From Dissimilar to Similar: Reverse Fading Assistance Improves Learning

By:

Jay JENNINGS

*A thesis submitted to the Faculty of Graduate and Post Doctoral Affairs in partial
fulfillment of the requirements for the degree of*

Master of Cognitive Science

Carleton University,

Ottawa, Ontario

©September 14, 2017

Jay JENNINGS

Abstract

This thesis investigated how worked examples could be used to fade assistance in the domain of algebra. The method of fading assistance was novel. It used similarity as the mode of assistance, with similar problem-example pairs providing high assistance and dissimilar pairs providing low assistance. Learning, performance, and gaze behaviours captured by an eye tracker were analyzed across three conditions – a faded assistance condition, a constant assistance condition, and a reverse faded assistance condition. Participant's personality traits and attitudes and behaviours regarding math were also collected and correlated with eye gaze sequences. We found that, contrary to our hypothesis, the reverse faded assistance condition resulted in the greatest learning gains. We analyzed the gaze behaviours to shed light on this finding and found that participants in this condition focused significantly more on the problem solution, suggesting more cognitive processing during problem solving than in the other conditions.

Acknowledgements

I would like to thank my parents, who forced their laid-back son to apply himself, and always encouraged me to think for myself. I would like to thank Vanessa Hendren for keeping my spirits up, for listening to me talk through problems, and for suggesting creative solutions. And thank you Kasia for being a well of knowledge, and for fostering a level of analysis and detail in both my thinking and writing I did not know I was capable of...

Contents

Abstract	i
Acknowledgements	ii
1 Introduction	1
2 Related Work	4
2.1 Cognitive Skill Acquisition	4
2.2 Using Examples: Copying and Self-Explanation	7
2.3 Problem-Example Similarity	10
2.4 Faded Assistance	15
3 Methods	18
3.1 Outline and Rationale	18
3.2 Apparatus and Materials	19
Problem-Example Pairs	19
Similar to Dissimilar	22
Similar to Similar	23
Dissimilar to Similar	24
3.3 Pre-Test and Post-Test	25

3.4	Math Surveys	26
	Math Behaviours and Interests Questionnaire	28
	Attitudes Towards Mathematics Inventory	28
	State-Trait Anxiety Inventory	29
	Achievement Goal Orientation	29
3.5	Problem Solving Interface	30
3.6	Measures	32
	Learning Measures	32
	Performance Measures	32
	Eye Gaze Measures	32
3.7	Participants	35
3.8	Design and Procedure	36
	3.8.1 Design	36
	3.8.2 Procedure	37
4	Results	39
4.1	Research Questions	39
4.2	Learning Results	41
	Regular Problems	42
	Transfer Problems	43
4.3	Performance Results	43
4.4	Dwell Time and Fixation Count on the Example and Problem	44
	Example Dwell Time	45
	Problem Dwell Time	45
	Example and Problem Fixation Count	46

4.5	Gaze Sequences	47
4.5.1	Attentional Shifts Between Lines of the Example, From Ex- ample to Problem, and From Problem to Example	48
4.5.2	Attentional Shifts Between Lines of the Problem	50
4.6	Influence of Mathematics Behaviours and Interests, Attitudes To- wards Mathematics, Anxiety, and Goal Orientation on Eye Gaze Se- quences	51
	Eye Gaze Sequences and Mathematics Behaviours and Inter- ests	55
	Eye Gaze Sequences and Attitudes Towards Mathematics: Enjoyment, Motivation, Self-Confidence, and Per- ceived Value	56
	Eye Gaze Sequences, State-Trait Anxiety, and Goal Orientation	56
	MBIQ,ATMI and Pre- and Post-Test Scores	59
5	Discussion	62
	Traits and Attitudes	69
	Limitations and Future Research	70
	Conclusions	72

List of Figures

3.1	The Interface Participants Used When Solving Problems	31
3.2	The Interface Participants Used When Solving Problems With Areas of Interest. <i>Note that the boxes around the example and problem lines were not visible to the participant and are present to illustrate the areas of interest (AOI) used for the analysis.</i>	34

List of Tables

2.1	<i>Similar and Dissimilar Structural Similarity</i>	11
3.1	<i>The Order of Rules Influences the Structural Similarity but Maintains the Conceptual Knowledge Needed to Solve the Problem</i>	21
3.2	<i>S=Similar, D=Dissimilar. Sequencing of Problems by Condition</i>	23
3.3	<i>Examples of Transfer in the Pre- and Post-Test</i>	27
4.1	<i>Means and Standard Deviations for Each Condition for Pre-Test, Post-Test, and Difference Scores (for All Problems, Regular Problems, and Transfer Problems)</i>	41
4.2	<i>Average Time to Complete the Learning Phase by Form of Assistance (min- utes)</i>	44
4.3	<i>Total Dwell Time for Example Area and Problem Area by Condition in Minutes</i>	45
4.4	<i>Fixation Count for Example and Problem by Condition</i>	47
4.5	<i>Attentional Shifts Between Lines of the Example, From Example to Prob- lem, and From Problem to Example by Level of Expertise and Form of As- sistance</i>	49

4.6	<i>Attentional Shifts Between Lines of the Problem for Level of Expertise and Form of Assistance</i>	50
4.7	<i>Correlations Between Individual Traits</i>	53
4.8	<i>Correlations for MBIQ and Sequence Fixation Count</i>	54
4.9	<i>Correlations for ATMI and Sequence Fixation Count</i>	58
4.10	<i>Correlations for STAI, Goal Orientation and Sequence Fixation Count . . .</i>	60
4.11	<i>Correlations of MBIQ, ATMI, and Pre-Test and Post-Test Scores . . .</i>	61

List of Abbreviations

ATMI	Attitudes Towards Mathematics Inventory
MBIQ	Math Behaviours (and) Interests Questionnaire
STAI	State Trait Anxiety Inventory Abbreviated
AOI	Area Of Interest

Dedicated to Vanessa Hendren...

Chapter 1

Introduction

Worked examples (hereafter referred to simply as examples) are problems where the solution is given step-by-step to the student. Examples help students learn underlying domain concepts (Richland, Zur, and Holyoak, 2007), improve performance on transfer tasks (i.e., problems that require application of knowledge in a new context) (Cooper and Sweller, 1987; Lee, Betts, and Anderson, 2015), and are more beneficial than reading general instructional text (Nokes-Malach et al., 2013). Moreover, examples can improve the speed at which students learn by as much as twenty percent (McLaren, Lim, and Koedinger, 2008; Sweller and Cooper, 1985). However, the benefits of examples also depend on how they are used and the degree of assistance they provide, as we describe shortly.

Early research focused on the benefits of examples in the absence of problem solving activities, where problem-solving activities were replaced by example study (Sweller and Cooper, 1985). Less research exists investigating which factors influence learning when examples are used to aid problem solving (where students are given problems to solve as well as examples to refer to when they require assistance). This latter context is the focus of the present thesis.

When students use examples to aid problem solving, two factors influence learning outcomes. The first factor is how students use examples. For instance, students learn less if they simply copy the example solution instead of trying to learn the underlying domain concepts. Copying directly from the problem allows a shallow procedure to develop for attaining the correct answer if the example and problem are highly similar, but does not encourage learning of the underlying domain rules (Chi et al., 1989; VanLehn, 1999; Rittle-Johnson and Schneider, 2014). The second factor is the similarity between the examples and problems. On the one hand, examples that are highly similar to the problem provide assistance to the student, but also promote copying, which is detrimental to learning. On the other hand, examples that include some differences to the problem challenge students to try to learn the underlying concepts instead of copying, which promotes learning (Muldner and Conati, 2010). Recently, when Lee, Betts, and Anderson (2015) presented students with superficially similar examples, the students directly compared them to the problem and copied liberally. However, if they were presented with superficially dissimilar examples, the students studied the example to understand it, and then applied that understanding to the problem, which ultimately led to better learning and transfer.

Thus, the level of similarity between a problem and an example impacts the level of assistance the example provides (i.e. low similarity corresponds to low assistance, high similarity to high assistance). There is some evidence that assistance works best when it is high at the beginning of the learning phase and low at the end of the learning phase, which is referred to as faded assistance (Kalyuga et al., 2003; Tullis, Goldstone, and Hanson, 2015). To date, however, it was an open

question of how varying the problem-example similarity within a faded assistance paradigm over time affected performance and learning.

Specifically, the present research evaluated how the systematic manipulation of problem-example similarity impacted learning (pre to post-test gains), performance (time on task), and behaviours during problem solving and example usage in the context of algebra problems. To obtain fine grain information on behaviours during problem solving, we used eye-tracking methodology. Prior to describing the proposed methodology, we present the related work.

Chapter 2

Related Work

2.1 Cognitive Skill Acquisition

When students are first learning to solve problems in a given domain they typically require some form of assistance. One conventional approach is to provide a lecture on the target topic and then have students solve problems related to that lesson. However, this approach provides little assistance for the problem solving aspect, as students who are unfamiliar with the domain resort to general search strategies, such as envisioning the goal state of the problem and taking steps to reach that goal (Tricot and Sweller, 2014). The issue with this strategy is there are many possible paths that can be taken in the search for the solution and therefore these search strategies are highly taxing on cognitive resources (Sweller and Cooper, 1985; Tricot and Sweller, 2014). To alleviate this cognitive load, a common type of assistance provided to students is in the form of worked examples, which are examples where the solution is given step-by-step to the student.

Studying examples and solving the associated problems can increase procedural knowledge, which is generally defined as knowing how to do something,

and characterized by the application of an algorithm or selecting from a subset of rules, that if applied in the correct order, will generate a correct solution to a problem (Rittle-Johnson and Star, 2007). On the other hand, conceptual knowledge is knowledge about facts and generalizations (Rittle-Johnson and Star, 2007). Early work considered procedural knowledge to be superficial and purely algorithmic whereas conceptual knowledge was considered to be deep and richly connected (Hibert and Lefevre, 1986). However, there is mounting evidence that conceptual and procedural knowledge influence each other extensively (Rittle-Johnson, 2006). In research into the influence of examples on learning, the conceptual and procedural distinction is rarely explicitly made. This may be due to the overlap between the two types of knowledge - for instance, one conceptual piece of knowledge in the context of algebra that is needed relates to the order of operations but this is also a procedure that you apply. Star (2005) argues that this is conceptual knowledge about procedures, highlighting the overlap and sometimes ill-defined boundary between the two types of knowledge. Furthermore, research has shown that increases in procedural knowledge can lead to increases in conceptual knowledge, and vice versa, suggesting that the procedural / conceptual knowledge dichotomy is not as rigid as previously believed (Star, 2005; Rittle-Johnson and Star, 2007; Star and Stylianides, 2013).

Regardless of whether examples are teaching procedural or conceptual knowledge, it is evident that they are highly beneficial for learning (McLaren, Lim, and Koedinger, 2008; Renkl, 2014; Richland, Zur, and Holyoak, 2007; Sweller and Cooper, 1985). However, this benefit is confined to the early and intermediate stages of cognitive skill acquisition (Atkinson et al., 2000; Kalyuga et al., 2003)

because students primarily need the most assistance early on when learning new concepts in a given domain (Chen, Kalyuga, and Sweller, 2015). In the early stages, when students first learn about the domain, studying worked examples is more beneficial than solving problems (Anderson, Fincham, and Douglas, 1997; Kalyuga et al., 2003; VanLehn, 1996). This is because studying examples reduces cognitive load by providing scaffolding for students to solve the problems and by directing students' attention to the important aspects of a problem. Thus, examples free cognitive resources needed to learn the relevant information that would otherwise be expended on general search strategies (Leahy, Hanham, and Sweller, 2015; Sweller and Cooper, 1985).

As students progress in their skill acquisition, problem solving becomes beneficial for learning as the examples have allowed the corresponding domain knowledge to be internalized, at which point the active application of that new found knowledge becomes important (Kalyuga et al., 2003). The transition from benefiting from pure example study to benefiting from problem solving is referred to as the expertise reversal effect, as the students now possess some background knowledge in the domain and they benefit from applying that knowledge (Kalyuga et al., 2003). However, at this intermediate point of skill acquisition students still require assistance, because they have not yet learned all the necessary concepts (i.e., they have knowledge gaps). Thus, examples are beneficial in the intermediate stage of skill acquisition by providing assistance to the problem solving process (Lee, Betts, and Anderson, 2015; Muldner and Conati, 2010; VanLehn, 1996). In the intermediate stage, students benefit from examples because the example provides an overall solution plan, allowing students to focus their attention on the essential aspects

of the problem rather than engage in ineffective search strategies. Examples also provide the student with the opportunity to refine their knowledge of the domain principles needed to generate the example solution, because they illustrate the application of these principles. When students use examples, they are learning the proper algorithms and sequences needed to solve other problems in that domain (Byrnes and Wasik, 1991). Furthermore, by practising these skills by solving problems, students are possibly gaining conceptual knowledge in the domain, which further increases their ability to apply their new knowledge to new problems (Baroody, 2003; Rittle-Johnson and Siegler, 1998; Rittle-Johnson, Siegler, and Alibali, 2001; Star and Stylianides, 2013).

To summarize, examples provide assistance to problem solving and are effective at bridging the gap between what students have learned in the early stage of cognitive skill acquisition and the application of that knowledge (Chen, Kalyuga, and Sweller, 2015; Renkl, 2014). However, the benefits of worked examples also depend on two factors: (1) how students use examples and (2) on the similarity between the problem and example.

2.2 Using Examples: Copying and Self-Explanation

The different strategies that students employ when solving problems in the presence of examples impacts their learning outcomes (Muldner and Conati, 2010; Roy and Chi, 2010; Chi et al., 1989). There are two major strategies that can impact

learning outcomes when worked examples are used, namely copying and self-explanation. Copying refers to using the example to solve the problem by transferring the example solution over to the problem solution. If the problem-example pair is similar, students can obtain the correct solution by copying, without learning the *if-then* rules needed to generate the steps they are copying on their own (Anderson, Fincham, and Douglas, 1997; VanLehn, 1999; Chi et al., 1989). Therefore, copying from an example can generate a correct solution to a problem, but the rules that are applied are not necessarily learned by the copier. On the other hand, self-explanation involves students explaining concepts to themselves by generating inferences that go beyond the surface features of the instructional material. Instead of directly copying, self-explainers describe why the rule they are applying makes sense, and then apply it. Thus, self-explanation has been shown to increase procedural transfer and improve conceptual knowledge in separate experiments (Berthold and Renkl, 2009; Rittle-Johnson, 2006).

The distinction between copying and self-explanation is critical, as example solutions typically include gaps (i.e. they fail to comprehensively explain the derivation of the target solution step) (Chi et al., 1989). The act of filling in the gaps with self-explained content increases memory retention (Roy and Chi, 2010), and self-explanation makes students more active in the learning process, which helps them learn the target concepts in the domain (Chi et al., 1989; Chi and Wylie, 2014), as well as the procedures needed to generate a correct solution in unfamiliar problems (Rittle-Johnson, 2006). In general, self-explanation is highly beneficial for learning (Atkinson, Renkl, and Merrill, 2003; Chi et al., 1989; Muldner, Burleson, and Chi, 2012). For instance, a recent study demonstrated that a self-explanation

group and an example group in an introductory physics course were significantly better at far transfer tasks than a condition that involved only reading the instructional material (Nokes-Malach et al., 2013). This is because a student using self-explanation can reinforce an existing procedure, and more importantly, can use it to infer a new rule. A student that explains why something happened, for example in an algebra problem, is essentially hypothesizing a rule. If an opportunity to apply the rule is encountered again, the student has already internalized it through self-explanation and can apply it to the new problem, thus strengthening their domain knowledge.

Unfortunately, many students lose learning opportunities because they copy too much and self explain too little (Chi et al., 1989; Muldner, Burleson, and Chi, 2012; Renkl et al., 1998; VanLehn, 1999). For instance, VanLehn (1999) evaluated a computational model named Cascade that simulated how students solved physics problems with access to examples. As part of that evaluation, VanLehn assessed protocols of students learning in this context (problem solving with access to examples). He found that out of 81 opportunities to learn, students only took advantage of 11 opportunities. Of the remaining opportunities, in 59 instances, students chose to copy the example solution and did not self-explain the corresponding steps, consequently missing opportunities to learn. Furthermore, students tended to copy from the example rather than solve the problem themselves, which did not foster deep processing or the acquisition of the underlying concepts. VanLehn (1999) concluded that if the students had used self-explanation rather than copy from examples, the amount of learning would be more than double.

Thus, the evidence for the use of examples as an effective learning tool is not

as straightforward as initially perceived. In particular, the benefit of examples depends on individual preferences for deep reasoning through self-explanation or a preference to copy instead (Chi et al., 1989; Chi and Wylie, 2014; VanLehn, 1999). In addition to student behaviours, another factor that impacts learning is the similarity between the problem and example, in part because it influences behaviours like copying and self-explanation, as we now describe.

2.3 Problem-Example Similarity

Problem-example similarity is characterized as superficial or structural (Muldner and Conati, 2010; Novick, 1988; Reed, Ackinclose, and Voss, 1990; Ross, 1987). Superficial similarity is characterized by the similarity between the constants making up the example and problem, but is not part of the knowledge needed to generate the solutions (see Problem 2 in Table 2.1). Structural similarity on the other hand refers to the type of conceptual knowledge (i.e., rules) needed to generate the solution, and the specific order of operations used in the example and the problem (see Problem 1 in 2.1). For instance, Problem 1 in Table 2.1 shows a problem-example pair where the superficial similarity between the problem-example constants is low (i.e., they are dissimilar) as the two problems only differ in terms of the variable names (e.g., in the problem the student solves for “ x ” vs. for “ a ” in the example). Furthermore, Problem 1 in Table 2.1 is structurally identical to the example because they both involve the same rules to generate their solutions and the same order of application of those rules. In contrast, Problem 2 in Table 2.1 is superficially similar to the example, but it is structurally dissimilar (i.e., the order of rules

are reversed).

TABLE 2.1: *Similar and Dissimilar Structural Similarity*

	Problem 1	Problem 2
Example	Solve for x :	Solve for a :
Solve for a :	Structurally Similar	Structurally Dissimilar
$b = (ad)/c$	$y = (xw)/z$	$b = d (a/c)$
$bc = ad$	$yz = xw$	$b/d = a/c$
$(bc)/d = a$	$(yz)/w = x$	$c (b/d) = a$

As illustrated in Table 2.1, similarity can vary in different ways. This varied similarity exerts a vital influence on how students use examples and what they learn from them. If examples are too similar to the problem in terms of both superficial and structural features, they do not improve learning, as they allow for copying and fail to encourage students to self-explain (Lee, Betts, and Anderson, 2015; Muldner and Conati, 2010; Reed, Dempster, and Ettinger, 1985). If examples are too different, this can also be problematic, as the learner may be unable to uncover the structural features embedded in the example (Gick and Holyoak, 1983; Richland, Holyoak, and Stigler, 2004). In particular, the impact of differences has to be considered in light of the type of difference (superficial or structural) and the impact on assistance provided by the example (Muldner and Conati, 2010). We describe work related to this issue as well as the broader question of the impact of similarity below.

In an early study by Novick (1988) that investigated the impact of problem-example similarity, both experts and novices were given examples to aid their problem solving in a learning phase and were then tested on transfer problems

in a testing phase. Experts performed better on the transfer tasks when they were given examples that were structurally similar to the problems, but superficially dissimilar. In contrast, novices performed the worst when the examples they were given were superficially similar but structurally dissimilar, indicating that they were persuaded by the superficial similarity of the examples to believe that the correct solution to the problem was present in the example. In an earlier study, Reed, Dempster, and Ettinger (1985) showed that students were unable to use the solution of an algebra word problem on equivalent problems when they had not seen how the problem was solved step-by-step. But when they were allowed to see more complete solutions or consult the solutions while solving the problem, students were better able to solve equivalent problems, but were still unable to solve slightly dissimilar problems. Thus, even when examples are similar to the problems, students may still have difficulty applying the solutions of the examples to solve the problems.

One type of problem-example structural difference is especially problematic – when the example contains fewer steps and corresponding knowledge than the problem requires (i.e., example solution is not inclusive of the problem solution). Reed, Ackinclose, and Voss (1990) compared this scenario against one where the example contained more solution steps and corresponding knowledge than the problem requires (i.e., example solution is inclusive of the problem solution). They found that students performed better when given inclusive examples because it was easier for the student to filter out the extra steps, as opposed to generate content they did not know if the example was not inclusive.

Regardless of inclusiveness, it had been hypothesized that any kind of difference hindered learning, but what was not clear from the early work on problem-example similarity was whether some differences were actually beneficial for learning, for instance because they discouraged ineffective behaviours like copying. To determine whether problem-example differences were beneficial, Muldner and Conati (2010) manipulated surface features between the examples and the problems but kept the similarity constant at the structural level. The superficial differences were introduced following an algorithm in a way as to block shallow strategies (e.g., transforming the example solution to one needed for the problem by replacing surface constants), thus discouraging copying and encouraging self-explanation. This mechanism was integrated into an intelligent tutoring system (EA-Coach), and tested by having the system present students with two types of problem-example pairs: (1) ones that were both superficially and structurally similar, vs. (2) ones that were superficially different but structurally similar. The results showed that superficial differences improved self-explanation and reduced copying.

Along the same line, Lee, Betts, and Anderson (2015) investigated the impact of superficial differences on students' ability to generate the solution to a problem (i.e., performance). The domain used in the study was similar to algebra, but was novel to all participants because it involved the use of algebra knowledge in the form of data-flow diagrams. Similar to the work described above, students were presented with problem-example pairs simultaneously so they had opportunities to compare problem-example solutions. Participants first solved problems with access to examples (Phase 1) and then were required to complete transfer

problems without the use of examples (Phase 2). The findings showed that when problem-example pairs were superficially similar, participants performed poorly on transfer tasks. It was argued that poor performance on transfer tasks was because the students could copy the material from Phase 1 with minimal changes. Furthermore, participants in the dissimilar example condition performed better on the transfer tasks, potentially using more self-explanation principles to understand the underlying concepts during the learning phase which could then be applied in the testing phase (although self explanation was not assessed directly in this study). Interestingly, this effect was not replicated when the example was presented after the problem instead of simultaneously as the problem. Therefore the simultaneous presentation of the problem and the example was a critical element driving the benefit of dissimilar examples. By presenting the example alongside the problem, the problem-example differences encouraged self-explanation via active comparison to understand how to apply the example solution to the problem (although this was not directly measured in this study).

To summarize, early work showed that problem-example differences hindered learning and performance – from an assistance standpoint, one could argue that differences diminish the assistance provided by examples. However, researchers have recently highlighted the potential benefits of certain problem-example differences. We now explore this concept in the context of faded assistance.

2.4 Faded Assistance

The review above has highlighted that examples are a critical element of cognitive skill acquisition, as students need assistance when learning and examples provide that assistance. As indicated earlier, assistance is necessary during the initial stages of cognitive skill acquisition, but becomes less critical as expertise is gained in a particular domain (Kalyuga et al., 2003), because students have internalized the conceptual knowledge and the procedural rules thereby rendering assistance redundant (Canobi, 2009). For instance, a recent study on tutoring used a novel mouse task to demonstrate that the type of assistance provided to participants, whether it was faded, constant, or no assistance, affected their performance (Tullis, Goldstone, and Hanson, 2015). Tullis, Goldstone, and Hanson (2015) found that assisting participants for the entirety of the knowledge acquisition phase resulted in fewer errors during learning, but during the testing phase in which assistance was removed, the participants made significantly more errors than all other groups. However, if assistance was provided initially and then slowly faded to no assistance during the learning phase, participants made less errors in the testing phase when assistance was removed. They found that faded assistance resulted in the best performance in the testing session.

Cognitive load theory can explain why fading assistance during the learning process had the highest performance gains in the testing phase. In the context of worked examples, we have already described that students initially experience high cognitive load during pure problem solving (without access to examples) due to the amount of new information (i.e. rules) that they must acquire. However, the expertise reversal effect indicates that at some point examples become redundant

and their assistance should be faded - Kalyuga et al. (2003) argue that this is because the redundancy introduces a new form of cognitive load that interferes with learning. While this is a possible explanation, an alternative one is that students at some point simply benefit from applying knowledge in the form of problem solving, rather than studying examples. Therefore, applying the procedures learned from the examples solidifies the conceptual principles, especially if the problems are conceptually sequenced (i.e. the order of the problems intentionally illustrates a conceptual rule, such as equivalence in mathematics) (Canobi, 2009). Thus, by fading assistance with examples, students are encouraged to gradually shift over to problem solving, thereby increasing their conceptual knowledge through applying their procedural knowledge (Star, 2005).

How should assistance from examples be faded? One option is to introduce gaps into examples that represent areas where students are encouraged to self-explain. This has been done previously by removing certain step(s) from an example, while the rest of the example is still present, similar to a fill-in-the-blank question (Renkl and Atkinson, 2003; Schwonke et al., 2007). With this method, students fill in the gaps with their own knowledge, essentially generating the solution steps corresponding to the gaps, and are guided by the previous and subsequent steps provided in the example. By sequentially increasing the number of gaps in the example, assistance is faded until the example is no longer present and students are problem solving unassisted.

An alternative, as yet unexplored method of fading assistance is to gradually alter the similarity of the example to the problem, over the course of multiple

problem-example pairs. The present research did this by simultaneously presenting the example and the problem together so that the examples provided a form of assistance. To address the finding that high levels of similarity hurt learning (Lee, Betts, and Anderson, 2015; Muldner and Conati, 2010), the level of similarity was faded, which required that we identified examples that provided some assistance. We now describe our research to test the benefits of problem-example similarity as a form of faded assistance.

Chapter 3

Methods

3.1 Outline and Rationale

As described above, it is an open question of how varying the problem-example similarity within a faded assistance paradigm over time affects learning, performance, and eye gaze behaviours. Based on prior findings, students with low domain knowledge may initially require high levels of assistance, in the form of similar problem-example pairs. However, once some domain specific concepts are acquired, the problem-example pairs should become less similar, challenging the student to self explain and extract the rules embedded in the examples, thereby increasing learning of domain specific conceptual knowledge. Because the focus of this research is on how students use examples while solving problems, to obtain attention-related information, this research relied on eye tracking. This made it possible to analyze gaze information that could be indicative of copying (many gaze shifts from example to problem), referencing the example (many gaze shifts

from problem to example), studying the example (long dwell times on the example and more gaze shifts between lines of the example), and problem solving behaviours (long dwell times on the problem and more gaze shifts between lines of the problem).

3.2 Apparatus and Materials

Problem-Example Pairs

Given the research goals, the target domain had to support a precise characterization of problem-example similarity, both in terms of superficial features and structural features. The latter required identifying the knowledge needed to solve the target problems and examples (i.e., to identify structural similarity between the two). A domain that met these requirements was algebra, and so was the chosen domain.

As a first step in preparing the materials (i.e., problems for students to solve and worked examples), a cognitive task analysis was performed that involved identifying the knowledge needed to solve a preliminary set of algebra problems and examples. We operationalized this domain knowledge at a high level in the form of if-then rules. The *if* part corresponded to an algebraic expression that represented the rule's antecedent (i.e., condition needed for the rule to fire) while the *then* part corresponded to the rule's consequent (i.e., a solution step produced by the rule by manipulating the antecedent expression). Thus, each rule that fired generated a step in the problem or example solution – with the application of the

sequence of rules firing resulting in the final solution. Table 3.1 shows an example of a problem and the corresponding domain rules. For instance, to solve for x in the problem $y = (xa)/z$, the student must first multiply both sides by z to get $yz = xa$, then divide both sides by a to get $(yz)/a = x$ (see Table 3.1 for a detailed explanation).

The previous problem was labelled as Multiply-Divide (MD), because its solution required two rules, one involving multiplication of the equation by the appropriate variable, and one involving division, applied in that order. Simply by using four different rules and varying the number of them needed to solve the problem (two step solutions, three step solutions, and four step solutions), as well as the ordering of the rules, we generated 60 different problems. To avoid a ceiling effect, we determined the difficulty of these problems by conducting a pilot study where eleven participants were given 26 of these problems to solve. The participant average was at fifty percent, with variability in individual responses ($M = 13.05$, $SD = 9.57$). Furthermore, for problems with two rules, the average was above sixty percent, but decreased to fifty percent for problems with three rules and to around forty percent for problems with four rules. Thus, to ensure that the problems were difficult enough to enable learning, the eye tracking portion of the experiment contained only three and four rule problems.

The rules and their order of application correspond to the structural aspect of problem-example similarity. Thus, once the algebra solutions were characterized by the rules needed to generate them, it was possible to manipulate structural similarity between the problems and examples. The present study focused on a specific

TABLE 3.1: *The Order of Rules Influences the Structural Similarity but Maintains the Conceptual Knowledge Needed to Solve the Problem*

Solve for x	Rules Needed
$y = (xa)/z$ $yz = xa$ $yz/a = x$	<p>Multiply: if the variable to solve for (x), or the compounded variable containing the variable to solve for, is divided by a variable (e.g. z), multiply both sides of the equation by the divided variable (z).</p> <p>Divide: if the variable to solve for (x), or the compounded variable containing the variable to solve for, is multiplied by a variable (e.g. a), divide both sides of the equation by the multiplied variable (a).</p> <p>Goal has been reached</p>
$y = z(x/a)$ $y/z = x/a$ $a(y/z) = x$	<p>Divide : if the variable to solve for (x), or the compounded variable containing the variable to solve for, is multiplied by a variable (e.g. z), divide both sides of the equation by the multiplied variable (z).</p> <p>Multiply : if the variable to solve for (x), or the compounded variable containing the variable to solve for, is divided by a variable (e.g. a), multiply both sides of the equation by the divided variable (a).</p> <p>Goal has been reached</p>

type of structural difference: the order of operations (i.e., rules) needed to generate the solutions. The first row of Table 3.1 shows an example that requires the exact same two rules as the example in the second row of Table 3.1. The difference between them is that the rules need to be applied in a reverse order in the second row to generate a correct problem solution. Table 3.1 illustrates how the structural similarity was manipulated in the present study, while holding constant the actual rules needed to generate the problem-example solutions – the latter is important because if the examples are not inclusive (i.e. they are missing information) students will not be able to use the example to learn the rules embedded therein and apply that knowledge to future problems (Reed, Ackinclose, and Voss, 1990).

We designed two sets of problem-example pairs (i.e., similar or dissimilar pairs) to be used in the three study conditions: similar to similar (Sim-Sim), similar to dissimilar (Sim-Dissim), and dissimilar to similar (Dissim-Sim). For all three sets, to avoid the example being exactly the same as the problem, the superficial similarity of the problem-example pairs were varied by using different names for the variables. Thus, the key difference between the three conditions corresponded to the level of structural similarity, as described below. A Dissim-Dissim condition was considered, but previous research in both the faded assistance literature and the productive failure literature agree that some form of assistance is needed to improve learning, thus it was not included as a condition (Kapur, 2014; Lee, Betts, and Anderson, 2015; Atkinson, Renkl, and Merrill, 2003).

Similar to Dissimilar

The Sim-Dissim set provided faded assistance, because the order of the rule application was the same in the initial portion (identical structural similarity), but was manipulated to be different towards the end portion (dissimilar structural similarity). Thus, structural similarity was faded in this condition as students progressed through the problems. The heuristic that was chosen to alter structural similarity and thus make the problem-example pairs structurally different was to simply reverse the order of rules in the example needed to solve the problem. For instance, if the *example* rules were ordered Multiply-Divide-Subtract-Add, the *problem* would be of the form Add-Subtract-Divide-Multiply. This simple reversal changed how the problems appeared structurally, while simultaneously containing the same rules. This meant that if the participant studied the example and learned the rules,

they could then apply them to the problem to generate the correct solution (in other words, all the information needed to solve the problem was available within the example). The ordering of similar to dissimilar problem-example pairs is shown in Table 3.2.

Similar to Similar

The Sim-Sim set provided constant, high assistance, because the problem-example pairs in this condition maintained the same order of operations between the example and the problem (see Table 3.2 below). Thus, the problem-example pairs were structurally identical in this condition, meaning that the order of operations that needed to be applied to solve the problem was identical to the order in the example. In this condition, participants only had to resolve the superficial similarity between the example and the problem to find a solution.

TABLE 3.2: *S=Similar, D=Dissimilar. Sequencing of Problems by Condition*

Condition	Type/Number	Ordering of the Problem-Example Pairs											
Sim-Dissim	Type of Problem	S	S	S	D	S	D	S	D	S	D	D	D
	Number of Rules	3	3	4	3	4	4	3	3	4	3	4	4
Sim-Sim	Type of Problem	S	S	S	S	S	S	S	S	S	S	S	S
	Number of Rules	3	3	4	3	4	4	3	3	4	3	4	4
Dissim-Sim	Type of Problem	D	D	D	S	D	S	D	S	D	S	S	S
	Number of Rules	3	3	4	3	4	4	3	3	4	3	4	4

Dissimilar to Similar

The Dissim-Sim condition was a reversal of the Sim-Dissim condition with regards to ordering, providing a reverse-faded assistance sequence in terms of problem-example similarity. Specifically, this condition started with structurally dissimilar problem-example pairs, and faded to structurally similar problem-example pairs. While this presentation may seem counter-intuitive at first, there is evidence that struggling initially in learning activities may be beneficial to learning (Kapur, 2014). Allowing students to struggle and explore unfamiliar problems can increase both conceptual and procedural knowledge, and this may be due to an increased level of self-explanation during the exploration phase (Schwartz et al., 2011). It can also be highly beneficial to allow students to try to figure out new problems on their own, and then provide a lesson on the proper way to do it after they have explored (DeCaro and Rittle-Johnson, 2011). On the other hand, most schools and other instructional activities start with simpler problems and move to more complex ones, with reasonable rates of success among the students. Thus, it is an open question as to how reverse fading similarity in this condition will impact learning, performance, and behavioural outcomes.

Each of the three conditions included 12 problem-example pairs, and within the faded assistance and reverse faded assistance conditions (Sim-Dissim and Dissim-Sim), an equal number of similar and dissimilar three and four rule problems were presented (three for Similar-3-rule problems, three for Similar-4-rule problems, three for Dissimilar-3-rule problems, and three for Dissimilar-4-rule problems). The Sim-Sim condition had an equal number of three and four rule problems. Table 3.2 shows the progression of problem-example similarity in each condition. This

number was determined by conducting a pilot study, and represents a balance between reducing participant fatigue from sitting still too long in an eye tracker chin rest, while providing sufficient problem solving opportunities.

3.3 Pre-Test and Post-Test

In addition to the problems and examples, the materials also included a pre-test and post-test that was used to assess learning gains. The pre and post-tests provided problems for participants to solve but did not include examples. This was done in order to assess the participants' ability to solve problems in the absence of assistance (i.e., examples) and ultimately to assess any potential learning outcomes. The two tests included the same problems, though variable names and the order of the problems were changed.

For both the pre-test and the post-test, two types of test problems were generated. The first set required the same rules as the problems and examples used in the learning, or eye tracking, phase of the experiment. Thus, these problems were similar to the learning phase, though were not exact duplications, and are referred to as regular problems below. Second, we added questions to assess the transfer of knowledge, referred to as transfer problems below. Transfer questions either required the application of rules a maximum of five times, rather than a maximum of four in the learning phase, or understanding that a combined term (such as $(a + z)$) could be treated as a single term. The latter portion of our definition of transfer additionally included cancellation problems that required the inference of the new rule (i.e., cancelling terms), as it was not seen in the learning phase and had to be

inferred from the underlying rules. This could have been accomplished by self-explaining the underlying concepts to oneself and realizing that the variables are simply place holders, which would allow the cancellation rule to become salient. Table 3.3 shows examples of each type of transfer.

Thus, we define transfer as problems that require additional application of previously seen rules (e.g., problems requiring five steps instead of four for their solution), treating a combined term as a single item (applying rules to $(x-y)$ just as one would to x when possible), and that a term could be removed if the opposite of it was present ($(x + y) - (x + y)$, i.e. cancelling). Also, the application of rules to a problem that had more than a single variable on the left side of the equation was included in the definition of transfer, as the participant had to understand that extra items did not change the domain rules ($yv = (ax + b) + c$ requires the same rules to solve if only y was on the left side of the equation).

3.4 Math Surveys

A composite questionnaire gathered data about math attitudes, behaviours, and character traits. It was comprised of the Math Behaviours and Interests Questionnaire (MBIQ), the Attitudes Towards Mathematics Inventory (ATMI), the State-Trait Anxiety Inventory - Abbreviated (STAI), and an Achievement Goal Orientation survey.

TABLE 3.3: Examples of Transfer in the Pre- and Post-Test

Initial State	Transfer State
<p>Application of 4 rules a maximum of 4 times (in the order Add-Multiply-Divide-Subtract).</p> $y = ((a (x + t)) / z) - c$ $y + c = (a (x + t)) / z$ $z (y + c) = a (x + t)$ $z (y + c)/a = x + t$ $z (y + c)/a - t = x$	<p>Application of 4 rules a maximum of 5 times (in the order Multiply-Divide-Add-Subtract-Multiply).</p> $yv = (a ((x/q) + t) - c) / z$ $zyv = a ((x/q) + t) - c$ $zyv/a = (x/q) + t - c$ $(zyv/a) + c = (x/q) + t$ $(zyv/a) + c - t = x/q$ $q ((zyv/a) + c - t) = x$
<p>Applying a rule moves an item from one side of the equation to the other (Notice the variable c after application of the rule Subtract).</p> $y = a(x/z) + c$ $y - c = a(x/z)$	<p>Applying a rule can move compound items from one side of the equation to the other (Notice the compound variable (w/z) after application of the rule Add).</p> $j + m = y ((aj - c) / g) - (w/z)$ $(j + m) / y = ((aj - c) / g) - (w/z)$ $(j + m) / y + (w/z) = (aj - c)/g$
<p>Variables are essentially place holders, and can be operated on as if they were numbers, if the variables are consistent.</p> $abc - abc = 0$	<p>If a variable is repeated twice on one side of the equation with opposite operators, then it can be cancelled out.</p> $y + (z - t) = ((x/a) + c) + (z - t)$ $y + (z - t) - (z - t) = (x/a) + c$ $y = (x/a) + c$ $y - c = x/a$ $a (y - c) = x$

Math Behaviours and Interests Questionnaire

The mathematics and personal information sections of the *Math Behaviours and Interests Questionnaire (MBIQ)* was used, as we omitted the reading and writing section. The survey asked five overarching questions about a participants experience and feelings about math (e.g. "How often do you avoid situations involving mathematics?"). The participants responded using a 7 point likert scale, which changed depending on the type of question (e.g. *Very Low* to *Very High* or *Almost Always* to *Almost Never*).

Attitudes Towards Mathematics Inventory

The short *Attitudes Towards Mathematics Inventory (ATMI)* is an abbreviated version of the ATMI (Lim and Chapman, 2013), which measures four subscales - *Enjoyment of Mathematics*, *Motivation to do Mathematics*, *Self-Confidence in Mathematics*, and *Perceived Value of Mathematics*. The short ATMI has been shown to be highly correlated with the full form ATMI overall ($r = .96$), and each subscale was highly correlated with the long-form as well ($r = .87$). The short ATMI consisted of a variety of statements that were rated by the participant on a 6 point Likert scale, from *Strongly Disagree* to *Strongly Agree*. The enjoyment subscale comprised 10 statements and evaluated how much a participant enjoys mathematics. The motivation subscale consisted of 5 statements and assessed the participants' interest in mathematics, and their desire to learn more mathematics. The self-confidence subscale consisted of 15 statements and assessed a participants confidence in their own abilities in mathematics. Lastly, the perceived value subscale consisted of 10 statements that

indicated the participants beliefs in the usefulness or relevance of mathematics in their lives.

State-Trait Anxiety Inventory

The *State-Trait Anxiety Inventory (STAI) Abbreviated* is a six-item short-form version of the STAI that measures state anxiety. The STAI-Abbreviated is highly correlated ($r = .91$) with the full-form STAI and has good internal reliability ($\alpha = .82$) (Marteau and Bekker, 1992). The survey consists of six statements, such as "I am worried" and "I feel content". Participants are then asked to indicate their level of agreement with each statement on a four-point scale. Half of the survey questions correspond to positive states, and the other half correspond to negative states. Scoring is done by assigning numbers to the participant responses, with 1 for *Not at all*, 2 for *Somewhat*, 3 for *Moderately*, and 4 for *Very Much*. These are then added for the three negative state scores, and the positive state scores are reverse scored. These are then combined to attain a total level of state anxiety. To be equivalent to the unabbreviated scale, total scores are multiplied by 20, then divided by 6. Scores range from 20 to 80 points, with a higher score indicating higher levels of anxiety (Marteau and Bekker, 1992).

Achievement Goal Orientation

Additionally, thirteen questions regarding achievement goal orientations were included in the questionnaire as well. They were modified from Middleton and Midgley (1997) to pertain to university students, and the questions were altered to refer to academic work in general, as opposed to only math. The three types of

achievement goals that are assessed in these questions include two performance oriented goals (Approach and Avoid), and a task oriented goal. The performance-approach orientation refers to a student's goal to perform well, be judged as knowledgeable by others, and ultimately to attain success in the subject. The performance-avoid orientation is a student's goal to avoid looking unintelligent, to not be judged poorly by others, and ultimately to avoid failure. The task-goal orientation is a student's intrinsic motivation to succeed and gain knowledge in the subject (this is also referred to as mastery goal orientation). Notice that the performance orientations are externally motivated, while the task orientation is internally motivated.

3.5 Problem Solving Interface

An SR Research Eye Link 1000 eye tracker was used to capture gaze and fixation data. To solve problems and refer to examples, students used a Java program created for the present research - as shown in Figure 3.1, the program presented the example on the left side of the screen and the unsolved problem on the right side of the screen. The program logged problem solving actions, and also allowed us to present the stimuli in a manner that minimized participant movement. In particular, because the study involved eye-tracking, it was important to avoid using a physical keyboard, as it would interfere with eye tracking data collection (e.g., due to participants looking down to see the keyboard). Thus, the interface included a virtual on-screen keyboard, which allowed the participant to enter the problem solution while maintaining their attention on the screen. The participant could move

FIGURE 3.1: The Interface Participants Used When Solving Problems

to the next line, move to the previous line, delete symbols (operators and variables), delete the entire line, and move on to the next problem. Once the student moved on to the next problem, they were unable to go back and revise. Furthermore, the problems the participant had to solve were generated using up to seven of the same letters. Thus, the on-screen keyboard contained only the necessary and sufficient buttons to solve every problem, though not every letter was used every time. Participants entered the letter by pressing the mouse button on the virtual keyboard.

3.6 Measures

The measures included learning, performance, and eye gaze behaviours. For the latter, we defined 10 areas of interest (AOIs) - one for each line of the example and problem (see each rectangle in Figure 3.2).

Learning Measures

The difference scores, calculated by subtracting the pre-test from the post-test, were used to determine which forms of faded assistance had an impact on overall learning outcomes. To achieve a more fine grained analysis of the learning results, we also calculated a difference score for regular problems (i.e., problems that were similar to the ones encountered by the participants in the learning phase of the experiment), and for transfer problems (i.e., problems that required additional steps or knowledge). Refer to Table 3.3 for a detailed explanation.

Performance Measures

The dependent variable for performance was the average overall time on the task per condition.

Eye Gaze Measures

The eye tracker related measures we extracted relate to a series of dwell times, fixation counts, and gaze sequence information using fixation counts. For this analysis, we drew a series of areas of interest (AOIs) around the problem and the example

solution lines, as shown in Figure 3.2, and then extracted a series of dependent variables based on those interest areas, as follows.

Dwell time refers to the amount of time a participant fixated within an interest area over the course of the entire learning phase (i.e., all 12 problem-example pairs). We extracted an overall dwell time for the problem and for the example, called Dwell Time_{Problem} and Dwell Time_{Example}: Dwell Time_{Problem} corresponded to the total amount of time participants spent looking at the problem area of interest (while solving the 12 problem-example pairs), and Dwell Time_{Example} corresponded to the total amount of time participants spent looking at the example area of interest (while solving the 12 problem-example pairs). We also obtained a fixation count, using the same methodology: Fixation Count_{Example} corresponded to the total number of fixations participants made within the example area of interest (while solving the 12 problem-example pairs) and Fixation Count_{Problem} was the total number of fixations participants made within the problem area of interest (while solving the 12 problem-example pairs).

Secondly, the overall gaze switches between the problem and example, example and problem, and within the example and problem were analyzed. To do this a fixation sequence report was generated, and the fixation counts of moving between AOIs were summed to get an overall score on example to example, example to problem, problem to example, and problem to problem for each participant. For the sequence analysis, fixation count usually refers to the number of times an interest area was entered from the preceding interest area. For instance, if a participant had three attentional shifts from one line of the example to the another line throughout the experiment, illustrated by arrow 1 in Figure 3.2, this was recorded

FIGURE 3.2: The Interface Participants Used When Solving Problems With Areas of Interest. *Note that the boxes around the example and problem lines were not visible to the participant and are present to illustrate the areas of interest (AOI) used for the analysis.*

as three example-example sequences. Such sequences (arrow 1 in Figure 3.2), may be indicative of greater self-explanation, as the participant is shifting between the example solution lines presumably trying to understand the example. The shifts from the example to the problem (i.e., example-problem sequences) are illustrated by arrow 2, which may be indicative of copying behaviour as the participant is presumably taking what is present in the example and seeing if they can apply it to the problem. Problem-example sequences are illustrated by arrow 4, which may be indicative of checking or referencing behaviours, as the participant is presumably checking the example in relation to the problem, or going back to the example in

preparation for the next line of the problem. Finally, shifts between solution lines of the problem are represented by arrow 3 in Figure 3.2, (i.e., problem-problem sequences), which may be indicative of problem solving and/or self-explanation during problem solving

Prior work on example based learning and copying behaviours seems to suggest that copying requires attention shifts between the problem and example in rapid succession, which are used by novices to determine how to make the example solution suitable for the problem (VanLehn and Jones, 1993). In contrast, an alternative strategy that should result in fewer gaze shifts between the example and problem is to study the example independently, and then apply that solution to the problem. In light of this, the level of expertise was included in this analysis, as different levels of skill may influence how participants used the examples in different conditions. A participants' level of expertise was determined by splitting the pre-test scores into thirds, so that there was a low scoring group (0-12), a medium scoring group (13-24), and a high scoring group (25-36), with a maximum score of 36 which only one participant attained.

3.7 Participants

A total of 60 undergraduates were recruited to participate from January, 2017, to June, 2017. Recruitment was through a variety of sources: the online recruitment program, SONA, at Carleton University in exchange for extra credit in a class, through flyers posted on bulletin boards throughout Carleton University for monetary compensation, and through social media. There were 34 females and 26 males.

The average year of university of the participants was 2.5. Participants were split into three groups: a Sim-Sim condition, a Sim-Dissim condition, and a Dissim-Sim condition. There were 20 participants for the Sim-Sim condition, 21 for the Sim-Dissim condition, and 19 for the Dissim-Sim condition. All participants had normal or corrected to normal vision.

3.8 Design and Procedure

3.8.1 Design

A between subjects design was used for this experiment, with each participant exposed to one of the three study conditions (Sim-Dissim, Sim-Sim and Dissim-Sim), which were variations of different forms of assistance (faded, constant, and reverse faded, respectively). Initially, participants were assigned to a condition in order (e.g., the first participant was assigned to the Sim-Dissim condition, the second to the Sim-Sim condition, and the third to the Dissim-Sim condition, and so on). Once 20 participants had completed the experiment, a matching criteria was used to assign them to each condition, so as to have similar average scores on the pre-test in all three conditions. This was done by having the experimenter quickly scan the pre-test for material to determine whether the participant appeared to be using the correct operations before choosing the condition for the participant. Three groups were chosen based on estimated performance - low (0 -12), medium (13 -24), and high (25 -36). For instance, if there were many low scoring participants in one condition, but not in another, the condition with less was selected for a low scoring participant.

3.8.2 Procedure

Each session was conducted individually in a quiet room. After signing the informed consent form, participants were given (1) the questionnaire compiled from the above described established surveys (i.e. ATMI, MBIQ, STAI, and Goal Orientation) and (2) after they finished the questionnaire, the paper and pencil pre-test. Participants were given up to twenty minutes to complete the pre-test, but were not required to take the full time. Time on the questionnaire was not recorded, but took approximately 5 minutes.

Once the pre-test was completed, participants were introduced to the problem solving interface by the experimenter. The same practice trial was loaded for each participant, which contained two simple problems (similar problem-example pairs, two rule problems). The experimenter first showed participants how to use the virtual key board to indicate solution variables (e.g. x or y), followed by the operators (e.g. $()$, $+$, $-$, $=$, and the functional buttons such as *Backspace* or *Clear Line*). Once the experimenter moved on to the final problem of the practice trial, they gave control to the participant and asked them to try the problem to get used to the interface. Note that the experimenter did not solve the first problem, but simply pressed all the buttons to show the participant how to use the keyboard.

After the participant had familiarized themselves with how to use the interface, they were calibrated in the eye-tracker. This involved the participant resting their head in a chin rest and looking at a series of dots as they moved about the screen. This happened twice, once for calibration, and a second time for validation. This process lasted approximately 2 minutes. Once calibrated, the participant was then presented with general instructions, both verbally and textually, on how the

experiment would proceed. Before the experiment began, the participants were verbally told "...the goal here isn't to be fast but rather to treat this as if you were practising solving problems to prepare for a test – so do what you would normally do when studying. You will have 40 minutes for 12 questions". Participants then worked through the problems (with access to examples) in their respective condition. When they pressed the Finish Experiment button, they were presented with a completion screen thanking them for their participation. They were then asked to complete a paper and pencil post-test which was identical to the pre-test in the knowledge needed to answer the questions, though variable names and the order of the problems were changed. The participant was given a maximum of twenty minutes to complete the post-test. After it was completed, participants were debriefed and granted credit or monetary compensation.

Chapter 4

Results

4.1 Research Questions

Our analysis was guided by the following research questions:

1. Learning

- How do different forms of faded assistance influence what participants learned?

2. Performance

- Are there differences in overall time to complete the learning phase of the study depending on which form of faded assistance the participant received?

3. Eye Gaze Behaviours

- Do participants look at the examples or problems for longer or more frequently depending on the form of assistance?

- How do condition and level of expertise affect eye gaze patterns, such as moving from one line of an example to another (example to example), moving from the example to the problem (example to problem), moving from the problem to the example (problem to example), and moving from one line of the problem to another (problem to problem)?

4. Traits and Behaviours

- How do math behaviours and interests, attitudes towards mathematics, state and trait anxiety, and goal orientation relate to how participants use the examples and problems?

Throughout the remainder of this thesis, the terms condition and form of assistance will be used interchangeably. For the statistical analyses, we conducted post-hoc tests when the overall effect was significant ($p < .05$) or marginal ($p < .10$), which can provide insight even in the presence of marginal overall effects (Brace, Kemp, and Snelgar, 2003). Furthermore, outliers were identified by SPSS generated boxplots (sometimes referred to as box and whisker plots). Boxplots were generated by plotting the interquartile range (IQR). Any values that were one-and-a-half times the IQR below the bottom quartile or one-and-a-half times the IQR above the third quartile were classified as outliers by SPSS, and were removed from the following analysis.

4.2 Learning Results

Table 4.1 presents the overall, regular, and transfer pre-test, post-test, and difference score averages as percentages for each condition. An analysis of variance (ANOVA) with pre-test scores as the dependent variable and condition as the independent variable demonstrated that there were no significant differences between conditions ($F(2, 57) = .24, p = .790, \eta_p^2 = .01$).

TABLE 4.1: Means and Standard Deviations for Each Condition for Pre-Test, Post-Test, and Difference Scores (*for All Problems, Regular Problems, and Transfer Problems*)

		Dissim-Sim $n = 19$		Sim-Dissim $n = 21$		Sim-Sim $n = 20$	
		M	SD	M	SD	M	SD
All Problems	Pre-Test %	32.89	37.34	36.25	37.36	41.11	37.89
	Post-Test %	72.94	26.90	57.53	36.88	70.83	31.61
	Difference %	40.05	27.58	21.28	26.62	29.72	28.74
Regular Problems	Pre-Test %	36.84	39.15	41.84	41.26	47.11	42.91
	Post-Test %	82.84	23.85	65.16	37.05	79.21	30.57
	Difference %	46.00	34.82	23.32	29.14	32.11	33.98
Transfer Problems	Pre-Test %	28.47	36.65	29.98	34.95	34.41	33.94
	Post-Test %	61.94	35.31	49.02	39.91	61.47	35.98
	Difference %	33.41	26.96	19.06	27.21	27.06	30.44

Learning was assessed by calculating a difference score between the pre-test and the post-test for each participant. A boxplot was generated to identify any outliers for difference scores, and one was removed from the Sim-Dissim condition. As can be seen in the top row of Table 4.1, the largest difference scores were observed in the Dissim-Sim condition followed by the Sim-Sim condition, with the lowest in the Sim-Dissim condition.

To determine if the form of faded assistance had a significant impact on overall learning gains, an ANOVA with the form of assistance as the independent variable

and the total difference scores as the dependent variable was conducted with significant results ($F(2, 56) = 3.46, p = .038, \eta_p^2 = .11$). Post hoc analysis using Tukey's Honest Significant Difference (HSD) test indicated that there was a significant difference between the Dissim-Sim and the Sim-Dissim condition ($p = .029$), indicating that participants in the Dissim-Sim condition had significantly higher total learning gains than those in the Sim-Dissim condition. No other differences were significant.

To understand how faded assistance influenced outcomes based on the type of question (regular or transfer), we conducted two separate ANOVAs with regular and transfer difference scores as the dependent variable. Regular problems were scored out of a total of nineteen marks, and transfer problems were scored out of a total of seventeen marks. A boxplot generated for both regular and transfer problems indicated there was one outlier in the Sim-Dissim condition, which was removed before analysis.

Regular Problems

For the regular problems, as can be seen in the middle row of Table 4.1, the Dissim-Sim condition had the largest learning gains, followed by the Sim-Sim, and the Sim-Dissim conditions. There was an overall significant effect of condition on the regular problem scores ($F(2, 56) = 3.40, p = .041, \eta_p^2 = .11$). A post hoc analysis using Tukey's HSD test indicated that the difference between the Dissim-Sim and the Sim-Dissim condition was significant ($p = .031$). No other differences were significant.

Transfer Problems

There was no overall significant effect for transfer problems between conditions, ($F(2, 56) = 2.10, p = .132, \eta_p^2 = .07$), though they demonstrated the same pattern of improvement as the regular and overall difference scores, with the Dissim-Sim condition being highest and the Sim-Dissim condition being lowest (See bottom row of Table 4.1).

4.3 Performance Results

As can be seen in Table 4.2, participants in the Dissim-Sim reverse faded condition took the longest to complete the eye-tracking portion of the experiment, followed closely by the Sim-Dissim condition, with the Sim-Sim condition completing it the fastest. To determine whether there were significant differences in overall time to complete the learning phase of the study, an ANOVA was conducted with the form of assistance as the independent variable. A boxplot identified three outliers which were removed from the Sim-Sim condition before analysis. The effect for form of assistance was significant ($F(2, 54) = 5.84, p = .005, \eta_p^2 = .18$), and post hoc tests using Tukey's HSD revealed that there were significant differences between the Dissim-Sim and Sim-Sim groups ($p = .004$). The other differences were not significant, though the difference between the Sim-Dissim and Sim-Sim conditions approached significance ($p = .055$). These results indicate that the reverse fading condition took significantly longer than the constant assistance condition.

TABLE 4.2: *Average Time to Complete the Learning Phase by Form of Assistance (minutes)*

Form of Assistance	Average Time	SD	<i>n</i>
Dissim-Sim	14.22	3.89	19
Sim-Dissim	13.05	3.79	21
Sim-Sim	10.24	3.13	19

4.4 Dwell Time and Fixation Count on the Example and Problem

To determine where participants were devoting time (e.g., to the examples vs. to the problems), we analyzed how long participants were looking at the example area of interest, summed over all 12 examples (Dwell Time_{Example}, which is the total example dwell time in a given session), and within the problem area of interest, summed over all 12 problems (Dwell Time_{Problem}, which is the total problem dwell time in a given session). We also analyzed how often participants looked within the example area of interest, summed over all 12 examples (Fixation Count_{Example}, which is the total example fixation count in a given session), and within the problem area of interest, summed over all 12 problems (Fixation Count_{Problem}, which is the total problem fixation count in a given session). This is a more fine grained measure of attention than the analysis of performance above because it can reveal how much time participants spent looking at either the example or the problem, and how many times participants looked at the example or the problem. We do not explicitly provide Dwell time per problem, but this could be attained by dividing the total results in Table 4.3 by the number of problems (i.e., 12). However, averaging across problems is not informative, at least in the Sim-Dissim and

TABLE 4.3: *Total Dwell Time for Example Area and Problem Area by Condition in Minutes*

	Form of Assistance	<i>M</i>	<i>SD</i>	<i>n</i>
Dwell Time _{Example}	Dissim-Sim	2.95	2.16	18
	Sim-Dissim	3.48	3.06	21
	Sim-Sim	2.81	2.35	20
Dwell Time _{Problem}	Dissim-Sim	9.15	2.76	19
	Sim-Dissim	7.99	2.90	21
	Sim-Sim	5.62	1.09	17

Dissim-Sim conditions, as the problem-example pairs require different amounts of attention depending on whether they were similar or dissimilar. Both dwell time and fixation count are indicative of greater processing (Holmqvist et al., 2011).

Example Dwell Time

To determine if the form of assistance influenced how long participants spent looking at the example area of interest (summed across all 12 examples), an ANOVA was conducted for Dwell Time_{Example}, which was the total amount of time participants spent looking at the example area per condition. Boxplots identified one outlier for Dwell Time_{Example} and it was removed from the Dissim-Sim condition. We did not find evidence that condition impacted how long participants looked at the example ($F(2, 56) = .38, p = .686, \eta_p^2 = .01$). The data on Dwell Time_{Example} is in Table 4.3.

Problem Dwell Time

To determine if the form of assistance influenced how long participants spent looking at the problem area of interest (summed across all 12 problems), an ANOVA

was conducted for Dwell Time_{Problem}, which was the total amount of time participants spent looking at the problem area per condition. Boxplots identified three outliers for Dwell Time_{Problem}, all three of which were removed from the Sim-Sim condition. The ANOVA produced an overall significant result for Dwell Time_{Problem} ($F(2, 54) = 9.58, p < .001, \eta_p^2 = .26$). Post hoc analysis using Tukey's HSD test revealed that there were significant differences between the Sim-Sim condition and the Sim-Dissim condition ($p = .012$) and between the Sim-Sim condition and the Dissim-Sim condition ($p < .001$) (Table 4.3, bottom). Thus, participants in the faded and reverse faded assistance conditions fixated on the problem for longer than participants in the Sim-Sim condition. This indicates that the differences between conditions in regards to total time is derived primarily from the time spent on the problem rather than studying the example, and that participants in the Sim-Dissim and Dissim-Sim conditions spent more time problem solving than the Sim-Sim condition.

Example and Problem Fixation Count

To determine if the form of assistance influenced the number of fixations participants made on the problem and example, we focused on the variables Fixation Count_{Example} and Fixation Count_{Problem} (see Table 4.4, which presents the total number of fixations participants made looking at the example or problem area per condition). There were no outliers for Fixation Count_{Example}; two outliers were identified for Fixation Count_{Problem} in the Sim-Sim condition and were removed. An ANOVA with Fixation Count_{Example} as the dependent variable and the form of assistance as the independent variable revealed that the effect for condition was

not significant ($F(2, 57) = .16, p = .852, \eta_p^2 = .01$). Therefore, mimicking Dwell Time _{Example}, participants did not fixate on the example a different number of times depending on condition.

TABLE 4.4: Fixation Count for Example and Problem by Condition

	Form of Assistance	<i>M</i>	<i>SD</i>	<i>n</i>
Fixation Count _{Example}	Dissim-Sim	993.00	756.92	19
	Sim-Dissim	965.86	861.46	21
	Sim-Sim	861.95	665.68	20
Fixation Count _{Problem}	Dissim-Sim	2621.00	610.299	19
	Sim-Dissim	2060.67	668.892	21
	Sim-Sim	1623.94	342.868	17

To determine if the form of assistance impacted fixation frequency on the problem, an ANOVA was conducted for Fixation Count _{Problem}. The ANOVA produced an overall significant result for Fixation Count _{Problem} ($F(2, 55) = 14.43, p < .001, \eta_p^2 = .34$). Post hoc analysis using Tukey's HSD test revealed that there were significant differences between the Dissim-Sim condition and the Sim-Dissim condition ($p = .008$) and between the Dissim-Sim condition and the Sim-Sim condition ($p < .001$). Thus, participants in the reverse faded assistance condition fixated on the problem more often than participants in the Sim-Sim and Sim-Dissim conditions, which may be indicative of more problem solving behaviours in the reverse faded assistance condition.

4.5 Gaze Sequences

To determine how the form of assistance and the level of expertise impacted different eye gaze patterns, a fixation sequence analysis report was generated from the

eye tracking file (refer to Section 3.6 for definitions of the eye gaze sequences). The participants' level of expertise was included in this analysis as differences in initial level of skill may affect how a participant behaves when examples are present (e.g., different levels of self-explanation or copying behaviours)(Inglis and Alcock, 2012). Level of expertise was operationalized by splitting scores on the pre-test into thirds, so that there was a low scoring group (from 0 -12), a medium scoring group (13 -24), and a high scoring group (25 -36). The level of expertise was included as an independent variable with three levels in the following analysis.

The gaze sequences were gathered by analyzing the number of fixations moving between lines of the the example, the problem, and between example to problem and problem to example. This provided information on the number of attentional shifts participants were engaging in between lines. The summed fixation sequences were entered as four dependent variables into separate ANOVAs with condition and level of expertise as the independent variables. Boxplots were generated for each gaze sequence pattern and outliers were removed.

4.5.1 Attentional Shifts Between Lines of the Example, From Example to Problem, and From Problem to Example

Since the results were similar for three of the eye gaze sequences, namely example to example, example to problem, and problem to example, we begin by presenting the results for these three. We began the analysis with the identification and removal of outliers: for the example to example sequence, three outliers were removed (one from the high expertise by Dissim-Sim condition, one from the low expertise by Dissim-Sim condition, and one from the high expertise by Sim-Sim

TABLE 4.5: *Attentional Shifts Between Lines of the Example, From Example to Problem, and From Problem to Example by Level of Expertise and Form of Assistance*

	Condition	M	SD	n
Example to Example				
Level of Expertise	High	249.77	197.47	13
	Medium	331.00	289.27	12
	Low	921.56	512.48	32
Form of Assistance	Dissim-Sim	639.24	431.15	17
	Sim-Dissim	682.52	625.19	21
	Sim-Sim	605.74	491.06	19
Example to Problem				
Level of Expertise	High	58.64	34.96	14
	Medium	71.50	60.99	12
	Low	226.33	135.29	33
Form of Assistance	Dissim-Sim	152.05	141.86	19
	Sim-Dissim	170.05	148.84	21
	Sim-Sim	141.47	108.11	19
Problem to Example				
Level of Expertise	High	53.00	35.49	14
	Medium	71.58	61.91	12
	Low	219.97	134.53	32
Form of Assistance	Dissim-Sim	127.39	108.01	18
	Sim-Dissim	169.10	156.09	21
	Sim-Sim	147.16	124.54	19

condition). For the example to problem sequence, one outlier was removed (high expertise by Sim-Sim condition). Lastly, for the problem to example sequence, two outliers were removed (one from the low expertise by Dissim-Sim condition, and one from the high expertise by Sim-Sim condition). There were no significant differences between conditions for example to example, example to problem, and problem to example sequences ($F(2, 48) = .06$, $p = .947$, $\eta_p^2 = .00$; $F(2, 50) = .20$, $p = .819$, $\eta_p^2 = .01$; $F(2, 49) = .42$, $p = .660$, $\eta_p^2 = .02$, respectively) (See Table 4.5 for average gaze sequences between conditions).

However, there was a main effect of expertise for the example to example, example to problem and problem to example sequences ($F(2, 48) = 14.92, p < .001, \eta_p^2 = .38$; $F(2, 50) = 15.63, p < .001, \eta_p^2 = .39$; $F(2, 49) = 16.05, p < .001, \eta_p^2 = .40$, respectively). Post hoc analysis using Tukey's HSD test demonstrated that there were significant differences between the high and low ($p < .001$), and between the low and medium levels of expertise ($p < .001$) for example to example, example to problem, and problem to example gaze sequences (Table 4.5). Thus, low expertise participants shifted their attention between lines of the example, from the example to the problem, and from the problem to the example, significantly more than participants in the medium and high expertise groups, which may be indicative of more self-explanation, copying, and referencing behaviours.

4.5.2 Attentional Shifts Between Lines of the Problem

TABLE 4.6: *Attentional Shifts Between Lines of the Problem for Level of Expertise and Form of Assistance*

	Condition	Mean	Std. Deviation	<i>n</i>
Level of Expertise	High	1541.80	600.66	15
	Medium	1606.92	535.20	12
	Low	1444.13	683.97	32
Form of Assistance	Dissim-Sim	2004.47	514.04	19
	Sim-Dissim	1450.90	588.68	21
	Sim-Sim	1056.21	383.87	19

For the problem to problem fixation sequences, two outliers were removed from the high expertise by Sim-Sim condition. There were no significant differences between levels of expertise for problem to problem gaze sequences ($F(2, 50) = 1.48, p = .237, \eta_p^2 = .06$).

In opposition to the other eye gaze patterns, there was a main effect of the form of assistance on the number of gaze shifts between the problem lines ($F(2, 50) = 15.09, p < .001, \eta_p^2 = .38$). Post hoc analysis using Tukey's HSD test demonstrated that there were significant differences between the Dissim-Sim and Sim-Sim conditions ($p < .001$), between the Dissim-Sim and Sim-Dissim conditions ($p < .003$), and between the Sim-Sim and Sim-Dissim conditions ($p = .041$). These findings indicate the Dissim-Sim group shifted their attention between lines of the problem more times than the Sim-Dissim and the Sim-Sim group, which may be indicative of more problem solving behaviours (see Table 4.6).

4.6 Influence of Mathematics Behaviours and Interests, Attitudes Towards Mathematics, Anxiety, and Goal Orientation on Eye Gaze Sequences

Responses on the questionnaire compiled from the MBIQ, ATMI, STAI, and Goal Orientation sections can provide a valuable insight into possible ways that character traits, transient states of emotion, or attitudes about math, could influence the way in which examples and problems are studied. Before conducting a bivariate correlation to investigate the effects of questionnaire items on eye gaze sequences, questionnaire items were correlated with themselves to verify that the participants responses followed the expected trend. The MBIQ, ATMI, STAI, and Goal Orientation variables are in Table 4.7.

The correlation results in Table 4.7 indicate the expected relationships found

in prior work (Goetz et al., 2008; Hannula, Maijala, and Pehkonen, 2004). For instance, how well a participant rated their own basic and complex math skill positively correlated with their ratings of enjoyment, motivation, self-confidence, and the value they perceive mathematics to have. Their self ratings were negatively correlated with how difficult mathematics was for them in high school, as well as how often they avoided and felt nervous around math. The same pattern was demonstrated for participants' self ratings of their skill in complex areas of math.

TABLE 4.7: *Correlations Between Individual Traits*

	Math-B	Math-C	Avoid	Nerv	Diff	Enj	Mot	S-Conf	Value	Anxiety	Per-Approach	Per-Avoid	Mastery
Math-B	1	.697**	-.467**	-.306*	-.366**	.465**	.424**	.527**	.307*	.161	.052	.037	-.015
Math-C		1	-.463**	-.353**	-.478**	.613**	.609**	.548**	.312*	.194	.014	.098	-.108
Avoid			1	.731**	.503**	-.588**	-.677**	-.723**	-.430**	-.009	-.106	.159	-.053
Nerv				1	.642**	-.481**	-.467**	-.672**	-.237	.121	.032	.146	-.038
Diff					1	-.569**	-.495**	-.605**	-.216	.008	.007	-.116	-.076
Enj						1	.782**	.806**	.413**	.078	-.076	-.103	.246
Mot							1	.712**	.556**	.146	.086	-.079	.278*
S-Conf								1	.318*	-.101	-.064	-.135	.217
Value									1	.163	.180	.118	.193
Anxiety										1	.367**	.263*	.176
Per-Approach											1	.423**	.377**
Per-Avoid												1	-.034
Mastery													1

** Correlation is significant at the 0.01 level (2-tailed)

* Correlation is significant at the 0.05 level (2-tailed)

Legend: Math-B = Self Rated Basic Math Skill; Math-C = Self Rated Complex Math Skill; Avoid = Avoidance of Math; Nervous = Nervousness Around Math; Diff = Difficulty of Math in High School; Enj = Enjoyment; Mot = Motivation; S-Conf = Self Confidence; Value = Perceived Value; Anxiety = Level of Anxiety; Per-Approach = Performance Approach; Per-Avoid = Performance Avoid; Mastery = Task Goal

TABLE 4.8: *Correlations for MBIQ and Sequence Fixation Count*

	Math-B	Math-C	Avoid	Nervous	Diff	Ex-Ex	Ex-Pr	Pr-Ex	Pr-Pr
Math-B	1	.697**	-.467**	-.306*	-.366**	-.547**	-.502**	-.501**	-.112
Math-C		1	-.463**	-.353**	-.478**	-.468**	-.492**	-.462**	-.056
Avoid			1	.731**	.503**	.326*	.302*	.291*	-.092
Nervous				1	.642**	.143	.134	.128	-.17
Diff					1	.099	.132	.106	.016
Ex-Ex						1	.896**	.891**	.262*
Ex-Pr							1	.989**	.094
Pr-Ex								1	.097
Pr-Pr									1

** Correlation is significant at the 0.01 level (2-tailed)

* Correlation is significant at the 0.05 level (2-tailed)

Legend: Math-B = Self Rated Basic Math Skill; Math-C = Self Rated Complex Math Skill; Avoid = Avoidance of Math; Nervous = Nervousness Around Math; Diff = Difficulty of Math in High School; Ex-Ex = Example to Example Fixation Count; Ex-Pr = Example to Problem Fixation Count; Pr-Ex = Problem to Example Fixation Count; Pr-Pr = Problem to Problem Fixation Count

Also, the more someone was motivated, the more likely they were to score high on the task goal orientation (sometimes referred to as the mastery goal orientation). Conversely, a participant with a higher level of anxiety was more likely to score high on performance-approach and performance-avoid orientations, suggesting they wanted to be seen as proficient, and simultaneously not be seen as deficient, at the task at hand. See Table 4.7 for significance and strength of the correlations.

Eye Gaze Sequences and Mathematics Behaviours and Interests

For the MBIQ, self-rated skill in basic mathematics was negatively correlated with example to example, example to problem, and problem to example fixation count, indicating that if a participant rated themselves as proficient in basic math, they spent less time studying the example, less time copying, and less time referencing the example. Self rated skill in complex mathematics showed the same pattern but with slightly less strength (See Table 4.8 for significance and strength of the correlations). This pattern echoes what was found for level of expertise and participants' allocation of attention in Section 4.5.1.

Alternatively, scores on the question "*How often do you avoid situations involving mathematics?*", were positively correlated with example to example, example to problem, and problem to example fixation count, indicating that if they avoided mathematics frequently, they were more likely to study the example, copy from the example to the problem, and reference the example from the problem more often. See Table 4.8 for the significance and strengths of the correlations.

Eye Gaze Sequences and Attitudes Towards Mathematics: Enjoyment, Motivation, Self-Confidence, and Perceived Value

The ATMI deals with attitudes towards math, and enjoyment of mathematics was negatively correlated with example to example, example to problem, and problem to example fixation count, indicating that if a participant reported experiencing enjoyment while doing math, they spent less time studying the example, less time copying, and less time referencing the example. This pattern was similar for motivation and self confidence, indicating that the more a participant was motivated and self-confident about mathematics, the less time they spent studying the example, copying, and referencing. See Table 4.9 for significance and strength of the correlations.

Eye Gaze Sequences, State-Trait Anxiety, and Goal Orientation

An overall anxiety score was attained from the six STAI questions, which was negatively correlated with example to example, example to problem, and problem to example fixation sequences, indicating that if a participant was anxious, they studied the example, copied from the example, and referenced the example less often. Oddly, this is a similar pattern for participants who were self-confident, motivated, and enjoyed doing mathematics (see Section 4.6).

Only the performance-approach goal orientation was significantly correlated with any eye gaze sequences; it was negatively correlated with example to example, example to problem, and problem to example fixation count, indicating that the more a participant had a performance-approach goal orientation style, the

smaller the number of fixations on these sequences (See Table 4.10 for significance and strengths).

TABLE 4.9: *Correlations for ATMI and Sequence Fixation Count*

	Enj	Mot	S-Conf	Value	Ex-Ex	Ex-Pr	Pr-Ex	Pr-Pr
Enj	1	.782**	.806**	.413**	-.296*	-.369**	-.343**	.114
Mot		1	.712**	.556**	-.350**	-.407**	-.379**	-.042
S-Conf			1	.318*	-.301*	-.352**	-.337**	.176
Value				1	-.227	-.210	-.202	.014
Ex-Ex					1	.896**	.891**	.262*
Ex-Pr						1	.989**	.094
Pr-Ex							1	.097
Pr-Pr								1

** Correlation is significant at the 0.01 level (2-tailed)

* Correlation is significant at the 0.05 level (2-tailed)

Legend: *Enj* = Enjoyment; *Mot* = Motivation; *S-Conf* = Self Confidence; *Value* = Perceived Value; *Ex-Ex* = Example to Example Fixation Count; *Ex-Pr* = Example to Problem Fixation Count; *Pr-Ex* = Problem to Example Fixation Count; *Pr-Pr* = Problem to Problem Fixation Count

MBIQ, ATMI and Pre- and Post-Test Scores

There were no significant correlations between the difference scores and the math survey variables, but the total pre-test and post-test scores revealed some significant correlations for the MBIQ and ATMI (See Table 4.11 for significance and strengths). Participants who rated themselves as self-confident and competent in basic and complex mathematics were more likely to have higher pre- and post-test scores, and while self-ratings of enjoyment of mathematics were positively correlated with pre-test scores, they were not significant for post-test scores. Furthermore, the more a participant reported avoiding math, the lower their score on the post-test.

TABLE 4.10: *Correlations for STAI, Goal Orientation and Sequence Fixation Count*

	Anxiety	Per- Approach	Per-Avoid	Mastery	Ex-Ex	Ex-Pr	Pr-Ex	Pr-Pr
Anxiety	1	.367**	.263*	.176	-.292*	-.297*	-.277*	-.097
Per- Approach		1	.423**	.377**	-.256*	-.284*	-.306*	-.083
Per-Avoid			1	-.034	-.082	-.128	-.092	-.138
Mastery				1	-.075	-.128	-.147	-.153
Ex-Ex					1	.896**	.891**	.262*
Ex-Pr						1	.989**	.094
Pr-Ex							1	.097
Pr-Pr								1

** Correlation is significant at the 0.01 level (2-tailed)

* Correlation is significant at the 0.05 level (2-tailed)

Legend: Anxiety = Level of Anxiety; Per-Approach = Performance Approach; Per-Avoid = Performance Avoid; Mastery = Task Goal; Ex-Ex = Example to Example Fixation Count; Ex-Pr = Example to Problem Fixation Count; Pr-Ex = Problem to Example Fixation Count; Pr-Pr = Problem to Problem Fixation Count

TABLE 4.11: Correlations of MBIQ, ATMI, and Pre-Test and Post-Test Scores

	Math-B	Math-C	Avoid	Nervous	Diff	Enj	Mot	S-Conf	Value	Pre-Test	Post-Test
Math-B	1	.697**	-.467**	-.306*	-.366**	.465**	.424**	.527**	.307*	.488**	.527**
Math-C		1	-.463**	-.353**	-.478**	.613**	.609**	.548**	.312*	.514**	.497**
Avoid			1	.731**	.503**	-.588**	-.677**	-.723**	-.430**	-.241	-.274*
Nervous				1	.642**	-.481**	-.467**	-.672**	-.237	-.172	-.219
Diff					1	-.569**	-.495**	-.605**	-.216	-.213	-.190
Enj						1	.782**	.806**	.413**	.276*	.192
Mot							1	.712**	.556**	.148	.238
S-Conf								1	.318*	.297*	.296*
Value									1	.085	.225
Pre-Test										1	.677**
Post-Test											1

** Correlation is significant at the 0.01 level (2-tailed)

* Correlation is significant at the 0.05 level (2-tailed)

Legend: *Math-B* = Self Rated Basic Math Skill; *Math-C* = Self Rated Complex Math Skill; *Avoid* = Avoidance of Math; *Nervous* = Nervousness Around Math; *Diff* = Difficulty of Math in High School; *Enj* = Enjoyment; *Mot* = Motivation; *S-Conf* = Self Confidence; *Value* = Perceived Value

Chapter 5

Discussion

Reverse faded assistance promotes learning. Prior work found that dissimilarity between problem and example pairs encouraged learning (Muldner and Conati, 2010; Lee, Betts, and Anderson, 2015). The present thesis work extends that research by exploring different ways to fade the similarity over a problem solving session. One of our main findings was that a reversed fading method, from dissimilar to similar problem-example pairs (the Dissim-Sim condition) resulted in the largest learning gains.

One explanation for this finding is that since participants in this condition were presented with dissimilar problem-example pairs first, they may have been encouraged to study the examples and extract the underlying rules, but once learned, the examples were attended to less. Thus, participants in the Dissim-Sim condition may be doing most of their cognitive processing while solving the problem (i.e., in the problem area), without help of the example. We do have some evidence of this flow of attention from the eye tracking data: participants in the Dissim-Sim condition looked at the problem (without looking at the example) significantly longer than the Sim-Sim condition, and with significantly more attentional shifts

between lines of the problem than in the other conditions. Grant and Spivey (2003) argued that more attention when problem solving is not a by-product of the problem solving taking place, but that attention is what makes problem solving possible. Through increased attention, participants may have been able to optimize their problem solving strategy leading to greater learning gains (Susac et al., 2014).

A type of cognitive processing that may be occurring in parallel with the increase in attention is self explanation, which involves a student explaining a concept to themselves and generating inferences that go beyond the surface features of the instructional material. Aleven and Koedinger (2002) demonstrated that the more self-explanation that took place during problem solving in a cognitive tutor, the higher the learning gains on a post-test. However, the cognitive tutor they used required that students explain their steps explicitly, whereas in our experiment there was no self-explanation prompting. One possibility for the increased learning gains in the Dissim-Sim condition is that participants relied on self-explanation because it was the only way they could solve the problem once they realized that copying was infeasible. While we do not have direct evidence of self-explanation, we do know from prior work that problem-example dissimilarity encourages it (Muldner and Conati, 1998). Moreover, it is established that self explanation increases the time on task (Muldner and Conati, 2010; Chi et al., 1989). Thus, the fact that participants took more time on task in this condition could be indicative of self-explanation occurring. On the other hand participants in the Sim-Dissim condition took a similar amount of time, but did not learn more, suggesting that perhaps they were self-explaining as well, but may not have been as successful in applying the inferred knowledge. We return to this issue when we discuss the

learning results for the Sim-Dissim condition.

Prior work on faded assistance with worked examples and cognitive load theory does not predict the result that reverse faded assistance would benefit learning better than faded assistance (Atkinson, Renkl, and Merrill, 2003; Tullis, Goldstone, and Hanson, 2015; Leahy, Hanham, and Sweller, 2015; Sweller and Cooper, 1985). This prior work suggests that students initially experience high cognitive load during problem solving and need to be provided scaffolding such as highly similar examples, of the type in the Sim-Dissim condition, so that they can focus on the relevant aspects of the problem. By providing assistance at the beginning of the learning stage, and then gradually shifting over to the problem solving stage, the student's cognitive load is lessened and they can build an internal model of the assistance (i.e., the domain rules) to be used at a later time.

Research on productive failure, however, does predict our results (Kapur, 2014). This work suggests that it is beneficial for students to explore an unfamiliar domain before they receive any instruction, typically in the form of a lesson. In our experiment, the Dissim-Sim condition structured the problem-example pairs so that initially, the examples looked quite different from the problem. This condition may be encouraging productive failure in the participants because the answer is less obvious – the “lesson” was provided by the three similar problem-example pairs at the end of the Dissim-Sim condition (see Table 3.2). Therefore, participants may have made more errors during the experiment, but once they were shown the similar problem-example pairs, these provided the feedback necessary to revise misconceptions. Thus, the mechanism that increased learning in the Dissim-Sim condition could be that fading from dissimilar to similar problem-example pairs

simultaneously provided examples for participants to learn from, opportunities to self-explain, and at the end of the session, a lesson on the correct way to solve the problem through similar problem-example pairs (Kapur, 2014; Pathak et al., 2011).

Impact of constant assistance. We did not find a significant difference between the constant high assistance condition (the Sim-Sim condition) and the faded assistance / reverse faded assistance conditions (Sim-Dissim and Dissim-Sim) in regards to learning gains. In fact, the Sim-Sim condition resulted in larger learning gains than the Sim-Dissim condition, albeit not significantly so. The Sim-Sim condition included only highly similar problem-example pairs, and prior work found that high levels of similarity hindered learning (Anderson, Fincham, and Douglas, 1997), as did high levels of assistance (Tullis, Goldstone, and Hanson, 2015). One reason that high similarity hinders learning is that it allows the learner to copy from the example to the problem without necessarily learning the underlying concepts through self explanation. We failed to find evidence that participants copied more in this condition as indicated by the example to problem gaze sequences. This may be because all conditions engaged with the examples equally but for different purposes. Since the examples were presented for every condition and for every problem, participants might have checked to see if they could use a given example line by line, which might result in a similar number of fixations for each condition. The Sim-Dissim and Dissim-Sim conditions may have copied at the beginning or end of the learning phase respectively, but they may have nevertheless checked for assistance from the example whenever they encountered a new problem, were unsure of an operation, or when they were struggling to understand a rule. The Sim-Sim condition may have struck a balance between both of these

conditions, where participants may have copied when they needed support and problem solved when they had enough domain knowledge to solve the problems alone.

However, the Sim-Sim condition did provide constant opportunities for the simple repetition of applying the rules compared to the faded / reverse faded assistance of the other conditions. Prior work on the iterative view of procedural and conceptual knowledge suggests that the relations between these two types of knowledge are bi-directional, with increases in one producing increases in the other (Star, 2005). Thus, it may be that through simple repetition of the procedural operations participants were able to gain some conceptual knowledge of the domain, as their transfer scores were higher than that of the Sim-Dissim condition, though not significantly so (See Table ??) (Baroody, 2003; Rittle-Johnson and Siegler, 1998; Rittle-Johnson, Siegler, and Alibali, 2001). Alternatively, it may be that examples in the Sim-Sim condition freed up cognitive resources, allowing the participant to learn the underlying domain concepts (Leahy, Hanham, and Sweller, 2015; Renkl, 2014). Contrarily, other evidence suggests that similar examples are only beneficial if the participant is actively trying to figure out the underlying domain concepts (i.e., self-explain), rather than just copy from the example. It is unlikely that participants were actively trying to learn the domain rules in the Sim-Sim condition because self-explanation takes time, and the time to complete the learning phase in this condition was lowest (Muldner and Conati, 2010; Chi et al., 1989). It may be that simple repetition improves learning indirectly and to a lesser degree than self-explaining, which would explain the modest learning gains in the Sim-Sim condition and the lower time to complete the task.

Faded assistance did not support learning. As to why participants in the Sim-Dissim condition had the lowest learning gains, in contrast to what some prior work would predict (Kalyuga et al., 2003; Tullis, Goldstone, and Hanson, 2015), it may be because this condition started with examples that were easily applied to the problem with minimal alteration. Thus, rather than using the assistance the similar examples provided, participants may have just superficially copied. Once participants were presented with a dissimilar problem-example pair towards the middle of the learning phase, several things may have happened. First, they may have been unable to recognize the dissimilarity between the example and problem and proceeded to copy, but now incorrectly, which may have negatively contributed to their learning of the material (Reed, Dempster, and Ettinger, 1985). Reed, Dempster, and Ettinger (1985) showed that when participants were given an example that included differences with the problem, they failed to resolve the differences correctly. The authors concluded that the participants were relying on superficial mapping between elements of the example to the problem, which did not allow them to achieve correct solutions or transfer their skills to slightly different problems. Comparably, Novick (1988) demonstrated that novices performed the worst when the examples they were given were superficially similar, but structurally dissimilar to the problem. This was because the students were persuaded by the superficial similarities between the problem-example pair to believe that the solution to the problem was present, when in reality, it was not. While in the present study all of the examples were superficially dissimilar to the problem, this type of superficial similarity is easily resolved through simple transformational analogy

(i.e., by mapping example-specific constants, like variable names, to problem constants, and using that mapping to guide the substitution of example constants with ones needed for the problem) (Muldner and Conati, 2010).

Secondly, the participants in the Sim-Dissim condition may have arrived at the dissimilar problem-example pairs, and attempted to learn the material, but since they had essentially missed the first three learning opportunities by transformational analogy, they received less feedback from the examples to learn from in the middle section of the experiment (See Table 3.2; there were three dissimilar and three similar alternating problem-example pairs before all the pairs became dissimilar). We have some evidence that participants in the Sim-Dissim condition were engaging in some form of cognitive processing, since participants in the Sim-Dissim condition had significantly longer dwell times on the problem than the Sim-Sim condition, and there was little difference in dwell time between the faded assistance and reverse faded assistance conditions. Longer dwell times have in prior work been associated with greater processing (Graesser et al., 2005; Jacob and Karn, 2003; Just and Carpenter, 1980) but in our study this did not translate to greater learning gains for the Sim-Dissim condition, which was the lowest of the three conditions. Thus, something was interfering with learning in this condition, and we have already presented several reasons above. Yet another explanation for their longer dwell times is that it may be indicative not of greater processing, but of bewilderment and confusion resulting from encountering the dissimilar problem-example pairs midway in the learning phase (Pachman et al., 2016).

Traits and Attitudes

Responses on the questionnaire indicated that compared to participants who avoided mathematics, participants who reported to enjoy and excel at mathematics had more eye gaze behaviours that may be indicative of self explanation, copying and example referencing. These results mimic what was found for the level of expertise and the allocation of attention between lines, with high scorers looking at the example, and between the example and problem, less often than low scorers. Thus, it appears that the more proficient a person is in mathematics, the less likely they are to need the example for assistance, as the more enjoyable someone finds an activity (i.e., solving algebra problems), the more likely they are to engage in it. However, participants who scored high on the anxiety scale, also looked at the example less, and copied and referenced less. This is the same pattern as participants who scored high on enjoyment, motivation, perceived value, and self-rated math skill. To explain this correlation, it could be that anxious participants simply tried to finish the experiment as soon as possible by speeding through, or it may be that they were trying to emulate the behaviours of participants who were proficient at mathematics. However, it may be more likely that the more anxious someone was, the less they wanted to do the problems, which translated into fewer fixations on the example, and between the example and problem. Likewise, someone who enjoyed mathematics and rated themselves as skilled, would also like to speed through the task, as they did not need the examples and finished the problems quickly. It would be informative to analyze the errors participants had in the learning phase to see if a high anxiety score correlated with poorer performance, which would suggest they rushed through the experiment.

Limitations and Future Research

The analysis of the eye tracking data was at a fairly high level of generality. For instance, our analysis was based on the solution lines of the problem and example pairs, but it might be informative to analyze more specific sequences such as eye gaze patterns that shifted between one variable and another within the same line, or gaze sequences that fixated from a bottom line to a top line of the example or problem. Furthermore, we did not perform analysis on a per problem basis, which could inform us about which problems participants struggled with and how they attempted to solve them. Thus, as of now we are unable to answer questions like: Are participants mapping from one variable in the example to the problem (i.e., superficial or transformational analogy), and does dwell time or the gaze patterns change over time depending on condition (i.e., does it get shorter or longer depending on the form of assistance)?

Another limitation is that participants had all experienced algebra before (although they were certainly not at ceiling as indicated by the pretest). Thus, we may not be testing the learning of a new skill, but rather how to use a skill that has been forgotten. This may imply that our study included less beginners and more participants at the intermediate level of skill acquisition. This might explain why more fixations between the lines of the problem resulted in greater learning gains through the expertise reversal effect, where problem solving becomes more beneficial than example study (Kalyuga et al., 2003). Future experiments using similarity as a method to fade assistance could use a domain not known to the participants, in order to accurately depict how similarity affects learning knowledge in a new domain (Lee, Betts, and Anderson, 2015).

Future studies could also administer a delayed post-test, as it would be valuable to determine if the conditions differed in the amount of retention over time. It has been shown that the more a student fills in their knowledge gaps by self-explaining, the better they retain the information, and if the Dissim-Sim condition was self-explaining more than they would be expected to have greater retention of the material over a longer period of time (Roy and Chi, 2010). Oddly, the Sim-Sim condition seemed to do surprisingly well on the post-test, but the results may have been due to simple repetition and the learning gains may have faded quickly. In contrast, participants in the Dissim-Sim condition could potentially have better knowledge retention, because they had to understand how to complete the problems before being shown (i.e., they filled in gaps by self-explaining).

In future research, it would be interesting to collect data on participants strategies for problem solving (i.e., self-explanation or copying), for instance through talk aloud protocols, and then correlate them with specific eye tracking behaviours. Yet another interesting direction could be to use similarity in conjunction with gaze direction to manipulate the way examples are presented, so that the participants' eye gaze patterns are encouraged to be beneficial for learning. For instance, the dissimilar example could be shown to the participant in segments, based not on lines of the problem, but on the relevant features, allowing them to see the underlying structural similarity first, then see the differences, and finally use them to solve the problem. There is some precedent for this in other instructional activities. For instance, some current eye tracking research has demonstrated that directing a participants gaze toward relevant features of a diagram, even if the participant is unaware of this movement, can increase a participants learning of the material

(Grant and Spivey, 2003; Thomas and Lleras, 2007; Jarodzka et al., 2013)

Conclusions

In our work, using problem-example similarity as a method to fade assistance proved to be beneficial, but only when problem-example pairs were faded from dissimilar to similar. This method also increased problem solving behaviour, as indicated by eye tracking data, and increased time on task. In contrast, using similarity in a traditional faded assistance paradigm did not improve learning in a way as anticipated by prior research. This may have important implications for how examples are presented to students in the classroom, as traditionally, students are taught in the traditional faded assistance manner, and more work is needed to determine if this is truly optimal.

Bibliography

- Aleven, V. A. and K. R. Koedinger (2002). "An effective metacognitive strategy: Learning by doing and explaining with a computer-based Cognitive Tutor". In: *Cognitive science* 26.2, pp. 147–179.
- Anderson, J. R., J. M. Fincham, and S. Douglas (1997). *The role of Examples and Rules in the Acquisition of a Cognitive Skill*.
- Atkinson, R. K., A. Renkl, and M. M. Merrill (2003). "Transitioning From Studying Examples to Solving Problems: Effects of Self-Explanation Prompts and Fading Worked-Out Steps." In: *Journal of Educational Psychology* 95.4, pp. 774–783. ISSN: 0022-0663. DOI: 10.1037/0022-0663.95.4.774.
- Atkinson, R. K. et al. (2000). "Learning from examples: Instructional principles from the worked examples research". In: *Review of educational research* 70.2, pp. 181–214.
- Baroody, A. J. (2003). "The development of adaptive expertise and flexibility: The integration of conceptual and procedural knowledge". In: *The development of arithmetic concepts and skills: Constructing adaptive expertise*, pp. 1–33.
- Berthold, K. and A. Renkl (2009). "Instructional aids to support a conceptual understanding of multiple representations." In: *Journal of Educational Psychology* 101.1, p. 70.
- Brace, N, R Kemp, and R Snelgar (2003). *SPSSfor psychologists*.

- Byrnes, J. P. and B. A. Wasik (1991). "Role of conceptual knowledge in mathematical procedural learning." In: *Developmental psychology* 27.5, p. 777.
- Canobi, K. H. (2009). "Concept-procedure interactions in children's addition and subtraction". In: *Journal of Experimental Child Psychology* 102.2, pp. 131–149.
- Chen, O., S. Kalyuga, and J. Sweller (2015). "The worked example effect , the generation effect and element interactivity". In: *Journal of Educational Psychology* 107.3, pp. 689–704. ISSN: 1939-2176. DOI: 10.1037/edu0000018.
- Chi, M. T. H. and R. Wylie (2014). "The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes". In: *Educational Psychologist* 49.4, pp. 219–243. ISSN: 0046-1520. DOI: 10.1080/00461520.2014.965823.
- Chi, M. T. H. et al. (1989). "Self-explanatirms: How students study and use examples in learning to solve problems". In: *Cognitive Science* 13, pp. 145–182. ISSN: 1551-6709. DOI: 10.1207/s15516709cog1302_1.
- Cooper, G. and J. Sweller (1987). "Effects of schema acquisition and rule automation on mathematical problem-solving transfer." In: *Journal of Educational Psychology* 79.4, pp. 347–362. ISSN: 0022-0663. DOI: 10.1037/0022-0663.79.4.347.
- DeCaro, M. S. and B. Rittle-Johnson (2011). "Preparing to learn from math instruction by solving problems first". In: *B. Rittle-Johnson & MS DeCaro (chairs), When are times for telling.*
- Gick, M. L. and K. J. Holyoak (1983). "Schema induction and analogical transfer". In: *Cognitive psychology* 15.1, pp. 1–38.

- Goetz, T. et al. (2008). "Antecedents of academic emotions: Testing the internal/external frame of reference model for academic enjoyment". In: *Contemporary Educational Psychology* 33.1, pp. 9–33.
- Graesser, A. C. et al. (2005). "Question asking and eye tracking during cognitive disequilibrium: Comprehending illustrated texts on devices when the devices break down". In: *Memory & cognition* 33.7, pp. 1235–1247.
- Grant, E. R. and M. J. Spivey (2003). "EYE MOVEMENTS AND PROBLEM SOLVING : Guiding Attention Guides Thought". In: 14.5, pp. 462–466.
- Hannula, M. S., H. Maijala, and E. Pehkonen (2004). "Development of Understanding and Self-Confidence in Mathematics; Grades 5-8." In: *International Group for the Psychology of Mathematics Education*.
- Hibert, J and P Lefevre (1986). "Conceptual and procedural knowledge in mathematics: An introductory analysis". In: *Conceptual and procedural knowledge; The case of mathematics*, pp. 1–23.
- Holmqvist, K. et al. (2011). *Eye tracking: A comprehensive guide to methods and measures*. OUP Oxford.
- Inglis, M. and L. Alcock (2012). "Expert and novice approaches to reading mathematical proofs". In: *Journal for Research in Mathematics Education* 43.4, pp. 358–390.
- Jacob, R. and K. S. Karn (2003). "Eye tracking in human-computer interaction and usability research: Ready to deliver the promises". In: *Mind* 2.3, p. 4.
- Jarodzka, H. et al. (2013). "Learning to see: Guiding students' attention via a model's eye movements fosters learning". In: *Learning and Instruction* 25, pp. 62–70.

- Just, M. A. and P. A. Carpenter (1980). "A theory of reading: From eye fixations to comprehension." In: *Psychological review* 87.4, p. 329.
- Kalyuga, S. et al. (2003). "The Expertise Reversal Effect". In: *Educational Psychologist* 38.1, pp. 23–31. ISSN: 0046-1520. DOI: 10.1207/S15326985EP3801_4. arXiv: arXiv:1011.1669v3.
- Kapur, M. (2014). "Productive failure in learning math". In: *Cognitive Science* 38.5, pp. 1008–1022.
- Leahy, W., J. Hanham, and J. Sweller (2015). "High Element Interactivity Information During Problem Solving may Lead to Failure to Obtain the Testing Effect". In: *Educational Psychology Review* 27.2, pp. 291–304. ISSN: 1573336X. DOI: 10.1007/s10648-015-9296-4.
- Lee, H. S., S. Betts, and J. R. Anderson (2015). "Not taking the easy road: When similarity hurts learning". In: *Memory & Cognition* 43.6, pp. 939–952. ISSN: 0090-502X. DOI: 10.3758/s13421-015-0509-3. URL: <http://link.springer.com/10.3758/s13421-015-0509-3>.
- Lim, S. Y. and E. Chapman (2013). "Development of a short form of the attitudes toward mathematics inventory". In: *Educational Studies in Mathematics* 82.1, pp. 145–164.
- Marteau, T. M. and H. Bekker (1992). "The development of a six-item short-form of the state scale of the Spielberger State—Trait Anxiety Inventory (STAI)". In: *British Journal of Clinical Psychology* 31.3, pp. 301–306.
- McLaren, B. M., S. J. Lim, and K. R. Koedinger (2008). "When is assistance helpful to learning? Results in combining worked examples and intelligent tutoring". In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial*

- Intelligence and Lecture Notes in Bioinformatics*) 5091 LNCS, pp. 677–680. ISSN: 03029743. DOI: 10.1007/978-3-540-69132-7-75.
- Middleton, M. J. and C. Midgley (1997). “Avoiding the demonstration of lack of ability: An underexplored aspect of goal theory.” In: *Journal of educational psychology* 89.4, p. 710.
- Muldner, K., W. Burleson, and M. T. H. Chi (2012). “Learning from Self-Explaining Emergent Phenomena Study : Student Self-Explanations in Emergent Domains”. In: 2009. ISSN: 18149316.
- Muldner, K. and C. Conati (1998). “Evaluating a Decision-Theoretic Approach to Tailored Example Selection”. In: *Interface*, pp. 483–488.
- Muldner, K. and C. Conati (2010). “Scaffolding meta-cognitive skills for effective analogical problem solving via tailored example selection”. In: *International Journal of Artificial Intelligence in Education* 20.2, pp. 99–136. ISSN: 15604292. DOI: 10.3233/JAI-2010-0004.
- Nokes-Malach, T. J. et al. (2013). “Coordinating principles and examples through analogy and self-explanation”. In: *European Journal of Psychology of Education* 28.4, pp. 1237–1263. ISSN: 02562928. DOI: 10.1007/s10212-012-0164-z.
- Novick, L. R. (1988). “Analogical Transfer, Problem Similarity, and Expertise”. In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 14.3, pp. 510–520. ISSN: 0278-7393. DOI: 10.1037/0278-7393.14.3.510.
- Pachman, M. et al. (2016). “Eye tracking and early detection of confusion in digital learning environments: Proof of concept.” In: *Australasian Journal of Educational Technology* 32.6.

- Pathak, S. A. et al. (2011). "Learning the physics of electricity: A qualitative analysis of collaborative processes involved in productive failure". In: *International Journal of Computer-Supported Collaborative Learning* 6.1, pp. 57–73.
- Reed, S. K., C. C. Ackinclose, and a. a. Voss (1990). "Selecting analogous problems: similarity versus inclusiveness." In: *Memory & cognition* 18.1, pp. 83–98. ISSN: 0090-502X. DOI: 10.3758/BF03202649.
- Reed, S. K., A. Dempster, and M. Ettinger (1985). "Usefulness of analogous solutions for solving algebra word problems." In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 11.1, pp. 106–125. ISSN: 0278-7393. DOI: 10.1037/0278-7393.11.1.106.
- Renkl, a et al. (1998). "Learning from Worked-Out Examples: The Effects of Example Variability and Elicited Self-Explanations". In: *Contemporary educational psychology* 23.23, pp. 90–108. ISSN: 0361-476X. DOI: 10.1006/ceps.1997.0959. URL: <http://www.ncbi.nlm.nih.gov/pubmed/9514690>.
- Renkl, A. (2014). "Toward an instructionally oriented theory of example-based learning". In: *Cognitive Science* 38.1, pp. 1–37. ISSN: 03640213. DOI: 10.1111/cogs.12086.
- Renkl, A. and R. K. Atkinson (2003). "Structuring the Transition from Example Study to Problem Solving in Cognitive Skill Acquisition: A Cognitive Load Perspective". In: *Educational Psychologist* 1520.38, pp. 15/22. ISSN: 0046-1520. DOI: 10.1207/S15326985EP3801.
- Richland, L., O. Zur, and K. J. Holyoak (2007). "Cognitive Supports for Analogies in the Mathematics Classroom". In:

- Richland, L. E., K. J. Holyoak, and J. W. Stigler (2004). "Analogy use in eighth-grade mathematics classrooms". In: *Cognition and Instruction* 22.1, pp. 37–60.
- Rittle-Johnson, B. (2006). "Promoting transfer: Effects of self-explanation and direct instruction". In: *Child development* 77.1, pp. 1–15.
- Rittle-Johnson, B. and M. Schneider (2014). "Developing Conceptual and Procedural Knowledge of Mathematics". In: 1. DOI: 10.1093/oxfordhb/9780199642342.013.014. URL: <http://oxfordhandbooks.com/view/10.1093/oxfordhb/9780199642342.001.0001/oxfordhb-9780199642342-e-014>.
- Rittle-Johnson, B. and R. S. Siegler (1998). "The relation between conceptual and procedural knowledge in learning mathematics: A review." In:
- Rittle-Johnson, B., R. S. Siegler, and M. W. Alibali (2001). "Developing conceptual understanding and procedural skill in mathematics: An iterative process." In: *Journal of educational psychology* 93.2, p. 346.
- Rittle-Johnson, B. and J. R. Star (2007). "Does comparing solution methods facilitate conceptual and procedural knowledge? An experimental study on learning to solve equations." In: *Journal of Educational Psychology* 99.3, p. 561.
- Ross, B. H. (1987). "This is like that: The use of earlier problems and the separation of similarity effects." In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 13.4, pp. 629–639. ISSN: 0278-7393. DOI: 10.1037/0278-7393.13.4.629.
- Roy, M. and M. T. H. Chi (2010). "The Self-explanation Principle". In: *Cambridge Handbook of Multimedia Learning* 0325054, pp. 271–286. DOI: 10.1016/0364-0213(89)90002-5.

- Schwartz, D. L. et al. (2011). "Practicing versus inventing with contrasting cases: The effects of telling first on learning and transfer." In: *Journal of Educational Psychology* 103.4, p. 759.
- Schwonke, R. et al. (2007). "Can Tutored Problem Solving Benefit from Faded Worked-Out Examples?" In: *European Cognitive Science Conference*, pp. 119–120.
- Star, J. R. (2005). "Reconceptualizing procedural knowledge". In: *Journal for research in mathematics education*, pp. 404–411.
- Star, J. and G. Stylianides (2013). "Procedural and Conceptual Knowledge: Exploring the Gap Between Knowledge Type and Knowledge Quality". In: *Canadian Journal of Science, Mathematics and Technology Education* 13.2, pp. 169–181. ISSN: 1492-6156. DOI: 10.1080/14926156.2013.784828.
- Susac, A. et al. (2014). "EYE MOVEMENTS REVEAL STUDENTS' STRATEGIES IN SIMPLE EQUATION SOLVING". In: *International Journal of Science and Mathematics Education* 12.3, pp. 555–577.
- Sweller, J. and G. A. Cooper (1985). "The use of worked examples as a substitute for problem solving in learning algebra". In: *Cognition and instruction* 2.1, pp. 59–89.
- Thomas, L. E. and A. Lleras (2007). "Moving eyes and moving thought: On the spatial compatibility between eye movements and cognition". In: *Psychonomic bulletin & review* 14.4, pp. 663–668.
- Tricot, A. and J. Sweller (2014). "Domain-Specific Knowledge and Why Teaching Generic Skills Does Not Work". In: *Educational Psychology Review* 26.2, pp. 265–283. ISSN: 1573336X. DOI: 10.1007/s10648-013-9243-1. arXiv: arXiv:1011.1669v3.

- Tullis, J. G., R. L. Goldstone, and A. J. Hanson (2015). "Scheduling Scaffolding: The Extent and Arrangement of Assistance During Training Impacts Test Performance". In: *Journal of Motor Behavior* March, pp. 1–11. ISSN: 0022-2895. DOI: 10.1080/00222895.2015.1008686. URL: <http://www.tandfonline.com/doi/abs/10.1080/00222895.2015.1008686>.
- VanLehn, K. (1996). "Cognitive skill acquisition." In: *Annual review of psychology* 47, pp. 513–539. ISSN: 0066-4308. DOI: 10.1146/annurev.psych.47.1.513.
- VanLehn, K. (1999). "Rule-Learning Events in the Acquisition of a Complex Skill: An Evaluation of Cascade". In: *Journal of the Learning Sciences* 8.1, pp. 71–125. ISSN: 1050-8406. DOI: 10.1207/s15327809jls0801_3.
- VanLehn, K. and R. M. Jones (1993). "Better learners use analogical problem solving sparingly". In: *Ann Arbor* 1001, pp. 48109–2110.