy-feedback after solving novel mathematical equivalence problems. After the tutoring session, children completed a posttest (immediately and after a 2-week delay) that assessed conceptual and procedural knowledge of mathematical equivalence (Rittle-Johnson et al., 2011). *Conceptual knowledge* is an understanding of the principles governing a domain and *procedural knowledge* is the ability to execute action sequences to correctly solve problems (e.g., Rittle-Johnson & Alibali, 1999). We also incorporated microgenetic methods, such as strategy reports (Siegler & Crowley, 1991) and cognitive load ratings (Hart & Staveland, 1988) to explore how feedback influenced learning.

The primary hypothesis was that children who received feedback would exhibit better procedural knowledge of mathematical equivalence than children who did not. However, I expected this effect to be larger for children with lower prior knowledge and to disappear or reverse for children with higher prior knowledge based on research on the expertise reversal effect (Kalyuga et al., 2003). Differences were predicted in procedural knowledge because the feedback was directed at children's problem solving. Based on the promise of strategy feedback in two previous studies (e.g., Luwel, et al., 2011), I tentatively predicted that strategy feedback would lead to higher performance than outcome feedback. I also explored why these differences in procedural knowledge might occur. Feedback was expected to influence the variability in children's strategy use, with feedback promoting the use of more diverse strategies relative to no feedback (Alibali, 1999). Finally, feedback was also expected to decrease cognitive load relative to no feedback for low-knowledge children, but not for higher-knowledge children, based on explanations for the expertise reversal effect (Kalyuga et al., 2003). The results from this study will help us understand not only if feedback is beneficial during exploratory problem solving prior to direct instruction, but also how and for whom it works. Two experiments were conducted with the same basic design to evaluate the replicability of the findings.

CHAPTER II

EXPERIMENT ONE

Method

Participants

Consent was obtained from 115 second- and third-grade children at a public school in middle Tennessee. Of those children 93 met criteria for participation because they scored below 80% on both a conceptual and procedural knowledge measure at pretest. A liberal inclusion criterion was used to examine learning outcomes for children who varied in terms of prior knowledge, but still had room for growth. Data from six additional children were excluded: one for failing to complete the intervention, one for failing to follow directions, and four for missing the retention test. The final sample consisted of 87 children (*M* age = 8 yrs, 6 mo; 52 girls; 35 boys; 44% Caucasian; 41% African American, 6% Hispanic, 9% Other). Approximately 47% received free or reduced price lunch.

Design

The experiment had a pretest – intervention – posttest design followed by a two-week retention test. For the brief tutoring intervention, children were randomly assigned to one of three conditions: strategy-feedback ($n = 25$), outcome-feedback ($n = 31$), or no-feedback ($n = 31$).

Procedure

Children completed a written pretest in their classrooms in one 30-minute session. Within one week, those who met the inclusion criteria completed a one-on-one tutoring intervention and immediate posttest in a single session lasting approximately 45 minutes. This session was conducted in a quiet room at the school with one of two female experimenters. Approximately two weeks after the intervention session (*M* = 14.0 days, *SD* = 2.7), children completed the written retention test in small-group sessions in their classrooms.

The tutoring intervention began with exploratory problem solving. Children were asked to solve 12 mathematical equivalence problems presented one at a time on a computer screen using E-Prime 2.0 (Schneider, Eschman, & Zuccolotto, 2002). Specifically, they were asked to figure out the number that went in the box to make the number sentence true. The problems increased in difficulty with two exceptions. The first four problems consisted of three- and four- addend problems (e.g., $10 = 3 + \square$, $3 + 7 = 3 + \square$, $3 + 7 = \square + 6$). These were followed by six five-addend problems with a repeated addend on either side of the equal sign (e.g., $5 + 3 + 9 = 5 + \square$, $9 + 7 + 6 = \square + 6$). Two additional problems (the seventh and tenth) were simple three-addend problems (i.e., $9 = 6 + \square$, $7 = 6 + \square$). These were included in the block with the six more difficult problems to ensure children in the two feedback conditions received some positive feedback and to ensure all children were paying attention. After each problem, children reported how they solved the problem and received different kinds of feedback based on their condition.

In the *strategy-feedback condition*, children received feedback on how they solved each problem (e.g., “Good job! That is one correct way to solve that problem.”/ “Good try, but that is not a correct way to solve the problem.”). The strategy feedback was based solely on the correctness of the child’s verbal strategy report and did not depend on the correctness of the numerical answer (though these differed on only 3% of trials). For example, if a child reported using a correct strategy but obtained an incorrect answer (e.g., due to an arithmetic error), we provided positive strategy feedback. In the *outcome-feedback condition*, children received feedback on their answer to the problem. This included a judgment about the correctness of the answer as well as the correct response (e.g., “Good job! You got the right answer, *X* is the correct answer.” / “Good try, but you did not get the right answer, *X* is the correct answer.”). The outcome feedback was based solely on the correctness of the child’s numerical answer and did not depend on the correctness of the strategy used (though these differed on only 4% of trials). We provided the correct answer because past work suggests this enhances the effects of outcome feedback (Kluger & DeNisi, 1996). For both conditions, feedback was presented verbally by the experimenter and visually on the computer screen. Finally, in the *no-feedback condition*, children did not receive any feedback after solving a problem and were simply told to go to the next one.

After the exploratory problem solving all children received brief conceptual instruction on the relational function of the equal sign, adapted from past research (DeCaro & Rittle-Johnson, 2011; Matthews & Rittle-Johnson, 2009). The experimenter provided a definition of the equal sign, using a number sentence as an example. Specifically, $3 + 4 = 3 + 4$ was displayed on the screen while the experimenter identified the two sides of the problem, defined the equal sign as meaning “the same amount as,” and explained how the left and right side of the problem were equal. The meaning of the equal sign was reiterated with four additional number sentences (e.g., $4 + 4 = 3 + 5$). Children were asked to answer simple questions and to identify the two sides of the number sentences to ensure they were attending to instruction. No solution procedures were discussed and children were not asked to solve any mathematical equivalence problems during the instruction.

Between the exploratory problem solving and instruction, children completed a brief form of the mathematical equivalence assessment (midtest) to gauge the immediate effects of exploration prior to instruction. Additionally, children rated their subjective cognitive load at this time and completed several other measures not relevant to the current results.

Assessments and Coding

The mathematical equivalence assessment, adapted from past work (Rittle-Johnson et al., 2011), was administered at pretest, posttest, and retention test. Two parallel forms of the assessment were used; Form 1 at pretest and Form 2 at posttest and retention test. The assessment included procedural and conceptual knowledge scales (see Table 1 for example items). The procedural knowledge scale assessed children’s use of correct strategies to solve eight mathematical equivalence problems (results are identical if based on correct numerical answers, as correct strategies and correct answers differed on less than 1% of all trials). The conceptual knowledge scale (8 items) assessed two key concepts: (a) the meaning of the equal sign, and (b) the structure of equations. A brief version of Form 1 (5 items—3 conceptual and 2 procedural) was used as a midtest during the intervention session. The more difficult items were included on the midtest, as they were similar in difficulty to the majority of problems presented during the intervention.

Table 1: Example Items from the Procedural and Conceptual Knowledge Scales on the Mathematical Equivalence Assessment

Item Type	Task	Scoring Criteria
Procedural		(α = .83 in Exp. 1) (α = .85 in Exp. 2)
Familiar problems	Solve 1 problem with operation on right side (8 = 6 + □)	Use correct strategy (if strategy is ambiguous, response must be within 1 of correct answer)
	Solve 3 problems with operations on both sides, blank on right (e.g., 3 + 4 = □ + 5)	Same as above
Transfer problems	Solve 3 problems with operations on both sides, blank on left or includes subtraction (e.g., □ + 6 = 8 + 6 + 5)	Same as above
	Solve 1 equation with an unknown variable ($y + 4 = 8 + 2$)	Same as above
Conceptual		(α = .64 in Exp. 1) (α = .71 in Exp. 2)
Meaning of equal sign	Define equal sign	1 point for relational definition (e.g., the same amount)
	Rate definitions of equal sign as good, not good, or don't know	1 point for rating "two amounts are the same" as a good definition
Structure of equations	Reproduce 3 equivalence problems from memory	1 point for correctly reconstructing all 3 problems
	Indicate whether 5 equations such as 3 = 3 are true or false	1 point for correctly recognizing 4 or more equations as true or false

Note. Cronbach alphas are for posttest. Alphas were somewhat lower at pretest, largely due to floor effects on some items.

We coded the conceptual knowledge items requiring a written explanation (see Table 1 for coding criteria). To establish inter-rater reliability, a second rater coded the written responses of 20% of the children. Inter-rater agreement was high (exact agreement = 95 – 97%; kappas = 90 – 95%). Kappas calculate inter-rater reliability adjusting for chance (Cohen, 1960). Values above 81% are considered excellent (Landis & Koch, 1977). We also coded children's problem solving strategies on the procedural knowledge assessment and on the intervention problems (Table 2). On the assessment, strategies were inferred from children's written work. For the intervention, strategies were based on children's verbal reports. A second rater coded the strategies of 20% of the children. Inter-rater agreement was high (exact = 88%; kappa = 86%). Although specific strategies were coded on the procedural knowledge assessment, scores were based solely on whether a strategy was correct or incorrect. So we also examined inter-rater agreement on correct strategy vs. incorrect strategy use (exact = 99%; kappa = 98%).

Table 2: Strategies Used to Solve Mathematical Equivalence Problems

Strategy	Sample explanation (4 + 5 + 8 = __ + 8)
Correct Strategies	
Equalize	I added 4, 5, and 8 and got 17, and 9 plus 8 is also 17.
Add-Subtract	I added 4, 5, and 8 and got 17, and 17 minus 8 is 9.
Grouping	I took out the 8's and I added 4 plus 5.
Ambiguous	8 divided by 8 is 0 and 4 plus 5 is 9.
Incorrect Strategies	
Add All	I added the 4, 5, 8 and 8.
Add-to-Equal	4 plus 5 is 9, and 9 plus 8 is 17.
Add-Two	I added the 5 and the 8.
Carry	I saw a 4 here, so I wrote a 4 in the blank.
Ambiguous	I used 8 plus 8 and then 5.

Note. Italicized strategies were demonstrated in the strategy evaluation task in Exp. 2.

Two items measuring *subjective cognitive load* were administered immediately after the exploratory problem solving. Two modified items from the NASA Task Load Index were used (Hart & Staveland, 1988): Effort ("I had to work hard to solve those problems.") and Frustration ("I was stressed and irritated when I solved those problems."). Children indicated the extent of their agreement with each item on a 5-point scale. Responses to the two items were averaged to form a single score for each child.

Data Analysis

A contrast analysis of variance model was used, as recommended by West, Aiken and Krull (1996). In this model, contrast codes represent a categorical variable with more than two levels, which in this case was condition. The condition variable had three groups (no-feedback, outcome-feedback, and strategy-feedback), so two coded variables were created. The primary goal was to determine whether no guidance or any guidance during exploration prior to instruction was more optimal. Thus, the first variable (feedback) compared no feedback to the two feedback conditions combined. I also explored whether the type of guidance mattered. Thus, the second variable (feedback type) compared outcome feedback to strategy feedback. We also included three covariates (children's age and procedural and conceptual knowledge pretest scores). Finally, to evaluate whether condition effects depended on prior knowledge two interaction terms were included: feedback by prior knowledge and feedback type by prior knowledge. The procedural knowledge pretest measure was used as the prior knowledge measure as it is the most relevant type of prior knowledge for learning during exploratory problem solving. Exploratory analyses indicated that conceptual knowledge pretest scores did not interact with condition. Thus, the statistical model was a contrast-based ANCOVA with two contrast-coded between-subject variables (feedback, feedback type), three covariates, and two prior knowledge interaction terms. All effects reported with this model are similar if the two contrast-coded variables are replaced with a single categorical "condition" variable. However, the contrast-based model allows me to test my specific predictions of interest using fewer statistical tests.

The assumption of homogeneity of variance for this model was largely supported. For the procedural knowledge variables, Levene's tests indicated equal variances on the posttest and retention test, F's < 1, though not at midtest, $F(2, 84) = 6.30$, $p = .003$ (likely due to limited number of midtest items). With all three time points in the same (repeated-measures) model, covariance matrices were also homogeneous, Box's M = 8.01, $F(12, 30952) = 0.63$, $p = .82$. For the conceptual knowledge variables, Levene's tests indicated equal variances at midtest, posttest, and retention test, F's < 2. With all three time points in the same (repeated-measures) model, covariance matrices were also homogeneous, Box's M = 7.78, $F(12, 30952) = 0.61$, $p = .83$. Overall, ANOVA models were appropriate for analyzing the data.

Results

On the pretest, children answered few procedural (M = 29%, SD = 22%) and conceptual (M = 27%, SD = 19%) items correctly. However, the scores ranged from 0 – 75% on both the procedural and conceptual knowledge scales. Importantly, there were no significant differences between conditions on either scale at pretest, Fs < 1.

To evaluate children's performance on the midtest, posttest and retention test I conducted repeated measures ANCOVAs with feedback (feedback vs. none) and feedback type (outcome vs. strategy) as between-subject variables and time (midtest, posttest, retention test) as the within-subject variable. The three covariates as well as the two interaction terms were included. The statistical conclusions remain largely unchanged when the midtest is removed from the model. Children's procedural and conceptual knowledge were examined as separate outcomes. Feedback was expected to lead to higher procedural knowledge scores than no feedback, but only for children with lower prior knowledge. For children with higher prior knowledge, I expected either no effect of feedback or a reversal such that feedback would lead to lower scores than no feedback. The effect of feedback on children's conceptual knowledge was examined, though there were no prior predictions. I also explored if strategy feedback led to higher scores than outcome feedback, and whether this effect interacted with prior knowledge.

Procedural Knowledge

Children's procedural knowledge increased from midtest to posttest and remained similar two weeks later (see Table 3), $F(2, 158) = 13.64$, $p < .001$, $\eta_p^2 = .15$. There were no main effects of feedback or feedback type, Fs < 1. However, there was a feedback by prior knowledge interaction, $F(1, 79) = 5.70$, $p = .02$, $\eta_p^2 = .07$. Consistent with the predictions, as prior knowledge increased, the benefits of feedback decreased ($B = -1.04$, $SE = 0.43$). Feedback type did not interact with prior knowledge ($p = .44$).[1]

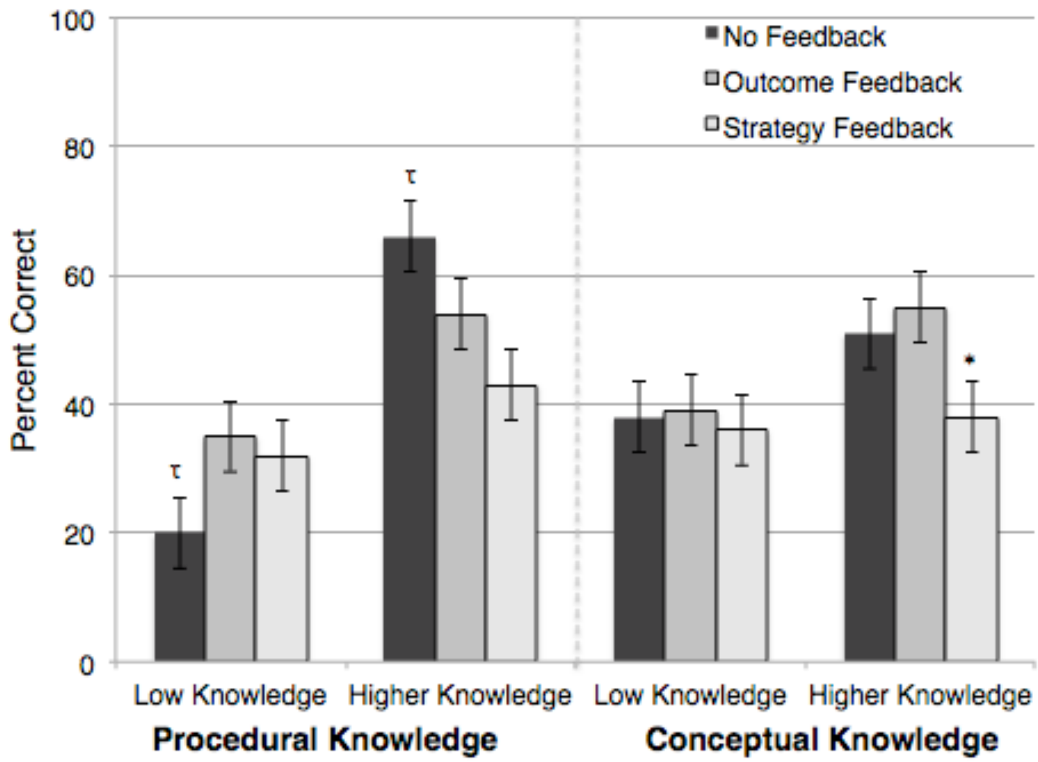
[1]The overall procedural knowledge results remain unchanged when the midtest is removed from the model.

Table 3: Procedural Knowledge Scores in Exp. 1 by Condition and Prior Knowledge

Time	Prior Knowledge	No Feedback	Outcome Feedback	Strategy Feedback
Pretest	Low	15 (9)	13 (9)	13 (9)
	Higher	51 (15)	54 (16)	48 (14)
Midtest	Low	0 (0)	22 (31)	23 (32)
	Higher	58 (42)	50 (46)	40 (46)
Posttest	Low	26 (31)	39 (34)	32 (27)
	Higher	71 (25)	61 (32)	50 (27)
Retention Test	Low	29 (32)	33 (32)	33 (36)
	Higher	81 (22)	71 (27)	41 (32)

Note. Raw percentage means are presented with standard deviations in parentheses. Children are categorized as low or higher prior knowledge based on a median split on the procedural knowledge measure at pretest; however, the primary analysis models treated prior knowledge as a continuous variable.

To help interpret the interaction, I categorized children as having higher prior knowledge (scored above the median on the procedural knowledge pretest) or low prior knowledge and examined the main effects of feedback for each group (see Figure 1). For the low-knowledge group, children who received feedback tended to exhibit higher procedural knowledge (M = 34%, SE = 5%) than children who did not (M = 20%, SE = 6%), $F(1, 79) = 3.28$, $p = .07$, $\eta_p^2 = .04$. For the higher-knowledge group, children who received feedback tended to exhibit *lower* procedural knowledge (M = 49%, SE = 6%) than children who did not (M = 66%, SE = 8%), $F(1, 79) = 3.66$, $p = .06$, $\eta_p^2 = .04$. Overall, the results resemble an expertise reversal effect (Kalyuga et al., 2003). Feedback during exploration was more beneficial than no feedback, but *only* for children with low prior knowledge. For children with higher prior knowledge, the reverse was true, with feedback actually hindering learning relative to no-feedback.



Note. Scores are estimated marginal means based on midtest, posttest, and retention test scores combined. For procedural knowledge, differences are between the no-feedback and the two feedback conditions combined. For conceptual knowledge, difference is between the outcome-feedback and the strategy-feedback condition. Error bars represent standard errors. $^{*}p < .05$, $^{\tau}p < .07$.

Figure 1: Percent Correct on Procedural and Conceptual Knowledge Measures in Exp. 1 by Condition and Prior Knowledge

Conceptual Knowledge

Children's conceptual knowledge also changed over time. Scores increased from midtest (M = 20%, SE = 2%) to posttest (M = 55%, SE = 2%) and remained similar at retention test (M = 51%, SE = 3%), $F(2, 158) = 89.73$, $p < .001$, $\eta_p^2 = .53$. There were no effects related to feedback versus no-feedback. There was a marginal effect of feedback type, $F(1, 79) = 3.56$, $p = .06$, $\eta_p^2 = .04$, which was qualified by a marginal feedback type by prior knowledge interaction, $F(1, 79) = 2.93$, $p = .09$, $\eta_p^2 = .04$. As prior knowledge increased, the benefits of outcome feedback increased relative to strategy feedback ($B = 0.67$, $SE = 0.39$).[2]

To help interpret this marginal interaction, the effect of feedback type was examined for low- and higher-knowledge children separately (based on a median split on procedural knowledge pretest scores, see Figure 1). For the low-knowledge group, there were no differences between the types of feedback, $F(1, 79) = 0.34$, $p = .56$. For the higher-knowledge group, children who received outcome-feedback had higher conceptual knowledge ($M = 55\%$, $SE = 5\%$) than children who received strategy-feedback ($M = 38\%$, $SE = 6\%$), $F(1, 79) = 5.07$, $p = .03$. These results suggest that outcome-feedback and strategy-feedback promoted similar levels of conceptual knowledge for low-knowledge children; but outcome-feedback promoted greater conceptual knowledge than strategy-feedback for higher-knowledge children. Note that outcome-feedback was not more effective than no-feedback for higher-knowledge children ($p = .56$).

Intervention Measures

To better understand *how* exploration impacted learning, I explored children's responses during the intervention. Recall, children reported the strategies they used to solve the problems as well as their subjective cognitive load after exploration.

I was interested in how feedback influenced children's strategy variability. Past work indicates that feedback's primary function is to identify errors and encourage the search for plausible alternatives (e.g., Kulhavy, 1977). Thus, promoting the use of

[2] The overall conceptual knowledge results remain similar when the midtest is removed from the model.

different strategies might be one mechanism by which feedback influences exploration. Overall, children used a variety of correct and incorrect solution strategies on the 12 practice problems (see Table 2). The number of different strategies used per child ranged from 1 to 6 (M = 2.9, SD = 1.4), and nearly 40% of children used four or more different strategies.

The primary ANCOVA model was used with feedback and feedback type as between-subject variables and number of different strategies used as the dependent variable. The number of correct and incorrect strategies were examined separately. There was a marginal effect of feedback on the number of *correct* strategies used, $F(1, 79) = 3.04$, $p = .09$, $\eta_p^2 = .04$, but no other effects were significant. Children who received feedback used a greater number of different correct strategies (M = 0.9, SE = 0.1) than children who did not (M = 0.6, SE = 0.1). A qualitative examination indicated that the effect was slightly stronger for low-knowledge children. For the higher-knowledge group, the number of different correct strategies barely differed in the feedback (M = 1.1, SE = 0.2) and no-feedback conditions (M = 1.0, SE = 0.2). For the low-knowledge group, children who received feedback generated a greater number of different correct strategies (M = 0.8, SE = 0.1) than children who did not (M = 0.3, SE = 0.2). The pattern of results was the same for both types of feedback.

For the number of different *incorrect* strategies used, there was a main effect of feedback, $F(1, 79) = 11.22$, $p = .001$, $\eta_p^2 = .12$, but no other effects were significant. Children who received feedback used a greater number of different incorrect strategies (M = 2.4, SE = 0.2) than children who did not (M = 1.6, SE = 0.2). The effect was similar in strength for all children. For the low-knowledge group, children who received feedback (*M* = 2.8, *SE* = 0.2) used roughly one more incorrect strategy than children who did not (*M* = 1.9, *SE* = 0.3). Similarly, for the higher-knowledge group, children who received feedback (*M* = 1.9, *SE* = .3) used roughly one more incorrect strategy than children who did not (*M* = 1.0, *SE* = 0.3). The pattern of results was the same for both types of feedback. Overall, for low-knowledge children, feedback promoted the use of both correct and incorrect strategies relative to no-feedback. For higher-knowledge children, feedback promoted the use of incorrect strategies more so than correct strategies relative to no-feedback, which may help explain why feedback hindered their performance.

There were also differences in *perseveration*—using the same incorrect strategy on all of the problems. More children perseverated in the no-feedback condition (23%) than in the strategy-feedback (8%) or outcome-feedback (0%) conditions, χ^2 (2, *N* = 87) = 8.73, *p* = .01. Moreover, the effect was more pronounced for children with low prior knowledge. For low-knowledge children, more children perseverated in the no-feedback condition (32%) than in the strategy-feedback (13%) or outcome-feedback (0%) conditions. For children with higher prior knowledge, few children perseverated at all (8% in no-feedback condition, 0% in strategy- and outcome-feedback conditions). Overall, low-knowledge children used more diverse strategies if they received feedback and tended to perseverate on the same incorrect strategy if they did not. Very few higher-knowledge children perseverated, regardless of feedback condition.

Researchers often explain the expertise reversal effect in terms of the learner's cognitive resources and the experience of *cognitive load* (e.g., Rey & Buchwald, 2011). Feedback may interact with prior knowledge because of differences in cognitive load.

Children's cognitive load ratings were variable, spanning the scale's range from 1 to 5 (*M* = 3.1, *SD* = 0.95). The primary ANCOVA model was used with feedback and feedback type as between-subject variables and mean cognitive load as the dependent variable. There was a main effect of feedback, $F(1, 79) = 4.63$, $p = .03$, $\eta_p^2 = .06$, but no other effects were significant. Children who received feedback experienced higher levels of cognitive load (*M* = 3.3, *SE* = 0.12) than children who did not (*M* = 2.8, *SE* = 0.17). A qualitative examination of the means indicates that, for low-knowledge children, those who received feedback experienced slightly higher levels of cognitive load than children who did not (see Table 4). For children with higher prior knowledge, the difference was more pronounced. Those who received feedback, particularly strategy feedback, experienced higher levels of cognitive load than those who did not. These results suggest that feedback during exploration increased cognitive load relative to no feedback.

Table 4: Children's Subjective Cognitive Load Ratings Across Experiments 1 and 2

Condition	Prior Knowledge	Experiment 1 NASA-TLX (out of 5)	Experiment 2 NASA-TLX (out of 5)	Experiment 2 New Item (out of 7)
No Feedback	Low	3.03 (.96)	3.28 (.93)	3.22 (1.7)
	Higher	2.46 (.58)	3.00 (.94)	4.08 (2.4)
Outcome Feedback	Low	3.36 (.94)	3.40 (.91)	5.45 (1.2)
	Higher	2.85 (1.1)	3.04 (.83)	4.85 (1.3)
Strategy Feedback	Low	3.23 (.90)	3.97 (.79)	5.40 (1.2)
	Higher	3.55 (.93)	4.00 (.71)	4.27 (1.9)

Note. Raw means are presented with standard deviations in parentheses

Discussion

In Experiment 1, the primary hypothesis was supported. Feedback led to higher procedural knowledge than no feedback, but only for children with low prior knowledge. For children with higher prior knowledge, feedback led to *lower* procedural knowledge relative to no feedback. The secondary analyses indicated that for low-knowledge children, feedback promoted the generation of both correct and incorrect strategies and also prevented perseveration relative to no feedback. For higher-knowledge children, feedback promoted the use of incorrect strategies relative to no feedback. Additionally, feedback led to higher subjective cognitive load relative to no feedback, particularly for higher-knowledge children. Overall, the benefits of providing feedback during exploration prior to direct instruction depend on prior knowledge. Children with low domain knowledge benefited more from receiving feedback, whereas children with higher domain knowledge benefited more from exploring without feedback. Feedback type had little effect in general, with the exception that outcome feedback tended to lead to higher conceptual knowledge than strategy feedback for children with higher prior children.

Despite supporting the primary hypothesis, several limitations remain. First, the condition manipulation was not as clean or as strong as it could have been. For example, all children were asked to report how they solved each problem. Though this resulted in detailed information regarding the strategies used, it inevitably guided all children's attention to some degree on their problem-solving strategies. The strategy-feedback manipulation would be stronger if only children in the strategy-feedback condition were encouraged to attend to their strategy use. Additionally, the feedback provided in both feedback conditions was relatively vague and not specific to the child's response. For example, in the strategy-feedback condition, incorrect strategies were referred to as "not a correct way," which may have been unclear to children. Further, children in both the strategy-feedback and outcome-feedback conditions were told if their target response (strategy or answer, respectively) was correct, but only children in the outcome-feedback were given additional correct information (i.e., the correct answer). The contrast between the two feedback conditions could be improved.

Second, I sought to clarify the influences of feedback type during exploration prior to instruction. Given the paucity of research comparing outcome-feedback to strategy-feedback and the slight variation in means for higher-knowledge children in these two conditions, I wanted to confirm that feedback type is not central to children's learning during exploration. To address these concerns, a second experiment was conducted similar to Experiment 1, but with several modifications intended to strengthen the design.

CHAPTER III

EXPERIMENT 2

Experiment 2 was designed to strengthen the condition manipulation from Experiment 1 and verify the results with an independent sample of children. The goal was to replicate the finding that low-knowledge children benefit from feedback during exploration prior to instruction, whereas children with higher prior knowledge benefit from no feedback. The condition manipulation was strengthened in three ways. First, to differentiate the conditions, only children in the strategy-feedback condition reported how they solved each problem. Children in the other conditions were asked to report different information to mimic the interaction with the experimenter (i.e., their answer in the outcome-feedback condition and their completion of the problem in the no-feedback condition). Second, the feedback was made more specific by having the experimenter revoice the child's response. In the strategy-feedback condition the child's strategy was restated and in the outcome-feedback condition the child's answer was restated. Finally, children in the outcome-feedback condition did not hear the correct answer. In Experiment 1, only children in the outcome-feedback condition received additional information (i.e., the correct answer). An alternative solution was to provide children in the strategy-feedback condition with additional information (i.e., a correct strategy). However, telling children how to solve a problem is a form of direct instruction, and I was interested in the guidance provided *prior* to direct instruction. So the correct answer in the outcome-feedback condition was eliminated to enhance parallelism across conditions. Consistent with Experiment 1, I expected low-knowledge children to benefit more from feedback relative to no feedback and higher-knowledge children to benefit more from no feedback, regardless of feedback type.

Method

Participants

Consent was obtained from 111 second- and third-grade children at two schools (one public, one parochial) in middle Tennessee. Of those children, 101 met criteria for participation because they scored at or below 80% on both a conceptual and procedural knowledge measure at pretest. Data from six additional children were excluded: two for failing to complete the intervention and four for missing the retention test. The final sample consisted of 95 children (*M* age = 7 yrs, 11 mo; 60 girls; 35 boys; 97% African American; 3% Caucasian). Approximately 61% received free or reduced price lunch.

Design

The design and procedure were identical to Experiment 1 with a few exceptions outlined below. As before, children were randomly assigned to one of three conditions for the intervention: strategy-feedback (n = 31), outcome-feedback (n = 32), or no-feedback (n = 29).

Procedure

Consistent with Experiment 1, all children received a tutoring session that began with exploratory problem solving followed by brief conceptual instruction. The 12 mathematical equivalence problems from Experiment 1 were used, but were presented in paper/pencil format rather than on a computer screen to simulate a more typical classroom activity. The computer program was still used by the experimenter to record information.

In the *strategy-feedback condition*, children reported how they solved each problem and received feedback on the strategy, which included a revoicing of the strategy report (e.g., "Good job! That is one correct way to solve that problem. [Child's strategy] is a correct way to solve it." / "Good try, but that is not a correct way to solve the problem. [Child's strategy] is not a correct way to solve it."). For example, if a child reported using the add-all strategy (see Table 2), the experimenter repeated the child's report: "Good try, but that is not a correct way to solve the problem. Adding all the numbers together is not a correct way to solve this problem." The experimenter revoiced the strategy just as the child stated it to ensure no additional information was given. If the child was unsure of the strategy used, the experimenter stated: "It is not clear if you used a correct way to solve this problem. Let's try another one. This time, try to remember how you solved the problem," though this occurred on only 2% of trials. In the *outcome-feedback condition*, children reported their numerical answer and received feedback on that answer, which included a revoicing of their answer but not the correct answer (e.g., "Good job! You got the right answer, [child's answer] is the correct answer." / "Good try, but you did not get the right answer, [child's answer] is not the correct answer."). Finally, in the *no-feedback condition*, children reported when they completed a problem and were told to go on.

Assessments and Coding

The mathematical equivalence assessment, modified slightly from Experiment 1 to improve psychometric properties, was administered at pretest, posttest, and retention test. Again, a brief version (5 items—3 conceptual and 2 procedural items) was used as a midtest during the intervention. To establish inter-rater reliability, a second rater coded the subjective responses of 30% of the children (see Table 1 for coding criteria). Inter-rater agreement was high for written explanations (exact agreement = 93 – 99%, kappas = 87 – 98%) and for strategy codes (exact agreement = 91%, kappa = 89%).

The two subjective cognitive load items from Experiment 1 were administered. A third item was also included. It was adapted from past cognitive load studies with children as young as 13-years-old (e.g., Ayres, 2006; Kalyuga, Chandler, Sweller, 2004): "How easy or hard was it to solve all of those problems?" Children responded on a 7-point scale ranging from very, very easy to very, very hard.

A strategy evaluation task was also added, which we administered after the posttest and after the retention test, to assess children's ratings of correct and incorrect strategies for solving mathematical equivalence problems (Rittle-Johnson & Alibali, 1999). Children were told that students from another school had solved these problems in different ways. They were presented with examples of the strategies used by these students (see Table 2 for the strategies demonstrated). Children were asked to evaluate each strategy as "very smart, kind of smart, or not so smart." This task was included to determine whether differences existed in terms of the *recognition* of correct strategies. However, preliminary analyses indicated no systematic differences between conditions; thus, I do not report results for this task.

Data Analysis

The same ANCOVA model from Experiment 1 was employed. I used a contrast-based ANCOVA with two contrast-coded between-subject variables (feedback, feedback type), three covariates, and two condition by prior knowledge interaction terms. The assumption of homogeneity of variance for this model was largely supported. For the procedural knowledge variables, Levene's tests indicated equal variances at midtest, posttest, and retention test, F's < 2.2. With all three time points in the same model, the variance-covariance matrices were also homogeneous, Box's M = 20.09, $F(12, 40695) = 1.60$, $p = .09$. For the conceptual knowledge variables, Levene's tests indicated equal variances at midtest and posttest, F's < 2, though not at the retention test, $F(2, 92) = 4.66$, $p = .01$. With all three time points in the same model, the variance-covariance matrices were also homogeneous, Box's M = 13.68, $F(12, 40694) = 1.09$, $p = .37$. Overall, ANOVA models were appropriate for analyzing the data.

Results

On the pretest, children answered few procedural (M = 20%, SD = 18%) and conceptual (M = 19%, SD = 18%) items correctly. However, the scores ranged from 0 – 75% on the procedural scale and from 0 – 80% on the conceptual scale. Importantly, there were no significant differences between conditions on either scale at pretest, Fs < 1.

To evaluate performance on the midtest, posttest and retention test a repeated measures ANCOVAs with feedback (feedback vs. none) and feedback type (outcome vs. strategy) as between-subject variables and time (midtest, posttest, retention test) as the within-subject variable was conducted. The three covariates and two interaction terms were also included. The statistical conclusions remain unchanged when the midtest is removed from the model.

Procedural Knowledge

Children's procedural knowledge scores increased from midtest to posttest and remained similar two weeks later (see Table 5), $F(2, 174) = 3.77$, $p = .03$, $\eta_p^2 = .04$. There were no main effects of feedback or feedback type, nor did feedback type interact with prior knowledge, *F*s < 1. However, consistent with Experiment 1, there was a feedback by prior knowledge interaction, $F(1, 87) = 4.67$, $p = .03$, $\eta_p^2 = .05$. As prior knowledge increased, the benefits of feedback decreased ($B = –1.06$, $SE = 0.49$).[3]

To help interpret this interaction, children were categorized as having higher prior knowledge (scored above the median on the procedural knowledge pretest measure) or low prior knowledge and examined the main effects of feedback for each group (see Figure 2). For the low-knowledge group, children who received feedback exhibited significantly higher procedural knowledge ($M = 33\%$, $SE = 4\%$) than children who did not ($M = 20\%$, $SE = 5\%$), $F(1, 87) = 4.00$, $p = .05$, $\eta_p^2 = .04$. For the higher-knowledge group, children who received feedback exhibited significantly *lower* procedural knowledge ($M = 28\%$, $SE = 5\%$) than children who did not ($M = 50\%$, $SE = 6\%$), $F(1, 87)$

[3] The overall procedural knowledge results remain unchanged when the midtest is removed from the model.

= 7.54, p = .007, η_p^2 = .08. The results replicated the expertise reversal effect found in Experiment 1. Feedback was more beneficial than no feedback for children with low prior knowledge, but for children with higher prior knowledge, the reverse was true. Feedback type did not matter, suggesting that both types of feedback were beneficial for low-knowledge children and both were detrimental for higher-knowledge children.

Table 5: Procedural Knowledge Scores in Exp. 1 by Condition and Prior Knowledge

Time	Prior Knowledge	No Feedback	Outcome Feedback	Strategy Feedback
Pretest	Low	8 (6)	9 (6)	11 (5)
	Higher	38 (19)	39 (19)	32 (10)
Midtest	Low	8 (19)	23 (26)	21 (34)
	Higher	43 (46)	29 (32)	40 (38)
Posttest	Low	24 (31)	38 (38)	40 (28)
	Higher	49 (41)	32 (23)	30 (27)
Retention Test	Low	24 (24)	31 (29)	29 (31)
	Higher	54 (36)	33 (28)	29 (29)

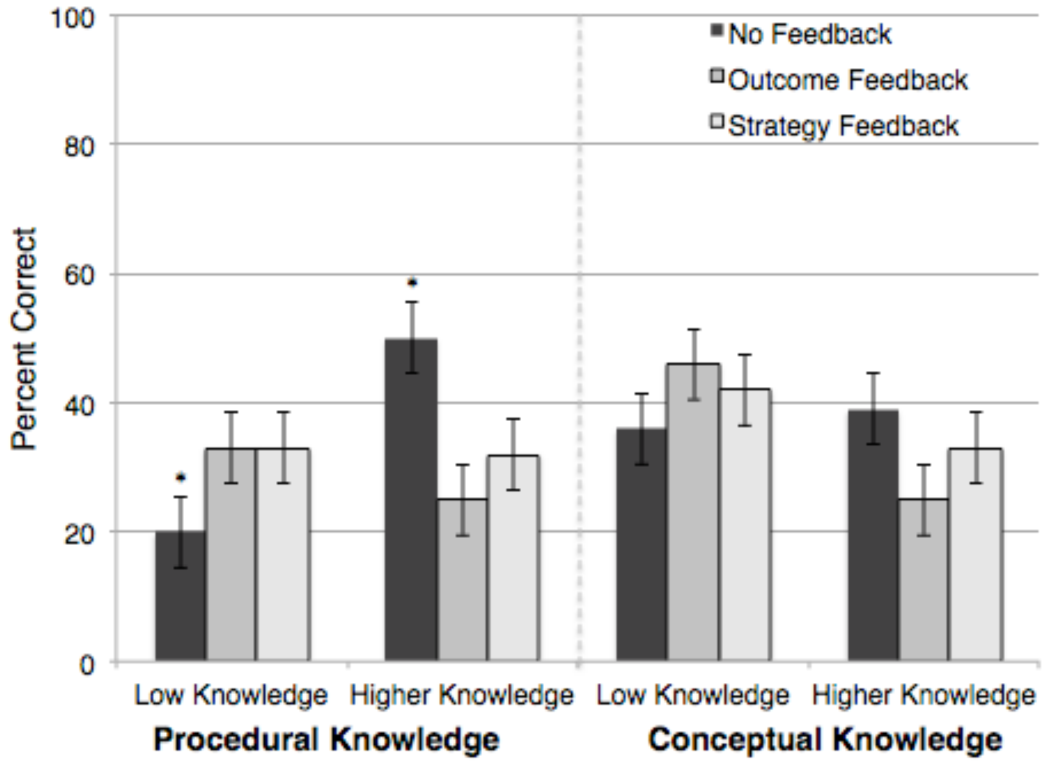
Note. Raw means are presented with standard deviations in parentheses. Children are categorized as low or higher prior knowledge based on a median split on the procedural knowledge assessment at pretest; however, the primary analysis models treated prior knowledge as a continuous variable.

Conceptual Knowledge

Children's conceptual knowledge scores also increased from midtest (M = 21%, SE = 2%) to posttest (M = 50%, SE = 2%) and stayed similar at retention test (M = 43%, SE = 2%), F(2, 174) = 67.13, p < .001, η_p^2 = .44. There were no main effects of feedback or feedback type, nor did feedback type interact with prior knowledge, F's < 1. There was a marginal feedback by prior knowledge interaction, F(1, 87) = 3.63, p = .06, η_p^2 = .04. As prior knowledge increased, the benefits of feedback marginally decreased (B = –0.70, *SE* = 0.37). Feedback also interacted with time, $F(2, 174) = 7.14$, $p = .001$, $\eta_p^2 = .08$, such that the benefits of feedback were stronger at the midtest and decreased over time.[4]

To help interpret the marginal interaction, I examined the effect of feedback for low- and higher-knowledge children separately (based on a median split on procedural knowledge pretest scores; see Figure 2). For the low-knowledge group, children who received feedback exhibited slightly higher conceptual knowledge (*M* = 44%, *SE* = 3%) than children who did not (*M* = 37%, *SE* = 4%), $F(1, 87) = 2.56$, $p = .11$, $\eta_p^2 = .03$. For the higher-knowledge group, children who received feedback exhibited slightly lower conceptual knowledge (*M* = 29%, *SE* = 3%) than children who did not (*M* = 39%, *SE* = 5%), $F(1, 87) = 2.60$, $p = .11$, $\eta_p^2 = .03$. Although not reliable, particularly when dichotomizing prior knowledge, these conceptual knowledge results resemble the pattern of findings found for procedural knowledge.

[4] The overall conceptual knowledge results remain unchanged when the midtest is removed from the model.



Note. Scores are estimated marginal means based on midtest, posttest, and retention test scores combined. For procedural knowledge, differences are between the no-feedback and the two feedback conditions combined. Error bars represent standard errors. $^{*}p < .05$.

Figure 2: Percent Correct on Procedural and Conceptual Knowledge Measures in Exp. 2 by Condition and Prior Knowledge

Intervention Measures

Recall that children were asked to report their subjective cognitive load after the exploratory problem solving. I employed the same measure as in Experiment 1 (mean ratings of the two NASA-TLX items) and also included a new measure for exploratory purposes. To explore children’s ratings, the primary ANCOVA model with feedback and feedback type as between-subject variables and cognitive load as the dependent variable was used. In Experiment 1, I analyzed children’s verbal strategy reports. However, in

Experiment 2 I only had detailed strategy reports from children in the strategy-feedback condition, so I could not perform a comparable analysis on children's strategy variability.

Similar to Experiment 1, for the mean rating on the NASA-TLX measure of subjective cognitive load, there was a main effect of feedback, $F(1, 86) = 5.79$, $p = .02$, $\eta_p^2 = .06$. Overall, children who received feedback experienced higher levels of cognitive load ($M = 3.6$ out of 5, $SE = 0.11$) than children who did not ($M = 3.1$, $SE = 0.16$). However, there was also a main effect of feedback type, $F(1, 86) = 10.52$, $p = .002$, $\eta_p^2 = .11$. Those who received strategy feedback experienced even higher levels of cognitive load ($M = 3.9$, $SE = 0.16$) than those who received outcome feedback ($M = 3.3$, $SE = 0.15$). Neither interaction was significant, F's < 1. A qualitative examination of the means (see Table 4) supports these conclusions. On average, children who received feedback, particularly strategy feedback, experienced higher levels of cognitive load than children who did not.

Children's cognitive load ratings on the new measure were similar to their ratings on the original measure, with minor differences. Again, there was a main effect of feedback, $F(1, 87) = 18.11$, $p < .001$, $\eta_p^2 = .17$. Children who received feedback experienced higher levels of cognitive load ($M = 5.1$ out of 7, $SE = 0.2$) than children who did not ($M = 3.6$, $SE = 0.3$). However, this was qualified by a significant feedback by prior knowledge interaction, $F(1, 86) = 4.75$, $p = .03$, $\eta_p^2 = .05$. Follow-up analyses indicated that, for the low-knowledge group, children who received feedback experienced much higher levels of cognitive load ($M = 5.4$, $SE = 0.3$) than children who did not receive feedback ($M = 3.2$, $SE = 0.4$), $F(1, 87) = 23.8$, $p < .001$, $\eta_p^2 = .22$. For the higher-knowledge group, children who received feedback experienced more similar levels of cognitive load (*M* = 4.7, *SE* = 0.3) to children who did not (*M* = 4.1, *SE* = 0.4), $F(1, 87) = 0.79$, $p = .37$, $\eta_p^2 = .01$. There were no effects related to feedback type. A qualitative examination of the means (see Table 4) supports these conclusions. Children who received feedback experienced higher levels of cognitive load than children who did not, and this was particularly evident for low-knowledge children.

Discussion

Results from Experiment 2 were consistent with Experiment 1 and supported the primary hypothesis. For children with low prior knowledge, feedback during exploratory problem solving led to higher procedural knowledge than no feedback. But, for children with higher prior knowledge, feedback led to *lower* procedural knowledge than no feedback. There was a similar, yet weaker effect for children's conceptual knowledge. Feedback type had little effect in general, providing evidence that both types of feedback hinder higher-knowledge children's performance. The exploratory findings on children's subjective cognitive load depended slightly on the measure, though the overall pattern suggested that feedback generally led to higher levels of cognitive load than no feedback. Overall, Experiment 2 replicated the findings from Experiment 1 with an independent sample of children and supported the primary conclusions.

CHAPTER IV

GENERAL DISCUSSION

Guided discovery facilitates deeper learning than discovery or instruction alone (e.g., Aflieri et al., 2010; Mayer, 2004). For example, providing exploratory activities with subsequent instruction can be beneficial (e.g., DeCaro & Rittle-Johnson, 2011; Schwartz & Bransford). But, the amount of guidance provided *during* the initial exploration has largely gone unstudied. I examined the effects of feedback during exploratory problem solving for children with various levels of prior knowledge. In two experiments, children solved unfamiliar mathematical equivalence problems, followed by conceptual instruction. Some children received feedback (on their answer or on their strategy), whereas others did not. In both experiments, the primary hypothesis was supported. Feedback led to higher procedural knowledge than no feedback, but *only* for children with low prior knowledge. For children with higher prior knowledge, feedback led to lower procedural knowledge relative to no feedback. Effects on children's conceptual knowledge were weak, suggesting feedback during exploration primarily impacts procedural knowledge. Feedback type (outcome vs. strategy) had little effect in general. I discuss these findings in light of past research on the expertise reversal effect and strategy generation. Finally, I consider the implications for guided discovery learning as well as potential future inquiries.

The Expertise Reversal Effect

The present findings are consistent with research demonstrating expertise reversal effects, in which an instructional technique that is effective for novices loses its benefits for learners with more experience (Kalyuga et al., 2003). Recall, the more experienced learners need not be experts and are often referred to as learners with higher prior knowledge (e.g., Kalyuga, 2007). The techniques that have been investigated continue to increase. For example, we know that low- but not higher-knowledge learners, benefit more from (a) studying a given solution rather than imagining it (Leahy & Sweller, 2005), (b) seeing worked examples rather than solving open-ended problems (Renkl & Atkinson, 2003), and (c) having multiple pieces of information presented together rather than in isolation (Kalyuga, Chandler, & Sweller, 1998). The current study extends the expertise reversal effect to the use of feedback during exploratory problem solving prior to instruction. Children with low prior knowledge need feedback to improve their knowledge of correct procedures. Children with higher prior knowledge, on the other hand, do not need feedback and actually perform better without it. This occurred even though the higher-knowledge children in our study were far from experts and still had a lot to learn.

As in the current study, low-knowledge learners typically profit from instructional methods that provide external support, whereas more knowledgeable learners typically profit from less structured methods (Kalyuga et al., 2003). The current findings highlight that learners with only moderate levels of prior knowledge can benefit from less structured methods.

Explanations for the effect stem from cognitive load theory, though expertise reversal effects have been found in studies outside of the cognitive load paradigm (see Kalyuga, 2007). Cognitive load theory proposes that cognitive resources influence the efficacy of instruction (Paas, Renkl, & Sweller, 2003). The goal of any learning task is to construct an accurate representation of the information and extract underlying concepts and procedures, which requires working memory resources (Sweller, 1988). To do so without overtaxing cognitive resources, one needs some form of guidance. For novices, this guidance comes from instructional supports, which help reduce the cognitive load associated with novel tasks. But for learners with more knowledge, guidance comes from pre-existing schemas. If further information is provided, it can result in the processing of redundant information (e.g., Kalyuga et al., 2003; Rey & Buchwald, 2011) and increased cognitive load. This theory-driven explanation is plausible, although direct evidence for a cognitive load explanation is limited (see Kirschner, Ayres, & Chandler, 2011).

Findings from the cognitive load measures partly support this explanation. The general pattern indicated that children who received feedback reported higher levels of cognitive load than children who did not, though there were slight differences across experiments and measures. This may help explain why higher-knowledge children did not benefit from the feedback. Providing them with additional guidance created a *more* cognitively taxing environment than letting them explore on their own. This increased load was not germane to the task; rather, feedback may have caused extraneous cognitive load (Sweller et al., 1998). However, why did children with low prior knowledge benefit from the feedback? These children also tended to report higher levels of cognitive load in the feedback conditions than in the no feedback condition. One possibility is that feedback caused germane cognitive load for low-knowledge learners – effortful learning that supports the refinement of schema (Sweller et al., 1998). Unfortunately, the distinction between germane and extraneous cognitive load is theoretical and current measurement techniques do not capture the distinction (Kalyuga, 2011). In addition, cognitive load scales have rarely been employed with children; and even in adults the ability to provide accurate self-reports of mental effort has been questioned (Schnotz & Kurschner, 2007).

Thus, cognitive load theory provides a plausible account of why feedback during exploration aids learning in low-knowledge learners, but harms learning in higher-knowledge learners. However, difficulties in gathering empirical evidence in support of the proposed cognitive load mechanism, particularly in children, makes this account difficult to verify empirically. Alternative explanations of the findings also remain (e.g., Schnotz, 2010). For example, Schnotz (2010) raises the possibility that children who are more motivated thrive in less structured, challenging environments whereas children who are less motivated do not. The current data point to the important role feedback can play in strategy generation and selection.

Strategy Variability During Exploration

One of feedback's primary roles is to help learners identify errors and search for more plausible alternatives (see Mory, 2004). Indeed, in Experiment 1 children who received feedback exhibited greater strategy variability than those who did not, and this effect was particularly true for the low-knowledge group. For children with low prior knowledge, feedback promoted the use of more different strategies relative to no feedback and also prevented perseveration on the same incorrect strategy. Past work indicates that cognitive variability is an impetus for cognitive change (e.g., Siegler, 1994; Siegler & Shipley, 1995). That is, thinking of or using a variety of strategies can help learners evaluate alternative approaches and be more receptive to subsequent instruction (e.g., Alibali & Goldin-Meadow, 1993; Perry, Church, & Goldin-Meadow, 1988). Feedback during exploration facilitated the use of diverse strategies, including more correct strategies, for children with low prior knowledge. Further, feedback prevented low-knowledge children from using the same incorrect strategy on all of the problems, which supports the idea that feedback's main function is identifying initial errors (Mory, 2004). This may help explain why these children ultimately learned more when they received feedback than when they did not.

Children's strategy use during the intervention may also provide insight into why feedback hindered the performance of higher-knowledge learners. Recall, for the higher-knowledge group, feedback promoted the use of a greater number of incorrect strategies relative to no feedback, but had a much weaker effect on correct strategies. Because feedback only facilitated the use of incorrect strategies, it may have had a negative impact on their procedural knowledge. If correct strategies compete against incorrect strategies for selection, increasing the number of incorrect strategies could reduce use of correct strategies (Siegler & Shipley, 1995).

Strategy variability use may also shed light more generally on how exploration prior to instruction impacts learning. Schwartz and Bransford (1998) suggest that exploratory activities facilitate the "development of differentiated knowledge" of the target problem space (p. 510). In past studies, exploration improved knowledge of the structure of the target problems (e.g., Schwartz & Martin, 2004) and the concepts underlying them (e.g., DeCaro & Rittle-Johnson, 2011). Consistent with this work, children in Experiment 1 generated a wide variety of problem solving strategies during the exploratory phase. Exploring a problem space may help learners acquire more nuanced knowledge, and thus prepare them to learn from subsequent instruction.

Guided Discovery Learning

The present study also has important implications for research on guided discovery learning. In particular, it suggests that prior knowledge (and other individual differences) should be considered when evaluating guided discovery methods. Too often researchers consider individual differences "an annoyance…to be reduced by any possible device," rather than a source of relevant information (Cronbach, 1957, p. 674). Future research should continue to examine these person characteristics to assess the generalizability of guided discovery methods and how they can be optimized for certain subsets of learners.

The results also highlight the need to evaluate different aspects of guided discovery. I examined the guidance provided *during* exploration prior to instruction and found that more was not always better. Unfortunately, even when researchers recognize the benefits of combining discovery and instruction, the usual recommendation is to include more guidance (e.g., feedback, hints) during exploratory activities (e.g., Alfieri et al., 2010; Mayer, 2004). For example, Mayer (2004) suggests learning may be best supported by "instructional guidance rather than pure discovery and curricular focus rather than unstructured exploration" (p. 14). However, the results from this study indicate that there is a time for just exploration. Combining unguided exploration with subsequent instruction improved higher-knowledge children's procedural knowledge to a greater extent than exploration with feedback. Thus, optimizing learning does not always require an increase in guidance; sometimes it requires the removal of unnecessary information.

Limitations and Future Directions

Despite the positive contributions of the current study, future research is needed. First, studies should continue investigating the effects of feedback type. In this study, there were few differences between outcome-feedback and strategy-feedback, and the differences that were there were weak and inconsistent. It is possible that outcome and strategy feedback influence children's problem solving in similar ways. However, past research suggests that is not the case. For example, Luwel et al. (2011) examined children's use of two different correct strategies for completing a numerosity judgment task and found that strategy feedback led to more adaptive strategy selection than outcome feedback. It may be that strategy feedback is more beneficial when choosing between known strategies as opposed to generating them for the first time.

More work is also needed to verify the generalizability of the results across domains and settings. For example, feedback may have a larger impact for low-knowledge learners in domains with misconceptions, such as mathematical equivalence, because feedback's role is to facilitate the correction of misconceptions and errors. In domains without misconceptions, feedback may be less necessary. Also, feedback may be most effective in one-on-one tutoring settings, in which feedback is immediate and can influence current performance. Future work should examine a more feasible application of feedback in a classroom setting, such as providing written feedback on a homework assignment or test.

Finally, additional clarifications regarding the distinction between levels of prior knowledge are necessary. For example, future work should address what counts as sufficient prior knowledge so as to know when feedback during exploration is no longer effective. In the current study, higher knowledge was fairly limited. More generally, studies that demonstrate treatment by prior knowledge interactions have not identified the precise level of prior knowledge at which the reversal occurs. As more and more research finds that the effectiveness of instruction depends on prior knowledge, instructors will need guidance on how to choose instructional techniques for particular children with particular levels of prior knowledge.

Conclusion

This study extends research on guided discovery methods in which exploratory activities are provided with subsequent direct instruction. I examined how and for whom a particular form of guidance, feedback, might enhance learning from the combination of exploration and instruction. Feedback during exploratory problem solving facilitates learning for children with low domain knowledge. However, children with higher prior knowledge benefit more from exploring independently without feedback. Thus, providing feedback may not always be optimal.

REFERENCES

Ahmad, M. (1998). The effect of computer-based feedback on using cognitive strategies of problem solving. *Proceedings of selected research papers, Association for Educational Communications and Technology, Research and Theory Division* (pp. 1-22). New Orleans, LA: AECT.

Alfieri, L., Brooks, P. J., Aldrich, N. J., & Tenenbaum, H. R. (2011). Does discovery-based instruction enhance learning? *Journal of Educational Psychology*, *103*(1), 1-18. doi: 10.1037/a0021017.

Alibali, M. W. (1999). How children change their minds: Strategy change can be gradual or abrupt. *Developmental Psychology*, *35*(1), 127-45.

Alibali, M. W., & Goldin-Meadow, S. (1993). Gesture-speech mismatch and mechanisms of learning: What the hands reveal about a child's state of mind. *Cognitive Psychology*, *25*, 468-523.

Anderson, R. C., Kulhavy, R. W., & Andre, T. (1972). Conditions under which feedback facilitates learning from programmed lessons. *Journal of Educational Psychology, 63*, 186-188.

Ayres, P. (2006). Impact of reducing intrinsic cognitive load on learning in a mathematical domain. *Applied Cognitive Psychology, 20*(3), 287-298. doi: 10.1002/acp.1245.

Baroody, A. J., & Ginsburg, H. P. (1983). The effects of instruction on children's understanding of the "equals" sign. *Elementary School Journal, 84*, 199-212.

Barrows, H. S., & Tamblyn, R. (1980). *Problem-based learning: An approach to medical education.* New York: Springer.

Birenbaum, M., & Tatsuoka, K. (1987). Effects of "on-line" test feedback on the seriousness of subsequent errors. *Journal of Educational Measurement*, *24*(2), 145-155.

Brainerd, C. J. (1972). Reinforcement and reversibility in quantity conservation acquisition. *Psychonomic Science, 27*, 114-116.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37-46. doi: 10.1177/001316446002000104.

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, *12*(11), 671-684. doi: 10.1037/h0043943.

Cronbach, L. J., & Snow, R. E. (1977). *Aptitudes and instructional methods: A handbook for research on interactions*. New York: Irvington.

DeCaro, M. S., & Rittle-Johnson, B. (2011, March). Preparing to learn from math instruction by solving problems first. In B. Rittle-Johnson & M. S. DeCaro (chairs), When are times for telling? Preparing students to learn from instruction. Symposium presented at the Society for Research in Child Development Conference, Montreal, Canada.

Hart, S. G. & Staveland, L.E. (1988). Development of NASA-TLX (Task Load Index): Results of experimental and theoretical research. In: P.A. Hancock & N. Meshkati (eds), *Human Mental Workload*, pp. 139-183. Amsterdam: North Holland.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81-112. doi: 10.3102/003465430298487.

Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 371-404). Charlotte, NC: Information Age Publishing.

Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, *42*(2), 99-107. doi: 10.1080/00461520701263368.

Homer, B. D., & Plass, J. L. (2009). Expertise reversal for iconic representations in science visualizations. *Instructional Science*, *38*(3), 259-276. doi: 10.1007/s11251-009-9108-7.

Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational Psychology Review, 19*(4), 509-539. doi: 10.1007/s10648-007-9054-3.

Kalyuga, S. (2011). Cognitive load theory: How many types of load does it really need? *Educational Psychology Review, 23*(1), 1-19. doi: 10.1007/s10648-010-9150-7.

Kalyuga, S., & Sweller, J. (2004). Measuring knowledge to optimize cognitive load factors during instruction. *Journal of Educational Psychology*, *96*(3), 558-568. doi: 10.1037/0022-0663.96.3.558.

Kalyuga, S., Chandler, P., & Sweller, J. (1998). Levels of expertise and instructional design. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *40*(1), 1-17. doi: 10.1518/001872098779480587.

Kalyuga, S., Chandler, P., & Sweller, J. (2004). When redundant on-screen text in multimedia technical instruction can interfere with learning. *Human Factors, 46*(3), 567-581.

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist*, *38*(1), 23-31. doi: 10.1037/a0022243.

Kieran, C. (1981). Concepts associated with the equality symbol. *Educational Studies in Mathematics, 12*, 317-326.

Kirschner, P. A., Ayres, P., & Chandler, P. (2011). Contemporary cognitive load theory research: The good, the bad and the ugly. *Computers in Human Behavior*, *27*(1), 99-105. doi: 10.1016/j.chb.2010.06.025.

Kluger, A. N., & DeNisi, A. (1996). Effects of feedback intervention on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, *119*(2), 254-284. doi: 10.1037/0033-2909.119.2.254.

Kluger, A. N., & DeNisi, A. (1998). Feedback interventions: Toward the understanding of a double-edged sword. *Current Directions in Psychological Science*, *7*(3), 67-72.

Kulhavy, R. W. (1977). Feedback in written instruction. *Review of Educational Research, 47*(2), 211-232. doi: 10.2307/1170128.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159-174. doi: 10.2307/2529310.

Leahy, W., & Sweller, J. (2005). Interactions among the imagination, expertise reversal, and element interactivity effects. *Journal of Experimental Psychology: Applied*, *11*(4), 266-76. doi: 10.1037/1076-898X.11.4.266.

Lorch, R. F., Lorch, E. P., Calderhead, W. J., Dunlap, E. E., Hodell, E. C., Freer, B. D., et al. (2010). Learning the control of variables strategy in higher and lower achieving classrooms: Contributions of explicit instruction and experimentation. *Journal of Educational Psychology*, *102*(1), 90-101. doi: 10.1037/a0017972.

Luwel, K., Foustana, A., Papadatos, Y., & Verschaffel, L. (2011). The role of intelligence and feedback in children's strategy competence. *Journal of Experimental Child Psychology*, *108*(1), 61-76. doi: 10.1016/j.jecp.2010.06.001.

Matthews, P., & Rittle-Johnson, B. (2009). In pursuit of knowledge: Comparing self-explanations, concepts, and procedures as pedagogical tools. Journal of Experimental Child Psychology, 104, 1-21. doi: 10.1016/j.jecp.2008.08.004.

Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? The case for guided methods of instruction. *The American Psychologist*, *59*(1), 14-9. doi: 10.1037/0003-066X.59.1.14.

McNeil, N. M. (2008). Limitations to teaching children 2 + 2 = 4: Typical arithmetic problems can hinder learning of mathematical equivalence. *Child Development, 79*, 1524-1537.

McNeil, N. M., & Alibali, M. W. (2005). Why won't you change your mind? Knowledge of operational patterns hinders learning and performance on equations. *Child Development, 76*(4), 883-899. doi: 10.1111/j.1467-8624.2005.00884.x

Mory, E. H. (2004). Feedback research revisited. In D. Jonassen, *Handbook of research on educational communications and technology: A project for the Association for Educational Communications and Technology* (2nd. pp. 745-783). Mahwah, NJ: Lawrence Erlbaum Associates.

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist, 38*, 1-4.

Perry, M. (1988). Problem assignment and learning outcomes in nine fourth-grade mathematics classes. *Elementary School Journal, 88*, 413-426.

Perry, M., Church, R. B., & Goldin-Meadow, S. (1988). Transitional knowledge in the acquisition of concepts. *Cognitive Development, 3*, 359-400.

Phye, G. D., & Bender, T. (1989). Feedback complexity and practice: Response pattern analysis in retention and transfer. *Contemporary Educational Psychology*, *14*, 97-110.

Renkl, A., & Atkinson, R. (2003). Structuring the transition from example study to problem solving in cognitive skill acquisition: A cognitive load perspective. *Educational Psychologist*, *38*(1), 15-22. doi: 10.1207/S15326985EP3801_3.

Rey, G. D., & Buchwald, F. (2011). The expertise reversal effect: Cognitive load and motivational explanations. *Journal of Experimental Psychology: Applied*, *17*(1), 33-48.

Rittle-Johnson, B., & Alibali, M. W. (1999). Conceptual and procedural knowledge of mathematics: Does one lead to the other? *Journal of Educational Psychology*, *91*(1), 175-189. doi: 10.1037/0022-0663.91.1.175.

Rittle-Johnson, B., Matthews, P. G., Taylor, R. S., & McEldoon, K. (2011). Assessing knowledge of mathematical equivalence: A construct modeling approach. *Journal of Educational Psychology, 103*(1), 85-104.

Rutherford, F. J. (1964). The role of inquiry in science teaching. *Journal of Research in Science Teaching, 2*, 80-84.

Schneider, W., Eschman, A., & Zuccolott, A. (2002). *E-Prime User's Guide*. Pittsburgh: Psychology Software Tools, Inc.

Schnotz, W. (2010). Reanalyzing the expertise reversal effect. *Instructional Science, 38*(3), 315-23. doi: 10.1007/s11252-009-9104-y.

Schnotz, W., & Kurschner, C. (2007). A reconsideration of cognitive load theory. *Educational Psychology Review, 19*, 469-508.

Schwartz, D., & Bransford, J. D. (1998). A time for telling. *Cognition and Instruction*, *16*(4), 475-5223. doi: 10.1207/s1532690xci1604_4.

Schwartz, D., & Martin, T. (2004). Inventing to prepare for future learning: The hidden efficiency of encouraging original student production in statistics. *Cognition and Instruction*, *22*(2), 129-184. doi: 10.1207/s1532690xci2202.

Schwartz, D. Lindgren, R., & Lewis, S. (2009). Constructivist in an age of non-constructivist assessments. In S. Tobias & T. Duffy (Eds.), *Constructivist Instruction: Success of Failure?* (pp. 34-61). New York: Routledge.

Schwartz, D. L., Chase, C. C., Oppezzo, M. A., & Chin, D. B. (2011). Practicing versus inventing with contrasting cases: The effects of telling first on learning and transfer. Journal of Educational Psychology, 103(4), 759-775. doi: 10.1037/a0025140.

Seo K.-H., & Ginsburg, H. P. (2003). "You've got to carefully read the math sentence…"; Classroom context and children's interpretation of the equals sign. In A. J. Baroody & A. Dowker (Eds.), *The development of arithmetic concepts and skills* (pp. 161-187). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Siegler, R. S. (1994). Cognitive variability: A key to understanding cognitive development. *Current Directions in Psychological Science, 3,* 1-5.

Siegler, R. S., & Crowley, K. (1991). The microgenetic method: A direct means for studying cognitive development. *American Psychologist*, *46*(6), 606-620.

Siegler, R. S., & Shipley, C. (1995). Variation, selection, and cognitive change. In T. J. Simon & G. S. Halford (Eds.), *Developing cognitive competence: New approaches to process modeling* (pp. 31-76). Hillsdale, NJ: Erlbaum.

Snow, R. (1978). Theory and method for research on aptitude processes. *Intelligence*, *2*(3), 225-278. doi: 10.1016/0160-2896(78)90019-3.

Snow, R. E., & Lohman, D. F. (1984). Toward a theory of cognitive aptitude for learning from instruction. *Journal of Educational Psychology*, *76*(3), 347-376.

Snow, R. E., & Swanson, J. (1992). Instructional psychology: Aptitude, adaptation and assessment. *Annual Reviews in Psychology, 43*, 583-626.

Steffe, L., & Gale, J. (Eds.). (1995). *Constructivism in education*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Stigler, J., & Hiebert, J. (1998, Winter). Teaching is a cultural activity. *American Educator*, 1-10.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*, 257-285.

Sweller, J., van Merrienboer J., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251-296.

West, S. G., Aiken, L. S., & Krull, J. L. (1996). Experimental personality designs: Analyzing categorical by continuous variable interactions. *Journal of Personality*, *64*(1), 1-48.

Wise, A. F., & O'Neill, K. (2009). Beyond more versus less: A reframing of the debate on instructional guidance. In S. Tobias & T. M. Duffy (Eds.), *Constructivist Instruction: Success of Failure?* (pp. 82-105). New York: Routledge.