2018

Eye Movements and Spatial Ability: Influences on Thinking During Analogical Problem Solving

Bradford Schroeder
University of Central Florida

EYE MOVEMENTS AND SPATIAL ABILITY: INFLUENCES ON THINKING DURING ANALOGICAL PROBLEM SOLVING

by

BRADFORD LOUIS SCHROEDER
B.A. University of Tennessee, Knoxville, 2012
M.A. University of Central Florida, 2016

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Department of Psychology
in the College of Sciences
at the University of Central Florida
Orlando, Florida

Fall Term
2018

Major Professor: Valerie K. Sims

ABSTRACT

Classic studies have examined the factors that influence the way in which people can solve difficult “insight” problems, which require creative solutions. Recent research has shown that guiding one’s eye movements in a pattern spatially congruent with the solution improves the likelihood of formulating a spatial solution. The authors in this line of research argued that guiding eye movements in a pattern spatially equivalent to the solution of the problem yields an embodied cognitive benefit that aids problem solving. Specifically, guiding eye movements leads to the generation of a mental representation containing perceptual information that helps a problem solver mentally simulate the problem features, increasing likelihood to generate a solution to the problem. However, evidence from a small but critically relevant area of research supports that this embodied effect may be more simply a creativity-priming effect. The proposed research aimed to disentangle these ideas while addressing other research questions of interest: do embodied problem solving benefits transfer to later problem solving? Do individual differences in spatial ability influence how people solve these problems? The present study combined previously established methodologies in problem solving and analogical problem solving to investigate these research questions. Results of the present work tentatively support the embodied priming effect, mediated by a creativity-priming effect that influences problem solving performance. Both effects emerged after manipulating problem solvers’ eye movements. There is also modest support for a link between spatial ability and analogical problem solving, but not initial problem solving. These results are interpreted through the lens of embodied cognitive theory, providing tentative support that guiding eye movements can influence reasoning through an enhancement of creativity.

ACKNOWLEDGMENTS

There are so many people who I have been fortunate to have in my life who supported me throughout this hefty endeavor, a multitude of mentors from my years as a younger student who instilled values that have led me to be successful, and a nearly countless number of beloved friends who have made this whole process bearable. Without these people in my life, I would not have made it to where I am today.

First, I would like to acknowledge my advisor, Dr. Valerie Sims, who supported me wholeheartedly even when I was not sure of myself. Although I did not always agree with her in each moment, in retrospect, her advice always steered me toward the right direction. The greatest gift she offered to me as a graduate student was the freedom to explore my own research interests and make my own mistakes. Hard lessons were certainly learned, but I have no doubt in my mind I could not have learned them better any other way. Her careful guidance throughout my entire graduate career allowed me to become the kind of scientist that I always aspired to be.

I would like to thank the members of my dissertation committee, Drs. Jim Szalma, Mark Neider, and Michele Gill, who were thoughtful and supportive in their critiques of my work. Each member of my committee offered suggestions to help me consider my work from a multitude of perspectives I had not considered, and I am grateful for their meaningful comments and interesting opinions throughout the entire process.

Several members of UCF's faculty were inspirational educators and helped me to learn many skills that I still use to this day. Drs. Valerie Sims, Florian Jentsch, Jim Szalma, Mark Neider, Dana Joseph, Mustapha Mouloua, Robert Kennedy, and Kenneth Vehec. This colorful cast of characters were all fantastic educators in their own unique ways and have left me with

many memorable moments from the classroom and beyond. I'll never forget their contributions to my education.

I would like to thank the graduate members of the ACAT lab who supported me during this process: Dr. Shannon Bailey for her helpful advice as a recent survivor of the dissertation process, Daphne Whitmer for her steadfast advice about the "best" dissertation, Ecem Olçum for her unconditionally positive support, Dr. Matthew Chin for his thoughtful comments aiding in refining my proposal and defense, and Jason Parker for always having a fresh perspective. I also effusively thank the undergraduate students who spent an exorbitant amount of time helping me collect data, as well as those who helped me with data coding: Gabriela Flores-Cruz, Griselda Alavez, Sydney Gass, Paulina Maxim, Angie Dekruif, Melodie Spiegel, Olivia Ojalvo, Jacquelyn Schreck, Keirstin Rudl, Ali Trombetta, and Sean Riegle. Without your help this dissertation would not have been possible. I wish all of you luck as each of you go off and do great things in your own graduate programs and in life.

During my time as an undergraduate student in Tennessee, I was fortunate to have many mentors who offered their valuable time and dedication to me. These people were inspiring professors and helped me learn and grow as a young student and researcher: Drs. Deborah Welsh, Katie Rowinski, Zac Wilcox, Shannon Morgan, Matthew Zagumny, and Anne Thurmond. I would especially like to thank Dr. Samantha Gray for taking me on as a research assistant seven years ago. She was an ideal mentor for me and offered many opportunities to expand my research skills through conference presentations, manuscript writing, and data analysis. Most importantly, she was always available for a candid conversation about graduate school. Many of her tips have been instrumental in my survival in graduate school. As someone who also mentored undergraduate students during my time at UCF, Dr. Gray was my model for

mentoring all my undergraduate research assistants. I hope that when my undergraduate students become graduate students and have research assistants of their own, that they also offer them all the opportunities that they had as my undergraduates, and that I had as Dr. Gray's undergraduate.

For the past two years, I have had the pleasure of working with the STRIKE lab at the Navy. I cannot thank them enough for challenging me every day, helping me hone my skills as a scientist. Drs. Wendi Van Buskirk, Cheryl Johnson, Carla Landsberg, Matt Marraffino, Shannon Bailey, Nick Fraulini, and Mrs. Daphne Whitmer have been steadfast sources of support and friendship during my time with them. Special thanks to Dr. Cheryl Johnson who served as my mentor during my Summer internship at the Navy, giving me some seriously awesome hands-on experience with virtual reality research which helped me to fully understand how psychologists use theory to solve real-world problems.

I have been fortunate to have many friends who were probably tired of hearing me talk about graduate school long ago, but still listened to me and offered their camaraderie regardless of how many times we all tried not talking about graduate school. First, I must acknowledge my cohort, who were ultimately my greatest source of peace during the most stressful times of graduate school: Drs. Ada Mishler, Victoria Claypoole, Kimberly Stowers, and Nicholas Kasdaglis, as well as future Drs. Kevin Leyva, Thien-An Le, Jessica Michaelis and Claas Boesser. There could not possibly have been a friendlier and more laid-back cohort than the 2013 entering class at UCF. My friends from my undergraduate years were always a text message away – Peter Robertson, Daniel Knecht, Kayla Jernigan, Shelza Rivas, Kelly Tomlinson, and Kristie Holt – and I appreciate them always keeping me in good spirits with thoughtful conversation or hilarious videos or memes.

Importantly, I would like to thank my parents, Frances Schroeder and Gregory Schroeder, as well as my grandparents, Peggy Schroeder and Winifred Johnson for supporting me throughout this entire process, even when they did not always understand exactly what kind of psychology I was studying. I hope that in my future, I can support my children's and future grandchildren's education as fully and unconditionally as my family had for me.

Lastly, I must thank my future wife, Karleah, for putting up with the long hours and unbridled stress that graduate school brought for me. She never left my side and always kept me company while I forged through my studies. No matter what went right or wrong over the past five years, I could always come back to her to calm myself down and settle my thoughts. Now, she is finishing a graduate program of her own, and it is my turn to return the favor.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE: INTRODUCTION

When faced with a novel and challenging problem, individuals often draw upon past experiences to generate a solution. These past experiences consist of mental representations of similar scenarios holding information that can aid in solving new problems. For example, students who must write an APA-style paper for a class assignment may remember another APA-style paper they had previously written and adapt that paper's format to satisfy their new assignment requirements. A plumber fixing a leaky toilet may draw upon past experiences fixing leaky sinks to aid in his repair job. An elementary school teacher dealing with an unruly student may recall prior experiences with unruly students and apply previously successful classroom management techniques to keep that student well-behaved. Young children learning to read may infer pronunciation rules for new words based on words they have already learned (e.g., pronouncing "cleat" and "meat" based on knowing the word "cheat"). These examples all highlight the usefulness of applying existing knowledge to new problems.

One way that psychologists have studied this transfer phenomenon is through an analogical problem solving paradigm (e.g., Gick & Holyoak, 1980). In this paradigm, participants are often presented with an initial story and its solution and then tasked with solving a new problem. This new problem is completely different at a surface level, but the underlying solution is the same. The studies using this paradigm have identified that, even when people have access to knowledge that is helpful for problem solving, they rarely retrieve it without an explicit hint to use it. Understanding when previously acquired knowledge is retrieved and how to facilitate its transfer to novel situations is the main problem in this body of research. To elaborate, researchers have been interested in answering research questions with the following

general themes: 1. What kinds of prior knowledge can be used to help solve new problems? 2. What characteristics of the individual influence whether prior knowledge is successfully transferred to new problems? 3. What perceptual factors influence whether prior knowledge is successfully transferred?

Solving Problems with Analogies

To answer these research questions, analogical problem solving studies have used Duncker's (1945) radiation problem as a testbed (text from Gick and Holyoak, 1980; pp. 307-308):

> "Suppose you are a doctor faced with a patient who has a malignant tumor in his stomach. It is impossible to operate on the patient, but unless the tumor is destroyed, the patient will die. There is a kind of ray that can be used to destroy the tumor. If the rays are directed at the tumor at a sufficiently high intensity, the tumor will be destroyed. Unfortunately, at this intensity, the healthy tissue that the rays pass through on the way to the tumor will also be destroyed. At lower intensities the rays are harmless to the healthy tissue, but they will not affect the tumor either. What type of procedure might be used to destroy the tumor with the rays, and at the same time avoid destroying the healthy tissue?"

One solution to this problem is to use multiple weak x-rays that converge on the tumor. The weak rays individually do not harm the healthy tissue, but their combined power is sufficient to destroy the tumor. Early research identified that analogies could improve the rate at which problem solvers provide this convergence solution (Gick & Holyoak, 1980; 1983). For example, before seeing the radiation problem, one might view an analogous story of a military general who must send his troops along multiple paths to capture a fortress, or a story of a firefighter who must put out an oil rig fire by surrounding it with fire hoses from multiple directions. In each story, the problem is solved using a converging forces solution, and these solutions can be retrieved to help solve the radiation problem.

In essence, the main task in an analogical problem solving paradigm represents a microcosm for observing analogical transfer of knowledge from one situation to another. Participants receive a source story, and they are tasked with solving an analogous problem. In the literature, the process of retrieving an existing mental representation and applying it to answer a new problem is known as analogical transfer (Gick & Holyoak, 1980). This process follows Sternberg's (1977) account of the component processes of analogical reasoning, where underlying structural features of previously solved problems are "mapped" to the problem at hand, and the novel problem is analyzed within the framework of the original problem. Essentially, the underlying structure is discovered by learning the solutions to analogous problems, which become synthesized into a problem schema. After recognizing corresponding structural features between the problem schema and the novel problem, this problem schema is applied to the novel problem. If the underlying structures match, a solution is generated.

Early research in analogical problem solving (Gick & Holyoak, 1980) investigated whether narrative source stories could aid in solving new (but analogous) target problems. The general idea from these studies was to provide problem solvers with prior knowledge (a source) and test whether they could apply that knowledge to a new problem (a target). This target problem was structurally analogous to the source story, but this work identified that problem solvers were not very likely to apply the solution from the source to the target (specifically, only about 10% were successful; Gick & Holyoak, 1980; 1983). Hinting problem solvers to consider using the source to solve the new problem resulted in increased success rates (to between 30% and 50%), but many problem solvers still remained unsuccessful (Gick & Holyoak, 1980; 1983). Gick and Holyoak argued that the resulting increase in success rates indicated that analogical

transfer was possible, but they and many others became interested in finding ways to improve the rate at which spontaneous (i.e., non-hinted) analogical transfer could be fostered.

Later studies spanning several decades continued this line of research by examining the attributes and quantity of source analogies (Gick & Holyoak, 1983; Holyoak & Koh, 1987), by providing illustrated diagrams (Beveridge & Parkins, 1987; Gick, 1985; Gick & Holyoak, 1983; Pedone, Hummel, & Holyoak, 2001), or by providing animations (Pedone, Hummel, & Holyoak, 2001; Kubricht, Lu, & Holyoak, 2017). These studies largely involved the use of visual or narrative aids and unveiled some general patterns that suggest the key to spontaneous analogical transfer is high quality encoding that results in a high quality mental representation of the problem.

Incorporating Embodied Theories of Cognition into Analogical Problem Solving

Barsalou's (1999) theory of perceptual symbol systems (often more broadly referred to as "embodied cognition" or "grounded cognition") sparked a paradigm shift to incorporate embodied aspects into analogical problem solving research. Pedone and colleagues (2001) tangentially discussed that animations could foster the acquisition of relevant perceptual information, but the earliest mention of an empirical test of embodied cognitive principles in analogical problem solving can be found in Craig, Nersessian, and Catrambone (2002). Below, I briefly review the relevant claims from Barsalou (1999), and discuss how they have been tested in studies on problem solving and analogical problem solving.

Barsalou (1999) offered a perceptual account of human cognition, arguing that cognition does not consist of purely abstract and amodal processes, but is heavily influenced by perceptual information and processes. Instead of viewing the brain as a processor of abstract symbols, he

argued that the brain instead processes perceptual symbols. He used the analogy that the brain serves as a kind of recording device, where the perceptual processes involved in human interaction with the world are recorded. When retrieving from long term memory, or processing information in working memory, this recorded perceptual information is processed through a modal simulator, which functions as a replay mechanism. To illustrate these ideas, he described an example of perceptually simulating a chair. When seeing the word “chair,” one might visualize a chair’s components (e.g., has four legs, a seat, and a back), and one might recall activities associated with chairs (e.g., sitting, standing, pushing). Both of these acts could be considered modal perceptual simulations. By visualizing a chair’s components, one draws upon the visual system to simulate the way the chair looks. By recalling activities associated with a chair, one draws upon the motor system to simulate kinesthetic interactions with a chair (e.g., how it “feels” to sit in a chair). Thus, one’s mental representation for a chair inherently contains relevant perceptual information. Characteristics and interactions with chairs are encoded within various modalities, and when they are retrieved from memory, those modalities are activated in the brain through simulation during the retrieval process. Barsalou noted that modal simulations are incomplete recreations of the original recorded perceptual information (e.g., motor system activation during simulation is a fraction of the activation during encoding), but that they are nonetheless inherent mechanisms for cognitive processing.

Shortly after Barsalou’s (1999) paper was published, researchers in the problem solving literature became interested in testing whether forcing problem solvers to encode relevant perceptual information could aid in problem solving. The general idea is that, when encoding problem features, encoding relevant perceptual information enriches one’s mental representation of that problem beyond what can be encoded from simply reading the problem. In particular,

Pedone and colleagues (2001) argued that a spatial and kinesthetic perceptual understanding of the radiation problem is critical to solving it. During the solution-forming phase of problem solving, mental representations consisting of perceptual information should be more likely to foster a perceptual simulation of the problem, which should improve problem solving success rates. For example, if one re-enacts a converging-forces analog by crashing several blocks together from multiple directions, one should be more likely to produce a convergence solution to the radiation problem than someone who only verbally recalls that converging-forces analog. By encoding the perceptual information in the source analogy, the problem solver would form a richer mental representation for the converging forces solution and be more likely to apply it to a new problem. The studies discussed below primed the motor system during the encoding process.

The earliest example of analogical problem solving research incorporating embodied cognitive themes is discussed in Craig, Nersessian, & Catrambone's (2002) book chapter, where they detail experimental results they would later publish in Catrambone, Craig, & Nersessian (2006). They supported that perceptual information encoded alongside the source analogy aided in solving the radiation problem. Specifically, they found that participants who physically re-enacted the source analogy by crashing blocks together at a central point were more likely to provide the convergence solution for the radiation problem, compared to those who verbally reenacted the source analogy, or those who sketched the source analogy. Craig and colleagues (2002) argued that physically re-enacting the converging forces solution accentuated the kinesthetic structure of the convergence solution (i.e., converging forces colliding at a central location), which led problem solvers to be more likely to generate that solution in an analogous problem.

Other studies have tested embodied cognitive research questions by using the radiation problem as a testbed. Grant and Spivey (2003) and Thomas and Lleras (2007; 2009) discovered that guiding problem solvers' eye movements around a simple diagram influenced solution success rates. Specifically, when participants' eyes were guided in a pattern spatially congruent with a problem's solution (e.g., repeated in-and-out eye movements crossing a central point of the diagram), they were more likely to solve the radiation problem. The main explanation for these results is that guiding problem solvers' eye movements serves to foster a perceptual simulation of the solution. In essence, the guided eye movements helped problem solvers to more easily imagine the solution of multiple beams of x-rays converging at a central point.

Similar research was inspired by these findings and tested embodied cognitive effects in an analogical problem solving paradigm with gesture (Cooperrider & Goldin-Meadow, 2014; Hostetter, Wieth, Foster, Moreno, & Washingon, 2016; Trowbridge, 2016). Their studies investigated whether gesturing a source solution (e.g., having hands coming together from opposite directions onto a central target) could aid in solving an analogous target problem, with the assumption that gesturing about a source story would foster a more comprehensive mental representation that could be more readily applied to a target problem. Unfortunately, none of these studies were able to directly support that gesture conferred any benefit to later problem solving or analogical transfer.

Other Considerations: Attention and Creativity

Altogether, the previously mentioned studies provide mixed evidence that embodying the solution aids in problem solving, but there is an alternative explanation that coincides with Grant and Spivey (2003) and Thomas and Lleras's (2007; 2009) eye movement studies. Thomas and

Lleras (2009) specifically were able to support that attention shifts (a component of the guided eye movements) were the main source of improvement in problem solving rates. If the embodied improvement is attentional in nature, it is worth considering correlates of attention that may relate to problem solving. In particular, creativity has been considered to be a predictor for success in solving insight problems (Gick & Holyoak, 1980; Dow & Mayer, 2004), and some research suggests that creativity is related to one's "breadth of attention" (Eysenck, 1995; Kasof, 1997; Mendelsohn, 1976; Schooler & Melcher, 1995). Further, other studies support that creativity can be primed by subjecting problem solvers to visual tasks that elicit broad visual attention (Friedman, Fishbach, Foerster, & Werth, 2003; Wegbreit, Suzuki, Grabowecky, Kounios, & Beeman, 2012).

Conceptually, these studies may seem only somewhat relevant to the present work, but they are important to consider as the eye movement manipulations in Grant and Spivey (2003) and Thomas and Lleras (2007; 2009) are undoubtedly visual attention tasks. Across these three studies, the experimental manipulations had problem solvers move their eyes in broad and narrow patterns, and these manipulations led to differences in problem solving success rates. Although Grant and Spivey (2003) and Thomas and Lleras (2007; 2009) volunteered a perceptual simulation explanation, it is possible that the effects they saw were due to a potential link between breadth of attention and creativity, which influenced problem solvers' ability to solve Duncker's (1945) radiation problem. The present study attempts to disentangle these competing accounts by modifying Thomas and Lleras's (2007; 2009) paradigm to assess whether their manipulation actually fostered a perceptual simulation of Duncker's (1945) radiation problem, or whether it resulted in a boost of creativity that improved problem solving success.

Other Considerations: Individual Differences in Spatial Ability

Lastly, it is worth mentioning that several studies have examined individual differences with analogical problem solving that may also be relevant for the initial stage of problem solving. Some have focused on cognitive ability (e.g., Antonietti & Gioletta, 1995; Corkill & Fager, 1995; Kubricht, Lu, & Holyoak, 2017) as a predictor for success in analogical problem solving. Antonietti & Gioletta (1995) and Kubricht et al. (2017) both used Raven's Progressive Matrices (e.g., Raven, 2000) as their measure of cognitive ability, but were unable to find a direct relationship between cognitive ability and success in analogical problem solving. One of the main research goals in Kubricht et al. (2017) was to assess the usefulness of animated diagrams as an analogue for problem solving. They found an indirect relationship; such that cognitive ability was predictive of analogical problem solving success only when animated diagrams were not presented. Corkill and Fager (1995) did find a relationship between both creativity and verbal ability independently predicting analogical problem solving success, and Antonietti and Gioletta (1995) found males and biomedical students to be more successful analogical problem solvers than females and non-biomedical students, respectively. Although there are only a few studies that found individual differences in analogical problem solving ability, they still must be assessed to better understand the nature of the emergent differences in problem solving success, both initially and through analogy.

That being said, there is a critical individual difference measure worth considering: spatial visualization ability. In the Grant and Spivey (2003) and Thomas and Lleras (2007; 2009) studies, solving Duncker's (1945) radiation problem is the main focus. This same problem is utilized in essentially all of the analogical problem solving studies (e.g., Antonietti & Gioletta,

1995; Beveridge & Parkins, 1987; Catrambone, Craig, & Nerssessian, 2006; Craig, Nerssessian, & Catrambone, 2002; Corkill & Fager, 1995; Gick, 1985; Gick & Holyoak, 1980; Gick & Holyoak, 1983; Holyoak & Koh, 1987; Kubricht, Lu, & Holyoak, 2017; Pedone, Hummel, & Holyoak, 2001), and it is noted for the spatial structure of its convergence solution (multiple beams focusing inward on a central location). In these studies, participants receive analogies in various forms (e.g., diagrams, animations, narratives) that match the radiation problem's spatial structure, with the assumption that they aid in participants' formation of a convergence solution schema to be applied when solving the radiation problem. The intricacies of these studies are discussed in more detail later, but it is worth recognizing that spatial ability has not been analyzed with respect to Duncker's radiation problem or in the context of analogical problem solving.

The studies using Raven's Progressive Matrices (Antonietti & Gioletta, 1995; Kubricht et al., 2017) perhaps proximally assessed spatial ability (as there are spatial components to the RPM task), but a purer measure of spatial ability may be more directly relevant to the spatial nature of solving Duncker's radiation problem and its analogues. Specifically, a spatial visualization measure of spatial ability may be particularly applicable, due to its role in mental animation (Hegarty & Sims, 1994). Hegarty and Kozhevnikov (1999), Kozhevnikov, Hegarty, and Mayer (2002), and Lohman (1996) argued that people with higher spatial ability are more likely to construct schematic or structured representations of problems they try to solve, which further supports the need to assess spatial ability when assessing initial problem solving and analogical problem solving. It is worth noting that mental imagery (of which spatial ability is a subset; Farah, Hammond, Levine, & Calvanio, 1988; Kosslyn, 1990; Logie, 1995) is considered to be a primary mechanism for mental simulation (Barsalou, 2008; Mandler & Cánovas, 2014),

and if perceptual simulation is key to solving the radiation problem, it follows that spatial ability may be a relevant factor that predicts problem solving success.

Other Considerations: Incubation Effect

The present study includes methodological features (by nature of replicating previous research) that may give rise to an incubation effect. The present work modeled previous studies (e.g., Thomas & Lleras, 2007; 2009) where participants were presented with the radiation problem, provided a seemingly irrelevant visual attention task, and then attempted to solve it again. This feature matches Smith and Blankenship's (1991) description of an incubation effect – an effect that arises when one is faced with a problem and cannot initially solve it, and after time away from the problem, the problem solver returns and is more successful in solving it. Segal (2004) claimed that insight problems are particularly susceptible to an incubation effect, due to the initial impasses they cause upon initial attempts. Segal (2004) further suggested that a period of incubation facilitates a mental restructuring of the problem representation and facilitates the problem solver's success by seeing the problem from a new perspective. To properly examine success rates on the present study's insight problem solving tasks, incubation effects must be assessed to verify they are independent of any potential embodied effects.

Although some have noted that studies on the incubation effect have returned mixed results (e.g., Segal, 2004), a meta-analysis of the incubation effect indicated that incubation effects tend to be small but generally significant (Sio & Ormerod, 2009). In particular, Sio and Ormerod noted that incubation effects tend to be larger for tasks requiring divergent thinking (a construct related to creativity; e.g., Piffer, 2012), and for tasks that require low cognitive effort. The attention-priming paradigms in Grant and Spivey (2003) and Thomas and Lleras (2007;

2009) included conditions that functioned essentially as incubation conditions of minimal cognitive effort, but they had no condition without an incubation. The present study used a modified version of the paradigm used in Thomas and Lleras (2007; 2009), including a condition where no incubation takes place to account for this effect. This permits the comparison of any potential embodied or attentional priming effects against smaller incubation effects.

The Present Study

The previous paragraphs have briefly introduced studies that have examined ways to increase the likelihood of solving Duncker's (1945) radiation problem. Traditionally, it has been used in analogical problem solving paradigms (e.g., Gick & Holyoak, 1980), where Duncker's radiation problem and analogous problems are presented to understand analogical transfer. At the time of this writing, there have been no investigations into whether guided eye movements or attention priming techniques influence analogical transfer in later problem solving. Do these manipulations only result in fleeting improvements in problem solving, or are there latent benefits for solving subsequent problems through analogy? According to those supporting embodied cognitive effects on problem solving, attention guidance congruent with the problem solution should facilitate higher quality encoding that leads to the generation of more robust mental representations that lead to better problem solving success (perceptual simulation hypothesis). The literature on visual attention and creative thinking would support that the same attention guidance would facilitate a boost in creative thinking that leads to better problem solving success (attentional priming hypothesis). It is entirely possible that any benefits seen from eye movement guidance on problem solving could be mediated by creativity, and this possibility is tested in Study 2. Although these accounts both represent embodied cognitive

effects, the former is more specific to the radiation problem, whereas the latter is more generalizable to a broader range of problems. Lastly, are these effects greater in magnitude than an underlying incubation effect? It is possible that previous studies' support for embodied attentional guidance could be more simply explained incubation (incubation hypothesis).

Furthermore, what is the role of spatial ability in solving spatially structured problems? Spatial ability should aid in the construction of spatially structured mental representations during problem solving, which are necessary to solving the radiation problem. Spatial ability also should be associated with a greater ability to mentally simulate the problem, which should result in increased problem solving rates. Lastly, no research has examined how spatial ability relates to analogical problem solving. If spatial ability aids in the generation of spatially structured mental representations, these mental representations should be easier to transfer to new problems with similar structure. The present study aims to combine the attention priming paradigms from previous work with analogical problem solving paradigms to address these research questions.

Research Questions and Hypotheses

Based on Grant and Spivey (2003) and Thomas and Lleras's (2007; 2009) studies, guiding eye movements in a broad pattern mirroring the spatial structure of the convergence solution fosters a perceptual simulation of the radiation problem, leading to increased problem solving success. Relating this more broadly to problem solving, this should aid mental representation construction during the encoding phase, adding perceptual information that can aid in solving the problem. With consideration to the idea that visual attention and creativity are linked (e.g., Friedman et al., 2003; Kasof, 1997; Wegbreit et al, 2012), there is a possibility that the guided eye movements of Grant and Spivey (2003) and Thomas and Lleras (2007; 2009)

primed creative thinking. Creative thinking is considered to be integral in solving problems like Duncker's (1945) radiation problem (Dow & Mayer, 2004; Gick & Holyoak, 1980), thus, it is possible that guiding eye movements primes creative thinking that leads to greater problem solving success. These effects could be additive, such that there is a basic improvement in problem solving success due to creativity priming, and an additional improvement in problem solving success due to fostering a perceptual simulation. It is also possible that these effects are one and the same; such that there is only an improvement in problem solving due to creativity priming, which is incidentally induced with guided eye movements. The major focus of the present study is to address the research question of whether these effects are distinguishable from one another, yielding two competing embodied cognitive hypotheses:

1) *Perceptual simulation hypothesis*: guided eye movements foster a perceptual simulation that improves problem solving success rates, above and beyond any effects due to an attentional priming effect of creativity, and beyond any benefits that can be attributed to incubation.
2) *Attentional priming hypothesis*: guided eye movements only prime creativity that improves problem solving success rates.
    a. *Flexibility hypothesis*: guided eye movements prime divergent thinking which results in a greater quantity of solutions generated, improving the likelihood of solving the problem.

Both of these effects can be considered embodied in nature (to the extent that eye movements influence thought), but Hypothesis 1 is a specific hypothesis that may only be applicable to Duncker's radiation problem and its analogs. Hypothesis 2 has much broader applicability and is more parsimonious. It is also possible that neither of these effects could

occur, which could suggest that the embodied effects from previous studies (Grant and Spivey, 2003; Thomas & Lleras, 2007; 2009) may not be as strong as they have been argued to be.

Because of the relationship between spatial visualization ability and problem solving processes (e.g., Hegarty & Kozhevnikov, 1999; Kozhevnikov et al., 2002), and the radiation problem's spatial emphasis, an additional goal of the present work is to examine how spatial ability is related to problem solving success for the radiation problem and its analogs. Furthermore, because of spatial ability's association with creative thinking (e.g., Shepard, 1978) and its association with mental simulation (e.g., Barsalou, 2008; Mandler & Cánovas, 2014), spatial ability may share variance with any embodied effects in predicting problem solving success. This research question yields the following hypothesis with two sub-hypotheses (a la Mayer & Sims, 1994):

3) *Spatial ability hypothesis*: due to the spatial nature of Duncker's radiation problem, problem solvers who achieve higher scores on a measure of spatial visualization ability should be more likely to solve the radiation problem and its analogs.
    a. *Spatial ability as enhancer hypothesis*: the aforementioned performance benefits of higher spatial visualization ability will be independent of any other effects (e.g., embodied/attentional priming; incubation), such that spatial ability serves as a direct covariate of problem solving success.
    b. *Spatial ability as compensator hypothesis*: the aforementioned performance benefits of higher spatial ability will be less pronounced in conditions where an embodied/attentional priming methodological aspect is present. Specifically, the relationship between spatial visualization ability and problem solving success will be stronger in conditions without guided eye movements. That is, higher spatial

ability will compensate for a lack of methodological aid (i.e., guided eye movements) in predicting problem solving success.

These hypotheses were tested at the initial problem solving stage and later problem solving of analogical problems. It should be noted that, due to the presence of hints throughout the analogical problem solving process, the predictive strength of spatial ability may diminish for the second and third problems. In these cases, performance may be better predicted by an individual's ability to make use of hints, rather than spatial ability. Hints were provided in Study 2, but Study 3 did not provide hints to better observe the effects of spontaneous transfer that may be explained by spatial ability.

Although not a major focus of the present study, the proposed methodology may result in problem solving success differences due to incubation. Previous research indicates that simple cognitive tasks can foster an "incubation effect," such that an initial exposure to a problem that results in an impasse is overcome by removing oneself from the problem, and then re-attempting that problem a few moments later (Segal, 2004; Sio & Ormerod, 2009; Smith & Blankenship, 1991). In order to more accurately assess the aforementioned hypotheses, it is prudent to ensure that any benefits that may be due to perceptual simulation are not confounded with incubation.

4) *Incubation hypothesis*: problem solvers who view the problem prior to receiving a guided eye movements (or fixed eye movements) intervention will show better problem solving success than those who do not view the problem before receiving such an intervention.

All four of these hypotheses have the potential to conflict with one another, or potentially combine with one another. The way these hypotheses were tested is discussed in chapter three and four.

CHAPTER TWO: LITERATURE REVIEW

The present studies aim to better understand the nature of an embodied aid to problem solving and seek to extend its investigation into the paradigm of analogical problem solving. This chapter reviews the literature on problem solving before discussing the research on analogical problem solving. In its review of the analogical problem solving literature, this chapter also highlights a variety of different ways of aiding analogical transfer. Over the course of the past 35 years, researchers have been interested in seeing what factors can improve attempts to solve Duncker's (1945) radiation problem. A modern translation of this problem, used in Gick & Holyoak's (1980; 1983) research, reads as follows:

> Suppose you are a doctor faced with a patient who has a malignant tumor in his stomach. It is impossible to operate on the patient' but unless the tumor is destroyed, the patient will die. There is a kind of ray that can be used to destroy the tumor. If the rays are directed at the tumor at a sufficiently high intensity, the tumor will be destroyed. Unfortunately, at this intensity, the healthy tissue that the rays pass through on the way to the tumor will also be destroyed. At lower intensities the rays are harmless to the healthy tissue, but they will not affect the tumor either. What type of procedure might be used to destroy the tumor with the rays, and at the same time avoid destroying the healthy tissue?

One possible solution to this problem is to use multiple separate rays of weaker intensity that converge on the tumor. The weaker intensity rays are weak enough that they do not destroy surrounding healthy tissue, but when they reach the tumor, their intensities are collectively strong enough to destroy the tumor. This solution has been coined the "convergence solution" (due to the theme of converging forces on a central point; Gick & Holyoak, 1983) and many studies have focused on ways to increase the likelihood that participants respond to the radiation problem with this solution. A variety of research over the past 15 years in this area has begun to examine embodied effects of problem solving. Specifically, they test the idea that by doing something physical (e.g., moving blocks, gesturing, moving eyes) that is relevant to the problem

(i.e., matching the solution spatially or perceptually) one can implicitly improve solution rates for the radiation problem.

Some recent work by Grant and Spivey (2003) and Thomas and Lleras (2007; 2009) identified a low-level implicit means of accomplishing this by guiding participants' eye movements in ways congruent with the convergence solution. The effect they found in their studies is the primary inspiration for the present work. Prior to these studies, research on this topic has identified that narrative analogies, diagrams, simple animations, and physical interactions could increase the likelihood of participants generating the convergence solution. The Grant and Spivey and Thomas and Lleras studies discussed an embodied cognitive explanation for their results, such that eye movements that embodied the convergence solution (i.e., repeated converging in-and-out eye movements crossing a diagram of the tumor problem) fostered a perceptual simulation of the radiation problem, which increased problem solvers' propensity to arrive at the convergence solution. The ending sections of this chapter review these studies in detail and the theoretical explanation behind their results.

The proceeding sections review the pertinent theoretical perspectives and studies leading up to the present work. First, I provide a brief review of literature on problem solving, and then discuss a specific type of problem solving – analogical problem solving, which includes a review of key literature on analogical reasoning ability. After reviewing the problem solving and analogical problem solving literature, I review three key studies that are directly relevant to the present research. Following that discussion, I review studies that pose an alternate explanation for embodied effects on problem solving – attentional priming of creativity. Finally, I discuss individual differences relevant for analogical problem solving and spatial thinking.

Problem Solving

Problem solving is one of the most complex expressions of human intelligence that occurs in everyday life (Chi & Glaser, 1985; Holyoak, 1995). Chi and Glaser (1985) described problem solving as a process that requires an initial encoding phase, where a problem solver "takes in" information about the problem at hand and attempts to form a mental model of that problem. Problem solvers construct these mental models by interweaving problem features with mental structures existing in their long-term memory, which may include knowledge relevant to the problem at hand, similar problems that may have been attempted previously, or other experiences that may emerge during mental model construction. How well a problem is encoded into a mental representation determines the steps a problem solver will take toward finding the solution. They argue that encoding is the most critical step of problem solving, and a key determinant of problem solving success because encoding determines the quality of the mental representation of the problem (Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Chi & Glaser, 1985).

How does one determine whether a situation is a problem that needs to be solved? Duncker (1945) defined problems as arising when one desires a goal but is unsure of how to reach it. Holyoak (1995) echoed this idea, claiming that problems emerge when one aims to reach a goal, but the means of reaching this goal are unclear, or not obvious. Chi and Glaser (1985) simply described a problem as a situation in which one must find a way of reaching a goal. Duncker further explained that when simple actions were insufficient for reaching this goal state, *thinking* was required before the goal could be reached. This process of thinking to bridge

the gap between the initial state and the goal state of a problem can be considered the act of problem solving.

Newell and Simon (1972; as cited by Holyoak, 1995) portrayed the problem solving process with the use of a spatial metaphors. They claimed problem solvers begin at a starting point – the ‘initial state’ – and must journey through a multitude of pathways to attain the ‘goal state’. Many of these pathways lead the problem solver to a dead end, but the goal state can only be reached by a few pathways – sometimes only one pathway. The pathways can be navigated through the use of ‘operators,’ which facilitate the journey from ‘initial state’ into the ‘goal state.’ These pathways can vary in complexity because of problem features that induce ‘path constraints,’ or conditions under which the problem must be solved. They described that these features, and the possible paths that may be taken toward reaching the goal state of the problem, were located within a ‘problem space.’ According to Holyoak (1995), the process of searching through this problem space for a way to reach the goal state is a metaphorical characterization of problem solving.

Chi and Glaser (1985) also described these components of problems, using similar terminology to Newell and Simon (1972) – claiming that all problems have an ‘initial state’ and a ‘goal’. Reaching the goal must be accomplished through ‘operations’, but there are generally ‘constraints’ that dictate how the operations are to be used. Chi and Glaser also described three types of problems that humans typically face: puzzles, classroom problems, and real-world problems. They defined puzzles as being problems that require little background knowledge to solve, tend to be simple and rule-based, but can still be very difficult. An example of a classroom problem would be a physics problem, whereas an example of a real-world problem

would be navigating a city. Unlike puzzles, classroom problems and real-world problems tend to require background knowledge.

Chi and Glaser (1985) further classified problems as being well-defined or ill-defined. They described that well-defined problems are identified by how easily a solution can be identified as correct, and thus, how easily the problem can be recognized as having been solved. Ill-defined problems, on the other hand, involve problems where one or more aspects of the problem are not clearly specified. Of relevance to the radiation problem, Chi and Glaser describe one feature of ill-defined problems as a general ambiguity with regard to how the problem is to be solved. In the radiation problem, the initial state, goal state, and operations are well-defined, but the means of using the operations to reach the goal state are unclear (i.e., it is unclear precisely how the rays should be used to destroy the tumor). Conversely, in a well-defined problem like the Tower of Hanoi (e.g., Lucas, 1882), the operations are well-defined (moving discs from one peg to another), and the means of reaching the goal state are also well-defined (moving discs from one peg to another in a specific order), the problem solver only needs to discover the specific order in which the discs must be moved. Notably, ill-defined problems are often solved when "looking at the problem in a new light," (Holyoak, 1995, p. 285) or by restructuring a mental representation of the problem (Grant & Spivey, 2003; Holyoak, 1995; Knoblich, Ohlsson, Haider, & Rhenius, 1999; Knoblich, Ohlsson, & Raney, 2001), which results in the solution "popping out" to the problem solver.

Incidentally, it is worth discussing that Duncker's (1945) radiation problem is often defined as an insight problem (e.g., Dow & Mayer, 2004; Grant & Spivey, 2003; Knoblich, Ohlsson, Haider, & Rhenius, 1999; Thomas & Lleras, 2007; 2009; Gick & Holyoak, 1980), which is an ill-defined problem where the solution is not immediately apparent to the problem

solver; rather, it "pops-out" to the problem solver in a moment of insight (Novick & Sherman, 2003). Schooler and Melcher (1995) argued that insight problems are particularly likely to create feelings of impasse in problem solvers (compared to well-defined, analytical problems), and Grant and Spivey (2003) and Knoblich and colleagues (2001) identified that one key way of overcoming these impasses is by restructuring one's mental representation of the problem (incidentally, insight is another term for this process; Metcalfe & Wiebe, 1987), or by retrieving a different mental representation that is more suited to solving the problem. Ormerod, MacGregor, and Chronicle (2002) identified that insight problems often lead problem solvers astray during their first attempts at problem solving. Similarly, Dow and Mayer (2004) argued that insight problems prime inappropriate solution representations that lead the problem solver astray, and that creative thinking was required to overcome these inappropriate representations and approach a solution. Thomas and Lleras (2007) cited Metcalfe and Wiebe (1987) and Weisberg and Alba (1981) to further specify that insight problems are particularly difficult, and a problem solver cannot use metacognitive strategies (e.g., monitoring their problem solving performance; Davidson & Sternberg, 1998) to help solve the problem. Instead, Thomas and Lleras argued that the impasse must be overcome before insight problems can be solved. In essence, an insight problem shares the characteristics of ill-defined problems, with the addition of an 'insight' component (i.e., sudden realization of the solution after other attempts have failed; Metcalfe & Wiebe, 1987; Novick & Sherman, 2003).

Duncker's radiation problem can be understood within Newell and Simon's (1972) spatial metaphor and Chi and Glaser's (1985) characterization of problem solving. The initial state is presented such that the problem solver is a surgeon and must destroy an inoperable tumor inside a patient's stomach. The operators are the variable-intensity rays that the surgeon can use

to destroy the tumor. The path constraints are that the tumor cannot be operated on (thus, the rays must be used), and the tumor can only be destroyed with a strong intensity of rays. Additionally, at this strength, the rays will destroy surrounding healthy tissue. If weaker rays are used, the healthy tissue will not be affected, but neither will the tumor. The goal state is where the tumor is destroyed without destroying healthy tissue. When faced with this information, problem solvers must encode all of this information, and begin to generate a mental representation of the problem. Once participants generate a mental representation of the problem, they proceed to attempt solving it, by searching the problem space and testing different pathways. The aforementioned descriptions of ill-defined and insight problems would support that problem solvers will not initially see the pathway to the goal state. In fact, Duncker (1945) identified that many problem solvers do not immediately reach the goal state, but those who eventually reached the goal state generally searched several different pathways before producing a solution.

In Duncker's (1945) study of the radiation problem, he employed think-aloud protocols in an effort to understand the problem solving process. While attempting to solve the radiation problem, participants described their thoughts by speaking them to the researcher. Duncker discovered that participants essentially generated propositions, which they then evaluated in reference the problem constraints, and generated another proposition if that evaluation revealed that the proposition did not reach the goal state. Duncker identified that participants tended to provide propositions in a variety of categories. One of these categories involved various ways to avoid the rays contacting the healthy tissue, and going straight to the tumor (for example, by sending rays through the esophagus into the stomach; by removing healthy tissue to irradiate the tumor, or by moving the tumor outside of the stomach tissue). These solutions, however,

violated the problem constraint that operating on the patient was impossible. The next category he identified concerned immunizing the healthy tissue or protecting it somehow from the rays (for example, through a chemical application that altered the state of the healthy tissue). However, chemical applications were not available as operators for the problem. The final category he observed involved reducing the intensity of the rays on their way to the tumor. Propositions in this category were acceptable within the problem constraints and were a correct solution to the problem if they incorporated multiple weak rays that centered on the tumor. Duncker described his participants' think aloud protocols such that they shifted repeatedly between these three categories of solution propositions until the goal state was reached. Considering Newell and Simon's (1972) spatial metaphor, this process of shifting between propositions reflects the process of searching the problem solving space.

Metacognitive Aspects of Problem Solving

In addition to the aforementioned depictions of the problem solving process, there are also metacognitive aspects worth noting that influence an individual's problem solving performance. To describe how metacognition influences the problem solving process, Davidson and Sternberg (1998) characterized problem solving in terms similar to Chi and Glaser (1985), Holyoak (1995), and Newell and Simon (1972). They described problem solving in terms of givens, goals, and obstacles, which could be mapped onto the operators, goal state, and path constraints, respectively. Davidson and Sternberg also noted that the problem solving process contains an initial encoding phase, during which a problem solver takes in the features of the problem and constructs a mental representation of the problem. Over the course of searching for

the solution, the problem solver may undergo representational change, where they change their mental representation to view the problem from a different perspective.

Although these processes may be natural for most problem solvers, Davidson and Sternberg (1998) note that the regulation of these processes (i.e., metacognitive aspect of problem solving) is not always automatic. Throughout the problem solving process, there are potential metacognitive aspects to consider that can benefit problem solving. When first faced with a problem, the problem solver may be *aware* of a *need* to effectively encode a problem, and further, they may monitor their encoding to judge whether or not it is satisfactory. When constructing mental representations, Davidson and Sternberg cited Newell & Simon's (1972) spatial metaphor to argue that mentally representing a problem consists of four steps. Problem solvers mentally represent the initial state, then the goal state. Next, they test the problem operators on the initial state to see how to transform the initial state into the goal state. The final step involves incorporating the path constraints into the mental representation before finding the solution. Metacognitive influences at this stage may involve monitoring and evaluating mental model construction, and awareness of the need for representational change when appropriate. In moving toward the goal state, problem solvers often generate plans to attain the goal state. Problem solvers employing metacognitive strategies may, for example, monitor these plans as they attempt to reach the goal state and modify them where appropriate.

The aforementioned metacognitive aspects of problem solving relate to the encoding and solution-searching components of problem solving. But Davidson and Sternberg also described several aspects that shape how these processes are carried out. They argued that problems tend to be domain-specific, and that a problem solver's domain knowledge strongly influences the quality of their mental representations of the problem. Specifically, they cite Chi, Feltovich, and

Glaser (1981) to clarify that experts tend to represent problems in a more abstract manner, whereas novices focus on concrete and potentially irrelevant details. Davidson and Sternberg noted that a lack of appropriate knowledge can often lead to an impasse, where a problem solver is completely unable to generate a solution.

When selecting strategies, good problem solvers (as well as those of greater intellectual ability; Sternberg, 1985) are often better able to notice the relevance of previously learned strategies and transfer them to novel problems. Poor problem solvers are less likely to transfer previously learned strategies, or to notice their relevance in novel problems. Strategy use during problem solving can sometimes result in stereotypy, which is where a problem solver gets "stuck" using a strategy that does not work. A metacognitive awareness of when one has reached stereotypy can lead to problem solvers freeing themselves from fixating on a failing strategy, but Davidson and Sternberg note that problem solvers with poor metacognition often fail to break such a fixation. In general, metacognition can influence how a problem solver attempts to solve a problem, but Davidson and Sternberg argue that a general metacognitive ability to monitor and evaluate one's knowledge about the problem, one's progress toward reaching the solution, and one's mental representations, strongly predicts success in problem solving.

Representing the Problem

Although Chi and Glaser (1985) use the terms "representation" and "mental model" relatively interchangeably, another term relevant to the mental representation of knowledge is "schema" (e.g., Lakoff & Johnson, 1980). Lakoff and Johnson's schema theory asserts that conceptual and perceptual knowledge about the world is organized into units called "schemata."

Schemata are constructed over time, through experience, and inform an individual how to process present information. Norman (1986) describes schemas as being flexible structures that constantly change with experience, allow for generalizations from past experiences, and shape one's interpretations of immediate circumstances. For the purposes of the present work (and to mitigate terminological confusion with the most closely relevant research), I use the term "schema" to describe mental representations of problems that participants construct during problem solving.

Research in the expertise literature has identified that problem schemas can vary as a function of individual differences. Chi et al. (1981) discovered that physics experts approached different physics problems in a distinctly different way from physics novices and argued that experts are more likely than novices to have developed schemas for the underlying problem structures of different types of physics problems. They argued that the knowledge structures of experts and novices are characteristically different – showing that experts were better able to categorize physics problems by their underlying structure, whereas novices tended to categorize physics problems by their surface features (and also tended to be less successful at solving problems). Though expertise appears to affect how problem solvers represent their knowledge, important work has examined how to increase the rate at which people generate structural schemas for problems. Considering Chi and Glaser's (1985) work, these studies have essentially examined either ways to influence the encoding phase of problem solving, or ways to influence the retrieval of prior knowledge, by providing analogical problems. Several characteristics of analogical reasoning have been discovered through problem solving paradigms where participants are provided some variation of a problem and its solution, then tasked with solving a novel problem containing the same structural solution as the previous problem. Ideally,

analogical transfer occurs, and the problem solver maps the analogy to the problem at hand and forms a solution. However, this is not always the case. The key research questions addressed by previous work involve understanding the ways in which this process, analogical transfer, can be facilitated, induced, or enhanced to improve solution rates for novel problems. Prior to discussing these studies, it is imperative to review theoretical work on analogical reasoning.

Analogical Reasoning

In general, analogical reasoning is claimed to be a multi-step process (Gentner, 1983; Sternberg, 1977) that is closely linked with creative thinking (Dow & Mayer, 2004; Gentner, Brem, Ferguson, Wolff, Markman, & Forbus, 1997). Craig, Nersessian, and Catrambone (2002) characterize it as a three step process involving mental representations – retrieval, alignment, and mapping. When faced with a novel problem, a problem solver undergoes a retrieval process, where multiple source mental representations are selected that are systematically relevant to the target problem. Next, these source representations are aligned with the target problem, such that their features are judged as to their relevance to the target problem. Lastly, features of the source representations are mapped onto the matching features of the target problem. During mapping, the existing source mental representation's structure is applied to the actively forming mental representation of the target problem. Applying the structure of a source representation to a target problem results in successful analogical transfer. It can be helpful to think of this process with a visual analogy. Below, an illustration of the process as described by Craig et al. (2002) is provided.

In the images below (Figures 1-4), assume each symbol represents an element of a mental representation. Each line connecting these elements represent structural relations between these

elements. Consider the following target problem: “How are the ‘T’ elements structurally related?”

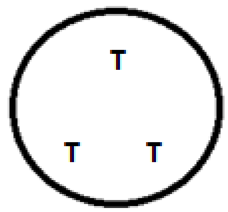

Figure 1. Analogical target problem example depiction.

Assume the solution for this problem is a structural relationship such that each of the elements are connected in the shape of a triangle with a point at its top. How would a problem solver carry out the analogical problem solving process? First, several relevant source mental representations are *retrieved*, which are assumed to contain underlying structure.

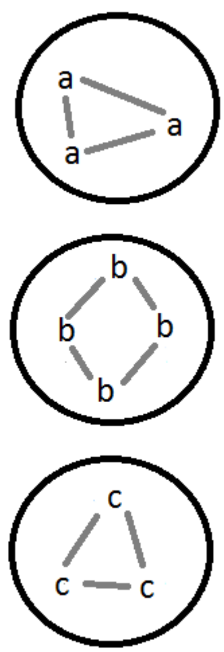

Figure 2. Depiction of retrieval process.

Next, source representations are then *aligned* to the target problem based on similar features.

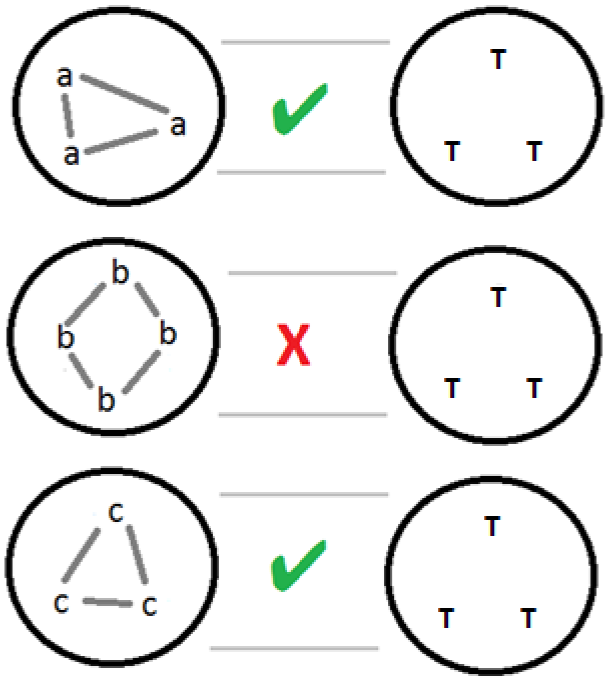

Figure 3. Depiction of alignment process.

In this example, the problem solver aligns source representations to the target problem on the basis that they also contain three elements.

Finally, in the *mapping* phase, the problem solver imposes the structural relationships of the aligned representations onto the target problem. When the underlying structure of a source representation is successfully mapped onto the target problem, the problem solver has discovered a solution to the target problem.

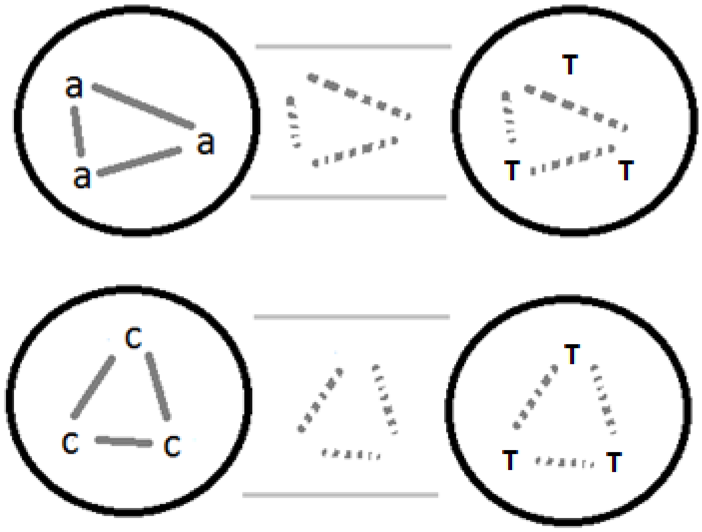

Figure 4. Depiction of mapping process.

This illustrated example is based on Gentner's (1983) structure mapping theory of analogy which, generally speaking, involves the retrieval of source representations that must be aligned to a target problem, and then mapped to generate a solution to the analogy. Sternberg (1977) offers a similar theoretical perspective describing similar processes but is cited more frequently in the analogical problem solving literature. In Sternberg's (1977) theory, solving an analogy requires an initial *encoding* phase, where the features of an analogy are encoded (i.e., the source and target analogies). Next, the problem solver must *infer* the relationship between the encoded features of the source analogy and the target analogy. Then, the problem solver *maps* the relationship between the source analogy's features and its solution. This mapped relationship is then *applied* to the target problem's features, and a solution to the target problem can be generated. These theories are similar to one another, but their differences can be more easily understood using a simple semidegenerate analogy metaphor (i.e., A_1:B_1::A_2:B_2; see Figure 5).

Sternberg (1977)		Gentner (1983)	
$A_1:B_1::A_2:B_2$ Source features : Source solution : : Target features : Target solution $A_1:B_1$ = Source problem $A_2:B_2$ = Target problem Analogy problem as presented: $A_1:B_1::A_2:$__			
1. *Encode* features of the analogy	$A_1:B_1::A_2:$__	$A_2:$__	1. *Encode* features of the target problem
2. *Infer* the relations between the encoded features of the source analogy and the encoded features of the target analogy	$A_1:A_2$	$A_1:B_1$	2. *Retrieve* source representations that may be useful for solving the target problem; among them may be the source problem representation
3. *Map* the relationship between the source problem features and its solution	$A_1:B_1$	$A_1:A_2$	3. *Align* source representation features to target problem features, filtering out source representations that do not align with target problem features
4. *Apply* the previously mapped and inferred relationship to the target problem features; evaluate whether a solution is feasible based on this applied relationship	$A_1:B_1$ ↓ $A_2:B_?$	$A_1:B_1$ ↓ $A_2:B_?$	4. *Map* the source representations to the target problem; if one source representation successfully maps to the target problem, then a solution to the analogy is formed
5. *Respond* to problem: provide a solution; assuming application stage was successful	$A_2:\mathbf{B_2}$	$A_2:\mathbf{B_2}$	5. *Respond*

Figure 5. Analogical problem solving steps for a semidegenerate analogy, according to Gentner and Sternberg.

Note: The steps for "Encode" and "Respond" were added to Gentner's (1983) theory.

The outline provided in Figure 5 illustrates minor differences in Gentner and Sternberg's theories, but ultimately, each theory leads to a similar result through a similar process (e.g., "map" in Gentner vs. "apply" in Sternberg). However, Gentner's structure mapping theory involves the retrieval of source representations, one of which may be the $A_1:B_1$ relationship. The features of the source representations are aligned to the target problem, at which point source representations that do not fit with the target problem features are discarded. Any remaining source representations are then mapped to the target problem. During the mapping process, if a source representation is successfully mapped to the target problem (e.g., mapping features *and* solutions), then a solution to the target problem can be generated.

Considering Sternberg's theory, the analogical problem solving process begins with encoding the features of the analogical relationship altogether. Next, the relationship between the features of the source problem and the features of the target problem are *inferred*. Then, the relationship between the source problem features and its solution is *mapped*. Finally, the mapped relationship is *applied* to the target problem, and the feasibility of this solution is evaluated. If the solution is deemed feasible, then the problem solver *responds* with their proposed solution.

Taken together, these theories both offer a similar account of how analogical thinking is used to solve problems. Gentner's theory is perhaps better suited to solving novel problems, where source representations may not be readily available, and must be retrieved from memory. For these novel problems, retrieved source representations are essentially tested within the frame of the target problem. On the other hand, Sternberg's theory is probably better suited more directly to solving problems by analogy, where a source problem and its solution is available to guide the search for a target problem's solution. With consideration to analogical problem solving paradigms (e.g., Gick & Holyoak, 1980), the concept of spontaneously transferring an analogy (i.e., solving a novel problem using a recently viewed analogy) can be better accounted for by the processes in Gentner's theory, whereas the concept of cuing analogical transfer or "mapping" the analogy (after hint provision, as per Gick & Holyoak, 1980; 1983) can be better accounted for by the processes in Sternberg's theory. It is critical to consider both of these theories, as they offer reasonable explanations for the component processes that underlie the pertinent outcomes (e.g., spontaneous transfer or cued transfer) in analogical problem solving paradigms.

It is worth mentioning that all the aforementioned steps are internal, mental actions, except for the final step of responding. While these mental operations are carried out, it is possible that the problem solver runs mental simulations of the source knowledge that they retrieve and submit the attributes of the source knowledge to another simulation when attempting to apply the solution to the target problem. As more studies have been conducted on analogical problem solving, the idea that mental simulations are key to successful problem solving has become more prevalent. The following section discusses key studies in analogical problem solving review these studies, as well as other pertinent studies that support and fail to support this idea.

Analogical Problem Solving

Analogical problem solving is considered to be a particular type of problem solving that involves components of analogical reasoning, where a problem solver retrieves analogically related source knowledge, and applies it toward solving a novel target problem. Extending from the literature of problem solving, Gick and Holyoak (1980) pioneered the early research into analogical problem solving. In their work, they studied whether or not narrative analogies could improve the rate at which people solved a structurally similar target problem, Duncker's (1945) radiation problem. Solution rates for this problem are abysmal without hints. Duncker (1945) identified 2 out of 42 participants were able to solve the radiation problem with hints, whereas more modern studies (e.g., Gick & Holyoak, 1980; 1983; Thomas & Lleras, 2007) have found non-hinted solution rates of around 10%. However, Gick and Holyoak's early work demonstrated that narrative analogies could be used successfully to increase the solution rate of Duncker's radiation problem, and further work has demonstrated the utility of other kinds of

analogies, such as diagrams (e.g., Beveridge & Parkins, 1987; Gick & Holyoak, 1983, Pedone, Hummel, & Holyoak, 2001) and manipulation of physical objects (e.g., Catrambone, Craig, & Nerssessian, 2006; Craig, Nerssessian, & Catrambone, 2002). The key ability studied in this body of research is *analogical transfer* – an individual's ability to retrieve a source analogy and apply its structure toward solving a novel problem.

In the real world, the process of analogical problem solving can be illustrated with a student-teacher metaphor. Consider students in a classroom learning a general principle during a lecture. As they encode the information from their lecturer, they construct a mental representation of that principle. During this encoding process, the teacher may demonstrate an application of this principle (e.g., by solving an example problem). When the students are tested with a new problem using the same principle, ideally, they will retrieve the mental representation they had constructed earlier (source) and identify corresponding features between the mental representation and the target problem. Once those features are identified, they should transfer the solution from the source representation and use it to solve the target problem. However, this is not always the case. A problem solver may have access to one or more source representations, but they are not always spontaneously retrieved when solving a new problem. Furthermore, if a source representation is successfully retrieved, it may not always be applied to the target problem correctly – particularly if the source was not correctly encoded or if the problem solver becomes distracted by irrelevant features in source representations. In traditional analogical reasoning problems (e.g., A_1:B_1::A_2:_?_), the source (A_1:B_1) is always present, and thus retrieval is not necessary. However, in analogical problem solving paradigms (e.g., Gick & Holyoak, 1980), the source (e.g., a narrative story) is typically unavailable during attempts to solve a target problem, and solution rates for target problems tend to be rather low without hints to retrieve the narrative

story. If the source is not retrieved while solving the target problem, the problem solver would perceive the problem as being completely new. If the source is retrieved, however, the problem solver would more readily notice the correspondences between source and target and have a greater chance of solving the target problem. Gick and Holyoak's (1980) work (and others that followed it) elaborates this process and is discussed in more detail in later sections.

Between the theoretical work of Sternberg (1977) and Gentner (1983), solving simple analogies is comprised of basic beginning and ending stages; encoding and responding, respectively. In all cases, a problem solver must encode the features of a problem before attempting to solve it. Similarly, once a problem solver has found a solution, a response is required to indicate that a solution has been found. With the contemporary literature, Sternberg (1977) identified that the in-between steps are generally disagreed upon. Gentner's (1983) steps are somewhat similar to Sternberg's (1977) steps, with the only differences being which aspects of the analogical relationship are linked together.

Task Analysis for Analogical Problem Solving Paradigm

In analogical problem solving studies, the usual experimental paradigm is based on Gick and Holyoak's (1980) experiments, where a problem solver is provided with a source story and its solution, then tasked with solving a target problem in a different domain. This target problem can be solved analogously to the source story and solving the target problem in such a way is considered to be analogical transfer (i.e., the problem solver saw the analogical relationship between source and target and transferred the source's solution to the target). With consideration to the literature on problem solving (e.g., Chi & Glaser, 1985; Davidson & Sternberg, 1998; Holyoak, 1995; Newell & Simon, 1972), as well as the literature on analogical reasoning (e.g.,

Gick & Holyoak, 1983; Gentner, 1983; Sternberg, 1977), one can construct a task analysis for the processes a problem solver experiences during an analogical problem solving paradigm (e.g., Gick & Holyoak, 1980). The example below outlines this process with the radiation problem as the source analogy and "Red Adair" as the target problem.

Receive Source Problem

1. Encode source problem features (radiation problem)
    - Encode initial state – Must remove a cancerous tumor from a patient
    - Encode operators – x-rays
    - Encode path constraints – cannot x-ray with full force or will destroy healthy tissue, cannot operate on tumor
    - Encode goal state: "How do you kill the tumor without harming the patient's healthy tissue?"
2. Encode solution to problem
    - Solution is to send multiple weaker x-rays to converge at the central tumor. The weak rays are weak enough to leave the healthy tissue unharmed, but their summative power at the central location is enough to destroy the cancerous tumor.

Throughout this process, the problem solver has generated a mental representation of the radiation problem.

Solve Target Problem

1. Encode target problem features (Red Adair)
    - Encode initial state – must put out a fire in an oil well
    - Encode operators – fire hoses

- Encode restrictions – cannot spray the fire all at once or it will spread to surrounding parts of the oil well
- Encode goal state – "How do you put out the oil fire without causing more damage to the oil well site?"

2. Devise solution(s) to the problem

- (Retrieval) search through long term memory for potentially relevant information
    - One piece of relevant information may be the mental representation of the radiation problem
    - Other relevant information may be domain-specific knowledge, such as how grease fires react to water (e.g., they spread easily)
- According to Gentner's (1983) steps, the problem solver will:
    - Attempt to align the retrieved source representation features to the target problem features. Assuming the radiation problem was retrieved, its features (e.g., initial state, operators, path constraints, and goal state) will be aligned with the features of the target problem. For example, the x-rays and fire hoses would be aligned with one another.
    - Next, assuming the alignment process was successful, the problem solver would map the source representation (which consists of a solution) to the target problem. In this mapping process, the solution from the source representation is tested on the target problem. If this stage is successful, a solution to the target problem will be formed.
        - Gick and Holyoak would consider this result *spontaneous analogical transfer*, where a problem solver spontaneously transfers the previously

encoded analogical story and volunteers an analogous solution to the target problem.

- If spontaneous analogical transfer does not occur, the participant may be cued to refer to the earlier encoded source problem. Upon retrieving that cued source representation, a problem solver would then follow Sternberg's (1977) analogical reasoning steps, as they now have access to the source problem and target problem for the analogical relationship.

- They must first infer the relations between the features of the source representation and the solution of the source representation.
- Then, they must map the features of the source representation to the features of the target problem. Mapped relationships may be…
    - (mapped initial state) Must get rid of cancerous tumor > must put out fire
    - (mapped operators) x-rays > hoses
    - (mapped path constraints) do not damage healthy tissue > do not spread fire
    - (mapped goal state) "How do you kill the tumor…" > "How do you put out the fire…?"
- If inference and mapping are successful, then the problem solver must apply the solution understood from the source representation to the target problem.
- However, Sternberg (1977) notes that these processes are repeated in a self-terminating fashion, until success is reached (or the problem solver gives up)

        - If inference is initially unsuccessful, the problem solver may need to re-evaluate the relationship between the features and solution of the source knowledge.
        - Similarly, if mapping is initially unsuccessful, the problem solver may need to retry based on newly inferred relations.
    - If application (and all subsequent steps) is successful, then the problem solver has discovered the solution to the target problem and recognized the analogical relationship between their source representation(s) and the target problem.
    - If application is unsuccessful, the problem solver may need to retry based on a newly mapped relationship between source knowledge features and target problem features.

It should be noted that Sternberg's (1977) steps could also be considered relevant for cases of spontaneous analogical transfer, but Gentner's (1983) steps were selected for this task analysis because Gick and Holyoak often argued that problem solvers were not always aware of the analogical relationship until they were cued to try and apply the source problem to the target problem. Thus, Gentner's steps more closely account for spontaneous analogical transfer (the analogical relationship is not laid out explicitly for problem solvers), but Sternberg's (1977) steps are better suited to cued analogical transfer (where the analogical relationship is more clearly explicated for problem solvers).

The following sections review the literature using this paradigm to explore factors that influence problem solving success through analogy.

Improving Problem Solving Success Rates

Around World War II, Duncker (1945) introduced the world to the radiation problem. Notably, this problem was incredibly difficult, and Duncker's analysis of 42 problem solvers found that only 2 of them were able to solve it successfully. This problem received very little attention in the literature for several decades, save for a brief mention in Resnick and Glaser (1976). After Gick and Holyoak (1980) reintroduced the radiation problem, the past few decades have seen its extensive use in experiments examining problem solving and analogical problem solving (incidentally, it maintained a baseline difficulty rate such that around 10% of problem solvers get it right on their first try; Bassok, 2003). With respect to the studies on problem solving, participants' solution rates increase when their attention or eye movements are guided in a spatial arrangement congruent with the convergence solution (Spivey, 2003; Thomas & Lleras, 2007; Thomas & Lleras, 2009). With respect to analogical problem solving, solution rates improve with the use of narrative analogies (e.g., analogous stories; Craig, Nersessian, & Catrambone, 2002; Gick & Holyoak, 1980, 1983; Holyoak & Koh, 1987), illustrated analogous diagrams (Beveridge & Parkins, 1987; Gick & Holyoak, 1983; Holyoak; Kubricht, Lu, & Holyoak, 2017; Pedone, Hummel, & Holyoak, 2001), and analogous physical interaction with objects (Catrambone, Craig, & Nersessian, 2006). The explanations as to why these features aid in problem solving is that they facilitate the formation of a "convergence schema" (an amodal mental representation of the convergence solution), or, the schema represented by a convergence of forces onto a single target. This schema is deemed critical to solving Duncker's (1945) radiation problem (e.g., Gick & Holyoak, 1983). It is worth noting that several of these studies refer to the convergence schema as a mental representation that can be induced (e.g., Gick &

Holyoak, 1983), whereas others refer to the convergence schema as being something that is simply activated during encoding and utilized during problem solving (e.g., Craig et al., 2002). The following sections review these studies.

Early Work using Analogies to Improve Problem Solving Success

The seminal work on analogical problem solving (Gick & Holyoak, 1980) outlined a paradigm that has been used in a multitude of studies, employing minor alterations to answer different kinds of research questions. The basic paradigm is to provide a problem solver with a source analogy of some kind – be it a story, illustration, or animation – and then present a target problem in a new domain that shares many of the underlying features of the source analogy. By varying features of the source analogies and examining solution success rates on the target problem, one can infer the underlying cognitive processes that relate to analogy use and analogical thinking. In a series of five experiments Gick and Holyoak (1980) employed Duncker's (1945) radiation problem as their target problem, which they note requires a degree of creativity to solve. Duncker (1945) originally identified that the radiation problem was difficult to solve without any hints, and Gick and Holyoak (1980) were interested in testing whether analogical reasoning processes could be used to aid in solving the radiation problem; a process they dubbed "analogical problem solving."

In these experiments, Gick and Holyoak (1980) utilized a story about a military general as the source analogy that they presented to problem solvers. The story concerns a military general that must overthrow a dictator by sending his army to capture his fortress. The general learns that the roads leading up to the dictator's fortress are laden with weight-sensitive land mines that would explode if he sent his full force down one road at once; however, the full force

of his army would be required to overthrow the dictator. All variations of this story had this same information (i.e., initial state, path constraints, goal state), but a variety of solutions were provided in the first experiment. The solutions were (1) that the general sent his army down an open (and mine-free) supply route, or (2) he dug a tunnel (beneath the mines) that led to the foot of the fortress, or (3) he separated his army into smaller units that travelled across multiple roads and met at the fortress. All of these solutions are feasible for the general problem, but only one of these is suitable as an analogous solution to the radiation problem. The third solution mirrors the convergence solution of the radiation problem – using multiple, weaker rays to attack the tumor without risking damage to the healthy tissue. The other solutions were crafted based on some of the typical responses Duncker (1945) identified – the open supply route and tunnel solutions most closely mirror the solution propositions that involve sending the rays down an alternative path (e.g., the esophagus or intestines), but these solutions violate the constraint that the patient cannot be operated upon.

Because different problem solvers received different solutions in their source analogies, Gick and Holyoak's (1980) first experiment examined differences in solution rates for the radiation problem based on the different solutions. In general, the solution participants received in the general problem determined which solution they provided for the radiation problem. For example, those who received the tunnel solution or the open supply route solution were more likely to offer solutions that required operation (e.g., cutting open the stomach; or sending rays through the intestines). A significantly greater amount of convergence solutions (using multiple weaker rays focused on the tumor from different angles) were provided by those who received the solution where the general divided his military along multiple roads. In the first experiment, it should be noted that participants were prompted to refer to the general problem that they had

previously received while attempting to solve the radiation problem. Within the frame of analogical reasoning, the previously mentioned results accounted only for analogical mapping (i.e., participants were prompted to try to use the general problem story they had read before), however, 2 of 42 problem solvers spontaneously volunteered the convergence solution, using the general problem as an analogy. These results largely suggested that the kind of information available in a source analogy strongly influences way target problems are solved, and more importantly, that analogical transfer from a recently learned source problem to a seemingly novel target problem was possible.

Their second experiment further explored how alterations in source analogy content influenced solution rates for the radiation problem. Like the first experiment, participants received a variety of stories to use as source analogues (as well as there being a control group who received no source analogy), but one of these stories had a completely different problem statement. One of the general problem stories was changed to be about throwing a parade for the dictator, rather than attacking his fortress. The result was the same (the dictator wanted a parade with portions of the army marching down all roads leading to his fortress), but because the problem was different, Gick and Holyoak (1980) were interested in assessing whether a source analogy with some elements of dissimilarity would lead problem solvers to the convergence solution on the radiation problem. Results indicated that both groups volunteered the convergence solution more frequently than a control group, and there were no differences between people who received the general story (with the separated army solution) and people who received the parade story. These results suggested that analogies with dissimilar elements could still foster the use of the convergence solution for solving a later problem, providing that the underlying structure of the solutions is analogous to the target problem.

In both experiments, the provided solutions held a strong influence on later solution generation; however, Gick and Holyoak (1980) noticed that that this influence was not absolute. Often, participants volunteered completely different solutions on their own. In fact, a few who did not receive the general's split army solution in their source story spontaneously volunteered the convergence solution in the radiation problem. For their third experiment, Gick and Holyoak were interested in exploring whether self-generated solutions for source problems could influence the solutions that are generated in later problems. As in both prior experiments, participants received the general problem, but in this experiment, participants had to come up with their own solutions. When faced with the radiation problem, they were hinted that the solutions they generated for the first problem may be helpful for the radiation problem, but that they were not necessarily needed to be able to solve the radiation problem. Gick and Holyoak identified that 22 of 45 subjects (49%) volunteered the convergence solution for the general problem. Of those 22, 9 (41%) produced the convergence solution for the radiation problem, and of the other 23 who did not initially provide a convergence solution, only 3 participants (8%) spontaneously produced the convergence solution on the radiation problem. Gick and Holyoak noted that, because these results were not experimental in nature, that there could be some underlying problem solving ability that led people to be more likely to volunteer a convergence solution. This would have emerged when comparing the two groups on the radiation problem – the ones who volunteered the convergence solution on the general problem may have also been more likely to volunteer the convergence solution on the general problem due to some underlying factor. However, there was another solution to the general problem that appeared to influence solutions for the radiation problem. Several participants (24 of 45) suggested that the general could send several small groups down the same pathway to avoid setting off the mines,

but still have the full force available at the dictator's fortress, albeit later. Interestingly, a large portion of these participants (10 of 24; 42%) volunteered a solution to the radiation problem that was analogous to this one. Their solutions involved shooting weaker rays at the tumor over an extended period. Although this solution would be infeasible for the radiation problem (the effects of the weaker rays would still accumulate on the healthy tissue, as well as the tumor), these results indicate that self-generated solutions to an initial problem do influence solutions generated for novel problems, even when the initial solutions are not particularly suitable for the novel problem.

To further explore the third experiment's findings, they empirically evaluated how problem solving was affected by the provision of multiple stories, each with different solutions. One story was the general problem with the convergence solution, and the other two were designed to be as mismatched to the radiation problem as possible, while still being similar in length and form to the general problem (i.e., describing problems, constraints, and providing solutions). After reading these stories, participants were divided into two conditions – one where they were hinted that one of the stories they had read would help in solving the radiation problem, and another where they received no such hint when attempting to solve the radiation problem. There were stark differences in performance; 11 of 12 in the hint condition generated a convergence solution to the radiation problem, but only 3 of 15 in the 'no-hint' condition generated a convergence solution. Gick and Holyoak were interested in whether the 'no-hint' participants had considered using the stories, and whether any participants considered any specific story to be useful for solving the radiation problem. Among the 'no-hint' participants, 12 of 15 had not considered using the previous stories, but 2 of those 12 participants spontaneously volunteered the convergence solution. Interviews with those 2 participants

indicated no clear link between the earlier problems and volunteering the convergence solution, overall indicating that spontaneously noticing an analogy between source and target problems to be a rare occurrence. For the participants who did receive hints, Gick and Holyoak argued that this experiment evinced that it was possible to search through memory to retrieve helpful information encoded from previous problems (even among distracting information) and use it as an analogy to help solve a novel problem.

Gick and Holyoak noted that the performance differences they saw in experiment 4's 'hint' and 'no-hint' conditions indicated that analogical problem solving is not usually a spontaneous process: it usually requires a nudge from an experimenter. The final experiment aimed to understand whether spontaneous analogy use under more optimal conditions. One group ("story first") received only the general problem and its solution prior to seeing the radiation problem, and another group saw the radiation problem prior to seeing the general problem and its solution ("story second"). In both cases, there were no distractor stories that could negatively impact retrieval of source knowledge from memory, but the "story second" condition emulates situations that occur in everyday life, where one happens upon useful information while trying to solve a different problem. In both of these conditions, when participants received the radiation problem, they were given a ten minute period to attempt solving it. Afterward, they received a hint to try using the general problem to solve the radiation problem. However, the key difference in the "story second" condition was that there was an additional 10-minute attempt to solve the radiation problem prior to ever seeing the general problem. Gick and Holyoak noted that having initial exposure to the radiation problem, and then solving it again a few moments later could be argued as an incubation effect, rather than emulating situations where one happens upon useful information during problem solving. To

address this, they added an incubation control condition, which was identical to the "story second" condition, except they received an unrelated distractor problem instead of the general problem. Interestingly, results indicated that both the "story first" and "story second" conditions nearly equally generated the convergence solution both before receiving a hint (41% and 35%, respectively) and after receiving a hint (35% and 30% respectively). Additionally, 10% of participants in the "story second" condition spontaneously produced the convergence solution after their first exposure to the radiation problem (prior to receiving the general problem). The "incubation only" group was drastically different: only one participant volunteered the convergence solution upon first viewing the radiation problem. The remaining participants in the "incubation only" group never provided the convergence solution. In essence, problem solvers who received a helpful source analogy were about 3 to 4 times as likely to spontaneously transfer that analogy to a novel problem, as compared to a control group that did not receive a helpful source analogy. In addition, these results indicated that incubation alone was simply not enough to increase the likelihood that someone would volunteer the convergence solution to the radiation problem.

All five of Gick and Holyoak's (1980) experiments demonstrated that problem solvers can retrieve knowledge from memory and apply it to solve analogous problems, even when both problems are from drastically different domains (e.g., military strategy vs. surgery). They defined this process as 'analogical transfer,' and explored under which circumstances analogical transfer could be fostered. Most simply, providing guidance to use previous knowledge as an analogy generally leads problem solvers to transfer source analogues toward solving a novel problem. However, there remained unanswered questions about how *spontaneous* analogical transfer could be fostered. Gick and Holyoak noted several times throughout their experiments

when participants spontaneously transferred their source analogy to the target problem, but only in experiment 5 did they have a compelling case for spontaneous transfer being fostered by initial source analogue presentation. Generally, most of their experiments explicitly stated to use a source story to solve the target problem, eliminating the chance for a problem solver to transfer the source story spontaneously. Further, in their closing remarks, they acknowledge the spatial nature of their source and target problems, and suggest that spatial representations (e.g., diagrams or illustrations) may be able to be used as analogies. They also note that a broad cognitive mapping process may play a role in noticing similarities across different domains, which could lead to improved performance on analogical problem solving. They further suggested that, if someone was exposed to a large number of analogous stories, they might develop a "convergence schema" that improves the likelihood of spontaneous analogical transfer. Their next major study in this research area explored ways to induce a convergence schema (by narrative as well as spatial means) that could increase the likelihood of spontaneous analogical transfer.

Exploring How to Induce the Convergence Schema

Gick and Holyoak (1983) produced follow-up work that expanded upon questions they had posed in their 1980 paper. In six experiments, they explored how best to induce a "convergence schema" to improve solution rates on Duncker's (1945) radiation problem, based on the aforementioned conjecture that forming a general schema should foster spontaneous analogical transfer. In this work, Gick and Holyoak (1983) varied how many source stories problem solvers received or the way they interacted with the source stories (i.e., summarizing, assessing underlying principles, or viewing diagrams of the source stories). Once problem

solvers were exposed to the source material, they were assessed on their performance of the radiation problem. In general, they were not able to support that diagrams were helpful for forming the convergence schema (c.f. Beveridge & Parkins, 1987), but that being exposed to two relevant source stories (instead of only one) was particularly helpful for performance on the radiation problem.

It should be noted that Gick and Holyoak (1983) acknowledged that “schema” is a vague term with a multitude of applications in cognitive psychology. For their purposes, they use “schema” to mean an abstract structural knowledge representation of some concept (e.g., the various convergence problems). In the frame of their analogical problem solving paradigms, they consider that a convergence schema may begin form during the encoding phase of the source story. Considering the general problem, one encodes the initial state (need to overthrow dictator), goal state (capture the dictator’s fortress), operators (army), and constraints (weight-sensitive land mines on all roads, but full force required to capture fortress). Lastly, the solution is presented (splitting army along multiple roads leading to fortress). Once these aspects are encoded, Gick and Holyoak argued that, if a problem solver were to abstract these features of the general problem, they would be able to form a “convergence schema” that could be applied to other analogous problems. Such a schema would consist of the same elements mentioned above in the general problem, but in an abstract form. For example, they argued that the initial state could be abstracted to “need to overcome a central target,” and the goal to “use force to overcome a central target.” The operators would be the “force” to be used, and that a large amount of that force is required to reach the goal state. The constraint would be that one cannot apply the force needed all along a single path, but instead, the solution would be to use multiple paths to converge on the central target with weaker forces.

In their experiments, Gick and Holyoak (1983) tested different circumstances hypothesized to foster the abstraction of this convergence schema, arguing that “mapping” (a la theories of analogical reasoning, e.g., Sternberg, 1977) the convergence schema onto a novel target problem (which could be solved with a convergence solution) would lead to success. Further, they argued that the schema itself could be a means by which analogical transfer could occur. If the schema is activated when attempting to solve the radiation problem, it may subsequently trigger the retrieval of other instances of that schema (e.g., source analogies). More simply, they also suggested that the schema itself may be sufficient to solve the problem. That is, instead of serving as a cue to retrieve concrete exemplars of that schema, the abstract qualities of the schema itself may be sufficient to discover the solution to the target problem. However, it is worth considering that many factors may affect an individual’s ability to abstract a schema, which may or may not also be affected by the quality of initial problem encoding. Similarly, when faced with a target problem, there may be individual difference factors that impact one’s ability to recognize that prior knowledge or schematic knowledge may be useful in the first place. Although individual differences likely play a role in analogical problem solving performance, they are beyond the scope of Gick and Holyoak (1983) but are discussed in later sections. In this work, Gick and Holyoak focused on methodological (i.e., stimulus presentation/interaction) means by which analogical transfer could be fostered. In this paper, the general methodology was mostly similar to Gick and Holyoak (1980), except that the initial source analogy presentation was disguised as a recall or story comprehension task. The second phase was deemed the problem solving phase, where they were faced with a target problem. After that initial attempt, they had a final attempt, with an explicit hint to try using the story from the recall task.

In the first experiment, Gick and Holyoak (1983) noted that there may be some differences in whether an abstract schema is formed based on whether one is tasked to recall a source story verbatim or summarize it at a higher level. The assumption was that a higher-level summary would be more likely to contain abstract elements that might lend themselves toward later transfer, whereas a verbatim summary would necessitate a more concrete representation that may prevent transfer to a novel problem. Thus, they tested these two conditions to assess analogical transfer on a different set of problems related to Maier's (1930; 1931) cord problem. In this problem, a problem solver must tie two cords together, but they are too far apart that they cannot be reached with both hands. There are several items in the room that may be used to help tie the two cords together: "poles, clamps, pliers, extension cords, tables, and chairs." There are a variety of solutions to the problem, but one solution was of particular interest for the present study: tying an object to one of the strings and throwing it toward the other string, like a pendulum. They note that this pendulum solution has a higher base success rate (39%) than they had found in their radiation problem and suggested that this problem might be more suited to testing the difference between verbatim recall and summary on analogical transfer (Gick, 1981 could not support that there was a difference for transfer on the radiation problem). They created an analogous "Birthday Party" story with two distant ribbons, where the solution was to tie a pair of scissors to one of the ribbons to use it as a pendulum to reach the other ribbon. They noted that this story had additional details to it to match the cord problem, and argued that, if verbatim recall hindered one's ability to abstract the pendulum solution, recalling these details would hinder performance on the cord problem. To assess how instructions dictating how to "encode" the source analogy, they compared three conditions: one where they had to recall the birthday party story verbatim; one where they had to summarize the birthday party story; and one where

they had to summarize an unrelated story (control). Results supported that summarization actually did not foster abstraction. Both the summary group and control group were relatively equal in performance (20-25% success rate), whereas the verbatim recall group had an initial 60% success rate. Providing a hint offered a further improvement to solution rates for the recall and summary groups (but not the control), but these improvements were statistically equivalent. They tentatively concluded that there was no difference for recall compared to summarizing but posed that ceiling effects may have been at play, given the close semantic relatedness to the two problems (ribbons and cords are similar to one another). Thus, in experiment 2, they returned to their traditional paradigm, but more directly tested a related assumption to that which was tested in experiment 1.

In experiment 2, Gick and Holyoak (1983) explained that understanding an abstract underlying principle to the convergence problems may be all that is necessary to foster analogical transfer. They noted that this should have emerged in the summary group of experiment 1, but by allowing participants to volunteer their own summaries, they might inadvertently miss the underlying abstract principle. In experiment 2, they sought to avoid this variability by providing one group with the abstract underlying principle to the general problem and assessed their performance on Duncker's radiation problem. The underlying principle was provided as such: "If you need a large force to accomplish some purpose, but are prevented from applying such a force directly, many smaller forces applied simultaneously from different directions may work just as well." They tested three groups: one that received the general problem with the underlying principle; one that received the general problem alone; and one that received the underlying principle alone. Interestingly, all groups performed relatively equal in both phases – before and after receiving a hint. Although they did not test a control condition

that solved the radiation problem cold, they noted that even the principle only condition had about three times greater performance than their previously identified base rate of about 10%. However, they tentatively concluded that encoding a basic underlying principle did not aid in analogical transfer.

In experiment 3, they posited that the abstract principle may be better conveyed in a spatial form (due to the spatial nature of the problems), as opposed to the narrative format used in experiment 2. So, Gick and Holyoak instead provided a visual diagram illustrated in two parts (see Figure 6).

Figure 6. Visual diagrams used in Gick and Holyoak (1983; p. 18).

The first part had one large arrow pointing in one direction, and the second had six smaller arrows, all pointing toward a central location. This diagram was intended to illustrate the principle that one large force could be split into smaller, converging forces. Similar to experiment 2, they tested three groups: one that received the general problem with the diagram, one that received the general problem only, and one that received the diagram only. Both groups with diagrams were expected to redraw the diagram from memory. The results were comparable to experiment 2, where groups were ultimately equivalent, with the exception of the diagram-only group. Their performance was significantly worse than the other groups at the before-hint stage (statistically equivalent to the aforementioned base-rate performance) but approached equality to the other groups in the after-hint stage. They concluded that an initial exposure to the

diagrams was insufficient to abstract a convergence schema from them, but when a hint was given, the drastic increase in performance for the diagram-only group (up to 67% successful) indicates that it is still possible to map the non-semantic properties of the diagram to the semantic properties of the target problem. So, diagrams may be useful as analogies for solving target problems, but only when prompted.

Gick and Holyoak's (1983) first three experiments largely provided null results for summaries, underlying principles, and diagrams' contribution toward solving the radiation problem. However, they noted that only providing a single analogue minimizes the likelihood for the development of a convergence schema. Such a schema would have to be developed on the fly while solving the target problem. Instead, in the next three experiments, they provided participants with two source analogies, based on the assumption that both sources could be mapped to one another, and an abstract convergence schema could be deduced from both of them.

In experiment 4, they tested performance on the radiation problem using a variety of analogies. They discussed that mapping two source analogies that were of different domains would be more likely to result in the formation of a more abstract schema, which could be more readily mapped to an analogous target problem. They utilized the general problem, as before, but created another military analogy (the commander problem) which involved a military capture of a fortress on an island, using multiple bridges. To utilize a different domain, they generated two additional stories related to firefighting: "Red Adair," where a firefighter had to suffocate a fire on an oil well with hoses from multiple directions to prevent it from spreading; and "The Fire Chief," where the same goal was accomplished using a team of firefighters with buckets. In this experiment, they tested three different groups: One received two similar analogues (either

two military or two firefighting); one received two dissimilar analogues (one from military and one from firefighting); and one received a random military or firefighting analogue, and a control story that was unrelated to all problems. In general, the results indicated that two useful analogues were better than one (solution rates were 45% vs. 20%, respectively; similar and dissimilar groups were statistically equivalent. After a hint, these numbers were 80% vs. 53%).

In this experiment, Gick and Holyoak also had participants describe similarities between the two source problems, in efforts to foster convergence schema development. The schemas were categorized into “good” (containing a description of forces converging from multiple directions, or multiple smaller forces and a centrally located target), “intermediate” (containing only one of the aforementioned descriptions), and “poor” (which contained none of the aforementioned descriptions). Results indicated that schema quality was a primary determinant of whether the problem would be solved successfully, and whether a hint was required. Of 51 participants, 11 were classified as having “good” schemas, 10 as “intermediate,” and 30 as “poor.” Prior to receiving a hint 10 (91%) of those with “good” schemas solved the problem, 4 (40%) of those with “intermediate” schemas solved the problem, and 9 (30%) of those with “poor” schemas solved the problem. These numbers increased after a hint was provided, and ultimately, only 1 participant with an “intermediate” schema, and 9 participants with “poor” schemas never solved the problem. These results appear to support that a higher quality schema is key to spontaneous analogical transfer, but Gick and Holyoak (1983) note that there may be some underlying individual difference at play. It is possible that an outside factor influences both schema quality and problem solving ability. In experiment 5, they attempted to mitigate the effects of poor schema quality construction (and therefore, mitigate the effects of a supposed

underlying individual difference in schema quality construction) by providing an underlying principle, as they did in experiment 2.

Experiment 5 sought to more firmly support the causal link between good schema construction and analogical transfer, as fostered by the presentation of two source analogies. In this experiment, they assessed solution rates before and after receiving a hint, and they also assessed solution rates by schema quality. They tested two groups – both receiving two analogies – but one group read an additional statement about the underlying principle of both of their analogies. Results indicated that reading the underlying principle significantly aided in success on the radiation problem, particularly before receiving a hint (62% were successful, compared to 40% who did not receive the principle). Further, those receiving the underlying principle generated significantly more "good" schemas (44% vs. 10%), and significantly fewer "poor" schemas (23% vs. 60%). Similarly, higher quality schemas led to a greater likelihood of solving the radiation problem, particularly for the condition that received the underlying principle. Thus, providing an underlying principle for the two analogues seems to improve schema quality that results in better performance on a target problem. Although there may still be individual differences at play that could explain additional variance in these results, these results lend more support to the causal link between schema quality and spontaneous analogical transfer.

In their final experiment, Gick and Holyoak (1983) sought to reevaluate the use of diagrams as a schematic aid. This experiment was designed similarly to experiment 5, where participants received two analogues either with or without the diagram used in experiment 3. Further, they varied source analogy similarity as in experiment 4, to test whether analogy similarity played any role in schema quality. As before, they mentioned that abstracting general

principles from dissimilar instances usually leads to a more robust schema that can be more readily applied to target problems, and they sought to test this assumption again in experiment 6. In general, analogy similarity did not seem to affect whether the radiation problem was solved successfully before or after a hint but providing diagrams did result in greater solution rates both before and after hints were provided. Interestingly, providing diagrams resulted in overall greater schema quality as compared to when diagrams were absent, nearly doubling the percentage of those producing “good” schemas, and halving the percentage of those producing “poor” schemas. Further, they noted that, in the absence of diagrams, analogy dissimilarity did not affect schema quality, but when diagrams were provided, those receiving dissimilar analogues produced a small but significantly greater quantity of “good” schemas (65% vs. 44%), but this did not necessarily translate to differences in performance on the target problem (“good” schemas were integral to solving the radiation problem, regardless of whether they were fostered with similar or dissimilar analogues). They note that, under optimal conditions, dissimilar analogues with an underlying principle provided should lead to the formation of a high-quality convergence schema, but providing dissimilar analogues also engenders the risk that a problem solver will not notice the correspondences between the two analogues. Such problem solvers would be more likely to form a poor schema, which would reduce their likelihood of solving the target problem. They attributed this factor to the lack of differences in solution rates between the similar and dissimilar analogues conditions.

Overall, Gick and Holyoak (1983) suggested that underlying principles and diagrams alone could not be useful analogues, rather, they should be used to help frame the relationships between two source analogies to foster high-quality schema development. By fostering high-quality schema development, the likelihood of solving an analogous problem in a different

domain increases drastically. At a more basic level, their results highlight the importance of having a good quality mental representation (or schema) for solving target problems through analogical transfer. Features that guide schema construction (e.g., diagrams; underlying principles) seem to aid in the mapping process between two source analogies, but these features do not seem to aid in spontaneous mapping from source to target without assistance. The results from Gick and Holyoak (1983) offer empirical evidence to support the link between schema quality and analogical transfer, and also support that a proper mental representation is necessary for successful problem solving. Gick and Holyoak (1983) suggested that a more abstract representation would be key to problem solving success, and this sentiment was echoed by Chi et al. (1981), who identified that expert problem solvers tended to represent their knowledge about problems in a more abstract form (compared to novices). In essence, it does not seem to matter how much source knowledge is available, but rather, how that source knowledge is represented.

Extensions of the Early Work in Analogical Problem Solving

Gick and Holyoak's (1980; 1983) two papers are perhaps the most influential works in the analogical problem solving literature (as indicated by Google Scholar citations at the time of this writing), but several studies have followed that expanded the methodology to examine other manipulations that can increase the likelihood of solving analogous target problems. Gick and Holyoak essentially established the analogical problem solving paradigm that these future studies employed, but further work has countered some of their results (e.g., Beveridge & Parkins, 1987; Pedone, Hummel, & Holyoak, 2001), and tested other avenues for increasing solution rates that Gick and Holyoak had not considered (Catrambone, Craig, & Nersessian, 2006; Cooperrider & Goldin-Meadow, 2014; Hostetter, Wieth, Foster, Moreno, & Washington, 2016; Trowbridge,

2016). Other important work has examined relevant individual differences in analogical problem solving (Antonietti & Gioletta, 1995; Corkill & Fager, 1995; Kubricht, Lu, & Holyoak, 2017), although there are some inconsistencies across this small set of studies. Furthermore, the Grant and Spivey (2003) and Thomas and Lleras (2007; 2009) studies could also be considered a tangentially related extension of Gick and Holyoak's early work if one considers guided eye movements to be an analogy. In the following section, I describe the aforementioned studies and their contributions to the literature in analogical problem solving.

Revisiting Diagrams and Exploring Animations as Analogues

Gick (1985) conducted a brief experiment intended to build on Gick and Holyoak's (1983) results and tested whether diagrams could be used as a retrieval cue for spontaneous analogical transfer. In this study, she presented participants with the same general problem and diagrams used in Gick and Holyoak (1983; see Figure 6 above), but included a cue referring to the diagram in some conditions. The diagrams were initially presented in two parts. Part A depicted a large arrow facing in a single direction, intended to represent the problem constraint (cannot use full force along a single path). Part B depicted multiple converging smaller arrows, intended to represent the convergence solution. Participants were cued while attempting to solve their target problem with diagram part A and the following written cue above: *"If the rays reach the tumor all at once at a sufficiently high intensity, as illustrated schematically below, the tumor will be destroyed"* (Gick, 1985 p. 462). Gick compared four conditions in a 2 (diagram only/diagram + story) by 2 (cued/uncued) design to assess whether the cues aided in analogical transfer, and whether they were useful in absence of an analogical story.

Participants who received the story read over the general problem, and beneath it were the diagrams used in Gick and Holyoak (1983). The solution text for the general problem was associated with the diagrams in this study, with part A described as the general's initial plan, and part B his solution. Those who did not receive the story only saw the diagrams. Before viewing the material, participants were instructed there would be a recall task following it, including redrawing the diagrams from memory and recalling the general's solution (if they received the general problem). After the recall task was complete, participants were faced with the radiation problem. Those in the cued conditions received part A of the diagram beneath the problem text alongside the previously quoted cue; otherwise, no cue was displayed. Results indicated that the cue helped drastically, roughly doubling success rates compared to those who did not receive cues (19/37; 51% successful cued participants vs. 10/44; 23% successful uncued participants). Interestingly, the differences were minimal for the main effect of diagrams vs. diagrams with the story, and there were no disproportionate benefits for any combination of cue and story presentation.

After providing answers to the radiation problem, participants were then asked a few questions about whether or not they considered using the prior information (i.e., story+diagram, or diagram only) when trying to solve the radiation problem. Afterward, they were instructed to use the prior information they had seen before to try answering the radiation problem again, even if they have to repeat their answer. After a hint was provided, the number of people volunteering the convergence solution for the radiation problem was statistically equal across conditions (approximately 80% of participants in every group were successful). These results indicated that even though performance was poor without a cue, providing a hint allowed people to transfer even the simplest forms of source knowledge (i.e., diagrams only) to be predominantly

successful at providing the convergence solution. In essence, it is possible to map visual features of a diagram to help solve transfer problems, but a strong cue is necessary. Gick (1985) notes that many educational techniques rely on the use of cues to refer to prior knowledge, but from her results, she argues that they may not be helpful for all learners. Instead, it appears that references to use prior knowledge must be explicit to foster the transfer of source knowledge to novel problems.

Critiques of Gick and Holyoak (1983)

In one of Gick and Holyoak's (1983) experiments, they were unable to find a beneficial effect of a diagram (presented alone) that represented the convergence solution. The diagram they used consisted of simple, solid-colored arrows converging on a central point. This diagram was spatially congruent with the convergence solution, but they could not support that the diagram was helpful and concluded that they could not be used as an analogy. However, Beveridge and Parkins (1987) criticized Gick and Holyoak's (1983) diagrams for being too simple and ambiguous. Instead, when using more detailed diagrams, Beveridge and Parkins found a beneficial effect. They included an intensity component in their analogous diagrams, which was missing from Gick and Holyoak's (1983) diagrams. They argued that, by failing to illustrate the multiple lowered and summating ray intensities, Gick and Holyoak's (1983) diagrams were failing to impart the critical element of intensity for the convergence solution. Beveridge and Parkins based their diagram features on the attributes that were included in Gick and Holyoak's (1983) classification of a "good" convergence schema and illustrated intensity in two different diagrams (see Figure 7).

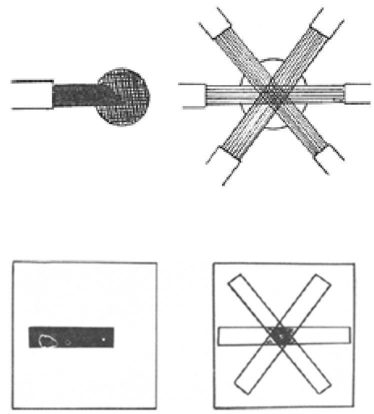

Figure 7. Two new sets of diagrams used in Beveridge & Parkins (1987).
The top set of diagrams depicts summative intensity with lasers, the bottom set of diagrams depicts the fanned translucent colored strips. Image source: Beveridge & Parkins (1987) pp. 231-232

One diagram had multiple rays shooting from multiple outside locations, converging on a central point, with lightly shaded rays that became very dark where their beams crossed at a central point. Another diagram was provided with blue colored semi-transparent strips. When closed together, the strips formed a dark blue rectangular shape, but when fanned apart, the blue strips appeared lighter on the outside, and very dark at their central crossing point. Beveridge and Parkins (1987) compared five groups' performance on the radiation problem after receiving one of their two summative intensity diagrams (rays and strips), Gick and Holyoak's (1983) converging arrows diagram, a narrative analogy, or a control condition (receiving no diagram or analogy). Their results indicated that radiation problem solution rates were roughly equal for the control group and those receiving Gick and Holyoak's (1983) simpler diagrams (approximately 60%), but there were significant improvements in solution rates for those who received the narrative analogy (75% successful), the rays diagram (80% successful), and the colored strips diagram (95% successful). Thus, although Gick and Holyoak (1983) concluded that diagrams

were not useful analogues for the radiation problem, Beveridge and Parkins (1987) demonstrated that diagrams could serve as useful analogues. Further, illustrating intensity in diagrams appears to be critical for inducing the convergence schema. Beveridge and Parkins compared their diagrams to Gick and Holyoak's (1983) diagrams which lacked an intensity component and found that the newer diagrams resulted in substantially higher success rates.

In effect, Beveridge and Parkins (1987) extended Gick and Holyoak's (1983) work to show that diagrams by themselves could be successfully used as analogies for solving a target problem. Although Beveridge and Parkins are quite critical of Gick and Holyoak's argument that visual diagrams alone are not useful, it is still worth recognizing that Gick and Holyoak (1983) supported that diagrams could be helpful when mapping relations between two source analogues and forming a mental representation of the two problems. Furthermore, Gick (1985; which Beveridge and Parkins did not cite), indicated that, in order to effectively use diagrams as analogies, fairly strong cues are required. It is possible that this recommendation may not apply if the more elaborate diagrams from Beveridge and Parkins (1987) were used. Gick's (1985) cue could have served the same functional purpose of a hint, as it was intended to cue problem solvers remember the diagram from earlier, which should have triggered recall for part B of the diagram. If recalling part B of the diagram was successful, it should have facilitated mapping of the diagram to the target problem, and the results support this.

Shortly after Beveridge and Parkins offered their critique, Gick (1989) published a book chapter reviewing much of the work with analogical problem solving and diagrams. In it, she acknowledges the potential usefulness of good diagrams like the ones in Beveridge and Parkins (1987) and describes that they have two major functions for analogical problem solving. One is that diagrams can be used as "schema-highlighting devices," which facilitate the schema

acquisition process (demonstrated in Gick & Holyoak, 1983). Another function is that they can serve as a recall cue for earlier encoded information, which she demonstrated in Gick (1985). Gick (1989) indicated that Beveridge and Parkins (1987) did not test their diagrams with stories as source analogues and contended that the comparatively more abstract diagrams from Gick and Holyoak (1983) would be better at facilitating convergence schema acquisition.

Several years later, Pedone, Hummel, and Holyoak (2001) expanded on the earlier work with diagrams, testing whether embedded perceptual properties within diagrams could facilitate spontaneous analogical transfer, and elaborating how diagrams are understood in analogical problem solving. With this approach, the diagrams themselves are considered to be source analogues, although they discussed that they schematically represent convergence, and could also be considered to be schemas. Nevertheless, they argued that the mapping process of a source or a schema toward a target problem is functionally the same (citing their earlier work, Hummel & Holyoak, 1997). That said, they mentioned that it is difficult to spontaneously notice the relevance of a diagram and map it to a target problem. One of the reasons they conjectured is that the diagrams are not interpreted with any semantic meaning. Therefore, when solving target problems in the semantic domain, there is no semantic meaning that can be easily mapped onto the target problem. However, when hinted to use a diagram to solve an analogous target problem, they argued that the mapping process happens “backward,” such that the features of the target problem are mapped onto the diagram, and in that way, the diagram is used to transfer the solution to the target problem. As with previous studies, Pedone et al. (2001) were interested in ways to improve spontaneous analogical transfer (i.e., without a hint) from diagrams by incorporating perceptual features (see Figure 8). They investigated how this could be

accomplished with a series of four experiments, adding features to the diagrams used in Gick and Holyoak (1983).

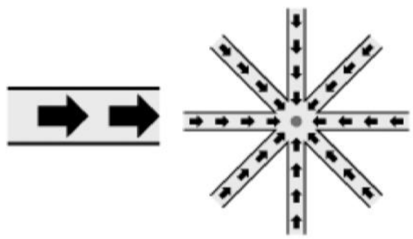

Figure 8. Diagrams used in Pedone, Hummel, & Holyoak (2001), incorporating perceptual features (arrows).
In several experiments, these arrows varied in which direction they pointed, and whether they were static or animated. Image source: Pedone et al. (2001), p. 215.

Pedone and colleagues' (2001) first experiment compared enhanced versions of Gick and Holyoak's (1983) diagrams, depicting multiple smaller arrows along the pathways. They used the original single, large pathway, and a final convergence diagram consisting of 8 converging pathways (instead of Gick & Holyoak's 6 pathways). They also included two intermediate steps – one containing two converging pathways, and one containing four converging pathways. The groups receiving the Gick and Holyoak (1983) diagrams and those receiving the intermediate diagrams did not fare any better than a control group in providing the convergence solution. After receiving a hint, both diagram groups had about 50% more participants volunteering the convergence solution, suggesting it possible to map diagrams to the target problem, but not providing evidence for a benefit of intermediate diagrams.

Their next experiment tested convergence and divergence depicted in the diagrams. Divergence was depicted with arrows pointing along the pathways, except pointing away from a central point. In addition to testing convergence against divergence, their participants were also assigned to either an animated or a static condition. The animated condition saw the arrows

along the pathways in their appropriate direction (either toward the center or away from the center). Pedone and colleagues suggested that animations may be particularly helpful for facilitating spontaneous transfer, and their results strongly suggested that this was the case. All conditions that received diagrams saw all intermediate steps of the diagrams (as described in the first experiment). The group who received the animated converging diagrams had a before-hint success rate of 55%, compared to all other groups, whose success rates were only between 5% and 15%. After receiving a hint, those receiving the converging diagrams improved their success rates substantially (55% additional problem solvers for the static group, and 35% for the dynamic group), but those in the divergence groups barely improved (between 5% and 10% additional participants were successful). Those in the control group received no hints and had to solve the problem immediately; only 2 of 19 participants in that group were successful (consistent with a solution rate of around 10% in previous studies). These results suggest that animations are a significant boon to spontaneous analogical transfer, and further, that misleading diagrams and animations do not appear to offer any improvements to solution rates. Additionally, providing a hint typically facilitates mapping source analogues to target problems, but this was only the case for the diagrams depicting convergence. Those receiving diagrams depicting divergence did not exhibit in a substantial increase in convergence solution rates after a hint was provided.

Pedone and colleagues' (2001) third experiment sought to examine whether a multimedia learning effect (a la Mayer & Gallini, 1990) might aid in spontaneous analogical transfer. They argued that providing a principle alongside the diagrams would help problem solvers abstract the convergence solution. Gick and Holyoak (1983) provided evidence that abstracting the convergence solution does improve spontaneous analogical transfer, at least when two analogies are provided, but Pedone and colleagues sought to test whether diagrams with principles aid in

abstraction. Additionally, they speculated that providing a principle would enhance spontaneous transfer for those receiving the static diagrams. The principle they provided to participants was the same as in Gick and Holyoak (1983): "If you need a large force to accomplish some purpose, but are prevented from applying such force directly, many smaller forces applied simultaneously from different directions may work just as well." To test these research questions, they provided this principle to one group receiving static diagrams against another group receiving animated diagrams. When comparing convergence solution rates, both groups were statistically equivalent for spontaneous transfer (45% vs. 50% successful), but the animated diagram condition was more successful after a hint was provided (25% vs. 45% additional successful participants). These results indicated that providing an abstract principle alongside the diagrams aided in spontaneous analogical transfer, presumably because it aids in convergence schema formation. Although spontaneous analogical transfer was statistically equivalent between groups, the animated condition showed greater additional success after providing a hint, indicating that animated diagrams are still superior to static diagrams overall.

In their final experiment, Pedone and colleagues were interested in testing diagram animations more purely. Whereas in the prior experiments, each of the diagrams contained arrows pointing in a particular direction, the final experiment eliminated this perceptual cue. Instead of arrows, they provided small blocks that were located along the converging pathways in the diagrams. By removing the perceptual affordances of arrows in the diagrams, they argued that they could test the benefits of animation more directly. They compared two groups receiving the diagrams with blocks instead of arrows, where one group received static diagrams, and the other group received animated diagrams. Results suggested a substantial benefit for animation; 30% of participants who received animated diagrams volunteered the convergence

solution compared to only 5% of participants who received the static diagrams. After a hint was provided, the increases were roughly equivalent (40% additional successful participants for the animated group; 30% additional for the static group). However, combining these numbers yields a stark difference in total success rates – 35% of participants in the static group ultimately volunteered the convergence solution, compared 70% of participants in the animated group. Based on these results, Pedone and colleagues argued that animations facilitated transfer more comprehensively than static diagrams.

Pedone and colleagues' (2001) experiments provide strong evidence that introducing perceptual features into diagrams improves their usefulness as an analogy for spontaneous transfer to target problems. Animations that depict the convergence solution seem to function as an implicit aid to spontaneous transfer. They noted that the perceptual qualities they tested in their experiments may explain why the more detailed diagrams in Beveridge and Parkins's (1987) studies were exceptionally successful in fostering spontaneous transfer. They also noted that, compared to previous research, the spontaneous transfer rates for their animated diagrams were greater than the spontaneous transfer rates in previous studies using distantly related analogues (e.g., Gick & Holyoak, 1983), indicating that the animations are more robust at facilitating spontaneous transfer than narrative analogues. It is worth noting that the narrative analogues have a greater quantity of features that could be mapped onto the target problem, but perhaps the perceptual information inherent in animations is simply a more salient means of fostering the convergence schema? Pedone and colleagues (2001) suggest that the diagrams may enable a problem solver to represent the perceptual information that is needed to solve the radiation problem. They further clarify that successfully solving the radiation problem (and other convergence schema problems) requires a physical and perceptual understanding of how

the converging forces interact with other objects. Thus, animations that foster these perceptual representations may prime problem solvers to represent the perceptual aspects of the problems they later try to solve.

Although Pedone and colleagues do not cite any embodied cognitive theoretical work, it is possible to consider their closing arguments as relating to theories of embodied cognition. In essence, the perceptual cues in the animated diagrams may serve to foster perceptual simulations (e.g., Barsalou, 1999; 2008; Wilson, 2002) of the convergence schema, particularly when a verbal principle is included. This is the main argument for Grant and Spivey (2003) and Thomas and Lleras's (2007; 2009) eye movements results, although it is worth noting that Pedone and colleagues did not introduce their participants to the problem before showing them the diagram, so the solution rates across these studies cannot necessarily be directly compared. It should also be noted that the nature of the diagrams in these studies are rather different – Grant and Spivey (2003) and Thomas and Lleras (2007; 2009) showed an abstract diagram of the tumor and surrounding healthy tissue and had participants move their eyes in a converging pattern. Although their arguments are similar, it is possible that the factors that led to increased solution rates (i.e., animated diagrams, Pedone et al., 2001; spatially congruent eye movements, Grant & Spivey, 2003; Thomas & Lleras, 2007, 2009) were due to a common underlying embodied mechanism. In that same vein, animated diagrams may inherently afford eye movements that facilitate perceptual representations of the convergence solution. Other studies have investigated embodied effects in analogical problem solving more directly, albeit with mixed degrees of success.

Embodied Cognition in Analogical Problem Solving

After Barsalou's (1999) theory of perceptual symbol systems was introduced, researchers became interested in revisiting the radiation problem to address embodied questions about problem solving. Similar to Grant and Spivey (2003) and Thomas and Lleras's (2007; 2009) studies that argued for an embodied cognitive effect in problem solving, Catrambone, Craig, and Nerssessian (2006), Craig, Nerssessian, and Catrambone (2002), and Pedone and colleagues (2001) supported that encoding perceptual information could also aid analogical problem solving. The theoretical explanations of Catrambone et al. (2006) are discussed more thoroughly in Craig et al. (2002), using the same data published in their 2006 study, alongside as other studies that investigated ancillary research questions. They discuss some of the preliminary theoretical rationale that led them to the results summarized in their 2006 study. They describe that when reading about problems, a problem solver is likely to spontaneously activate various schemas while they encode the problem. For example, reading the phrase "hammering a nail into the fence" might elicit a schema for linear force. Along these lines, they sought to test whether differences in the wording of narrative analogies and their solutions may differentially elicit converging forces schemas (e.g., "pinching" or "squeezing" schemas) during encoding. In this initial study, they provided one group of participants Gick and Holyoak's (1980; 1983) "General Problem" as the source analogy (an analogous problem used in much of the analogical problem solving work following Gick & Holyoak, 1980), and compared them to a group of participants who received an analogous problem about beehives. In the "General Problem," a military general must overtake a fortress by sending his army in small groups on multiple roads so as to avoid setting off weight-sensitive landmines that would explode if the full force was sent

down a single road. In this beehive problem, one must destroy the queen at the center of the hive without disturbing too many bees in the process. The solution was to use multiple sticks and press them all toward the center of the beehive, so the queen would be killed without disturbing too many other bees (e.g., if one large stick was used). Craig et al. (2002) argued that the wording of this narrative analogy of the beehive would be more likely to elicit a converging forces schema (as opposed to the general problem, which had less concrete wording), and this argument was supported by their results. Approximately twice as many people who received the beehive problem (20/49) spontaneously produced the solution to the radiation problem, as compared to those who received the general problem (9/44). These results supported that perceptual information encoded from narrative source analogies may be helpful when solving novel analogous problems.

Catrambone et al. (2006) reiterated that success in analogical problem solving is dependent on successfully forming a mental representation of the problem's structure, which could then be mapped onto analogical problems to solve them. However, they suggested that these mental representations could be enhanced with the encoding of perceptual information, with the assumption that perceptually representing the problem structure would improve solution rates for analogical problems. Specifically, they suggest that perceptually encoding the problem may activate a converging force schema (e.g., "squeezing" or "pinching"), which should affect how the problem is represented, and may further affect how that problem representation is applied in later problem solving. In another experiment, they tested this using Gick and Holyoak's (1980; 1983) "General Problem" as the source analogy (an analogous problem used in much of the analogical problem solving work following Gick & Holyoak, 1980), and Duncker's (1945) radiation problem as the target problem. They compared two traditional kinds of analogy

(verbal analogies and sketched analogies) and compared them to an analogy that would impart perceptual information to the problem solver (manipulating blocks to enact the analogy). All participants received the general problem (as mentioned previously), and they had to re-enact the problem either by verbally explaining it, by verbally explaining it while sketching a diagram of it, or by verbally explaining it while moving blocks to demonstrate the solution. They argued that the perceptual information that would be encoded by moving the blocks would enhance participants' likelihood of spontaneously mapping the solution to the general problem when faced with the radiation problem. The results largely supported this argument, with roughly twice as many problem solvers successfully solving the radiation problem when the general problem was re-enacted with blocks (17/33), as opposed to the verbal (4/21) and sketch (10/33) conditions. They argued that these differences signified that those re-enacting the problem with blocks were more likely to spontaneously apply the general problem's solution to the radiation problem, and the relevant mechanism that could explain these differences was perceptual simulation. That is, the participants who perceptually represented the problem structure (i.e., the re-enactment with blocks group) were more successful in transferring that problem representation to an analogous problem. Thus, perceptual representation during encoding of a problem appears to be useful for analogical problem solving.

Interestingly, Grant and Spivey (2003), and Thomas and Lleras (2007; 2009) did not cite Craig et al. (2002) or Catrambone et al. (2006), when their primary manipulations concerned very similar theoretical underpinnings – that perceptually representing problem features would aid in problem solving. One key distinction could be that Craig et al. (2002) and Catrambone et al. (2006) explicitly provided source analogies to be used in later analogical problem solving, but it could be argued that Grant and Spivey (2003) and Thomas and Lleras (2007; 2009) implicitly

provided a perceptual source analogy by means of guided eye movements or guided attention. If this is indeed the case, then the perceptual information encoded in a source analogy would no doubt be a significant determinant of success in later analogous problem solving. Taken together, these studies strongly suggest that embodied cognitive effects (i.e., perceptually representing/simulating a problem) largely influence analogical problem solving. However, one of the main embodied cognitive findings from Catrambone et al. (2006) could have been confounded by arousal. Solution rates increased as participants had more interaction when recalling the general problem (i.e., verbal only vs. verbal+sketch vs. verbal+blocks). Thus, it is possible that the perceptual representation effects discussed in Catrambone et al. (2006) and Craig et al. (2002) may in fact be due to arousal, and not due to an embodied cognitive effect. Incidentally, there are a few studies that have tested embodied influences of analogical problem solving and failed to demonstrate a beneficial effect on schema induction and problem solving.

Expanding from the embodied cognitive explanations posited by Catrambone et al. (2006), Craig et al. (2002), Grant and Spivey (2003), and Thomas and Lleras (2007; 2009), several researchers attempted to assess whether these embodied effects could be fostered through the use of gesture. Of relevance to the present work, Alibali, Bassok, Olseth, Syc, and Goldin-Meadow (1995), Hostetter and Alibali (2008), and Kirsh (2010) argue that gestures may function as a means of externally representing internal mental processes and serve to guide cognition. Although the previously mentioned studies argued that an embodied cognitive effect exists that improves solution rates (either via enacting the convergence solution with blocks or mimicking the shape of the convergence solution with eye movements), the following studies were unable to find such an effect when using gestures.

Cooperrider and Goldin-Meadow (2014), Hostetter, Wieth, Foster, Moreno, and Washington (2016) and Trowbridge (2016) all conducted experiments examining whether gesturing while summarizing source analogies (or comparing source analogies to target problems) aided when attempting to solve an analogous, structurally similar problem. In these studies, participants were instructed to provide the solution to the problem either by gesturing with a verbal response, by specifically withholding gestures while verbally responding, or by providing a verbal response without explicit instructions to gesture. Both Cooperrider and Goldin-Meadow (2014), and Hostetter et al. (2016) found detrimental effects – if participants gestured while summarizing or recanting the source analogy stories, they were actually less likely to produce the convergence solution to the radiation problem.

Theories of embodied cognition would suggest that incorporating gesture would enhance how the schema for the convergence solution was encoded, for example, by adding a perceptual dimension to a narrative to be encoded. The results of Catrambone et al. (2006), Craig et al. (2002), Grant and Spivey (2003) and Thomas and Lleras (2007; 2009) would suggest that participants were more successful when they embodied the convergence solution, either through physical interaction with blocks, or through simple eye movements in the spatial arrangement of the convergence solution. It is interesting that these effects did not manifest in Cooperrider and Goldin-Meadow (2014), Hostetter et al. (2016), and Trowbridge (2016), as their studies employed physical enactments of the convergence schema. If the effects found in previous literature are truly embodied cognitive effects, then the use of gesture should certainly enhance encoding of the convergence schema. However, this was not the case, and it is possible the effects Catrambone et al. (2006) and Craig et al. (2002) observed were simply due to arousal – the embodied condition interacted with physical objects, and the conditions to which their

solution success was compared only used verbal summaries or drawings to summarize the source analogy. The benefits conferred by the interaction with physical blocks may have only emerged because that condition was the most engaging of the three conditions. Additionally, it is possible that the effects that Grant and Spivey (2003) and Thomas and Lleras (2007; 2009) observed were due to attentional priming, rather than an embodied cognitive mechanism. Because Grant and Spivey and Thomas and Lleras's work is of more direct relevance to the present study, it is worth exploring why the benefits they saw may have only been caused by an attentional priming mechanism, rather than an embodied cognitive one.

Review of Eye Movement Studies with Problem Solving

The following sections will review in detail the most pertinent studies for the present work.

Grant and Spivey (2003) – Eye Movements and Problem Solving: Guiding Attention Guides Thought

Grant and Spivey (2003) conducted two experiments analyzing eye movements during problem solving, arguing that an eye movement analysis could reveal underlying aspects of visual attention that may be lost in traditional measures of problem solving success – solution time and accuracy. They presented participants with a version of Duncker's (1945) radiation problem, including a very simple diagram (see Figure 9) representing a cross-section of a stomach with a tumor represented inside. This diagram consisted of a circular outline, inside of which was a smaller, concentric, dark-filled circle. The circular outline represented the skin, the smaller dark-filled circle represented the tumor, the space between them represented the healthy tissue, and the space outside the circular outline represented space outside of the stomach.

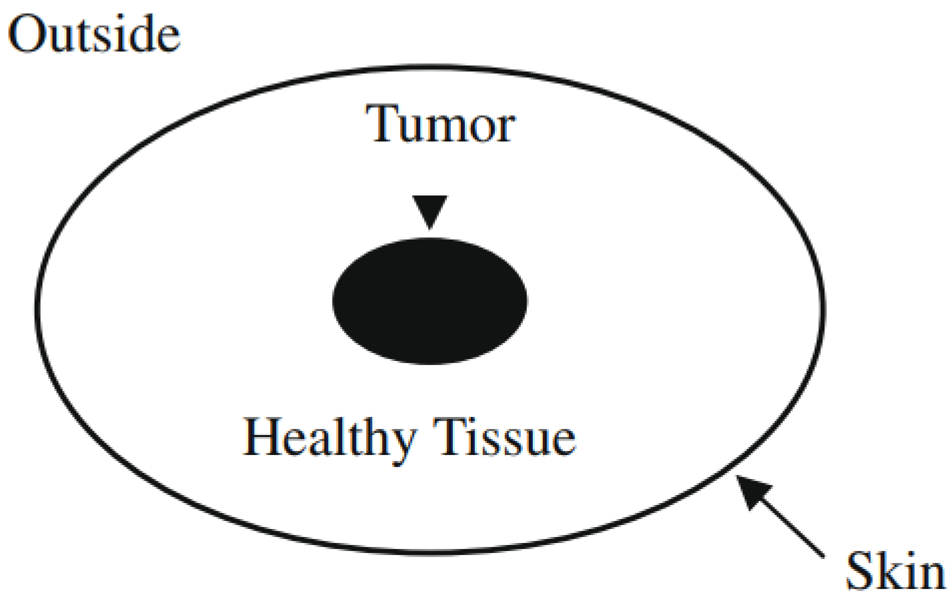

Figure 9. Simple diagram of the radiation problem presented to participants in Grant and Spivey (2003); as well as Thomas & Lleras (2007; 2009).
Image source: Thomas & Lleras (2009, p. 169)

While participants tried to solve the problem, Grant and Spivey (2003) were interested in observing their eye movements about this diagram. In the first experiment, they hypothesized that those who would be successful would spend more time looking in the outside area of the diagram, as that is where multiple lasers would have to be imagined; a key feature of the convergence solution of the problem. They recruited fourteen college student participants who had not seen Duncker's radiation problem and had normal or corrected-to-normal vision. Participants were instructed to provide verbal protocols for their solution to the problem, and to draw out their solution on the diagram. The experimenter read the problem to participants but remained silent while the participants attempted to solve the problem. If participants had not provided the convergence solution within 10 minutes, they were provided one or more hints to allow them to complete the task. Hint one was, "What if you could adjust the intensity of the

lasers?" and hint two was, "What if you had more than one laser?" Participants were only coded as being successful if they produced the solution without hints. The authors argued that non-hinted solutions represented solutions that participants implicitly embodied. Although eye movements were recorded during the entirety of this solution process, Grant and Spivey only analyzed the 30 seconds of eye movements after participants received the instructions, and the 30 seconds of eye movements immediately prior to participants providing their final solution to the problem. In these recordings, independent coders classified the eye movements as being located in one of five diagram locations: the tumor, the healthy tissue, the skin, outside the skin, or irrelevant (e.g., participant looking away from the screen, or system error).

Using these codes, the authors sought to identify differences in eye movement patterns between those who solved the problem (5 out of 14 participants) and those who did not solve the problem (9 out of 14, who required hints). During the initial 30 seconds of the problem, there were no differences between successful and unsuccessful problem solvers. However, during the 30 seconds prior to providing a solution, successful participants spent a significantly greater proportion of time looking at the skin, compared to those who were unsuccessful (there were no other significant differences). Contrary to their initial hypothesis, both groups spent a statistically similar amount of time looking at the outside of the diagram – they concluded that this must not be a critical area where the multiple lasers are imagined. Based on the results from experiment 1, they argued that the skin region of the diagram must be the critical area for participants to infer the convergence solution. In experiment 2, they tested the hypothesis that repeatedly drawing participants' attention to this critical area of the diagram would increase the likelihood that participants would provide the convergence solution to the radiation problem.

To test this hypothesis, 81 college students were randomly assigned to one of three groups: view diagram without any attentional cues; view diagram with attentional cues to skin (which they deemed a critical region); or view diagram with attentional cues to the tumor (which they deemed a non-critical region). Aside from the attentional cuing of the diagram, the method was identical to experiment 1. The animations were facilitated by causing the diagram to "pulse" three times per second. In these "pulses," the relevant diagram feature would increase in size by 1 pixel, then immediately return to the initial size.

Although there were more participants per group in experiment 2, the rates for solving the problem were nearly identical – 10 out of 27 who viewed the non-animated diagram were successful, which is only 1 percentage point greater than those who solved the problem in experiment 1. Key differences emerged in the animated conditions. Those who viewed the diagram with the pulsing tumor were roughly equally as likely to generate the solution (9 out of 27 participants) than those who viewed the non-animated diagram. However, those who viewed the skin-animated diagram were nearly twice as likely to generate the solution (18 out of 27 participants) than all other conditions; confirming their hypothesis that drawing attention to this critical area would increase their likelihood of solving the problem.

Their explanation for this phenomenon was largely that the eye movements themselves guided the problem solving process, by mirroring the eye movements of those who would normally solve the problem without any help. They also argued that it may not necessarily have been fixations on the skin that mattered, but rather, eye movements across the skin from the outer to the inner sections of the diagram. To assess this, they re-examined the eye movement data from experiment 1 to assess whether or not this was the case. Indeed, participants who were destined to solve the problem made significantly more saccades across the skin than those who

were not successful. Interestingly, when the unsuccessful participants received hints, their skin-crossing saccades significantly increased just prior to providing the convergence solution. It is worth noting that these skin-crossing saccades tended to be saccades that were triangular in fashion, such that they began outside the skin, moved across the skin toward the tumor, and then moved back out, across the skin, to a different area outside of the diagram (in the shape of a triangle).

Grant and Spivey (2003) interpreted their results within the framework of an embodied cognitive perspective. Citing seminal work on embodied cognition (Barsalou, 1999) they argued that the eye movements of participants who were ultimately successful might have fostered a perceptual simulation of the multiple lasers firing inward toward the tumor. They indicated that these eye movements were spatially congruent with the spatial nature of the solution, and that they engendered sensorimotor processes that enhanced participants' mental representation of the problem. In essence, they claimed that eye movements themselves facilitated problem solving success by introducing perceptual information that was spatially congruent to the solution.

Building upon this work, Thomas and Lleras (2007; 2009) more specifically tested the hypothesis that guided eye movements facilitated perceptual simulations of the radiation problem by explicitly guiding eye movements around a similar diagram. In the 2007 paper, Thomas and Lleras tested several different combinations of guided eye movements, and in the 2009 paper, they examined whether similar benefits could be observed with covert attention (i.e., by directing attention in and out of the diagram without moving one's eyes).

Thomas and Lleras (2007) – Moving Eyes and Moving Thought: On the Spatial Compatibility between Eye Movements and Cognition

Directly extending from Grant and Spivey's (2003) work, Thomas and Lleras (2007) tested participants' ability to solve Duncker's radiation problem while guiding their attention in various related and unrelated patterns to the convergence solution via a digit-tracking task. This digit-tracking task required participants to attend to letters and numbers appearing at various locations inside and outside of a diagram of the tumor problem (similar to the one used by Grant & Spivey; 2003). The letters and numbers appeared eight times over a period of four seconds, and in this sequence, a random letter would instead be replaced by a number. Whenever participants saw a number appear, they would have to hit a button to indicate that they had seen the number appear. After this four second sequence had passed, participants had a "free-viewing period" for 26 seconds, during which no letters or numbers would be presented. The digit-tracking portion and free-viewing portion were repeated a total of 20 times, for a duration of 10 minutes.

Prior to undergoing the digit-tracking task, participants were presented with the radiation problem, and a diagram of the skin and tumor. After verifying whether participants previously had seen the radiation problem, they began the digit tracking task in one of four randomly assigned groups with differing letter and number sequences. The "embodied-solution" group tracked letters and numbers that appeared in and out of the diagram in a triangular fashion – appearing in areas outside of the skin of the diagram, then appearing in the center of the diagram, on the tumor, and then again appearing outside of the skin in a new location. The "areas-of-interest" group tracked letters and numbers outside and inside of the diagram, except the letters and numbers appeared outside of the diagram four times, before appearing inside of the diagram

for the remaining four times. The “repeated-skin-crossing group” tracked letters and numbers repeatedly appearing outside and inside of the diagram, except without changing location (i.e., similar to the “embodied-solution” group, except without appearances at other outside locations). Finally, the “tumor-fixation” group conducted their digit tracking task by tracking letters and numbers appearing in the center of the diagram, solely on the location of the tumor.

Participants were instructed to try to come up with a solution for the radiation problem while they performed the digit-tracking task. Thomas and Lleras (2007) argued that, by analyzing differences in problem solving success among groups, they could assess the extent to which the guided eye movements (facilitated by the digit-tracking task) influenced cognition. Specifically, they argued that, if skin-crossing saccades from multiple outer locations were key to solving the problem (as Grant & Spivey suggested), then the “embodied-solution” group would show the best performance. They noted that it was also possible that participants only needed to view key areas to help infer the solution – if this was the case, then the “areas-of-interest” group and the “embodied-solution” group would show better performance than the other two groups. They also argued that, if skin-crossing saccades were the only necessary guidance to improve solution rates, then the “repeated-skin-crossing” group and the “embodied-solution” group would have better performance than the other groups. It should be noted that one of the intermediate assertions from Grant and Spivey’s (2003) studies was that the skin was the critical area of the tumor diagram, which the “repeated-skin-crossing” group would be guided to repeatedly view. Lastly, they argued that if eye movements had no influence on problem solving, that all groups would show equivalent performance.

To test these predictions, Thomas and Lleras (2007) recruited 99 undergraduate students to participate in this study. A number of participants were excluded for either failing the

tracking task (i.e., not responding when the number appeared), or indicating that they saw a relationship between the digit tracking task and the radiation problem. Participants who saw a relationship between the digit tracking task and the radiation problem were excluded so as to analyze participants who were not explicitly aware of the purpose of the experiment (only 6 of 99 participants saw the relationship, and none of those six were in the tumor-fixation group). On average, there were roughly 18 participants analyzed in each condition. As previously mentioned, participants were provided instructions for Duncker's radiation problem and instructed to attempt to solve it while undergoing the digit-tracking task. This task could last up to ten minutes, but participants were allowed to stop this task at any time that they thought they had devised a solution to the radiation problem. To provide their solution, they had to draw it on a piece of tracing paper held over the computer screen, in front of a display of the radiation problem diagram. If participants provided a correct solution after interrupting the digit-tracking task, the experiment finished; but if they were incorrect, they resumed the digit-tracking task where they left off. Thus, the experiment ended whenever participants provided a correct solution, or when participants completed the digit-tracking task and made a final attempt at solving the problem. Solutions were coded as correct if the drawings contained at least two lines drawn from the outside of the diagram, crossing the skin, and converging on the tumor. After this was complete, participants indicated whether or not they noticed a link between the digit-tracking task and the radiation problem.

Results indicated that the different eye-movement pattern groups differed significantly with regard to solution success rates. The embodied-solution group had the highest success rate (at 50%), which was significantly different from the other groups, as indicated by a chi-square test (areas-of-interest: 33%; repeated-skin-crossing: 19%; tumor-fixation: 22%). These results

supported their prediction that eliciting skin-crossing saccades from multiple outer locations would aid in participants generating the solution to the radiation problem. The guided eye movements in the embodied-solution group influenced cognition to the extent that it improved solution rates compared to the other groups. They also conducted a survivor analysis to compare the solution rates for each group over time and supported that the embodied-solution group had statistically better solution rates over time than the tumor-fixation group and the repeated-skin-crossing group (but was statistically equivalent to the areas-of-interest group).

In sum, the results largely supported the hypothesis that guided eye movements that were spatially congruent with the solution to the task provided the best aid to problem solving. Interestingly, a majority of participants indicated that they did not recognize a link between the digit-tracking task and the solution to the problem, and several of them believed it to be an overt distraction from the radiation problem. Thus, Thomas and Lleras (2007) argued that the benefits that the guided eye movements conferred toward problem solving were implicit, rather than an explicit aid to problem solving.

Relating back to Grant and Spivey’s (2003) assertion that skin-crossing saccades from multiple outer locations were the key to improving solution rates for the radiation problem, Thomas and Lleras (2007) also assessed the number of skin-crossing saccades that participants made during the tracking task and during the free-viewing period. Although by nature of the design of the experiment, the embodied-solution group and repeated-skin-crossing group had significantly greater skin-crossing saccades compared to the areas-of-interest group and the tumor-fixation groups, which had relatively low skin-crossing saccades during the tracking task. However, skin-crossing saccades during the free-viewing period were statistically equivalent across all groups, indicating that eye movements elicited during the digit-tracking task were

more likely the driver of any group differences. Thus, the particular pattern of skin-crossing saccades may have been critical for the group differences that were observed (assuming no confounding variables were to blame). Thomas and Lleras (2007) argued that guided eye movements were most effective when they *embodied* the solution to the radiation problem. Further, they asserted that skin-crossing eye movements initiated from different areas outside of the diagram bring rise to a perceptual simulation of the solution of the problem. That is, the embodied eye movement patterns led participants to mentally simulate lasers firing on the tumor from multiple points, leading them to be more likely to solve the problem. Aligning with embodied cognitive perspectives (e.g., Wilson, 2002), they argued that the eye movements and diagram allowed problem solvers to off-load their cognition onto the environment, which aided in problem solving. They further suggested that their results provide evidence that physical interactions with the environment may not only influence cognitive processes but may also generate cognitive processes. In essence, Thomas and Lleras (2007) demonstrated that low-level guided movement of the eyes in a simple environment (the tumor diagram) enhanced cognition as indicated by problem solving success on Duncker's radiation problem.

Thomas and Lleras (2009) – Covert Shifts of Attention Function as an Implicit Aid to Insight

In this follow-up study, Thomas and Lleras essentially replicated their 2007 study, but were interested in testing an additional research question: are physical movements required to influence cognition, or could covert processes generate the same effect? To test this research question, they modified their earlier methodology to include a condition similar to the "embodied-solution" condition (a digit-tracking task with letters and numbers appearing repeatedly outside and inside of the diagram from multiple points), except where participants

were instructed not to move their eyes – to stay fixated on the tumor. They were still required to respond to the digit-tracking task as before by hitting a button when a number appeared in place of a letter. By not moving their eyes but still having to respond in accordance with the digit-tracking task, Thomas and Lleras (2009) argued that they were directing participants' covert attention movements, rather than their overt eye movements. In doing so, they were able to test whether attention movements drove the embodied effect observed in Thomas and Lleras (2007), rather than the physical eye movements themselves.

They acknowledged that there are many different views that attempt to disambiguate attention movements from eye movements, citing some that have suggested attention movements to be a precursor to saccades, with others suggesting attention movements to function as a separate mechanism that does not necessarily drive saccades. They further cited other arguments that eye movements tend to move in a smooth, analog fashion, whereas attention shifts occur in a more discrete fashion, regardless of the distance between attention shift-points. A detailed discussion of these perspectives is beyond the scope of the present work, but it is pertinent to acknowledge the concomitant attentional aspects of eye movements. Recognizing the disparities between eye movements and attention shifts, Thomas and Lleras (2009) admitted that it was unclear whether the embodied eye movement effects claimed in Thomas and Lleras (2007) would emerge through shifts of attention when eye movements were restricted.

Using similar rationale from their previous study, they sought to answer their research question – whether eye movements (as opposed to only attention shifts) were necessary for embodied effects to emerge – by examining solution rates to the radiation problem among different groups. As in their previous study, they utilized the digit-tracking task with a 4 second tracking period and a 26 second free viewing period, but the groups were slightly different. The

four groups were “eye-movement” (identical to the “embodied-solution” group from Thomas & Lleras 2007), “Attention-shift” (identical to the “eye-movement” group, except with a restriction for participants to keep their eyes on the center of the diagram during the tracking task), “tumor-fixation” (all items for the digit-tracking task appeared in the center of the diagram; this group was identical to the “tumor-fixation” group from Thomas & Lleras 2007), and “no-eye-movement” (same as the “tumor-fixation” group, except participants were instructed to never move their eyes from the center of the diagram). Each group consisted of 23 participants, and no participants needed to be excluded for tracking task failures or for noticing a relationship between the tracking task and the problem’s solution.

As before, participants received instructions for Duncker’s radiation problem prior to beginning the digit-tracking task. The same solution procedure was employed (where participants were permitted to interrupt the digit-tracking task to provide a solution), and the digit-tracking task lasted 10 minutes if they did not provide a correct solution before the 10-minute mark. After participants provided their solutions, they were asked whether they noticed any relationship between the tracking task and the solution to the radiation problem (for this study, no participant noticed such a relationship).

Results indicated that both the “eye-movement” and “attention-shift” groups had superior problem solving success (39% and 30%, respectively), compared to the “tumor-fixation” and “no-eye-movement” groups (13% and 9%, respectively), supported by chi-square and survivor analyses. Thomas and Lleras noted that, because the “eye-movement” and “attention-shift” groups were not statistically different from one another, attentional mechanisms were the main driver behind the embodied effect they observed. That said, their results also indicate that shifts of attention are not completely divorced from eye movements, at least with respect to the

benefits they confer for solving spatial problems like Duncker's radiation problem. In essence, Thomas and Lleras (2009) claimed that physical action is not necessary for an embodied effect to emerge, and that something as small as shifts in attention can influence how people think.

Relating Previous Studies to the Present Work

The general explanation for the results among all three previously reviewed studies (Grant & Spivey, 2003; Thomas & Lleras 2007; 2009) is that the guided eye movements (and concomitant attentional shifts) fostered a perceptual simulation where participants imagined lasers firing on the tumor from multiple outside points. Perceptual simulations, and the interplay between body movements, the environment, and cognition, are key components of theories of embodied cognition (e.g., Barsalou, 1999; Barsalou, 2008; Wilson, 2002). The guided eye movements (and to a similar extent, the guided attention shifts) were argued to cause participants to embody the solution which improved their solution rates. When shifting attention or moving their eyes around the environment (the diagram) in a pattern that embodied the convergence solution, participants were purportedly more likely to perceptually simulate the convergence solution to the problem. It is conceivable that problem solvers run mental simulations of the problem features as they search for the solution. With regard to the Grant and Spivey (2003) and Thomas and Lleras's (2007; 2009) studies, those who received embodied guidance for their eye movements arguably had their mental simulations enhanced or guided by the concomitant perceptual information. This embodied effect resulted in the increased problem solving success.

However, there is a possible alternative explanation. It is possible that the effects observed in these previous studies were due to an attentional priming of creativity effect (e.g., Friedman, Fishbach, Foerster, & Werth, 2003). According to this effect, it is possible that the

eye movements and attention shifts that guided participants to attend to a broader area led to improved creativity (which increased solution success rates), rather than engendering a perceptual simulation of the radiation problem. Creativity is suggested to be a requirement for solving Duncker's radiation problem (e.g., Gick & Holyoak, 1980; Dow & Mayer, 2004), and for solving insight problems in general (Schooler & Melcher, 1995; Wegbreit, Suzuki, Grabowecky, Kounios, & Beeman, 2012). If Grant and Spivey (2003) and Thomas and Lleras's (2007; 2009) manipulations gave rise to a creativity priming effect, it may explain why they failed to see significant differences between the "embodied-solution" and "areas-of-interest" groups in their first study (2007), as well as the lack of significant differences between the "eye-movement" and "attention-shift" groups in their second study (2009). As they claim in Thomas and Lleras (2009), it was the guided attention shifting component that influenced cognition, not specifically the embodied eye movements themselves. This may also relate to a limitation that Thomas and Lleras (2007; 2009) mentioned: that their guided task does not specifically convey the direction of the rays. That is, the eye movements could just as equally convey inward movement toward the tumor from multiple points, or outward motion originating from the tumor and exiting at multiple different points. Thus, it is possible that the effect observed in all three of these studies was not driven by an embodied perceptual simulation effect, but rather, an attentionally-primed increase in creativity that led to increased solution rates.

There are other possible effects and relationships to consider with respect to this embodied or attentional priming effect. Duncker's radiation problem has been used in classic research on analogical problem solving (e.g., Gick & Holyoak, 1980), but only a few of the studies examining analogical problem solving have examined embodied effects, with mixed results. To date, there are no studies that have tested whether guided eye movements can aid in

analogical problem solving, and the present work aims to test this research question. Further, by testing whether Grant and Spivey (2003) and Thomas and Lleras's (2007; 2009) effects are due to fostering perceptual simulation or priming creativity with attention, the present work may be able to clarify why other studies that have examined embodied mechanisms in analogical problem solving have been largely unsuccessful in demonstrating an embodied effect (e.g., Cooperrider & Goldin-Meadow, 2014; Hostetter, Wieth, Foster, Moreno, & Washington, 2016; Trowbridge, 2016). The present work expands on Thomas and Lleras's (2007; 2009) methodology to address these missing pieces in the literature, and additionally combine their methodology with a traditional analogical problem solving paradigm to assess whether embodied effects (or attentional priming of creativity effects) confer any benefits for novel problems with analogous solutions.

Attention, Creativity, and Problem Solving

Since creativity has been associated with success in solving the radiation problem, as well as with analogical problem solving (Corkill & Fager, 1995; Dow & Mayer, 2002; Gick & Holyoak, 1980), and Grant and Spivey (2003) and Thomas and Lleras (2007; 2009) employed attentional techniques for their "embodied" conditions, it is worth considering the interrelatedness of attention and creativity. Previous work has identified a relationship between attention and creativity (e.g., Vartanian, 2009; Zabelina, Saporta, & Beeman, 2016) such that those who are more creative are more likely to possess a "broad attentional capacity," which enables them to combine two or more concepts with one another (Eysenck, 1995; as cited in Vartanian, 2009) or similarly, a "wide breadth of attention," which increases the probability that distant concepts or stimuli will be brought together (Kasof, 1997). Empirical evidence has

supported that people who are more creative tend to suffer when performing simple tasks where attentional distractors are in place due to their wider breadth of attention, whereas those who are less creative show lesser performance decrements in the face of attentional distractors due to their narrower breadth of attention (e.g., Kasof, 1997; Vartanian, Martindale, & Kwiatkowski, 2007).

Understanding the existence of a link between attention and creativity, one line of research has examined whether broad or narrow attentional focuses can be primed, leading to a difference in creative performance. Friedman, Fishbach, Foerster, and Werth (2003) conducted a series of experiments that identified that priming someone by having them focus in a relatively narrow visual field, compared to a relatively broad visual field, facilitated participants' creativity as indicated by later tasks requiring participants to come up with creative uses for bricks and generating unusual category exemplars. Those who focused on a broader visual field (as compared to a narrower visual field) were more likely to come up with a creative use for a brick, or a more unusual exemplar for a number of categories. They further identified that this effect held when participants were asked to contract their muscles to raise their eyebrows (as if looking in a broad visual field) compared to furrowing their brow (as if focusing on a small visual field) when coming up with as many creative uses for scissors as they could think of – those who held their eyebrows in a raised position while completing the task generated more creative uses for scissors than those who furrowed their brows. Friedman and colleagues' (2003) effects were supported in a traditional laboratory visual attention experiment (a visual search task with broad and narrow search areas) and a more applied visual attention task (viewing Rand McNally maps, focus either on an entire state or a specific city within that state).

Another study examined how visual attention affects problem solving strategies. Wegbreit, Suzuki, Grabowecky, Kounios, and Beeman (2012) examined problem solvers' strategies for solving Compound Remote Associate (CRA) problems before and after performing a visual attention task. These problems are unique in that they can be solved either through a search strategy, or through a more "creative" strategy: insight (Wegbreit et al., 2012). For example, a problem solver may be presented with the words "wheel" "hand" and "shopping," and they must identify a word that can go along with the words presented to form compound words or phrases. In this example, "cart" would be the remote associate that fits all three presented words, yielding the compounds "cartwheel" "handcart" and "shopping cart." Adopting a search strategy, problem solvers could test different words through trial-and-error to identify the compound remote associate that fits with all three presented words. Alternatively, a problem solver could wait for a moment of insight, where the solution word suddenly "pops-out" to the problem solver (Novick & Sherman, 2003).

Wegbreit and colleagues tested their participants with several CRA problems, and then submitted them to either a narrow or broad visual attention task. The narrow visual attention task was a simple flanker task (e.g., identify whether a T or S was in the center of a computer screen, with two distractor stimuli presented to the left and right). The broad visual attention task was a rapid identification task, where problem solvers saw a blurred greyscale image of an animal, which they had to identify quickly. This was assumed to require broad attentional focus, as opposed to the flanker task, which required narrow attentional focus. Both groups endured 80 of these trials in 4 separate blocks. Afterward, they solved a new set of CRA problems. Interestingly, self-reported strategy use during the initial set of CRA problems was statistically equivalent across groups. However, after undergoing the visual attention task, the groups

showed different patterns of strategy use, such that the group receiving the broad visual attention task reported using insight strategies significantly more frequently.

In short, these studies demonstrated that visual attention tasks can influence creative thinking and creative problem solving, with the general idea that visual attention tasks that require broader visual attention are more likely to elicit a behavioral change toward creativity. Interestingly, these studies are not cited in the analogical problem solving literature (at the time of this writing), which seems a missed opportunity of sorts. This literature is worth mentioning due to the association of analogical reasoning and creativity, but it is particularly relevant in light of the effects observed in Grant and Spivey (2003) and Thomas and Lleras (2007; 2009), whose experimental manipulations explicitly involved the use of narrow and broad visual attention. Their results may be more accurately explained by an attention-modulated creativity effect.

Lastly, it is worth mentioning how creativity is measured as a construct. Piffer (2012) claimed that creativity as a broad construct is incredibly difficult to measure, but one sub-construct of creativity (divergent thinking) is of relevance to the aforementioned studies and the present work. Divergent thinking as a construct is based on Guilford’s (1967) perspective of creativity which emphasizes the generation of multiple ideas, originality of ideas, and detail of ideas. Each experiment in Friedman and colleagues’ (2003) work employed a divergent thinking task measuring either fluency or originality. Vincent, Decker, & Mumford (2002) identified that divergent thinking was associated with creative problem solving in a military context, which may be relevant to the creative aspects of solving the radiation problem and its analogies. For the present study, participants were instructed to provide as many solutions to the radiation problem and its analogies within a time limit. For this reason, this solution-generation aspect could be

considered a divergent thinking task, and the number of solutions provided by participants was analyzed to answer hypotheses about creativity priming and problem solving.

If creativity may be a key to successful problem solving or seeing the relationships between problems (e.g., Corkill & Fager, 1995), it is possible that the priming effect Grant and Spivey (2003) and Thomas and Lleras (2007; 2009) observed was not entirely due to a facilitation of a perceptual simulation of the radiation problem. Based on Friedman and colleagues (2003) and Wegbreit and colleagues' (2012) work, it is possible that the aforementioned embodied effects were due to an attentional priming of creativity that led to an increase in divergent thinking during problem solving, rather than facilitating participants perceptually simulating the radiation problem. The present work aims to disentangle these effects.

Individual Differences in Analogical Problem Solving

Corkill & Fager (1995) quote Anderson (1990), who claimed "it requires a little sophistication to use analogy correctly…" arguing that 'sophistication' likely refers to individual differences. Including Corkill & Fager (1995), only a few studies have examined analogical problem solving from an individual differences standpoint, with evidence for differences between gender, educational discipline, and cognitive style (Antonietti & Gioletta, 1995), and various aspects of creativity and cognitive ability (Corkill & Fager, 1995; Kubricht, Lu, & Holyoak, 2017). Interestingly, two of these studies have identified conditional relationships where individual differences seem to make a bigger impact. For example, Corkill and Fager (1995) discovered that most of the effects of individual differences to predict problem solving success disappeared when participants received a hint when trying to solve the problem.

Kubricht et al. (2017) observed that individual differences in cognitive ability were only predictive in the absence of a visual animation. These effects suggest that there are ways in which disparate levels of individual differences (e.g., cognitive ability) can be equalized through instructional features or interventions. It is worth noting, that although Kubricht et al. (2017) found evidence for a relationship between cognitive ability and problem solving, Antonietti & Gioletta (1995) used the same measure (Raven's Progressive Matrices, e.g., Raven, 2000) and did not find such a relationship. The contributions of individual differences toward analogical problem solving success are still not well understood, but it is possible that more basic measures of cognitive ability are predictive of analogical problem solving success. Another aspect of cognitive ability, spatial ability, has been associated with creative thinking (e.g., Shepard, 1978), and schematic thinking (e.g. Hegarty & Kozhevnikov, 1999; Kozhevnikov, Hegarty, & Mayer, 2002; Lohman, 1996), which are of relevance to the present work.

Spatial Ability

Spatial ability has been classified as an independent component of intelligence that involves the mental encoding, processing, and manipulation of spatial representations (Eliot & Smith, 1983, as cited by Carroll, 1993). Spatial ability has been classified as a component of mental imagery abilities, complementary to visual imagery (e.g., Farah, Hammond, Levine, & Calvanio, 1988; Kosslyn, 1990; Logie, 1995). Farah et al. (1988) classify visual imagery as the pictorial component of mental imagery abilities, responsible for the pictorial details of imagery, whereas the spatial subcomponent of mental imagery relates to visualizing spatial relationships among mentally imaged spatial components. Further, they defined visual imagery as being

exclusively modal (e.g., the visual modality), whereas spatial imagery could be modal or amodal, and represent more abstract relationships within mental images.

Within the broad domain of spatial ability, factor analytic studies have supported that there are several sub-abilities which comprise spatial ability as a construct (Carroll, 1993). A multitude of works have categorized spatial ability into multiple different sub-constructs (e.g., Carroll, 1993; French, 1951; Pellegrino, Alderton, & Shute, 2009; Voyer, Voyer, & Bryden, 1995), but there is one sub-construct that exists in nearly every discussion: spatial visualization. Spatial visualization is the most pertinent to the present work and has been defined as the mental manipulation of spatial relationships in two-dimensional and three-dimensional space, involving imagined movements and manipulations of imagined objects (French, 1951, as cited by Carroll, 1993).

Several works have uncovered associations of spatial ability which may be integral for problem solving. Lohman (1996), in a review of spatial ability and general intelligence, identified that people with lower spatial ability tended to have difficulty constructing "systematically structured images," or, a spatial mental representation. Regarding problem solving, Sternberg and Gardner (1983) suggested that spatial mental representations are helpful when solving problems with spatial components (e.g., word problems where one must figure out who is the shortest out of several comparative descriptions – "Bill is shorter than Bob, but taller than Bryan. Bob is taller than Bryan. Who is the shortest?").

In a study examining whether spatial ability was associated with successful problem solving, Hegarty and Kozhevnikov (1999) explored the distinction between visual and spatial imagery regarding children's math problem solving. Those with higher spatial ability tended to adopt more schematic, spatial strategies for solving the math problems, and tended to be more

successful as compared to those with lower spatial ability, who tended to adopt a pictorial strategy. Hegarty and Kozhevnikov's work is perhaps the earliest that identified the role of spatial ability in the formation of schematic mental representations, which aid in problem solving. Later, Kozhevnikov, Hegarty, & Mayer (2002) identified that students who had higher spatial visualization ability tended to construct schematic mental representations of learning material they were presented, based on interviews about how students understood the material. Due to the spatial, schematic nature of Duncker's (1945) radiation problem and all of its analogous problems, spatial ability should be a meaningful predictor of performance on these problems and may prove beneficial for analogical problem solving as well.

Summary

This chapter has reviewed a multitude of factors that influence analogical problem solving (e.g., narrative stories, diagrams, animations, enactments) and problem solving in general (e.g., guided eye movements, visual attention tasks). Additionally, the rationale and justification for alternative factors (e.g., attentional priming of creativity, spatial ability) have been considered. The present study aims to combine the methodologies from the eye-movement studies (Grant and Spivey, 2003; Thomas and Lleras, 2007; 2009) with traditional analogical problem solving paradigms (e.g., Gick and Holyoak, 1980; 1983) to explore the nature of how embodiment and spatial ability influence problem solving success and analogical transfer.

CHAPTER THREE: STUDY 1

Method

To assess the factors that contribute to improvements in problem solving, the following experiments investigated the effects of guided attention on problem solving, in addition to evaluating the influence of cognitive abilities. This section describes three studies. The purpose of the first study is to determine an appropriate duration of induction time that results in a notable increase in problem solution rates, as previous research has not prescribed a set duration for induction. Largely, the first study is a simplified version of Study 2, with the purpose of providing a justification for methodological components of Study 2 (e.g., how long of an induction period to use), and conducting preliminary testing of hypotheses to be assessed in the second study. Study 2 delved deeper into the effects from previous research and systematically examined whether the perceptual simulation effect occurred, or whether this effect was more simply explained by an attentional priming effect of creativity. In addition, it also examined whether there were any unique contributions or interacting effects of spatial ability on the various experimental combinations of this study. It also expanded on the initial study by including an analogical problem solving component, where participants attempted to solve problems with structurally analogous solutions. This served to evaluate whether any initial problem solving benefits confer latent effects on analogical problem solving, or if they are merely temporary benefits for solving a single problem. The final study expanded on the results of Study 2 to examine additional research questions relating to perceptual priming, drawing solutions, and embodied cognition.

The primary purpose of Study 1 was to establish the methodology for the forthcoming studies. Participants underwent an induction phase with a varying duration of 4 minutes or 7 minutes, and performed one of three tracking tasks: Embodied eye movements, Fixed eye movements, or Free Viewing (without tracking). These times were selected based on two timepoints where Thomas & Lleras (2007; 2009) saw the most gains in success rates during their problem solving task. These factors were to serve as an experimental vs. control comparison to provide justification for the design of Study 2. In addition, in line with previous research, this experiment also included a baseline condition, where no induction phase was performed. This served to further differentiate the effects of induction by allowing a comparison from a control group that attempts the problem in absence of induction or incubation. Thus, the major conditions can be conceived as experimental (digit tracking with guided movements), incubation (digit tracking with no guided movements), and control (attempt problem cold). Having these major conditions also permits a more clarified analysis of the unique contributions of cognitive abilities to problem solving success.

Participants

For the first study, 170 participants volunteered their time. The mean age was 19.38 years (*SD* = 2.24), and there were 70 males, 99 females, and 1 who preferred not to respond. The ethnic makeup was 47% Caucasian/White, 30.3% Hispanic, 14.6% Asian, 11.7% Black/African American, and 5.8% Other. Regarding major, 69% of participants in this sample were majoring in STEM (Science, Technology, Engineering, Mathematics) areas, and 34.5% of participants were majoring in a medical-related concentration. The class rank makeup of the sample was 64.1% Freshmen, 20.5% Sophomores, 11.8% Juniors, and 3.7% Seniors.

Materials

Digit Tracking Task

Past research only used a 10 minute induction phase (Thomas & Lleras, 2007; 2009), where there was a 30 second tracking task repeated 20 times. However, participants were able to stop the tracking task whenever they felt prepared to solve the problem. Since needed induction time could vary significantly among participants (not considering those who would not solve the problem), one of the major aims of this initial study was to determine how many stimulus sets would be suitable for a general sample to have sufficient induction. For this task, participants saw a simple diagram representing a tumor and healthy skin tissue, related to the radiation problem they are trying to solve. Around this diagram, participants saw letters appearing inside and outside of the diagram in a particular orientations, intended to guide their eye movements. While participants keep their focus on these appearing letters, at random points during the task, a number appeared in place of a letter. Participants were instructed to click the mouse on the screen when they see a number pop up instead of a letter, and their response time to this digit appearance was recorded. This digit tracking task is nearly identical in design to the task used in Thomas and Lleras (2007; 2009), in efforts to more closely replicate their methodology. In Thomas and Lleras (2007), participants sat 58 cm away from the computer, and viewed a large image of the diagram (subtending a horizontal visual angle of 20.3° and a vertical visual angle of 14°). Each letter and number for the digit tracking task was comparatively small (subtending a horizontal visual angle of 0.3° and a vertical visual angle of 0.5°). To ensure a similar visual angle for all stimuli is subtended, participants were seated approximately 60 cm away from the computer screen, and their chair was guided into place with reference marks on the floor. Those

who are assigned to the experimental groups tracked letters and digits that appeared inside and outside of the center of the diagram in a triangular fashion, which is congruent with the spatial solution to the problems participants solved. Those who are assigned to the incubation groups tracked letters and digits that appeared only in the center of the screen, in absence of a diagram.

Problem Solving

Participants attempted to solve Duncker's (1945) Radiation Problem, as described in Gick and Holyoak (1980, pp. 307-308):

> "Suppose you are a doctor faced with a patient who has a malignant tumor in his stomach. It is impossible to operate on the patient, but unless the tumor is destroyed the patient will die. There is a kind of ray that can be used to destroy the tumor. If the rays reach the tumor all at once at a sufficiently high intensity, the tumor will be destroyed. Unfortunately, at this intensity the healthy tissue that the rays pass through on the way to the tumor will also be destroyed. At lower intensities the rays are harmless to healthy tissue, but they will not affect the tumor either. What type of procedure might be used to destroy the tumor with the rays, and at the same time avoid destroying the healthy tissue?"

All participants viewed this problem with a simple diagram illustrating the tumor and skin, as in previous research (e.g., Grant & Spivey, 2003; Thomas & Lleras, 2007;2009). Previous research classifies this problem as an insight problem, which is traditionally very difficult to solve using typical problem solving techniques (e.g., heuristics, trial-and-error, etc.) (Weisberg & Alba, 1987). In line with previous research, participants also had scratch paper available to them when attempting to solve the problems for their convenience. As previous research has indicated that there may be issues with forcing participants to respond in a certain way (e.g., gesturing, Cooperrider & Goldin-Meadow, 2014), the scratch paper was provided for participants' spontaneous use at their discretion. Although using the scratch paper was optional, their responses were graded based on a written solution to the problem. All participants' scratch

paper was collected whether or not they used it. For this study, only two participants used the scratch paper, and none of them used it on the radiation problem.

Spatial Ability

Participants completed two measures of spatial ability: the card rotations test (CRT), hypothesized to assess basic spatial processing, and the paper folding test (PFT), hypothesized to assess spatial visualization. Due to the spatial and schematic nature of the radiation problem's solution, and considering prior research suggesting that spatial ability (in particular, spatial visualization) is associated with the use of schematic representations during problem solving, these measures were selected to assess whether these aspects of spatial ability contribute to problem solving success, and whether they interact with a schematic induction phase.

The CRT (Ekstrom, French, Harman, & Dermen, 1976) is a 3-minute test containing ten sets of items. For each set of items, participants see a prototypical "card." Next to this card are eight test cards, which are rotated or mirrored and rotated versions of the prototypical card. Participants must make a judgment about whether each of these eight test cards are the same as the prototypical card (e.g., rotated), or different (e.g., a mirrored image). Participants' scores are calculated by summing the number of correctly identified test items and subtracting the number of incorrectly identified test items.

The PFT (Ekstrom et al., 1976) is also a 3-minute test containing ten sets of items. Participants are presented an image of a square piece of paper that undergoes a series of one to three folds, before a hole is punched through it on the final step. Participants must select from among five choices representing what the punctured piece of paper would look like after being

completely unfolded. Participants' scores are calculated by summing the number of correctly selected items and subtracting 20% of the number of incorrectly selected items.

Although the literature suggests that spatial visualization ability (a construct measured by the PFT) is related to spatial thinking during problem solving (Hegarty & Kozhevnikov, 1999; Kozhevnikov et al., 2002), a more basic measure of spatial ability may be of use in more broadly assessing the relationship between spatial ability and problem solving success. Conceptually, each measure assesses qualitatively different aspects of spatial ability, which may be disparately related to problem solving success, or predict complementary aspects of problem solving success which may not be detectable with only a single measure. Although spatial visualization specifically is hypothesized to be related to problem solving success, a composite measure may prove useful in predicting variance in problem solving success if elements of spatial relations play a complementary role in problem solving.

Demographic Questionnaire

This questionnaire included basic demographic questions, including as age, school classification (e.g., freshman, senior, etc.), gender, and ethnicity, as well as various questions hypothesized to be related to problem solving success. Specifically, participants were asked questions about their relevant experience, due to previous research suggesting that medical students are more likely to solve the radiation problem due to its medical-procedural features. In a similar vein, participants were also asked to provide their major, or the subject they were most interested in if they had not yet declared a major. Participants also were asked if they had previously seen any of the problems presented in the study (which naturally increases the

likelihood of solving the problems). Lastly, participants were asked to self-report on a 5-point Likert-type scale the degree to which they liked solving puzzles.

Design

To determine the appropriate length for the induction phase, this study was a 3 (condition) by 2 (digit tracking duration) between subjects design, with an additional baseline condition that does not undergo an induction phase (for a total of 7 groups). These groups were compared between each other at each time point, and individually against the baseline condition, to determine which induction duration results in the most significant increase in solution success rates. These results were intended to provide justification for the induction duration for the proceeding study.

Procedure

For study 1, participants were provided a general overview of the study, using the informed consent as a guide. They had an opportunity to ask questions before the study began. Once any questions were answered to participants' satisfaction, they were directed to a computer where they began the study. First, they were presented with an informed consent form, after which they were prompted to indicate whether they provide consent.

After participants have consented, all conditions viewed the radiation problem. They were permitted to read it for as long as they liked and clicked a "next" button when they were ready to proceed. Those in the guided eye movements, fixed eye movements, and free-viewing conditions proceeded to do their digit-tracking tasks for an 'induction phase' of 4 or 7 minutes. Those in the baseline condition did not undergo an induction phase but were immediately

prompted to generate solutions for the radiation problem. Those in the other conditions proceeded to the solution phase after they complete their induction phase.

For all conditions, during the solution phase, they saw the problem as described above, with the same simple diagram they saw during the initial problem presentation. They did not have a time limit to solve the problem. Once participants submitted their solutions to the problem, they received a hint statement, saying "What if you had multiple lasers? It's okay to repeat a solution you gave earlier." After they received this hint statement, they saw the problem one final time, and were instructed to provide the solution again (even if they have to repeat themselves). Afterward, they were shown the written form of the convergence solution and queried if they noticed a similarity between the induction phase and the solution to the problem.

After the problem solving portion of the study has concluded, the experimenter collected participants' scratch paper, and they proceeded to complete their spatial tests Once the cognitive tests have been completed, participants completed the demographic questionnaire. This questionnaire was presented last to minimize potential stereotype threat effects that may occur as a result of having participants identify their gender or race prior to taking cognitive tests (e.g., Steele & Aronson, 1995). The average time participants spent in this study was approximately 40 minutes, for which they received course credit.

A figure outlining the procedure of Study 1 is presented below (see Figure 10).

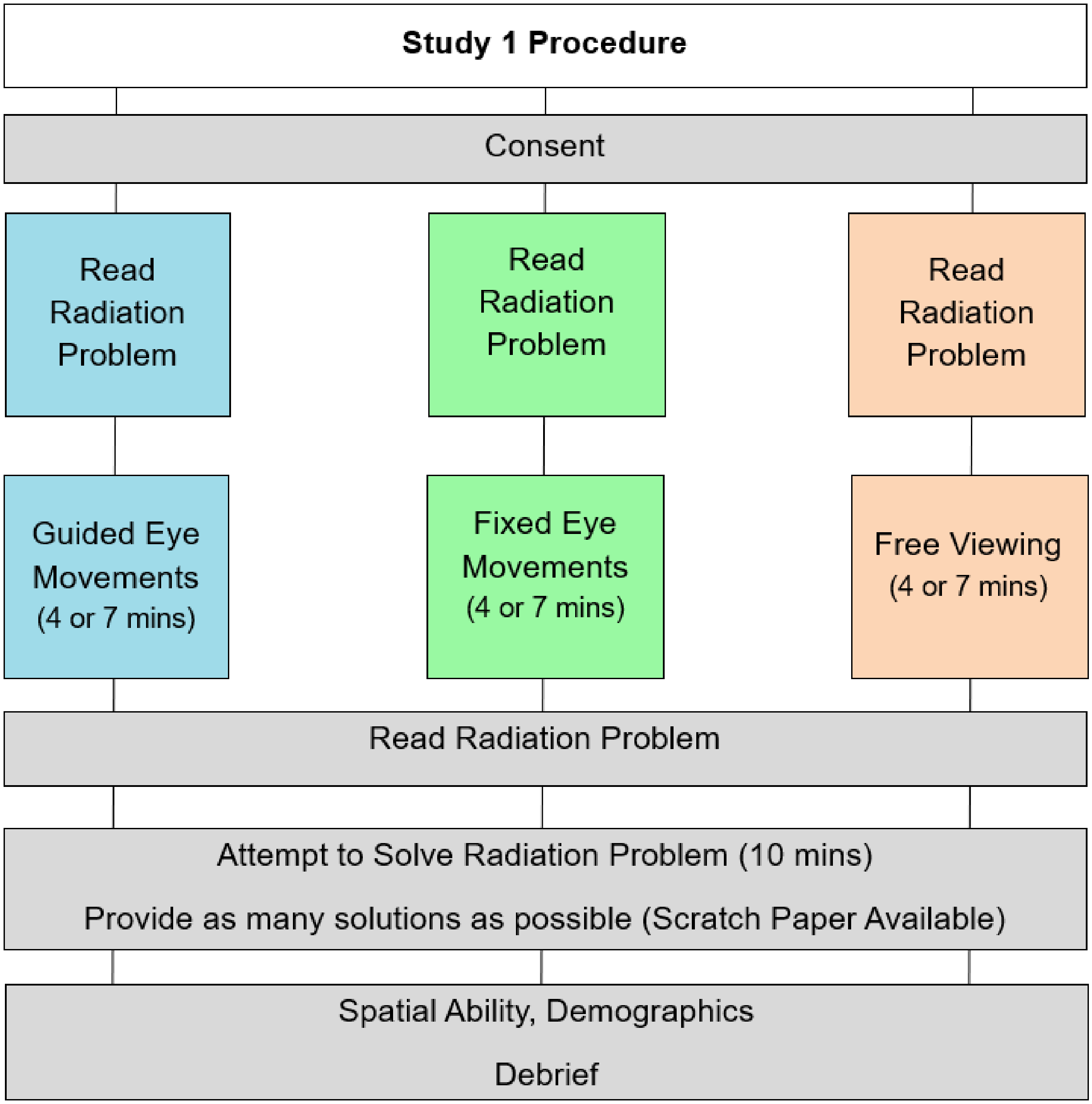

Figure 10. Outline of the procedure for study 1.

Different conditions are denoted by differently colored boxes. Note that the "Guided Eye Movements," "Fixed Eye Movements," and "Free Viewing" conditions consist of 2 groups each, separated by duration. Additionally, a baseline group went straight to attempting to solve the radiation problem without performing any eye movement or free viewing task.

Results

For an overview of the results addressing each study's research questions, please refer to Appendix D.

In this section, each hypothesis is listed with an analysis plan. A power analysis (through G*Power; Faul, Erdfelder, Lang, & Buchner, 2007) is provided for each analysis to be conducted. As a general note, response times on the digit-tracking task were utilized as an exclusionary criterion from any analyses. Participants whose response times were greater than 3 standard deviations above the mean response time on the digit tracking task were excluded from analyses. Similarly, those whose accuracy was less than 75% on the digit tracking task were also removed.

Data Prescreening

Participants Removed from Analyses

For this study, two participants were removed from analyses for the following reasons: One participant Christmas-treed their responses on the individual difference measures, and both participants did not complete the tracking task correctly. There were no participants who reported having seen the radiation problem before. All other participants were retained for analyses (remaining N = 168).

Tracking Task Reaction Times

Before analyzing Study 1's data, participants' performance on the tracking task was assessed. Only those whose tracking accuracy was 75% or greater were retained for analyses. Accuracy was determined by excluding participants whose reaction times were greater than three

standard deviations above the mean value, which was roughly 1000 milliseconds after the stimulus was presented. Participants were also excluded if their reaction was recorded prior to the onset of the stimulus. As mentioned before, two participants were excluded from analyses for this reason.

For the digit-tracking task, participants were required to react by clicking the mouse when they saw a number appear on a diagram of the radiation problem. Their response times were recorded when they clicked their mouse. In Study 1, there were only four groups that underwent a tracking task that required reaction time (the free-viewing and baseline groups did not have a tracking component in their version of the study). A two-way ANOVA on tracking times with duration (4 minutes or 7 minutes) and eye movement guidance (embodied or fixed) was performed but showed no significant differences across groups (F (3, 102) = 2.08, p = .107, η^2_p = .06). These reaction times were consistent in pattern and duration with previous studies (Thomas & Lleras, 2007; 2009), where the fixed conditions showed quicker reaction times than the guided conditions, likely due to the guided tracking task requiring more eye movements which increased their reaction time. Although the differences in the present study were not significant, they were similar in direction from previous research. Tracking times for each of the relevant conditions are provided below in Table 1.

Table 1. Tracking Task Reaction Times (Means and Standard Deviations) for Study 1

Condition	M	SD
Guided, 7 min	603ms	105ms
Guided, 4 min	625ms	131ms
Fixed, 7 min	537ms	162ms
Fixed, 4 min	582ms	121ms

Hypothesis Testing

The purpose of Study 1 and preliminary hypothesis tests are described in this section, based on the conditions outlined in Table 2 below.

Table 2. Groups in Study 1 with sample sizes and solution rates (provided as percentages)

	4 min. tracking task	**7 min. tracking task**
Guided Eye Movements	**(1)** *n* = 28; 17.9%	**(2)** *n* = 27; 11.1%
Fixed Eye Movements	**(3)** *n* = 26; 26.9%	**(4)** *n* = 25; 8.0%
Free Viewing	**(5)** *n* = 20; 20.0%	**(6)** *n* = 21; 23.8%
(7) Baseline Group *n* = 21; 14.2%		

Note. These sample sizes were calculated after removal of participants, described in the results section of this chapter.

In general, the purpose of Study 1 was to determine an appropriate tracking task duration that results in notable differences in solution rates between the embodied eye movements and fixed eye movements groups. Other measures were collected during Study 1 (e.g., spatial ability), and were used to conduct preliminary tests of major study hypotheses.

The main comparison is a six-group chi-square test, comparing the solution rates of groups (1) – (6) (this procedure is in-line with traditional comparisons of problem solving success; Gick & Holyoak, 1980). The baseline group (7) was collected to identify a general rate of solving the radiation problem without any aids. Using default values from G*Power (Faul et al., 2007), a power analysis suggested a sample size of 127 is applicable, requiring roughly 20-25

participants for each of the 6 groups (see Figure 11). The number of participants analyzed in Study 1 was 168 total (147 when excluding the baseline group from analyses).

χ^2 tests – Goodness-of-fit tests: Contingency tables

Analysis:	A priori: Compute required sample size		
Input:	Effect size w	=	0.35
	α err prob	=	0.05
	Power (1–β err prob)	=	0.95
	Df	=	2
Output:	Noncentrality parameter λ	=	15.5575000
	Critical χ^2	=	5.9914645
	Total sample size	=	127
	Actual power	=	0.9514452

Figure 11. Power analysis for chi-square test used in Study 1.

Solution Rates by Condition

The solution rate for the baseline group was 14.2% (3 correct, 18 incorrect). A chi-square test comparing groups (1) through (6) showed that no group was statistically different from any of the others; χ^2 (5) = 4.55, p = .47. The chi-square tests comparing condition (Guided, Fixed, Free Viewing) and duration (4 minute, 7 minute) were also not statistically significant; χ^2 (2) = 0.86, p = .64 for condition; χ^2 (1) = 1.58, p = .21 for duration. No single group was statistically better than the baseline group (lowest p = .29).

Spatial Ability and Incubation

Because spatial ability measures were collected in Study 1, it is possible to preliminarily test Hypotheses 3, 3a, and 3b. Hypothesis 3 (that spatial ability is associated with performance) was tested using a logistic regression analysis, with convergence solution rates as the outcome variable and spatial ability (as measured by the PFT or the CRT) as the predictor variable. The result with the PFT was not statistically significant (χ^2 (1) = 1.16, p = .28), and the result with the CRT was also not significant (χ^2 (1) = 0.05, p = .83).

Hypotheses 3a and 3b were also tested using logistic regression, allowing the continuous variable of spatial ability to interact with the categorical variable eye condition. Because there were no significant differences between the 4 or 7-minute induction groups, this analysis was performed inclusive of both durations. There were no significant main effects for spatial ability or condition, and the interaction was not significant (overall model with PFT χ^2 (3) = 3.68, p = .30; overall model with CRT χ^2 (3) = 1.46, p = .69).

Hypothesis 4 (incubation) was also preliminarily tested by comparing the fixed eye movement groups against the baseline group. The fixed eye movement manipulation should convey no benefit other than incubation of the problem. By comparing it to the baseline group, an incubation effect can be estimated. This result was not statistically significant (χ^2 (2) = 3.41, p = .18), although this effect was in the hypothesized direction, with the fixed eye movement groups showing slightly higher solution rates than the baseline group.

Creativity in Written Solutions

Aside from simple solution rates, there are also a variety of ways that creativity can be assessed. Adapting Guilford's (1967) ideas of divergent thinking, participants' written responses were analyzed in four different ways. Fluency, or the correctness of an answer; Originality, or the idea that a response is rare compared to other responses; Flexibility, or the amount of different kinds of solutions provided; and Elaboration, described as the amount of detail provided in a response. Fluency was assessed by whether the participant provided a convergence solution (results provided in previous section). Originality was scored using Guilford's (1967) guidelines, where answers provided by 5% of the sample earn 1 point, and answers provided by 1% of the sample earn 2 points. Higher points correspond to greater originality. Flexibility was

determined based on the number of unique categories of solutions provided as indicated by an answer coding scheme. Lastly, Elaboration was determined by counting the character length of the typed responses.

In addition to these categories based on Guilford's (1967) work, two other categories were created to expand the present work's ability to test how the experimental manipulations may have affected participants' problem solving behaviors. These variables were the total number of solutions each participant generated (Total Solutions), and the total number of redundant solutions generated (Redundant Solutions). Redundant solutions were calculated by subtracting the number of unique solution types (Flexibility) from the total number of solutions. A high number of redundant solutions could indicate that a problem solver has become fixated on a type of solution and may represent a negative side to attentional priming. These variables were included in the forthcoming analyses to better understand the potential effects of attentional priming on creativity.

Prior to analyzing creativity dependent variables, a coding scheme for the written solutions to the radiation problem was generated with one other independent coder. This coding scheme was developed with both coders naïve to which condition each participant was randomly assigned (see Table 3). Based on recent guidelines for interrater reliability from McHugh (2012), Cohen's κ scores of interrater reliability were excellent for each of the three studies (mean $\kappa = .82$, .88, and .96 for Studies 1, 2, and 3, respectively).

Table 3. Coding Scheme for Solution Types to the Radiation Problem.

		Rarity of solution		
Coding Scheme		Study 1	Study 2	Study 3
A	Convergence solution	**4.63%**	9.54%	7.74%
B	Cut/Open up patient, then fire at tumor	10.69%	15.05%	13.81%
C	Consume something that alters tissue/tumor weakness/strength	**4.63%**	**1.83%**	**3.77%**
D	Strengthen/protect healthy tissue	**4.15%**	**2.57%**	**4.18%**
E	Variable intensity/High and Low intensities	**3.99%**	8.07%	5.44%
F	Let the tumor grow before attempting anything	**0.64%**	**0.37%**	**0.21%**
G	Go full intensity and hope patient recovers	13.72%	15.96%	19.67%
H	Medium Intensity	6.38%	7.71%	7.95%
I	Send something through the body that shoots rays at tumor	5.42%	**3.49%**	**4.18%**
J	New technology (nanorobots, CRISPR, gene editing, etc.)	**1.12%**	**0.73%**	**0.42%**
K	Move/Replace healthy tissue	7.18%	7.71%	6.28%
L	Vague/nondescript/find another way	10.53%	5.32%	3.56%
M	Ignore/don't operate/do nothing	1.44%	1.65%	1.26%
N	Timing/Multiple sessions (brief high intensity or brief low intensity)	10.37%	9.17%	11.30%
O	Remove the tumor/move the tumor	5.42%	**3.30%**	**3.97%**
P	Send rays at areas/angle with least healthy tissue	5.42%	5.69%	**2.93%**
Q	Organ/Tissue Transplant	**1.44%**	**0.55%**	**0.63%**
R	Reflect/Refract rays using lenses/mirrors	**0.80%**	**0.55%**	**0.42%**
S	Chemotherapy	**1.44%**	**0.55%**	**1.88%**
T	Split the tumor up and then fire at it	**0.32%**	**0.18%**	**0.42%**
U	Sample the tumor to create a drug to fight it	**0.16%**	**0.00%**	**0.00%**
V	Use electric current	**0.16%**	**0.00%**	**0.00%**

Note. Solution rarities were determined by dividing each solution's frequency by the number of all solutions provided in each sample. "Rare" solutions (<5% occurrence rate as per Guilford, 1967) are bolded. Note that solution category L was not counted as a "rare" solution because it served as a "catch-all" category for solutions that were incomplete or did not make sense. Solution category M was also not counted as a "rare" solution because it represents a lack of attempt for solving the problem. Participants were instructed to provide as many different solutions as they could think of to solve the radiation problem.

To assess whether duration or condition resulted in differences in creativity, five 2 (Duration) x 3 (Condition) factorial ANOVAs were performed on scores of Originality, Flexibility, Elaboration, Total Solutions, and Redundant Solutions (see Appendix E for a

MANOVA of these analyses). For Originality, there were no significant differences among any of the variables or their interaction, $F(5, 141) = 0.70$, $p = .63$, $\eta^2_p = .024$ (see Figure 12).

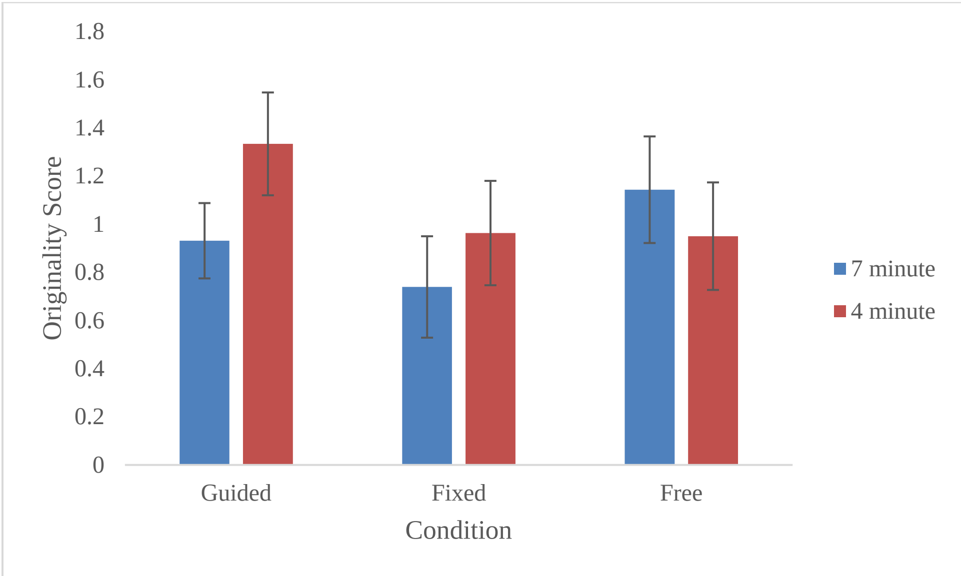

Figure 12. Originality scores for each condition and duration. Error bars represent the standard error of the mean. There were no significant differences among the groups.

For Flexibility, there also were no significant differences among any of the variables or their interaction, $F(5, 141) = 1.38$, $p = .24$, $\eta^2_p = .047$ (see Figure 13).

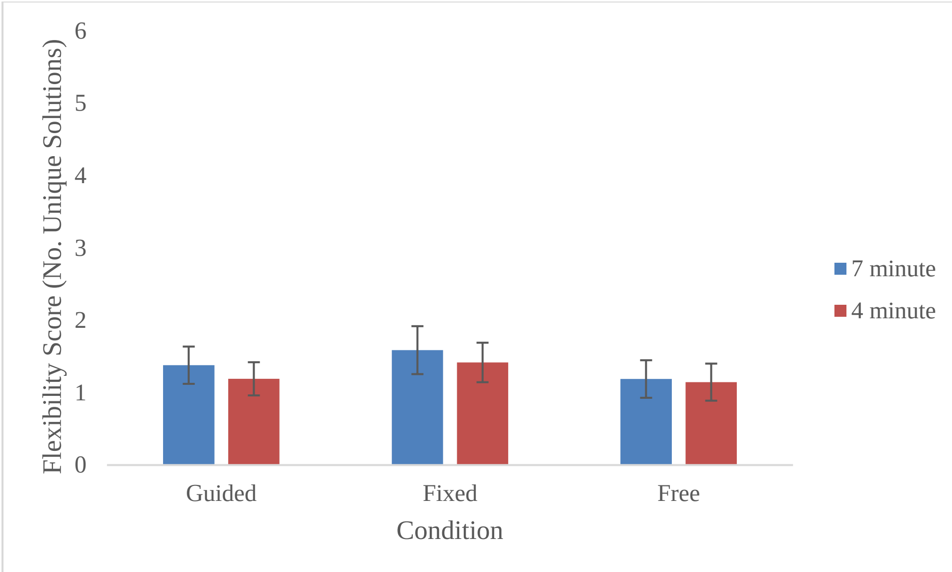

Figure 13. Flexibility scores for each condition and duration. Error bars represent the standard error of the mean. There were no significant differences among the groups.

For Elaboration, there also were no significant differences among any of the variables or their interaction, $F(5, 141) = 1.18$, $p = .33$, $\eta^2_p = .040$ (see Figure 14).

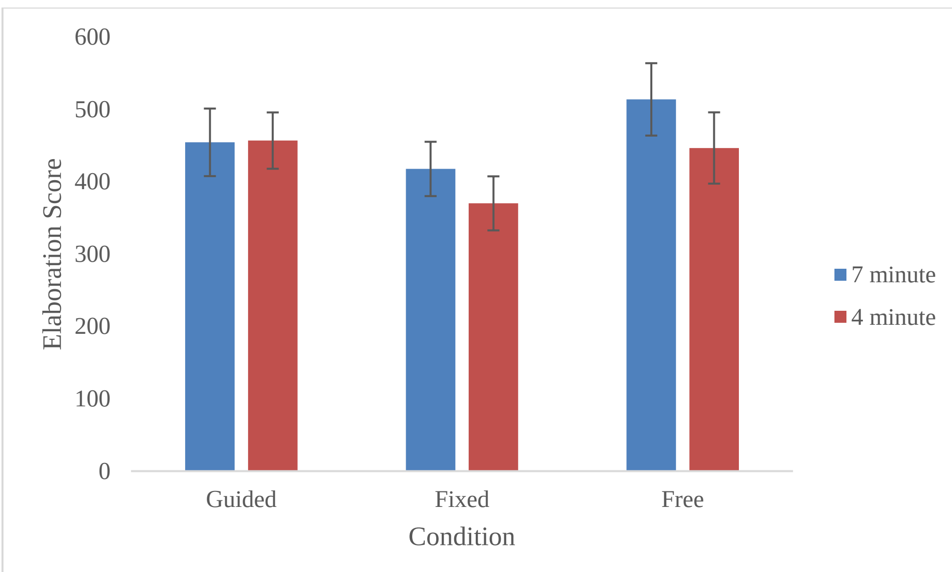

Figure 14. Elaboration scores for each condition and duration. This score was measured using character count length in participants' responses. Error bars represent the standard error of the mean. There were no significant differences among the groups.

There also were no significant differences for Total Solutions among any of the variables or their interaction, $F(5, 141) = 1.09$, $p = .37$, $\eta^2_p = .037$ (see Figure 15).

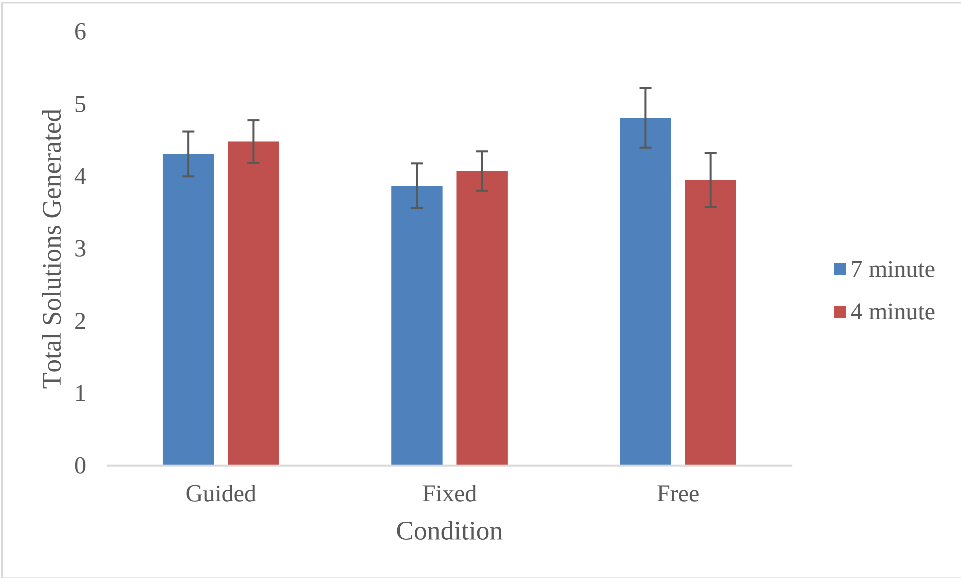

Figure 15. Total Solutions generated for each condition and duration. Error bars represent the standard error of the mean. There were no significant differences among the groups.

Lastly, there was a significant difference in Redundant Solutions across the groups, $F(5, 141) = 3.59$, $p = .004$, $\eta^2_p = .113$. This difference was driven by a main effect for condition ($F(1, 141) = 4.55$, $p = .01$, $\eta^2_p = .061$), as well as an interaction between duration and condition ($F(2, 141) = 3.46$, $p = .03$, $\eta^2_p = .047$). The Free Viewing group had significantly more Redundant solutions than all other groups (Least Significant Difference (LSD) *p*s for mean differences < .05), and this was particularly more pronounced for the 7-minute duration (LSD *p*s for mean differences < .01; see Figure 16).

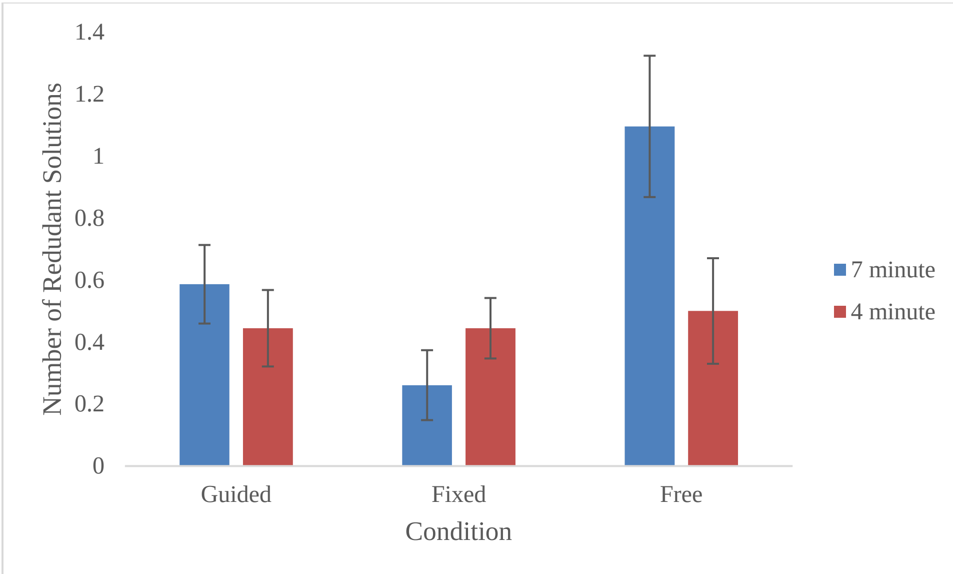

Figure 16. Number of Redundant Solutions for each condition and duration. Error bars represent the standard error of the mean. The Free-Viewing condition had significantly more redundant solutions than the other two conditions, which was driven by those in the 7-minute free-viewing condition.

It should be noted that the baseline group was excluded from the previous analyses because they did not undergo a digit-tracking or free-viewing task. When compared against all other conditions on the previously analyzed creativity variables, the baseline group was not significantly different than any other condition on any variable (all *p*s > .15) except for Originality and Redundant Solutions (LSD *ps* < .03). The baseline group had significantly lower Originality than the 4 minute guided condition and the 7-minute free-viewing condition, as well as significantly fewer Redundant Solutions than the 7-minute free-viewing condition (this is parallel with the previous factorial ANOVA showing the 7-minute free-viewing condition had disproportionately the most Redundant Solutions compared to all other conditions).

Discussion

In general, these results do not support any proposed hypotheses, and the data do not support any specific induction times for the digit-tracking task. Furthermore, there were no differences in creativity scores observed with any of the conditions. Prior to proceeding with Study 2, the methodology was re-evaluated in the context of the previous studies. There were two main differences in Study 1's methodology that may have contributed to why the effects seen in Grant and Spivey (2003) and Thomas and Lleras's (2007; 2009) studies did not replicate in Study 1's similar conditions. These methodological factors are discussed below, with the overall goal to more closely align the initial problem-solving aspect of the present work to that of Grant and Spivey (2003) and Thomas and Lleras's (2007; 2009) studies. In doing so, the next two studies should be better able to verify the existence of an embodied effect on problem solving.

The first methodological factor is the impact of implicit feedback. In Thomas and Lleras's (2007; 2009) studies, participants underwent the digit-tracking task while attempting to solve the radiation problem simultaneously. When a participant attempted to generate a solution, the digit-tracking task was interrupted. During this interruption, participants either exited that phase of the study if their solution was judged to be correct, or they resumed the digit-tracking task if their solution was incorrect. In cases where their solution was incorrect, this return to the digit-tracking task could have served as a kind of implicit feedback. A participant might think "my solution must be wrong, so I should think of a different solution." The implicit feedback inherent to their task design likely facilitated a metacognitive reflection of each participant's problem solving process, which could reasonably improve solution success rates (Davidson &

Sternberg, 1998). The present study did not include this methodological feature, because including it would preclude any possibility of answering relevant research questions about incubation or embodied priming. Specifically, it would not make sense to have a participant undergo a digit-tracking task while simultaneously attempting to solve a problem if they had never seen it before.

The second methodological factor is the impact of drawing solutions. In all three previous studies, participants' only way of providing the solutions was by tracing their solution on the computer screen in which they performed the digit-tracking task. Successful solutions involved multiple lines converging on the tumor from different angles. Although participants in Study 1 were permitted to use scratch paper at their convenience, they were not required to do so. This methodological aspect was adopted based on previous research that indicated that forcing problem solvers to respond in a certain way could restrain their responses, risking that they would not express a solution that they could generate (Cooperrider & Goldin-Meadow, 2014; Hostetter et al., 2016). Incidentally, even though participants in Study 1 had scratch paper available to them, only three of them used it while solving the radiation problem. In Grant and Spivey (2003) and Thomas and Lleras's (2007; 2009) studies, their participants only solved the radiation problem through drawing their solutions, so it is possible that the embodied effects they found were spatial in nature and may not emerge through written answers. To address this, all participants were required to draw their solutions on scratch paper in the forthcoming studies.

Prior to beginning data collection for Study 2, the digit-tracking task's duration and type was also re-evaluated. Although there were no statistical comparisons that supported the 4-minute or 7-minute durations compared to the Baseline Group, the 4-minute digit-tracking task duration was selected for the proceeding studies. Thomas and Lleras (2007; 2009) saw the

largest increases in solution rates at the 4-minute mark of their digit-tracking tasks, and those benefits diminished thereafter for the remainder of the 10-minute task. Taken together, these criteria were used to justify the use of a 4-minute tracking task in Study 2. Regarding the eye movement guidance, participants anecdotally appeared to experience frustration and discomfort when doing the free-viewing task. Those in the free-viewing condition stared at a diagram of the radiation problem for 4 or 7 minutes with no eye movement guidance or other interaction with the computer. It is possible that this lack of interaction did nothing to help participants incubate their solutions to the radiation problem, but rather increased boredom due to a lack of interaction. Furthermore, its primary purpose in Study 1 was to serve as a preliminary comparison condition against the pure baseline group to test the incubation hypothesis. It was not supported as being any better than the baseline group, nor was it statistically better than any of the other conditions. For these reasons, and because the incubation hypothesis was tested in a different fashion for Study 2, the free-viewing condition was not included in any proceeding studies.

CHAPTER FOUR: STUDY 2

Method

Considering the null results of Study 1, Study 2's design sought to match previous work (i.e., Grant & Spivey, 2003; Thomas & Lleras, 2007; 2009) to more accurately verify the existence of any embodied effects on problem solving. This study's design also expands Study 1's procedure, extending the embodied problem solving paradigm into analogical problem solving to answer research questions regarding whether embodied problem solving effects transfer into later problem solving.

The primary research questions of Study 2 are an extension of those in Study 1. First, does embodied priming lead to an increase in problem solving success rates, and furthermore, does this effect also emerge when solving structurally similar transfer problems? Second, is incubation a better explanation for the effects observed in previous work? Additionally, are these effects reliably predicted by spatial ability? And lastly, are any of these effects better explained by an attentional priming of creativity? Although some of these research questions were not supported in Study 1, Study 2's criteria for evaluating problem solving success changed to allow drawings to be judged for correctness. Therefore, this change allows for these research questions to be tested in a new way.

Participants

For the second study, 173 participants volunteered their time. The mean age was 19.64 years (*SD* = 1.97), and there were 94 males and 79 females. The ethnic makeup was 49.1% Caucasian/White, 32.3% Hispanic, 9.9% Asian, 13% Black/African American, and 6.1% Other. Regarding major, 70.8% of participants in this sample were majoring in STEM (Science,

Technology, Engineering, Mathematics) areas, and 28% of participants were majoring in a medical-related concentration. The class rank makeup of the sample was 62.1% Freshmen, 13.7% Sophomores, 16.1% Juniors, 7.5% Seniors, and 0.6% 5th year or higher. Participants were randomly assigned to 1 of 4 groups, described below in Table 4.

Table 4. Groups and Sample Sizes in Study 2

	Problem Before + After	**Problem After**
Embodied Eye Movements	n = 39 **(1)** Perceptual Simulation Group	n = 39 **(2)** "Attentional Priming" Group
Fixed Eye Movements	n = 40 **(3)** Simple Incubation Group	n = 43 **(4)** "Control" Group

Note. These sample sizes were calculated after removal of participants, described in the results section of this chapter.

Materials

All materials used in Study 1 were used in Study 2, with additional problems for participants to solve, going beyond assessing problem solving to assess analogical problem solving. For intermediary tasks between these new problems, participants completed measures of verbal ability. It also included additional variations to the digit tracking task to more precisely examine differences among different induction phase characteristics.

Verbal Ability

Participants completed a word fluency test (WFT) and a verbal comprehension test (VCT). Previous research indicated that these measures are important for analogical problem solving (e.g., Corkill & Fager, 1995), and it might be possible that successfully solving the

problem may differ as a function of verbal ability. The scores provided by these measures were included as a control variable when analyzing problem solving success.

The WFT (Ekstrom et al., 1976) consists of two three-minute tests that require participants to produce as many words as they can beginning or ending with a certain string of letters (e.g., valid words starting with “re-” might include “reduce,” “reprimand,” or “regal”; valid words ending in “-ate” might include “rate,” “equate,” or “sedate”). Participants’ WF score is the number of correct words that they produce in the time limit.

The VCT (Ekstrom et al., 1976) assesses participants’ knowledge of word meanings by presenting a test word with five multiple choice words. Among these five words, participants must select the most closely related choice to the test word. For example, if the test word “jovial” was presented, and choices “refreshing, scare, thickset, wise, and jolly” are available, a correct response would be to select “jolly.” The VC test is 6 minutes in length and contains 24 items.

Digit Tracking Task

For Study 2, the digit tracking task had two possible levels of the independent variable – guided eye movements (guided or fixed). Those who are randomly assigned to the guided level of the guided eye movements variable did their digit tracking task with letters and numbers appearing in the center of the diagram, and at various points at the outside in a triangular fashion with repeated visits to the center of the screen. Those randomly assigned to the fixed level of the guided eye movements variable did their digit tracking task with letters and numbers appearing only in the center of the diagram. The procedure is largely the same as Study 1 with regard to

digits appearing and a number randomly appearing in each set. Participants were instructed to click the mouse whenever they saw a number appear, and this response time was recorded.

Problem Solving

In addition to the radiation problem, participants were tasked with solving the problems entitled "Red Adair" and "The General." Their text are as follows (as written in Gick and Holyoak 1983; excerpts from pp. 36-38; last sentence added to each to prompt problem solving process):

> **Red Adair:** "An oil well in Saudi Arabia exploded and caught fire. The result was a blazing inferno that consumed an enormous quantity of oil each day. After initial efforts to extinguish it failed, famed firefighter Red Adair was called in. Red knew that the fire could be put out if a huge amount of fire retardant foam could be dumped on the base of the well. There was enough foam available at the site to do the job. However, there was no hose large enough to put all the foam on the fire fast enough. The small hoses that were available could not shoot the foam quickly enough to do any good. It looked like there would have to be a costly delay before a serious attempt could be made. Using the available hoses and flame retardant, how might Red Adair put out the fire?"

> **The General:** "A small country was ruled from a strong fortress by a dictator. The fortress was situated in the middle of the country, surrounded by farms and villages. Many roads led to the fortress through the countryside. A rebel general vowed to capture the fortress. The general knew that an attack by his entire army would capture the fortress. He gathered his army at the head of one of the roads, ready to launch a full-scale direct attack. However, the general then learned that the dictator had planted mines on each of the roads. The mines were set so that small bodies of men could pass over them safely, since the dictator needed to move his troops and workers to and from the fortress. However, any large force would detonate the mines. Not only would this blow up the road, but it would also destroy many neighboring villages. It therefore seemed impossible to capture the fortress. After learning of the mines, what strategy can the general take to overtake the dictator's fortress?"

Although the radiation problem includes a diagram when presented to participants, the additional problems in Study 2 did not include a diagram so as not to implicitly prime

participants to retrieve or recall the radiation problem. Including diagrams could unintentionally initiate the analogical reasoning process (as per Gick, 1985; Gick & Holyoak, 1983), and the present study aimed to investigate analogical reasoning independent of any cuing effects. Thus, these problems were presented in narrative form with a new piece of scratch paper provided for each problem. Since there were three problems in this study, participants were provided with three pieces of scratch paper; one for each problem. After each problem was solved, participants were instructed to deposit their scratch paper into an envelope matched to their participant number. This was for data storage purposes, as well as to prevent their scratch paper from one problem being used on another problem and potentially priming their responses.

Design

To examine the influence of context, embodied aid, and spatial ability, the present study used a 2 (problem context) x 2 (guided eye movements) completely between design, using spatial ability as a continuous moderating variable for additional analyses. Problem context had two levels: whether the participant sees the radiation problem before or after the induction phase. Guided eye movements also had two levels: participants either had guided eye movements or fixed eye movements during the digit tracking task.

There were several dependent variables. For each problem, there are multiple points where participants are prompted to provide a solution to the problem. Each of these points generated a binary dependent variable – did the participant solve the problem successfully? Did the participant solve the problem successfully after receiving a hint? On the transfer problems, participants were asked to describe similarities between the radiation problem and one or more transfer problems. These narrative descriptions provided qualitative data that were analyzed to

judge whether participants noticed the convergence similarities among problems (e.g., they noticed that the problems all required a combining of weaker forces at a central target).

In line with the analysis techniques of previous studies (e.g., Gick & Holyoak, 1980; 1983), all groups were compared using a Chi-Square test to examine whether any of the groups significantly differ with respect to solution rates before and after hints on each of the problems. To further address research questions with spatial ability, these data were analyzed using regression techniques to examine the influences of each variable and the interactions among the variables on solution success. These analyses were conducted for each step of the procedure where solutions are generated (e.g., initial solutions, hinted solutions). Because the DV is binary (a solution is either provided or not provided), logistic regression analyses were used to examine problem solving success. The regressions were analyzed in several blocks as outlined below (see Table 5).

Table 5. Regression analysis plan for study 2.

Block 1 (Main effects and Moderator)
IV1: Problem Context
IV2: Embodied Aid
M: Spatial Ability
Block 2 (2-way interactions/moderations)
Context * Embodied Aid
Spatial Ability * Context
Spatial Ability * Embodied Aid
Block 3 (3-way interaction/moderations)
Embodied Aid * Context * Spatial Ability

Logistic regression was the main form of analysis for the present study, and one of the most interpretable effects from a logistic regression analysis is the odds-ratio (or Exp(B) value).

In essence, the odds-ratio provides a multiplier of sorts for a given effect. Grant and Spivey (2003) and Thomas and Lleras (2007; 2009) identified that their guided eye movement manipulations resulted in problem solving success rates that were between 2 to 4 times greater than those with fixed eye movements. To err on the conservative side, a power analysis is provided with a low-end estimate odds-ratio of 2, with all other values set to default (see Figure 17). This required approximately 40 participants per group, for a total required sample size of 160 for Study 2.

Options:	Large sample z-Test, Demidenko (2007) with var cc		
Analysis:	A priori: Compute required sample size		
Input:	Tail(s)	=	One
	Odds ratio	=	2
	Pr(Y=1|X=1) H0	=	0.2
	α err prob	=	0.05
	Power (1−β err prob)	=	0.95
	R^2 other X	=	0
	X distribution	=	Normal
	X parm μ	=	0
	X parm σ	=	1
Output:	Critical z	=	1.6448536
	Total sample size	=	148
	Actual power	=	0.9502750

Figure 17. Power analysis for logistic regression analysis used in Study 2.

Procedure

The initial procedure was similar Study 1, except there were no free-viewing or baseline groups. As in Study 1, all participants performed a simple training task showing them how to perform the tracking task on their computer before beginning the study. Participants were randomly assigned to a "problem before" or a "problem after" digit-tracking condition (both levels of the problem context variable). Participants randomly assigned to the "problem before" conditions received the radiation problem before the digit-tracking task. They were permitted to read it for as long as they liked and clicked a "next" button when they were ready to proceed to

the digit-tracking task. Those assigned to the "problem after" conditions began their study with their randomly assigned digit-tracking task.

For the digit tracking task, participants were randomly assigned to one of two possible conditions – guided eye movements (where the digit tracking task occurs with letters and digits appearing in and out of the center of the radiation problem diagram, in a triangular fashion), or unguided movements (where the digit tracking task occurs with a diagram present, but letters and numbers only appear in the center of the diagram). This task lasted for a duration of 4 minutes, consisting of eight 30-second tracking tasks.

After completing the digit-tracking task, all participants saw the radiation problem and attempted to solve it, providing as many solutions as possible within a 10-minute limit. This procedure was identical to the solution phase in Study 1, including provision of the convergence solution at the conclusion of the solution phase.

Once participants completed the solution phase, they completed the VCT to serve as an intermediary task before proceeding to attempt solving "Red Adair." Once they attempted solving "Red Adair," they were asked to describe whether they noticed any differences and/or similarities between "Red Adair" and the radiation problem. At that point, they were instructed to attempt providing a solution again to "Red Adair," even if it was the same one they provided before. According to theories of analogical reasoning, prompting participants to recall the radiation problem serves to facilitate the "retrieval" process of analogical reasoning, and should increase the rates at which participants solve the problem.

After providing solutions for "Red Adair," participants completed the two parts of the WFT as another intermediary task and then proceeded to solve "The General." After providing their solution to the general, participants were asked to describe whether they noticed any

differences and/or similarities among “The General,” “Red Adair,” and the radiation problem. At this point, they were instructed to attempt providing a solution again to “The General,” even if it was the same one they provided before. After they provided this solution, they were presented with the solution to the radiation problem (“convergence solution”). Then, they were instructed to attempt to apply the radiation problem’s solution to “The General,” even if they had to repeat any of their previous solutions. In doing so, this step facilitated the “retrieval” and “mapping” processes of analogical reasoning, which should further increase the rates at which participants solve the problem. A figure outlining the procedure of Study 2 is presented below (see Figure 18). The average time participants spent in this study was approximately 1 hour 20 minutes, for which they received course credit.

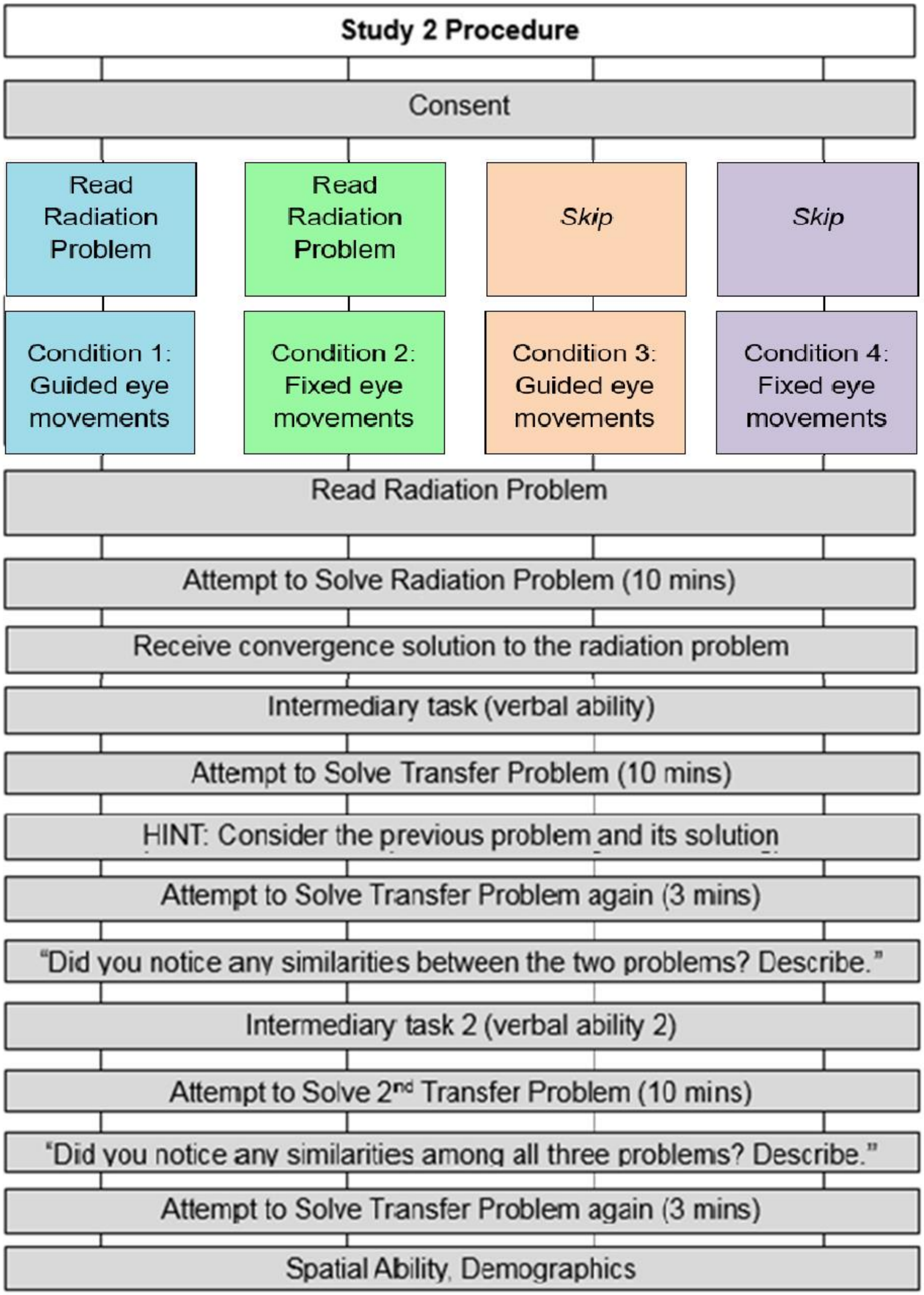

Figure 18. Outline of the procedure for study 2. Different conditions are denoted by differently colored boxes.

Results

For an overview of the results addressing each study's research questions, please refer to Appendix D.

Data Prescreening

Participants Removed from Analyses

To summarize the forthcoming sections, a total of 161 participants were retained for analysis. Five participants were removed because they saw an episode of Grey's Anatomy that featured a problem similar to the Radiation Problem, 12 participants were removed due to poor tracking task performance, and four were removed for christmas-treeing their responses to the spatial ability measures. The total number of participants excluded for Study 2 was 12 (some participants failed multiple pre-screening checks).

Grey's Anatomy

On March 1st, 2018, season 14 episode 13 of Grey's Anatomy was aired by the American Broadcast Company. In this episode, one of the characters developed a brain tumor. The doctors debated over how best to treat it using lasers. They initially proposed using a single laser to kill the tumor but were worried about damaging the patient's brain tissue. Later in the episode, another doctor proposes using multiple weaker lasers to attack the tumor while sparing the brain tissue (Fremont, 2018). The events of this episode very clearly mirror the solution to the radiation problem, and this episode happened to be aired during the middle of data collection for Study 2. Several participants indicated that they were fans of Grey's Anatomy, but only 5 participants reported seeing the problem on that specific episode of Grey's Anatomy. Those 5

participants were excluded from all analyses, and an additional logistic regression was performed on the remaining participants to assess whether viewers of Grey's Anatomy fared any better on solving the radiation problem than non-viewers. This regression was not statistically significant, $\chi^2(1, N = 140) = 1.20$, $p = .27$. Note that 21 cases were not included in this analysis because their data were collected prior to March 1st, 2018, and they were not asked whether they watched Grey's Anatomy.

Tracking Task Reaction Times

Before analyzing Study 2's data, participants' performance on the tracking task was assessed. Only those whose tracking accuracy was 75% or greater were retained for analyses. Accuracy was determined by excluding participants whose reaction times were greater than three standard deviations above the mean value, which was roughly 1000 milliseconds after the stimulus was presented. Participants were also excluded if their reaction was recorded prior to the onset of the stimulus. Twelve participants were removed from analyses based on these criteria.

A two-way ANOVA was performed using eye movement guidance and problem context as the predictors of tracking task reaction time. The overall model was not statistically significant ($F(3,157) = 2.55$, $p = .058$, $\eta^2_p = .047$). Digit tracking task means and standard deviations are provided in Table 6. These tracking times are similar to those found in previous studies which had participants perform a similar tracking task (Thomas & Lleras, 2007; 2009).

Table 6. Tracking Task Reaction Times (Means and Standard Deviations) for Study 2

Condition	M	SD
Problem First, Guided	633ms	161ms
Skip to Guided	627ms	156ms
Problem First, Fixed	565ms	160ms
Skip to Fixed	559ms	153ms

Coding the Convergence Solution

As in Study 1, participants' written responses were coded following the coding scheme in Table 3. However, participants were required to draw their solutions on scratch paper to more closely match the methodology of Thomas & Lleras (2007; 2009), where their participants traced their solution to the radiation problem on a diagram of the radiation problem. These drawings were coded using the same scoring criteria as in Thomas & Lleras (2007; 2009). Participants' scratch paper drawings were considered correct if they depicted two or more lines converging on a central point. Participants were judged as having solved the problem with the convergence solution if either of their written or sketched responses met the criteria described above (for drawings) and in Appendix B (for written solutions). Examples of sketches deemed acceptable according to this coding scheme are provided below in Figure 19.

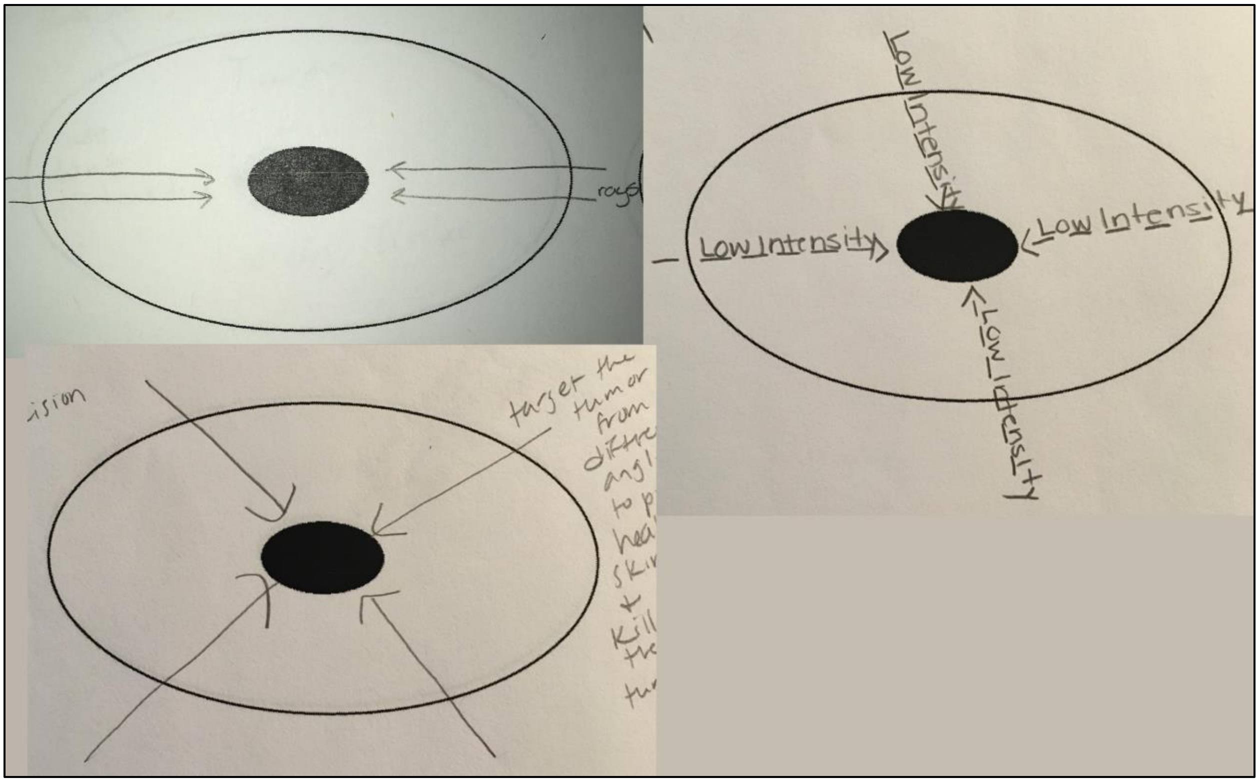

Figure 19. Examples of drawings that would be coded as having satisfied the requirements for the convergence solution as per Thomas & Lleras's (2007; 2009) guidelines.

Hypothesis Testing

To assess the hypotheses pertinent to Study 2, a logistic regression analysis was performed to predict the likelihood of a participant providing the convergence solution based on when they saw the problem (either before or after digit-tracking), and what kind of digit-tracking task they performed (either with embodied eye movements or fixed eye movements). To first evaluate these effects independently, two logistic regressions were performed with each main variable (problem context or eye movement guidance) as a singular predictor. The results were not significant for problem context (χ^2 (1, N =161) = 0.46, p = .50, Nagelkerke R^2 = .004), but they were significant for eye movement guidance (χ^2 (1, N =161) = 5.86, p = .016, Nagelkerke R^2

= .05). The significant result for eye movement guidance favored the guided eye movements conditions, such that those whose eye movements were guided in a pattern that embodied the convergence solution were twice as likely to solve the radiation problem than those who were in the fixed eye movements condition (Wald = 5.774, p = .017, Exp(B) = 2.19, Observed Power = .96).

To test whether there was an interacting effect between the two experimental variables, a logistic regression was performed in two blocks, with each variable (problem presentation and eye movement guidance) entered in the first block and their product term entered in the second block. In the first block, the overall model was statistically significant (χ^2 (2, N =161) = 6.27, p = .043, Nagelkerke R^2 = .052), and the variable for eye movement guidance did significantly improve prediction odds (Wald = 4.81, p = .028, Exp(B) = 2.1), indicating that those whose digit-tracking task guided their eye movements in a pattern physically similar to the convergence solution had roughly 2 times greater odds of generating that solution to the problem when controlling for problem presentation. In the second block, the overall model was not statistically significant χ^2 (3, N =161) = 6.37, p = .095, Nagelkerke R^2 = .05, and adding the product term of the conditions did not result in a significant increase of variance explained ($\Delta\chi^2$ (1, N =161) = 0.10, p = .76, Nagelkerke ΔR^2 = .00). Regression coefficients are provided in Table 7.

Table 7. Logistic Regression Coefficients for Experimental Variables Predicting Solution Rates for the Radiation Problem.

	Variable	B	S.E.	Wald	Exp(B)
Block 1	Constant	-0.89	0.29	9.50**	0.41
	Eye movement Guidance	0.78	0.33	5.70*	2.19
	Problem Context	0.21	0.33	0.42	1.24
Block 2	Constant	-0.84	0.33	6.34*	0.43
	Eye movement Guidance	0.68	0.46	2.18	1.98
	Problem Context	0.11	0.47	0.05	1.11
	Guidance*Context	0.20	0.66	0.10	1.23

Note. *$p < .05$; **$p < .01$. Note that Guidance was coded 0 = Fixed eye movements, 1 = guided eye movements. Problem Context was coded 0 = skip straight to tracking task, 1 = read problem before tracking task.

Two of the conditions in this study were nearly identical to the conditions in Thomas and Lleras's (2007; 2009) studies. Although there were slight changes to their methodology to allow other research questions to be explored in the present study, it is worth comparing the effects of their nearly-equal conditions to those in the present study. These two conditions of interest were the Perceptual Simulation Group and the Simple Incubation group in the present work. A logistic regression revealed that those in the Perceptual Simulation group were significantly more likely to produce a solution to the radiation problem than the Simple Incubation group ($\chi^2(1, N = 79) = 5.124$, Nagelkerke $R^2 = .11$, p = .024), and they had roughly three times greater odds of answering the problem correctly (Wald = 4.91, $p = .027$, Exp(B) = 3.35, Observed Power = .97). The comparison between these two groups appears to replicate the embodied effect found in previous work in terms of effect size and direction (Grant and Spivey, 2003; Thomas and Lleras, 2007; 2009).

The prior analyses do not statistically support an incubation effect increasing solution success rates, and they do not support any one condition having a greater success rate when

compared against all others. However, when examining the eye movement guidance variable alone, those who performed guided digit-tracking tasks performed significantly better than those who performed the fixed digit-tracking task. These results fail to support the perceptual simulation hypothesis but lend provisional support to the attentional priming hypothesis. In the next section, spatial ability variables are analyzed to test these hypotheses further.

Spatial Ability

Prior to including spatial ability as a covariate or interacting variable in any of the previously analyzed logistic regression models, it is prudent to examine the independent effects of spatial ability. Neither spatial measure alone significantly predicted a greater likelihood of solving the radiation problem (Paper folding test χ^2 (1, N =161) = 0.53, p = .47, Nagelkerke R^2 = .01; Card Rotations Test χ^2 (1, N =161) = 0.01, p = .94, Nagelkerke R^2 = .00). Neither of these results supports the spatial ability hypothesis. These models were also tested with Red Adair, where PFT performance was a significant predictor of success, but CRT performance was not (PFT χ^2 (1, N =158) = 4.20, p = .041, Nagelkerke R^2 = .04; Card Rotations Test χ^2 (1, N =158) = 0.01, p = .91, Nagelkerke R^2 = .00). Lastly, these models were tested with The General, where PFT performance was a significant predictor of success, but CRT performance was not (PFT χ^2 (1, N =161) = 3.97, p = .046, Nagelkerke R^2 = .05; Card Rotations Test χ^2 (1, N =161) = 0.74, p = .39, Nagelkerke R^2 = .01). These results lend partial support to Hypothesis 3, that those with higher spatial ability will tend to perform better on each problem.

Following the structure of the previously proposed analysis plan in Table 5, two more logistic regression analyses (one for each spatial ability measure) were performed in three blocks for each of the problems. For this analysis with the paper folding test, the model was not

significant at any of the blocks (final model χ^2 (7, N =161) = 9.32, p = .23, Nagelkerke R^2 = .08). When running the same model with the card rotations task as the spatial ability measure, this model exhibited similar null effects (final model χ^2 (7, N =161) = 9.254, p = .24, Nagelkerke R^2 = .08).

The previous models were also tested with solution rates for Red Adair. Both models exhibited similar null effects as the analyses with the radiation problem (PFT final model χ^2 (7, N =158) = 10.14, p = .18, Nagelkerke R^2 = .09; CRT final model χ^2 (7, N =158) = 6.34, p = .50, Nagelkerke R^2 = .06). Lastly, these models were tested with solution rates for The General. Both models replicated the null effects of both previous problems (PFT final model χ^2 (7, N =161) = 10.84, p = .15, Nagelkerke R^2 = .13; CRT final model χ^2 (7, N =161) = 7.87, p = .34, Nagelkerke R^2 = .10). As a whole, these analyses fail to support Hypotheses 3a and 3b, showing that spatial ability and its interactions did not have much impact on problem solving performance in the context of the experimental variables.

Noticing the Analogy

In Study 2, participants were also asked if they noticed similarities among the problems at two separate timepoints in the study. This was intended to facilitate the analogical transfer process by prompting participants to assimilate multiple solutions together. However, it also offers a means of assessing spatial ability hypotheses. Due to the spatial nature of the analogical solutions, those with higher spatial ability should see the structural and spatial similarities between analogues more readily. Also, participants with higher verbal ability should be more likely to recognize the similarities between the problem (based on results from Corkill & Fager, 1995). To test this idea, forward logistic regressions were performed examining the predictive

strength of spatial and verbal ability measures at two timepoints: when participants were prompted to describe similarities between the Radiation Problem and Red Adair (after attempting Red Adair), as well as when participants were prompted to describe similarities among all three problems. Participants' responses were judged as correct if they referred to all problems having a similar underlying theme of combining multiple weaker forces into one larger force.

Fifty-two out of 161 participants noticed the convergence analogy between the Radiation Problem and Red Adair, but the logistic regression analysis did not support that any of the spatial or verbal ability measures were significant predictors of noticing the convergence analogy between these two problems. Ninety-seven out of 161 participants noticed the convergence analogy across all three problems, but as before, the logistic regression analysis did not support that any of the spatial or verbal measures were significant predictors for noticing the convergence analogy among all three problems.

Transferring to Analogical Problems

Prior analyses for Study 2 showed a significant effect of digit-tracking eye movement guidance, such that embodied eye movements improved the rate at which a participant would be likely to solve the radiation problem. An additional aim of Study 2 was to test whether this embodied effect transferred when problem solvers attempted to solve analogically similar problems. Although the solution rates were significantly different for the radiation problem, solution rates were roughly equivalent for Red Adair ($\chi^2(3, N = 161) = 1.865, p = .601$) and The General ($\chi^2(3, N = 161) = 5.186, p = .159$). The solution rates for each group are depicted graphically in Figure 20.

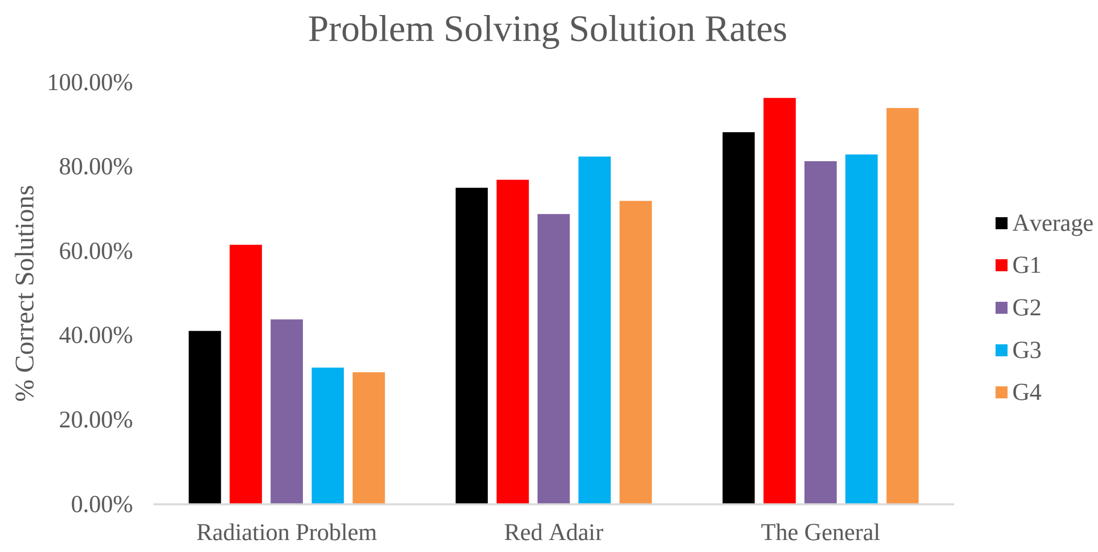

Figure 20. Problem solving solution rates for each group and each problem.

The radiation problem was the initial problem, attempted after the digit-tracking task. Red Adair and The General were the two transfer problems. Note: G1 = Problem first + guided eye movements, G2 = Skip to guided eye movements, G3 = Problem first + fixed eye movements, G4 = Skip to fixed eye movements

To assess whether embodied priming effects transferred into analogical problem solving, several conditional analyses were conducted for all four groups for success rates for each problem dependent on every combination of problems solved. For example, someone who does not solve the radiation problem, but was in an embodied eye movements group, should have better performance on later problems if the embodied effect aids in transfer as hypothesized. Additionally, those who have more time to think about the problem may have a better likelihood of constructing the convergence schema to transfer to later problems, evincing a benefit for incubation. To evaluate this idea, problem solving success rates were compared across all four groups with all possible combinations of solution patterns analyzed. In all cases, there were no significant differences among the groups (all *p*s > .10). Solution rates and chi square tests for

each of the four groups is provided in a bracket format in Figure 21. In no case did any one group appear to perform better than any other group, suggesting that there were no benefits for eye movement guidance or incubation, or their interaction.

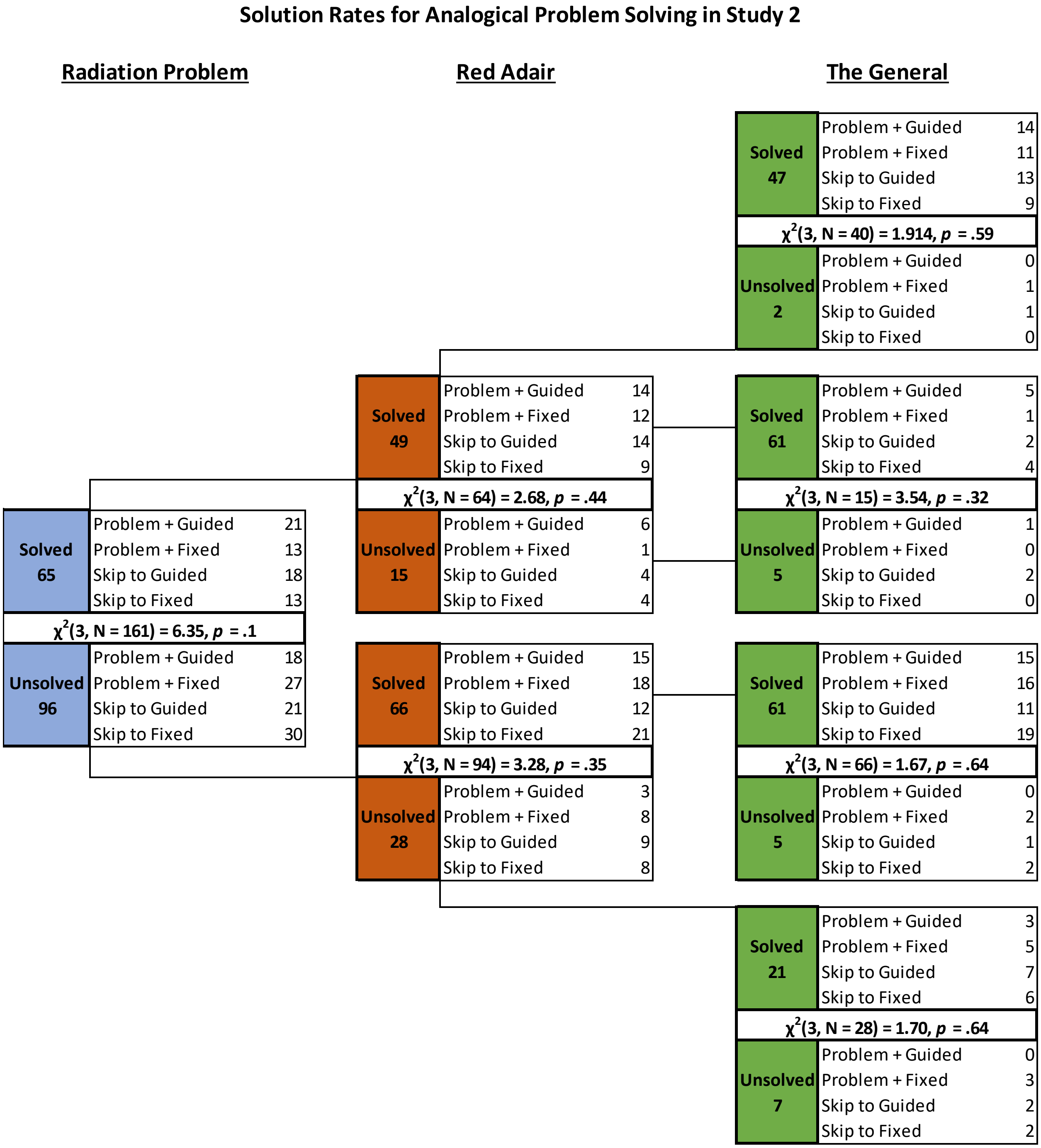

Figure 21. Solution rates for the Radiation Problem and both transfer problems for each combination of solution rates. No group had an advantage over any other group for each problem. Total N = 161, but 3 participants did not attempt problem 2; therefore, N = 158 for results in Red Adair and The General.

Creativity in Written Solutions

To assess whether eye movement guidance or problem presentation context resulted in differences identifiable through creativity variables, five 2 (Eye movement guidance) x 2 (Problem Context) factorial ANOVAs were performed on scores of Originality, Flexibility, and Elaboration, Total Solutions, and Redundant Solutions (see Appendix E for a MANOVA of these analyses). For Originality, there were no significant differences among any of the variables or their interactions, $F(3, 157) = 0.87$, $p = .32$, $\eta^2_p = .022$ (see Figure 22).

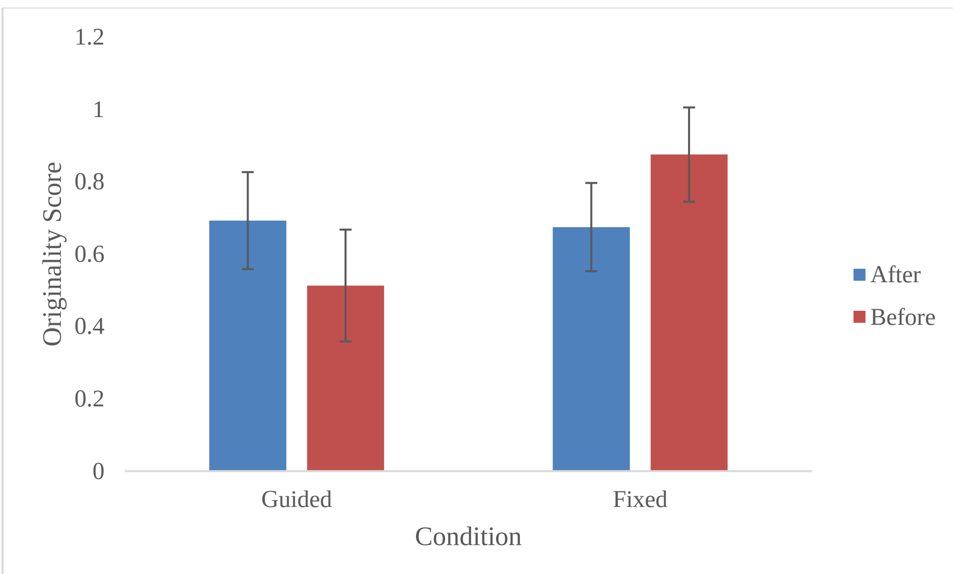

Figure 22. Originality scores for each condition. Error bars represent the standard error of the mean. There were no significant differences among the groups.

For Flexibility, there also were no significant differences among any of the variables or their interactions, $F(3, 157) = 1.95$, $p = .12$, $\eta^2_p = .036$ (see Figure 23).

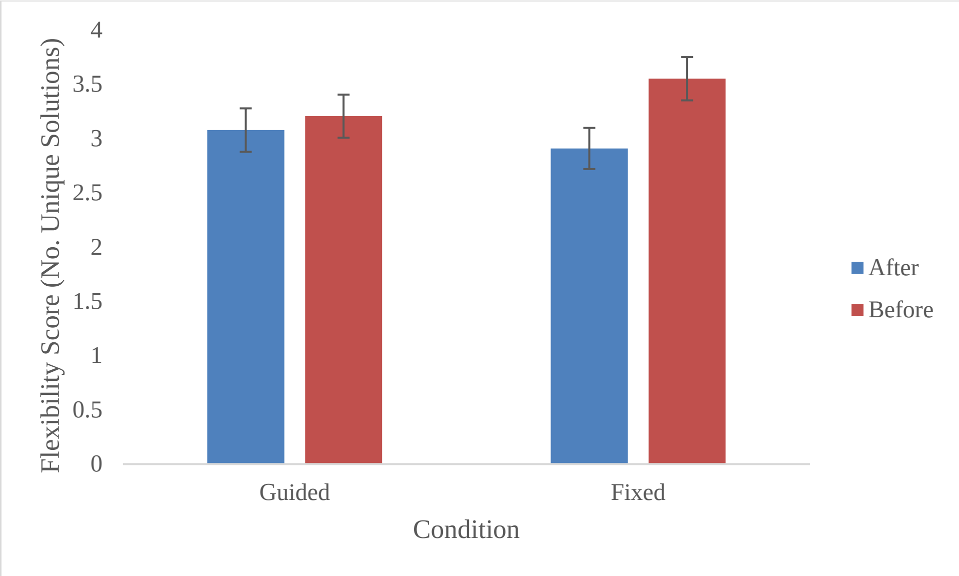

Figure 23. Flexibility scores for each condition. Error bars represent the standard error of the mean. There were no significant differences among the groups.

For Elaboration, there was a significant difference attributable to both main effects, but not their interaction, $F(3, 157) = 6.27$, $p < .001$, $\eta^2_p = .11$. The main effect for eye movement guidance favored those who received the guided digit-tracking task, whose solutions contained significantly more detail (character length $M = 516.91$, $SD = 276.38$) than those who received the fixed digit-tracking task ($M = 403.51$, $SD = 214.72$; $F(1, 157) = 8.52$, $p = .004$, $\eta^2_p = .05$). The main effect for problem context favored those who saw the problem before their digit-tracking task, whose solutions contained significantly more detail (character length $M = 517.92$, $SD = 266.38$) than those who received the fixed digit-tracking task ($M = 401.15$, $SD = 224.87$; $F(1, 157) = 8.99$, $p = .003$, $\eta^2_p = .05$; see Figure 24). These results lend partial support to the attentional priming hypothesis, as well as the incubation hypothesis.

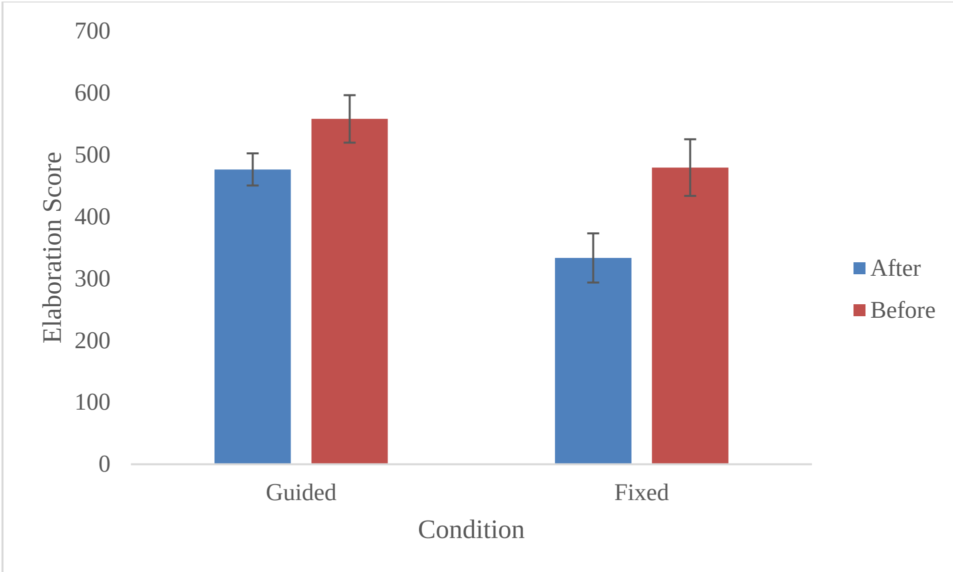

Figure 24. Elaboration score each condition. Error bars represent the standard error of the mean. The guided condition had significantly higher elaboration scores than the fixed condition, and those who saw the problem before their tracking task had greater elaboration scores than those who only saw the problem after their tracking task.

For Total Solutions, the overall ANOVA model was not significant, $F(3, 157) = 2.24$, $p = .09$, $\eta^2_p = .041$ (see Figure 25).

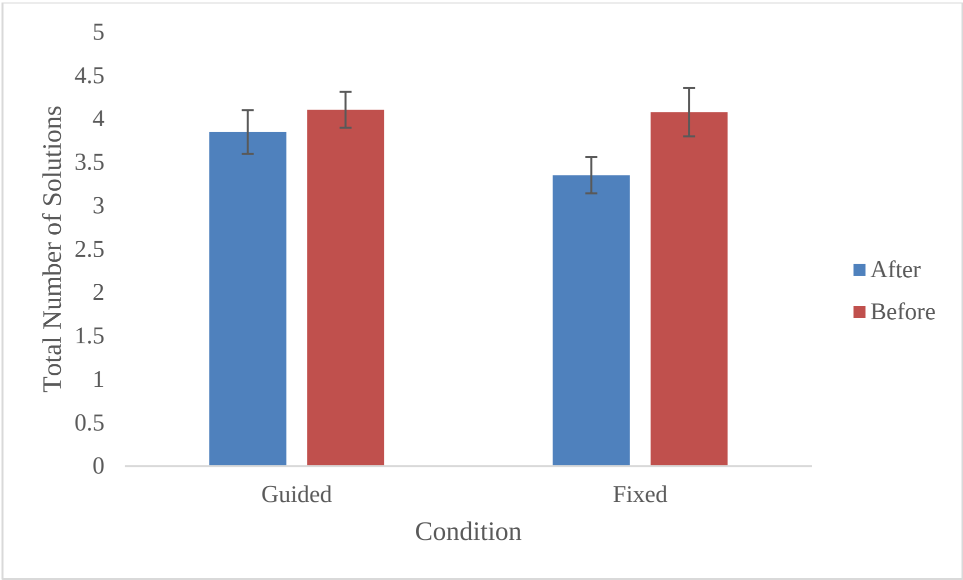

Figure 25. Total Solutions provided for each condition. Error bars represent the standard error of the mean. Those who saw the problem before their tracking task generated significantly more solutions than those who only saw the problem after their tracking task.

For Redundant Solutions, the overall ANOVA model was not significant, $F(3, 157) = 1.54$, $p = .21$, $\eta^2_p = .028$ (see Figure 26).

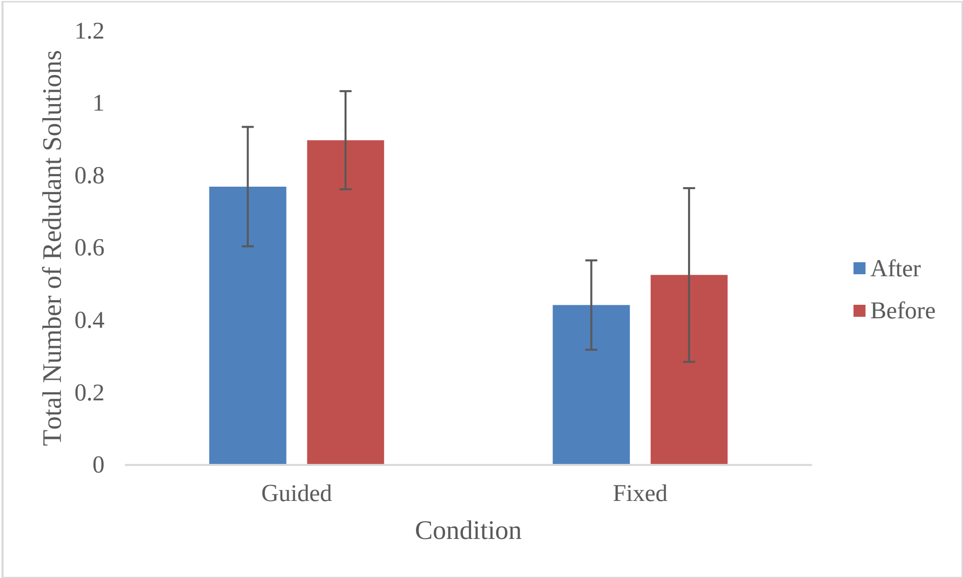

Figure 26. Redundant Solutions provided for each condition. Error bars represent the standard error of the mean. Those who in the guided condition generated significantly more redundant solutions than those in the fixed condition.

Mediating Effects

Several individual difference and performance variables were collected that may relate to problem solving success. If any of these variables shows a relationship with problem solving success, it is possible that the effects of the experimental manipulations could be better explained by a confounding variable. To assess this possibility, several independent-samples *t*-tests were conducted to examine whether any of the following variables were higher among those who solved the problem and those who did not: Uniqueness, Number of Unique Solutions, Elaboration, Total Number of Solutions, Age, Gender, Class Rank, Self-reports of liking to solve puzzles, whether a participant's major was a STEM (Science, Technology, Engineering, Mathematics) major, whether a participant's major was medical related, Verbal Performance score (word beginnings, word endings, and vocabulary), and Spatial Performance score (on the PFT and the CRT).

The results of the *t*-tests are provided in Table 8. Four variables were identified that were significantly different between those who solved the problem and those who did not: Elaboration, Age, STEM Major, and self-reports of liking to solve puzzles.

Table 8. Means, Standard Deviations, *t*-test Results, and Cohen's *d* Effect Sizes for Variables of Interest Between Those Who Solved the Radiation Problem and Those Who Did Not.

	Solved Radiation Problem?	*M*	*SD*	*t(159)*	*d*
Originality	No (n = 96)	0.79	0.94	1.95[a]	0.30
	Yes (n = 65)	0.54	0.71		
Flexibility	No	3.26	1.32	0.98	0.16
	Yes	3.06	1.17		
Elaboration	No	417.04	245.28	-2.58*	0.41
	Yes	519.60	251.75		
Total No. Solutions	No	3.86	1.55	0.33	0.05
	Yes	3.78	1.49		
Redundant Solutions	No	0.60	0.91	-0.68	0.10
	Yes	0.72	1.32		
Age	No	19.92	2.23	2.42[a]*	0.37
	Yes	19.22	1.42		
Gender	No	1.41	0.49	-0.69	0.11
	Yes	1.46	0.50		
Class Rank	No	2.76	1.05	0.78	0.13
	Yes	2.63	0.99		
"I like solving puzzles"	No	3.51	0.98	-2.34*	0.37
	Yes	3.89	1.06		
STEM Major	No	0.66	0.48	-2.06*	0.33
	Yes	0.80	0.40		
Med. Related Major	No	0.30	0.46	-0.74	0.13
	Yes	0.36	0.48		
Word Beginnings Score	No	14.58	5.28	-1.60	0.25
	Yes	16.02	6.00		
Word Endings Score	No	16.13	4.70	-1.61	0.26
	Yes	17.38	5.12		
Vocabulary Score	No	7.42	3.20	-0.70	0.11
	Yes	7.79	3.34		
CRT Score	No	52.82	19.07	0.31	0.05
	Yes	51.89	18.76		
PFT Score	No	9.88	4.15	-0.54	0.09
	Yes	10.25	4.53		

Note. *p < .05, [a]adjusted *df* were used as Levene's test indicated a violation of the equality of variance assumption. *d* = Cohen's *d* effect size (provided as absolute values). Significance of statistical tests was determined based on Bonferroni corrections for familywise error rate in each family of tested variables, where appropriate (i.e., creativity variables, verbal ability variables, spatial ability variables).

To further test whether these significantly different variables confounded the main effects observed before, a mediation analysis was conducted for each variable. Because eye movement guidance was the only significant experimental predictor of solution rates in the radiation

problem, it is included as the predictor in each model. Indirect effects were assessed using Hayes's PROCESS macro version 2 for SPSS (Hayes, 2012), which generated bias corrected and accelerated bootstrap confidence intervals using 5,000 samples. Logistic regression models were tested by entering eye movement guidance in the first block, and the mediating variable of interest in the second block.

A mediation model where elaboration mediates the relationship between eye movement guidance and problem success rates tests the idea that eye movement guidance engenders a creativity boosting effect, which may incidentally increase solution rates if participants are primed to generate more unique solutions. The logistic regression model was significant in block 1 (χ^2 (1, N = 161) = 5.86, p = .016, Nagelkerke R^2 = .05), with eye movement guidance as a significant predictor for solution rates. The model was also significant in block 2 after the addition of the elaboration variable (χ^2 (2, N = 161) = 10.15, p = .006, Nagelkerke R^2 = .08), with elaboration as a significant predictor for solution rates, but eye movement guidance no longer a significant predictor. The indirect effect of condition on problem solving success was estimated at B = 0.16, 95% CI (.02, .44). Because the 95% confidence interval did not include zero, this effect was judged significant at $p < .05$. This mediation model suggests that the improvement in solution rates due to eye movement guidance is better explained by a concomitant increase in creativity (as measured through elaboration). Logistic regression coefficients are provided below in Table 9, and a diagram illustrating this relationship is depicted in Figure 27.

Table 9. Logistic Regression Coefficients for Model Testing Elaboration as Mediator

	Variable	B	S.E.	Wald	Exp(B)
Block 1	Constant	-0.79	0.24	11.00**	0.46
	Eye movement Guidance	0.79	0.33	5.74*	2.19
Block 2	Constant	-1.36	0.38	13.09**	0.26
	Eye movement Guidance	0.65	0.34	3.67	1.91
	Elaboration	0.14	0.07	4.08*	1.15

Note. *$p < .05$; **$p < .01$. To simplify regression coefficient estimation, Elaboration was calculated in units of 100.

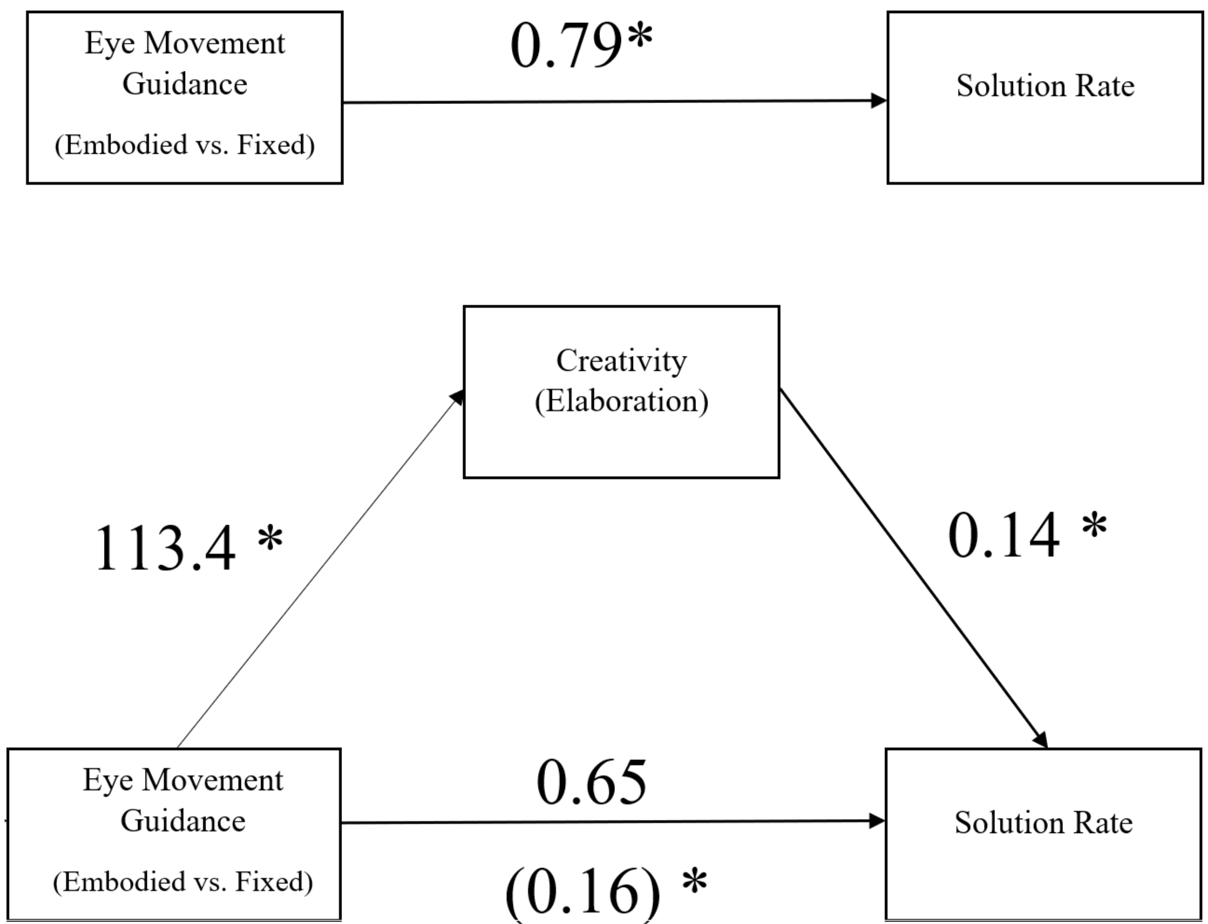

Figure 27. Unstandardized regression coefficients for the model of eye movement guidance predicting solution rate, mediated by creativity. The top shows the total effect of eye movement guidance on solution rate. The bottom shows the mediation model, with the indirect effect in parentheses and the direct effect above it. The indirect effect was estimated using 5,000 sample bias-corrected and accelerated bootstrapping procedure. *$p < .05$.

A mediation model where age mediates the relationship between eye movement guidance and problem success rates tests the idea that the problem solving benefits of guiding eye movements only arose due to an age difference between the fixed and guided eye movements conditions. Both eye movement guidance and age independently predicted variance in problem solving success (χ^2 (2, N = 161) = 11.26, p = .004, Nagelkerke R^2 = .09), but no indirect effect was supported (95% CI: -0.23, 0.12). It should be noted that age was not different by condition, but it was a significant negative predictor of problem solving success rates on the radiation problem, such that younger participants were more likely to generate the convergence solution to the radiation problem (see Table 10).

Table 10. Logistic Regression Coefficients for Model Testing Age as Mediator

	Variable	B	S.E.	Wald	Exp(B)
Block 1	Constant	-0.79	0.24	11.00**	0.46
	Eye movement Guidance	0.76	0.33	5.34*	2.14
Block 2	Constant	3.66	2.01	3.31	38.79
	Eye movement Guidance	0.80	0.34	5.78*	2.24
	Age	-0.23	0.10	4.87*	0.80

Note. *p < .05; **p < .01.

Next, a mediation model where STEM Major mediates the relationship between eye movement guidance and problem solving success rates tests the idea that students in STEM majors are more likely to solve the problem, regardless of how their eyes are guided during the digit-tracking task. The overall model was significant in block 1 (χ^2 (1, N = 161) = 5.86, p = .016, Nagelkerke R^2 = .05), as well as block 2 after the addition of STEM major as a variable (χ^2 (2, N = 161) = 9.70, p = .008, Nagelkerke R^2 = .08). However, STEM major was not a significant predictor of problem solving success when controlling for eye movement guidance,

indicating the presence of a mediating effect to be unlikely. It should be noted that Hayes's (2012) PROCESS macro cannot estimate indirect effects with dichotomous mediating variables, but the fact that eye movement guidance maintained its statistical significance from block 1 to block 2 leaves little evidence for the presence of a mediating effect of STEM major (see Table 11).

Table 11. Logistic Regression Coefficients for Model Testing STEM Major as Mediator

	Variable	B	S.E.	Wald	Exp(B)
Block 1	Constant	-0.79	0.24	11.00**	0.46
	Eye movement Guidance	0.79	0.33	5.74*	2.19
Block 2	Constant	-2.33	0.70	11.11	0.10
	Eye movement Guidance	0.78	0.33	5.55*	2.18
	STEM Major	0.74	0.38	3.67	2.08

Note. *$p < .05$; **$p < .01$.

Lastly, a mediation model where self-reports of liking to solve puzzles mediates the relationship between eye movement guidance and problem solving success rates tests the idea that the degree to which someone likes solving puzzles explains more of the difference in solution rates than eye movement guidance. The overall model was significant in block 1 (χ^2 (1, N = 161) = 5.86, p = .016, Nagelkerke R^2 = .05), as well as block 2 after the addition of their self-reports as a variable (χ^2 (2, N = 161) = 11.99, p = .002, Nagelkerke R^2 = .10). Like the previous mediation analysis with age, both variables uniquely predicted variance in problem solving success rates, but no mediating effect was supported (95% CI: -0.21, 0.11; see Table 12).

Table 12. Logistic Regression Coefficients for Model Testing Self-Reports of Liking to Solve Puzzles as Mediator

	Variable	B	S.E.	Wald	Exp(B)
Block 1	Constant	-0.79	0.24	11.00**	0.46
	Eye movement Guidance	0.79	0.33	5.74*	2.19
Block 2	Constant	-2.33	0.70	11.11	0.10
	Eye movement Guidance	0.85	0.34	6.33*	2.33
	Self-Reports	0.41	0.17	5.78	1.51

Note. $*p < .05$; $**p < .01$. Self-Reports of the degree to which one likes solving puzzles was rated on a Likert-type scale of 1-5.

Discussion

In this section, I briefly describe results from Study 1 and Study 2 and describe how they impact the design of Study 3. In study 1, there were no group differences in solution rates for the radiation problem and no group differences in metrics for creativity. In short, the data from Study 1 yielded no results of significance. This prompted a methodological re-evaluation to ensure that Study 2 could more accurately test the embodied effect from previous studies. By requiring participants to draw their solutions, in addition to describing them, the embodied effect was successfully replicated in Study 2. Additionally, there were significant effects for creativity variables that lent support to the Incubation and Attentional Priming hypotheses.

Guiding Eye Movements

The results of Study 2 support the existence of the embodied effect found in previous studies (Grant & Spivey, 2003; Thomas & Lleras, 2007; 2009), as evidenced by the significantly greater solution likelihoods for those who were in the guided eye movement conditions compared to the fixed eye movement conditions. This effect was similar in magnitude from previous studies and would suggest that guiding eye movements aids in problem solving by

fostering a perceptual simulation of the convergence solution. However, solution rates appeared to be related to one creativity variable that was significantly different between the guided and fixed eye movements conditions. Those in the guided eye movements had significantly higher Elaboration scores on their written solutions to the radiation problem, and Elaboration was higher in those who solved the problem compared to those who did not.

To ensure that the improved solution rates from eye movement guidance were uniquely due to guidance, and not a concomitant increase in creativity, the mediation analysis with Elaboration revealed that these effects were in fact highly related. Because eye movement guidance no longer significantly predicted solution rates after inclusion of the mediating variable, it can be argued that this creativity boosting effect mediates the benefits of eye movement guidance on performance. In broader terms, guiding eye movements improved creativity which led to an increase in solution rates for the Radiation problem. These effects lend evidence to support the Attentional Priming hypothesis, which states that guiding eye movements does not foster a perceptual simulation of the radiation problem but increases creativity that leads to better problem solving performance. Additionally, the suppressing effect of this mediating relationship provides evidence against the Perceptual Simulation hypothesis, which states that the improvements in problem solving success are due to guiding eye movements independent of any creativity effects. Lastly, it is worth mentioning that the effect of eye movement guidance was no longer significant after accounting for creativity, but this reduction in effect size was about 20% (Unstandardized coefficient from 0.79 to 0.65). Although the mediated direct effect was not deemed significant in statistical analyses, it may still represent meaningful variance in problem solving success explained from eye movements.

Incubation

The general results from the experimental variables did not support that an incubation effect aided in problem solving. Participants who saw their problem before the tracking task and then saw the problem again during the solution phase did not fare any better than participants who performed their tracking task and proceeded to the solution phase. These results suggest that there was no benefit of incubation for problem solving success rates. However, it should be noted that there was an incubation effect for creativity measures, which is discussed in the next section.

Examining Creativity

In Study 2, there were significant effects for creativity that supported the Incubation hypothesis, as well as the Attentional Priming and Flexibility hypotheses (partially). It is interesting to note that these effects did not emerge in Study 1, which had a few conditions that overlapped with Study 2. Despite nearly similar conditions, the stark difference in effects may have alternative explanations. Because participants were required to draw their solutions in Study 2, it is possible that this additional feature in the problem solving process may have engendered a more descriptive written approach to problem solving. These effects are revisited in Study 3.

Analogical Problem Solving

Although there was a noticeable difference in performance on the radiation problem, solution rates were practically equal for the two analogical transfer problems. This lack of differences in transfer problem performance does not support that embodied priming transfers to later analogical problem solving. Even when comparing solution rates based on prior solution

success, there still appears to be no benefit of embodied priming on analogical transfer. However, there are some methodological aspects of Study 2 that may have limited the ability to test whether embodied priming transfers to later analogical problems.

Specifically, participants were provided answers to each problem after they attempted to solve it. Having access to the answer (or an example of a convergence solution) is key to transferring the problem to an analogically related one (e.g., Gick & Holyoak, 1980; 1983), so it is possible that providing answers to participants may have clouded any embodied effects that might have aided in analogical transfer. Embodied theories of cognition would suggest that perceptual simulations should enhance schema formation, so those whose eye movements are guided should have a convergence schema that is more readily accessible for problem solving than those whose eye movements are not. This greater ease of access should manifest itself in greater solution rates for later problem solving. To test this possibility, participants did not receive solutions to any problem in Study 3.

The Importance of Drawing Solutions

Comparing the results of Study 1 and Study 2, it would appear that requiring participants to draw their solutions is key to replicating the embodied effect from prior research. However, the effect of drawing solutions raises new research questions related to embodied cognition. Although the data from study 2 mirrored the embodied problem solving effect from previous studies, the data do not fully confirm that it is exclusively an embodied effect. Aside from the potential mediating effects of creativity, the fact that participants in previous work drew their solutions on a diagram of the radiation problem (which their eyes were guided around) could be a cue or a demand characteristic that led them to draw the pattern they had been tracing with their eyes. One research question to be addressed in study 3 is whether this is a cue or demand characteristic – would problem solvers fare just as well by drawing their solutions on blank scratch paper? If this is an embodied effect as claimed in previous studies, eye movement guidance would add perceptual information into a problem solver's mental representation, leading them to be better able to solve the problem. If truly an embodied effect, it should be independent of the presence of a diagram when problem solvers draw their solutions.

Embodying the Solution to Improve Mental Representations

Beyond only scratch paper, there is an additional way of testing whether this effect is embodied that was assessed in Study 3. The analogical problem solving literature suggests that a robust mental representation is key to successfully solving problems (Chi et al., 1989; Chi & Glaser, 1985). Embodied theories of cognition would support that adding perceptual information (as could be conferred by guided eye movements) to a mental representation should make it more robust than what written information alone could engender. Thus, those whose eye

movements are guided in the form of the convergence solution should have a more robust mental representation that leads to a greater likelihood of analogical transfer for later problem solving.

In Study 2, problem solvers were significantly more likely to draw a convergence solution if their eye movements were guided in that pattern. However, when solving analogical problems thereafter, there were no significant group differences for drawn or written solutions. If this effect is embodied, it should have led to greater solution rates for analogical problems. But Study 2's design may have limited my ability to assess this perspective. In study 2, participants received solutions to each problem after their solution attempts. The analogical problem solving literature suggests that knowing the solution to a problem is a necessary prerequisite before one can successfully transfer it to an analogical problem. It is possible that providing the solution may have masked any potential benefits of "embodying" the solution through eye movements. Thus, for Study 3, no solutions were provided to participants to assess whether the embodied effect carries through to analogical problems.

CHAPTER FIVE: STUDY 3

Method

For study 3, the general procedure is identical to that of Study 2 with minor changes to the design. Instead of manipulating problem presentation context, this study sought to better understand the effects of scratch paper on solution rates. Specifically, scratch paper was manipulated in conjunction with eye movement guidance. In Study 2, the scratch paper had simple diagrams of the radiation problem on which participants drew their solutions, aligning with previous studies who had participants trace their solutions on diagrams of the radiation problem (Grant & Spivey, 2003; Thomas & Lleras, 2007; 2009). In Study 3, this variable was manipulated such that participants either received the same scratch paper (with diagrams), or they received blank scratch paper upon which they drew their solutions. This methodological change was implemented in Study 3 to better understand the effects of scratch paper on problem solving.

The other major methodological change is the removal of solution provision. To better estimate whether diagrams or eye movement guidance aided in spontaneous analogical transfer, participants no longer received solutions to the problems after attempting to solve them.

Participants

For the third study, 158 participants volunteered their time. The mean age was 19.13 years (*SD* = 3.02), and there were 59 males, 96 females, and 3 who preferred not to respond. The ethnic makeup was 47.6% Caucasian/White, 22.4% Hispanic, 10.5% Asian, 22.4% Black/African American, and 7% Other. Regarding major, 67.8% of participants in this sample were majoring in STEM (Science, Technology, Engineering, Mathematics) areas, and 31.4% of

participants were majoring in a medical-related concentration. The class rank makeup of the sample was 74.6% Freshmen, 9.9% Sophomores, 8.5% Juniors, 5.6% Seniors, and 1.4% 5th year or higher. Participants were randomly assigned to 1 of 4 groups, described below in Table 13. Although this sample size was smaller than the sample from Study 2, this smaller sample still provided sufficient statistical power (Power = .95) to detect an effect for eye movement guidance at the same level of significance as was observed in Study 2 (at $p = .016$)

Table 13. Groups and Sample Sizes in Study 3

	Diagrams on Scratch Paper	**Blank Scratch Paper**
Embodied Eye Movements	n = 37 **(1)** Guided + Diagrams	n = 34 **(2)** Guided + Blank
Fixed Eye Movements	n = 36 **(3)** Fixed + Diagrams	n = 36 **(4)** Fixed + Blank

Note. These sample sizes were calculated after removal of participants, described in the results section of this chapter.

Materials

All materials for Study 3 are identical to those in Study 2, with the addition of two different kinds of scratch paper for solving the radiation problem. These pieces of scratch paper were either blank, or contained the diagrams used in the radiation problem to assess whether the presence of diagrams influences problem solving success. Additionally, to better test research questions from an embodied perspective, participants no longer received solutions after they attempted to solve each problem, and they were no longer required to re-evaluate their solutions after their first attempt.

Design

To answer additional research questions about the influence of diagrams, embodied aid, and spatial ability, Study 3 had a 2 (diagram presence) x 2 (guided eye movements) completely between design, using spatial ability as a continuous moderating variable. Diagram presence had two levels: whether the participant draws their solutions on scratch paper using a diagram, or whether the participant draws their solutions on completely blank scratch paper when solving the radiation problem. Guided eye movements also had two levels as in Study 2.

Procedure

The procedure was largely the same as in Study 2, with a few differences. Participants were randomly assigned to one of four conditions and underwent all the same procedures as in Study 2, except they did not receive hints, and they did not receive solutions to the problems after they completed them. This change was made to better estimate whether embodied effects transfer to later analogical problem solving by eliminating the analogical transfer benefit from the provision of prior solutions. The average time participants spent in this study was approximately 1 hour 5 minutes, for which they received course credit.

Results

For an overview of the results addressing each study's research questions, please refer to Appendix D.

Data Prescreening

Participants Removed from Analyses

A total of 15 participants were removed from the forthcoming analyses. Five participants failed the criteria for tracking task accuracy (described in the next section), nine participants were removed for Christmas-treeing in the spatial and/or vocabulary measures, two participants were removed for having seen the radiation problem before, and one participant was removed for having seen the radiation problem on Grey's Anatomy. Note that these numbers are greater than the total of 15 because a few participants met multiple removal criteria.

Tracking Task Reaction Times

Before analyzing Study 3's data, participants' performance on the tracking task was assessed. Only those whose tracking accuracy was 75% or greater were retained for analyses. Accuracy was determined by excluding participants whose reaction times were greater than three standard deviations above the mean value, which was roughly 1000 milliseconds after the stimulus was presented. Participants were also excluded if their reaction was recorded prior to the onset of the stimulus. Five participants were removed from analyses based on these criteria.

A two-way ANOVA was performed using eye movement guidance and diagrams as the predictors of tracking task reaction time. The overall model was not statistically significant ($F(3,139) = 1.36$, $p = .144$, $\eta^2_p = .038$). Digit tracking task means and standard deviations are provided in Table 14. These tracking times are also like those found in previous studies which had participants perform a similar tracking task (Thomas & Lleras, 2007; 2009).

Table 14. Tracking Task Reaction Times (Means and Standard Deviations) for Study 3

Condition	M	SD
Guided, Blank	614ms	116ms
Guided, Diagrams	625ms	179ms
Fixed, Blank	564ms	143ms
Fixed, Diagrams	562ms	135ms

Hypothesis Testing

A logistic regression was employed to test the Embodiment hypothesis – comparing those whose eye movements were guided to those whose were not. This effect was not statistically significant, (χ^2 (1, N =143) = 0.00, p = .985, Nagelkerke R^2 = .00), suggesting that the guided digit-tracking task conferred no benefit to problem solving compared to the fixed digit-tracking task (Wald = 0.20, p = .655, Exp(B) = 1.16, Observed Power = .07). Next, a logistic regression was performed to test the Visual Aid hypothesis, which compares the effect of diagrams versus blank scratch paper on solution success rates. This effect was significant (χ^2 (1, N =143) = 6.48, p = .011, Nagelkerke R^2 = .06), such that those who drew their solutions on scratch paper containing diagrams were about two-and-a-half times more likely to have generated the convergence solution to the radiation problem (Wald = 6.29, p = .012, Exp(B) = 2.45, Observed Power = .99).

To test the demand characteristic hypothesis, a logistic regression combining the two previously analyzed variables and their product term was conducted. This hypothesis asserts that the embodied effect found in previous studies arose out of demand characteristics. By guiding one's eyes in a pattern tracing the convergence solution on the diagram which was to be sketched upon, solution rates should increase due to an apparent visual correspondence, rather than embodiment alone. This regression was tested in two blocks, with the main effects entered

simultaneously in the first block, and their product term entered in the second block. This hypothesis was not supported, but the overall model was significant due to a significant effect for diagrams (χ^2 (3, N =143) = 8.42, p = .038, Nagelkerke R^2 = .08). The effect of diagrams was the only significant predictor of convergence solution success. Regression coefficients for all variables are provided in Table 15.

Table 15. Logistic Regression Coefficients, Wald Statistics, and Odds Ratios for Demand Characteristic Hypothesis Test

	B	S.E.	Wald	Exp(B)
Guidance	0.55	0.54	1.03	1.73
Diagrams	1.40	0.52	7.23**	4.05
Diagrams*Guidance	-1.00	0.72	1.92	0.37
Constant	-1.29	0.40	10.40**	0.28

Note. All terms consumed 1 degree of freedom in the analysis. Guidance was coded 0 = fixed eye movements, 1 = guided eye movements. Diagrams was coded 0 = blank scratch paper, 1 = scratch paper with diagrams. *p < .05, **p < .01

Spatial Ability

As before, both spatial ability measures were entered into their own logistic regression model to predict solution success rates. Neither spatial measure alone significantly predicted a greater likelihood of solving the radiation problem (Paper folding test χ^2 (1, N =142) = 0.92, p = .34, Nagelkerke R^2 = .01; Card Rotations Test χ^2 (1, N =142) = 1.25, p = .26, Nagelkerke R^2 = .01). For Red Adair, the Paper Folding Test did not significantly predict a greater likelihood of success (χ^2 (1, N =142) = 2.19, p = .14, Nagelkerke R^2 = .02); However, performance on the Card Rotations Test was a significant positive predictor of success in solving Red Adair (χ^2 (1, N =142) = 6.25, p = .012, Nagelkerke R^2 = .06). Lastly, for The General problem, neither spatial measure significantly predicted a greater likelihood of success (Paper folding test χ^2 (1, N =142) = 0.02, p = .90, Nagelkerke R^2 = .00; Card Rotations Test χ^2 (1, N =142) = 0.61, p = .44,

Nagelkerke R^2 = .01). These results offer limited support for an effect of spatial ability, although it is worth noting that success on Red Adair was predicted by one measure of spatial ability.

As in the previous studies, higher-order analyses of spatial ability interacting with the experimental variables were performed next to assess Hypotheses 3a and 3b. For the radiation problem, no models were significant for the PFT or the CRT (PFT final model χ^2 (7, N =142) = 13.11, p = .07, Nagelkerke R^2 = .12; CRT final model χ^2 (7, N =142) = 11.07, p = .14, Nagelkerke R^2 = .10).

For Red Adair, the models with CRT were not statistically significant (CRT final model χ^2 (7, N =142) = 8.44, p = .30, Nagelkerke R^2 = .09), however, there was a significant model with the PFT. The model with all possible interactions was not statistically significant (PFT final model χ^2 (7, N =142) = 8.72, p = .27, Nagelkerke R^2 = .09), but there was a simpler model that evinced a significant interacting effect for PFT scores and the diagrams variable. This regression was not statistically significant at Block 1 (χ^2 (2, N =142) = 3.02, p = .22, Nagelkerke R^2 = .03), but in Block 2, there was a significant increase in variance explained as a function of adding the product term of Diagrams and PFT scores, ($\Delta\chi^2$ (1, N =142) = 4.61, p = .03, ΔNagelkerke R^2 = .05), yielding an overall model that was significant χ^2 (3, N =142) = 7.82, p = .05, Nagelkerke R^2 = .08). Regression coefficients for this analysis are provided in Table 16 below. To better understand the nature of this interacting effect, the conditional effect of PFT scores on Red Adair solution rates was tested at 1 SD below, Mean, and 1 SD above the mean values of PFT scores. This effect was only significant at higher levels of PFT scores (z = -2.12, p = .03, 95% CI: -2.67, -0.11). The nature of this effect was such that high levels of spatial ability predicted a greater likelihood of solving Red Adair, but only in the condition that did not receive diagrams see Figure 28 for a graphic depiction of these effects).

Table 16. Logistic Regression Model for Paper Folding Test Score and Diagrams Predicting Solution Rates for Red Adair

	Variable	B	S.E.	Wald	Exp(B)
Block 1	Constant	0.57	0.59	0.93	1.76
	PFT Score	0.08	0.05	2.05	1.08
	Diagrams	-0.36	0.40	0.83	0.70
Block 2	Constant	-0.75	0.85	0.79	0.47
	PFT Score	0.22	0.09	5.93*	1.25
	Diagrams	1.86	1.12	2.75	6.45
	Diagrams*Score	-0.24	0.12	4.36*	0.79

Note. *$p < .05$; **$p < .01$. PFT = Paper Folding Test. The diagrams variable was coded 0 = blank scratch paper, 1 = scratch paper with diagrams.

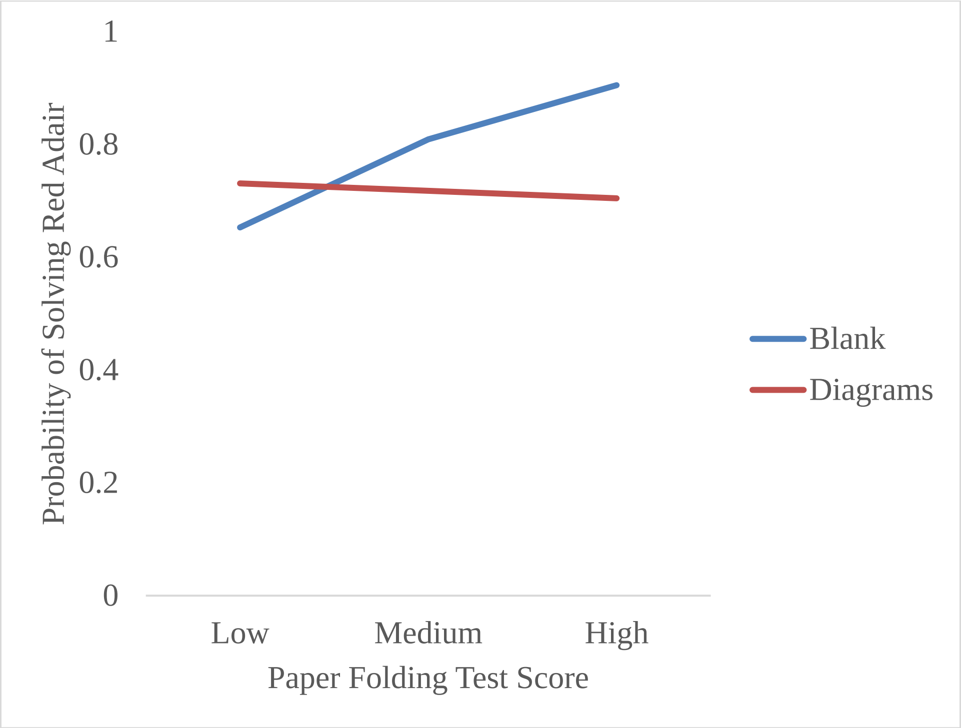

Figure 28. Slope estimates for probability of solving Red Adair as a function of scratch paper condition (either blank scratch paper or scratch paper with diagrams) and paper folding test score. The conditional difference in solution probability was significant at high paper folding test scores.

Lastly, these models were also tested for The General problem. Neither spatial measure was a significant variable for any tested models (PFT final model χ^2 (7, N =142) = 4.78, p = .69, Nagelkerke R^2 = .05; CRT final model χ^2 (7, N =142) = 1.59, p = .98, Nagelkerke R^2 = .02).

Noticing the Analogy

In Study 3, participants were not prompted to summarize similarities between the Radiation Problem and Red Adair. This was intentional to withhold from aiding in the analogical problem solving process and to better assess spontaneous analogical transfer. However, participants were asked if they noticed any similarities among all three problems after attempting the final problem. It was hypothesized that spatial ability and verbal ability may play a role in the formation of the convergence schema. To test this hypothesis, a logistic regression was performed on all measures of spatial ability and verbal ability using the forward method. Unfortunately, not one of them was statistically significant. Fifteen out of 139 participants recognized the common convergence schema across all three problems (3 participants did not attempt this question), so it is possible that there simply were not enough people who noticed the similarities to adequately make predictions with the cognitive ability measures.

Transferring to Analogical Problems

Prior analyses for Study 3 showed a significant effect of diagram presence, such that those who drew their solutions on diagrams with scratch paper would be more likely to solve the radiation problem. An additional aim of Study 3 was to continue to test whether this embodied effect transferred when problem solvers attempted to solve analogically similar problems. Since Study 3 did not provide answers or hints for any of the problems, a clearer test of whether embodied effects carry through to transfer problems can be performed. A priming task that is

supposed to facilitate a perceptual simulation of a problem's structure should enrich one's mental representation of the problem, which leads it to be a more accessible analogy to be transferred to structurally similar problems. Although the solution rates were significantly different for the radiation problem, solution rates were roughly equivalent for Red Adair ($\chi^2(3, N = 143) = 0.87$, p = .834) and The General ($\chi^2(3, N = 143) = 0.93$, $p = .817$). These data do not support that the embodied effect of eye movement guidance transferred to later analogical problems (Analogical Boost Hypothesis). The solution rates for each group are depicted graphically in Figure 29.

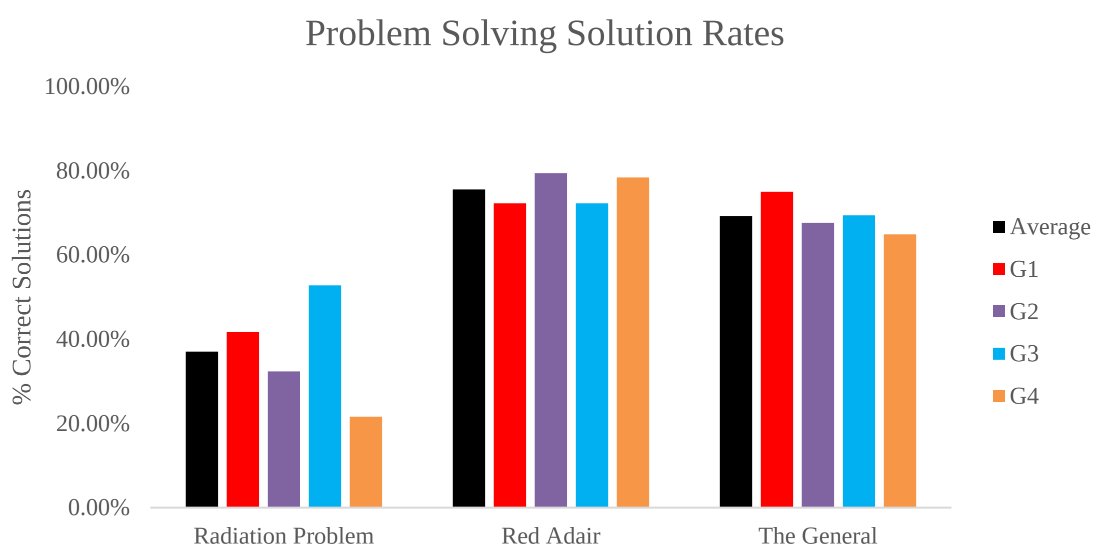

Figure 29. Problem solving solution rates for each group and each problem.

The radiation problem was the initial problem, attempted after the digit-tracking task. Red Adair and The General were the two transfer problems. Note: G1 = Guided + Diagrams, G2 = Guided + Blank, G3 = Fixed + Diagrams, G4 = Fixed + Blank.

As in Study 2, participants' solution rates were analyzed for every solution combination of the analogical problems. As previously mentioned, there was a significant difference for the

Radiation Problem; however, there were no significant differences among the groups at any stage of the problem solving process for any combination of solution patterns (see Figure 30).

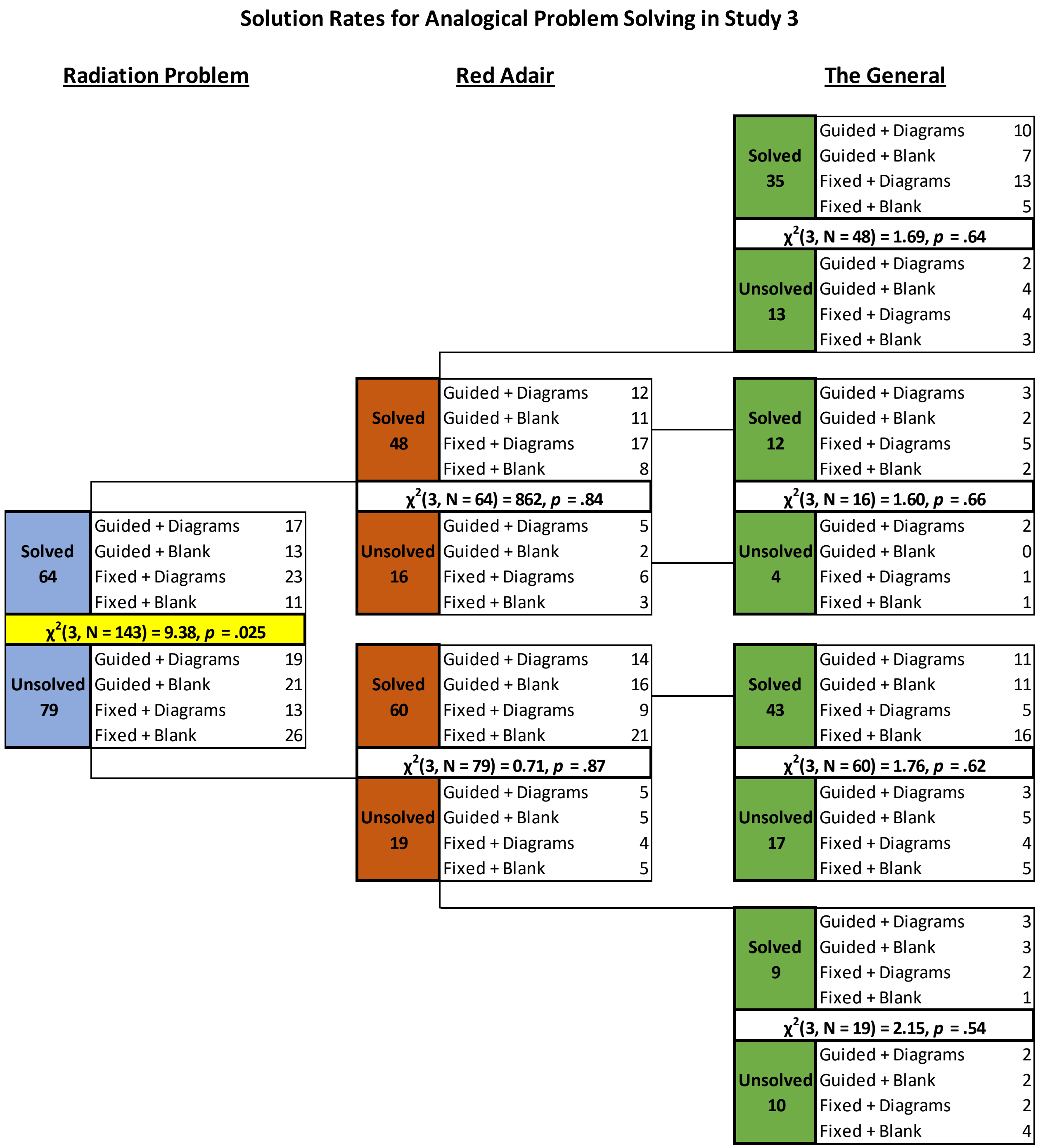

Figure 30. Solution rates for the Radiation Problem and both transfer problems for each combination of solution rates. Significant χ^2 tests are highlighted. No group had an advantage over any other group for each transfer problem. Total N = 143.

Creativity in Written Solutions

To assess whether eye movement guidance or diagrams resulted in differences identifiable through creativity variables, five 2 (Eye movement guidance) x 2 (Diagrams) factorial ANOVAs were performed on scores of Originality, Flexibility, and Elaboration, Total Solutions, and Redundant Solutions (see Appendix E for a MANOVA of these analyses). For Originality, the overall model was significant, $F(3, 139) = 5.03$, $p = .002$, $\eta^2_p = .098$, and it was driven by a significant main effect of diagrams, $F(1, 139) = 12.02$, $p = .001$, $\eta^2_p = .080$ (see Figure 31). In this case, those in the diagrams condition were more likely to have higher scores of Originality.

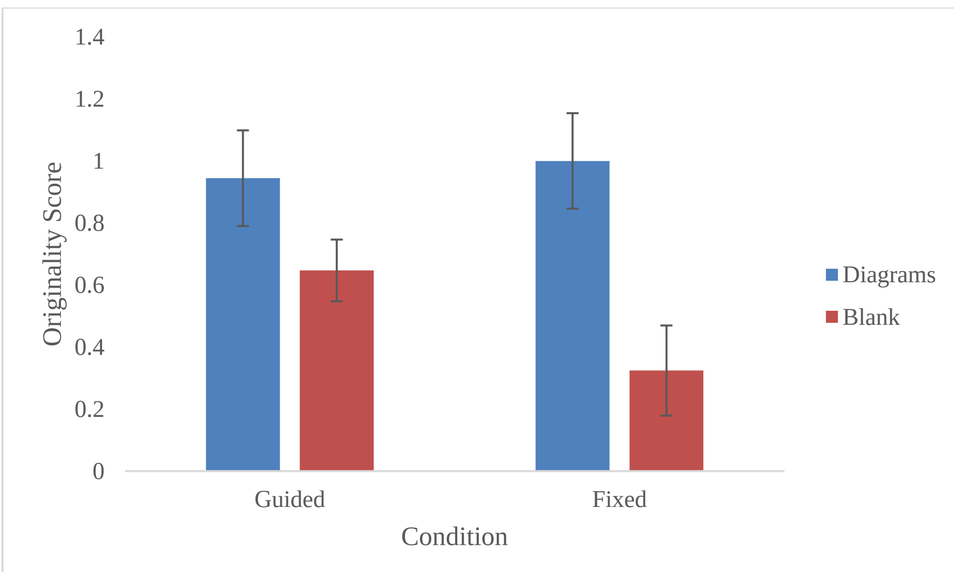

Figure 31. Originality scores for each condition. Error bars represent the standard error of the mean. Those in the scratch paper with diagrams conditions had significantly higher originality scores than those in the blank scratch paper conditions.

For Flexibility, the overall model was significant, $F(3, 139) = 2.67$, $p = .050$, $\eta^2_p = .054$, and it was driven by a significant main effect of diagrams, $F(1, 139) = 6.97$, $p = .009$, $\eta^2_p = .048$.

This effect was in favor of the scratch paper with diagrams condition, who showed higher Flexibility scores than those in the blank scratch paper conditions (see Figure 32).

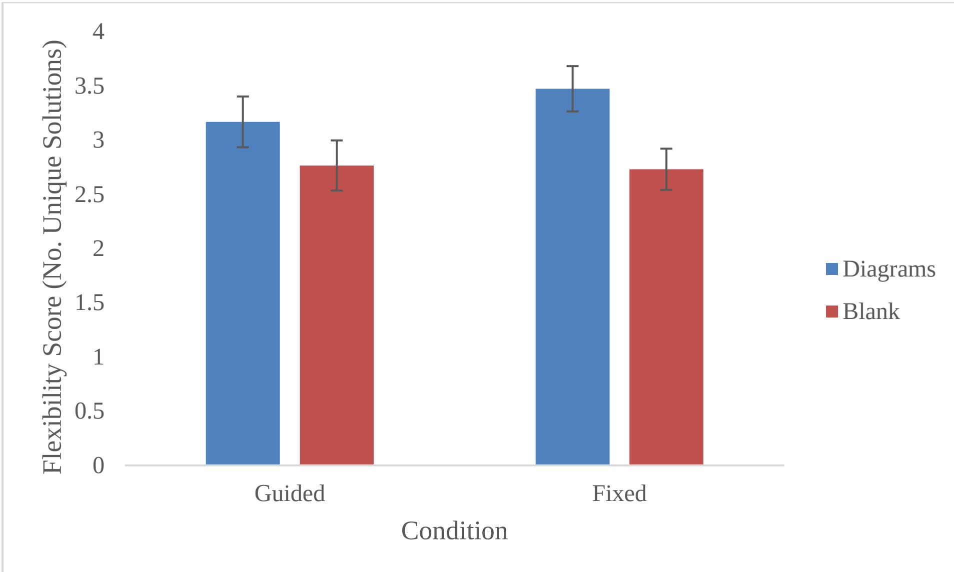

Figure 32. Flexibility scores for each condition. Error bars represent the standard error of the mean. The diagrams condition had significantly higher Flexibility scores than the blank scratch paper conditions.

For Elaboration, the overall model was not statistically significant ($F(3,139) = 1.66$, $p = .179$, $\eta^2_p = .035$), and there were no significant differences among the main effects or their interaction (see Figure 33).

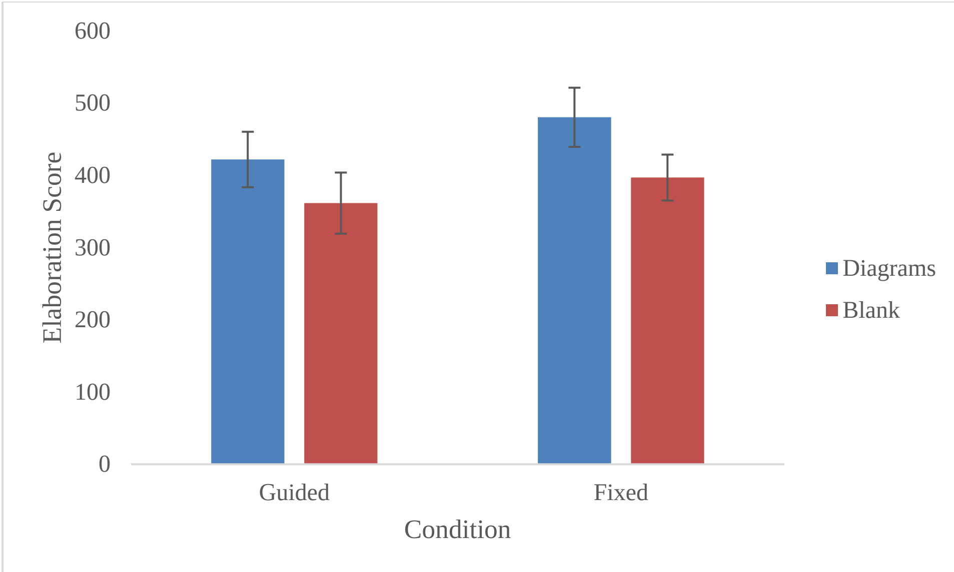

Figure 33. Elaboration score each condition. Error bars represent the standard error of the mean. There were no significant differences across the groups.

For Total Solutions, the overall ANOVA model was statistically significant, $F(3, 139) = 3.67$, $p = .014$, $\eta^2_p = .073$, and there was a significant main effect for Diagrams ($F(1, 139) = 10.87$, $p = .001$, $\eta^2_p = .073$). Those whose scratch paper contained diagrams ($M = 3.93$, $SD = 1.61$) generated significantly more solutions than those who only saw the problem after their tracking task ($M = 3.08$, $SD = 1.42$; see Figure 34).

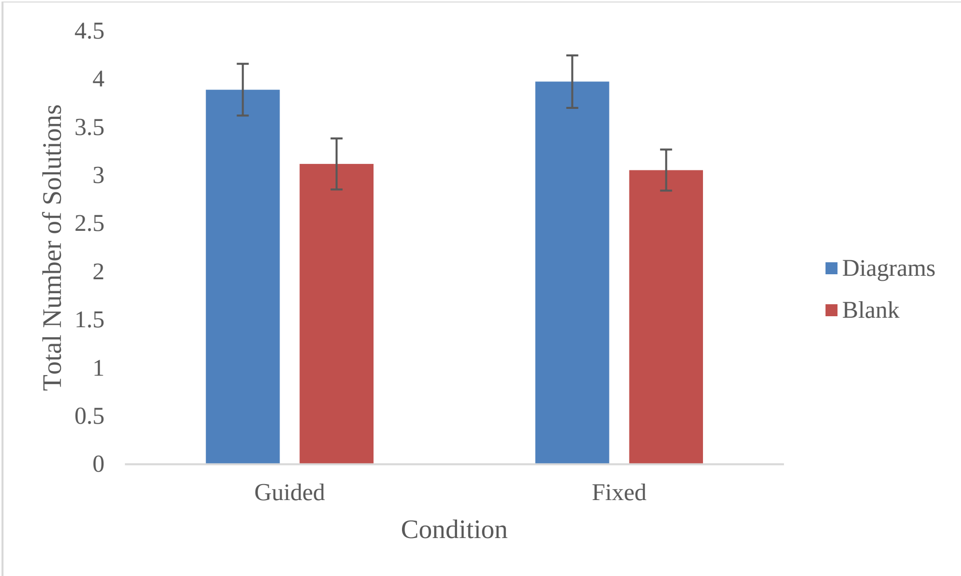

Figure 34. Total Solutions provided for each condition. Error bars represent the standard error of the mean. Those whose scratch paper contained diagrams generated significantly more solutions than those whose scratch paper was blank.

For Redundant Solutions, the overall ANOVA model was not significant, $F(3, 139) = 2.29$, $p = .081$, $\eta^2_p = .047$ (see Figure 35).

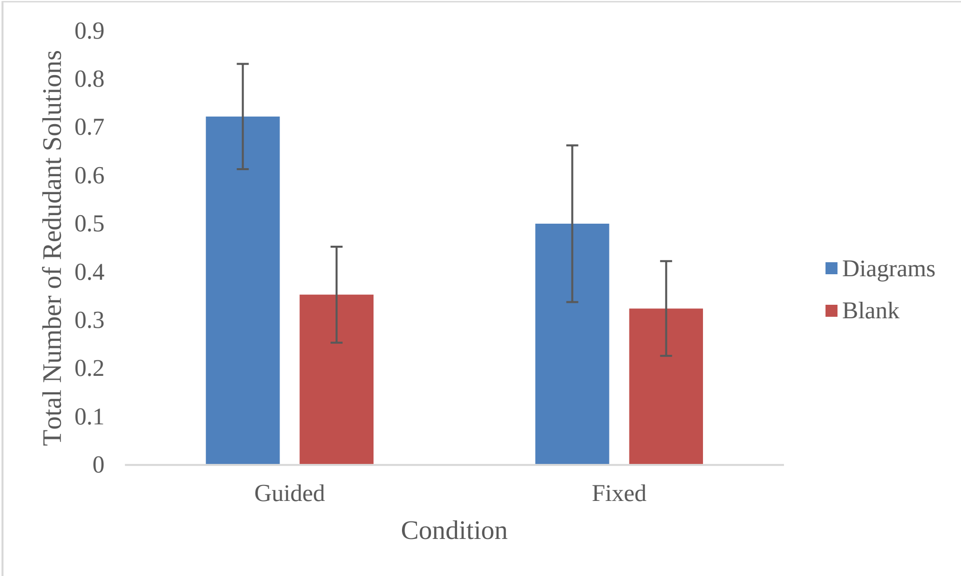

Figure 35. Redundant Solutions provided for each condition. Error bars represent the standard error of the mean. Those who in the guided condition generated significantly more redundant solutions than those in the fixed condition.

Mediating Effects

As in Study 2, the same individual difference and performance variables were collected that may relate to problem solving success. The previously mentioned independent-samples *t*-tests were conducted to identify whether those who solved the radiation problem were different on any variable than those who did not solve it. The results of the *t*-tests are provided below in Table 17. Three variables were identified that were significantly different between those who solved the problem and those who did not: Flexibility, Elaboration, and Total Solutions provided.

Table 17. Means, Standard Deviations, *t*-test Results, and Cohen's *d* Effect Sizes for Variables of Interest Between Those Who Solved the Radiation Problem and Those Who Did Not.

	Solved Radiation Problem?	*M*	*SD*	*t(141)*	*d*
Originality	No (n = 79)	0.61	0.74	1.84	0.30
	Yes (n = 64)	0.88	1.00		
Flexibility	No	2.73	1.24	3.12**	0.52
	Yes	3.41	1.33		
Elaboration	No	354.75	193.72	3.54**	0.60
	Yes	491.05	254.25		
Total No. Solutions	No	3.13	1.44	3.36[a]**	0.56
	Yes	3.98	1.62		
Redundant Solutions	No	0.39	0.63	1.48[a]	0.25
	Yes	0.58	0.83		
Age	No	18.84	1.75	1.19[a]	0.21
	Yes	19.49	4.08		
Gender	No	1.58	0.50	1.03[a]	0.17
	Yes	1.67	0.48		
Class Rank	No	2.47	0.93	0.34	0.06
	Yes	2.52	1.01		
"I like solving puzzles"	No	3.19	1.06	1.08	0.19
	Yes	3.37	0.81		
STEM Major	No	0.66	0.48	0.57	0.10
	Yes	0.70	0.46		
Med. Related Major	No	0.41	0.49	-0.56	0.09
	Yes	0.36	0.48		
Word Beginnings Score	No	13.23	4.56	0.44	0.07
	Yes	13.59	5.31		
Word Endings Score	No	15.53	5.45	0.43	0.07
	Yes	15.92	5.46		
Vocabulary Score	No	5.88	3.37	1.36	0.23
	Yes	6.61	2.91		
CRT Score	No	46.62	17.49	0.91	0.15
	Yes	49.29	16.98		
PFT Score	No	9.48	3.67	1.20	0.20
	Yes	10.21	3.55		

Note. *p < .05, **p < .01, [a]adjusted *df* were used as Levene's test indicated a violation of the equality of variance assumption. *d* = Cohen's *d* effect size (provided as absolute values). Significance of statistical tests was determined based on Bonferroni corrections for familywise error rate in each family of tested variables, where appropriate (i.e., creativity variables, verbal ability variables, spatial ability variables).

As in Study 2, a mediation analysis was conducted for each variable to assess potential confounding effects on experimental variables. Because diagram presence was the only

significant experimental predictor of solution rates in the radiation problem for Study 3, it is included as the predictor in each model.

A mediation model where Flexibility mediates the relationship diagram presence and problem success rates tests the idea that diagrams facilitate a creativity boosting effect that leads problem solvers to generate more solutions, which may incidentally increase the likelihood that participants generate the convergence solution. The logistic regression model was significant in block 1 (χ^2 (1, N = 143) = 6.90, p = .009, Nagelkerke R^2 = .06), with diagrams as a significant predictor for solution rates. The model was also significant in block 2 after the addition of the Flexibility variable (χ^2 (2, N = 143) = 13.75, p = .001, Nagelkerke R^2 = .12), with Flexibility and Diagrams both as significant predictors for solution rates. The indirect effect of condition on problem solving success was estimated at B = 0.21, 95% CI (.04, .53). Because the 95% confidence interval did not include zero, this effect was judged significant at $p < .05$. This mediation model suggests that the improvement in solution rates due to Diagrams is still meaningful, but also partially explained by a concomitant increase in creativity (as measured through Flexibility). Logistic regression coefficients are provided below in Table 18, and this relationship is depicted graphically in Figure 36.

Table 18. Logistic Regression Coefficients for Model Testing Flexibility as Mediator

	Variable	B	S.E.	Wald	Exp(B)
Block 1	Constant	-0.67	0.25	7.18**	0.51
	Diagrams	0.90	0.35	6.72**	2.45
Block 2	Constant	-1.70	0.49	12.10**	0.18
	Diagrams	0.73	0.36	4.22*	2.08
	Flexibility	0.36	0.14	6.43*	1.44

Note. **p* < .05; ***p* < .01.

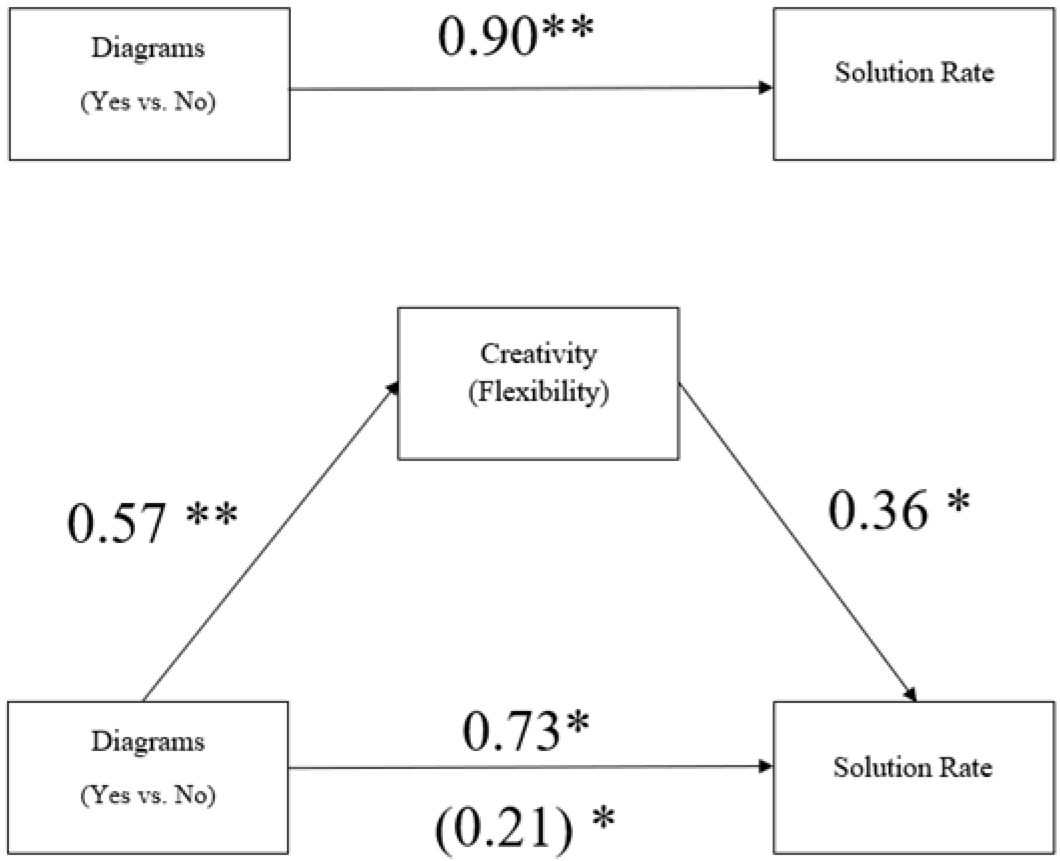

Figure 36. Unstandardized regression coefficients for the model of diagrams predicting solution rate, mediated by creativity (Flexibility). The top shows the total effect of diagrams on solution rate. The bottom shows the mediation model, with the indirect effect in parentheses and the direct effect above it. The indirect effect was estimated using 5,000 sample bias-corrected and accelerated bootstrapping procedure. $^*p < .05$, $^{**}p < .01$.

Next, a mediation model was tested where Elaboration was the mediator. This model tests the idea that a boost in creativity (as indicated by more detailed responses) better explains an increase in solution rates than diagrams. As before, this model was significant in block 1 (χ^2 (1, N = 143) = 6.90, p = .009, Nagelkerke R^2 = .06), with diagrams as a significant predictor for solution rates. The model was also significant in block 2 after the addition of the Flexibility variable (χ^2 (2, N = 143) = 13.75, p = .001, Nagelkerke R^2 = .12), with Flexibility and Diagrams

both as significant predictors for solution rates. The indirect effect of condition on problem solving success was estimated at B = 0.18, 95% CI (.01, .51). Because the 95% confidence interval did not include zero, this effect was judged significant at $p < .05$. As with the Flexibility model, this model suggests that the improvement in solution rates due to Diagrams is still meaningful, but also partially explained by a concomitant increase in creativity (as measured through Elaboration). Logistic regression coefficients are provided below in Table 19, and this model is depicted graphically in Figure 37.

Table 19. Logistic Regression Coefficients for Model Testing Elaboration as Mediator

	Variable	B	S.E.	Wald	Exp(B)
Block 1	Constant	-0.67	0.25	7.18**	0.51
	Diagrams	0.90	0.35	6.72**	2.45
Block 2	Constant	-1.70	0.42	15.65**	0.19
	Diagrams	0.78	0.36	4.76*	2.19
	Elaboration	0.26	0.08	9.52**	1.29

Note. *$p < .05$; **$p < .01$. To simplify regression coefficient estimation, Elaboration was calculated in units of 100.

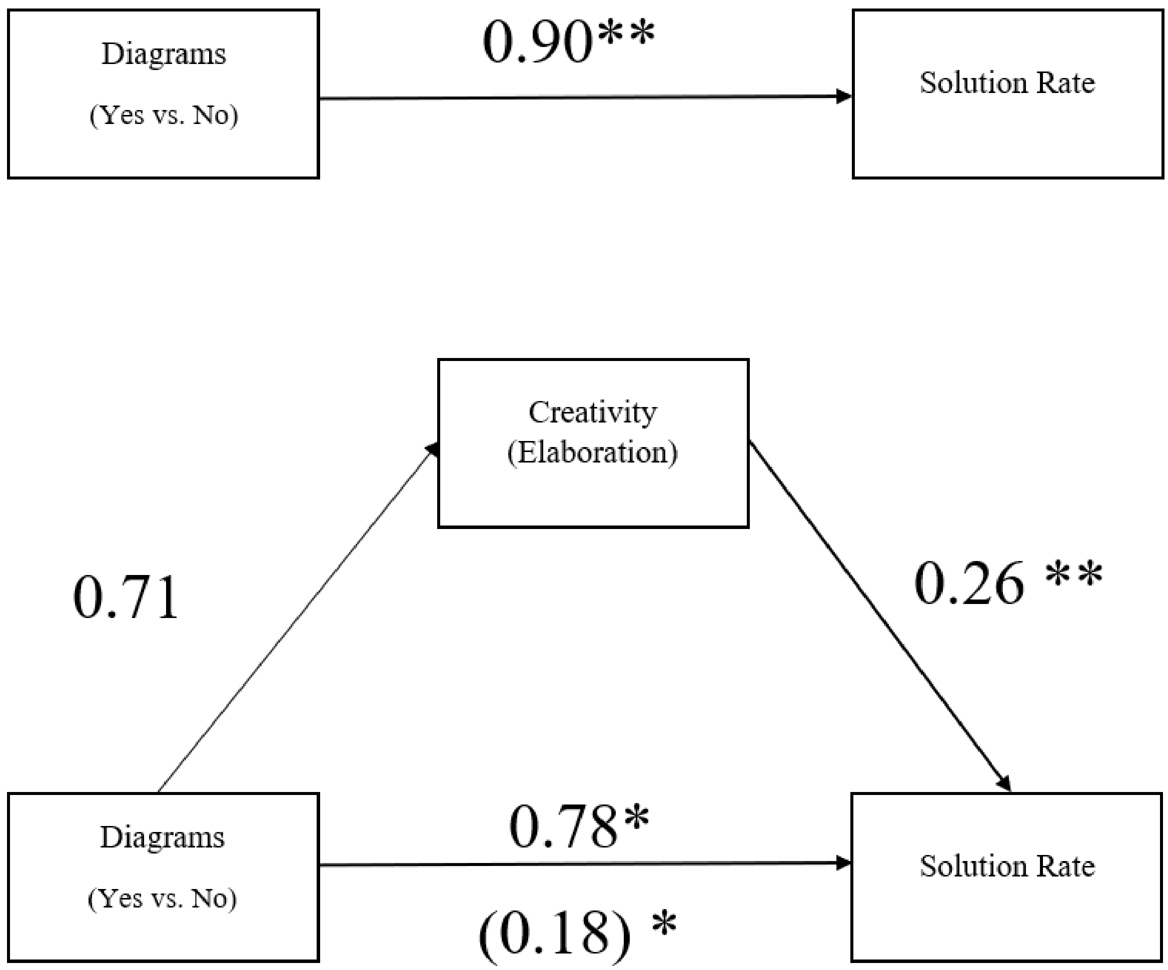

Figure 37. Unstandardized regression coefficients for the model of diagrams predicting solution rate, mediated by creativity (Elaboration). The top shows the total effect of diagrams on solution rate. The bottom shows the mediation model, with the indirect effect in parentheses and the direct effect above it. The indirect effect was estimated using 5,000 sample bias-corrected and accelerated bootstrapping procedure. *$p < .05$, **$p < .01$.

Finally, a mediation model was tested where Total Solutions was entered as the mediator. This tests the idea that one's likelihood of solving the radiation problem is better explained by a boost in number of solutions generated, rather than from any benefit of diagrams. As before, this model was significant in block 1 (χ^2 (1, N = 143) = 6.90, p = .009, Nagelkerke R^2 = .06), with diagrams as a significant predictor for solution rates. The model was also significant in block 2

after the addition of the Total Solutions variable (χ^2 (2, N = 143) = 14.33, p = .001, Nagelkerke R^2 = .13), with Total Solutions as a significant predictor for solution rates, but diagrams no longer a significant predictor. The indirect effect of condition on problem solving success was estimated at B = 0.27, 95% CI (.07, .63). Because the 95% confidence interval did not include zero, this effect was judged significant at $p < .05$. Because the inclusion of Total Solutions into this model reduced the significance of Diagrams as predictor, this model supports Total Solutions as a better explanation of the effect of Diagrams on solution success rates. Regression coefficients are provided in Table 20 below, and this model is depicted graphically in Figure 38.

Table 20. Logistic Regression Coefficients for Model Testing Total Solutions as Mediator

	Variable	B	S.E.	Wald	Exp(B)
Block 1	Constant	-0.67	0.25	7.18**	0.51
	Diagrams	0.90	0.35	6.72**	2.45
Block 2	Constant	-1.69	0.47	12.88**	0.19
	Diagrams	0.67	0.36	3.45	1.96
	Total Solutions	0.32	0.12	6.99**	1.38

Note. *p < .05; **p < .01. To simplify regression coefficient estimation, Elaboration was calculated in units of 100.

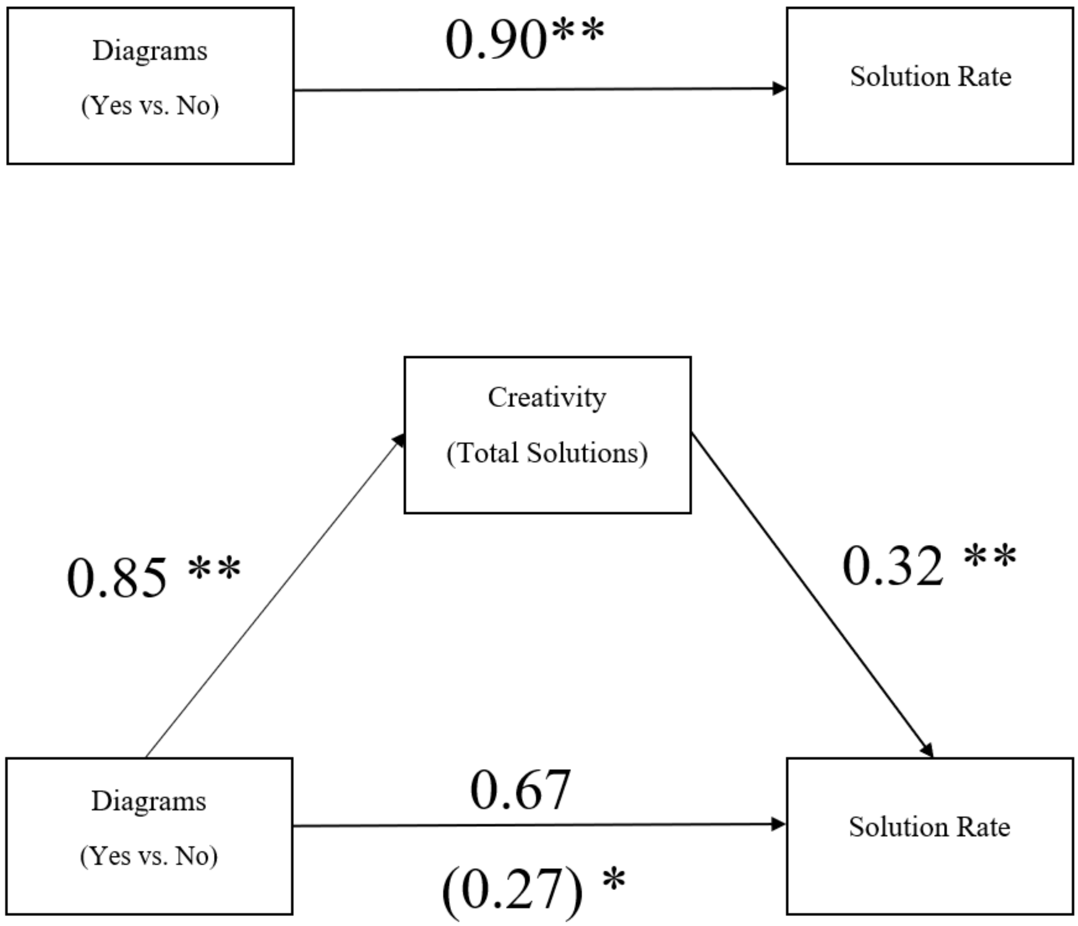

Figure 38. Unstandardized regression coefficients for the model of diagrams predicting solution rate, mediated by creativity (Total Number of Solutions). The top shows the total effect of diagrams on solution rate. The bottom shows the mediation model, with the indirect effect in parentheses and the direct effect above it. The indirect effect was estimated using 5,000 sample bias-corrected and accelerated bootstrapping procedure. *p < .05, **p < .01.

Discussion

Failure to Replicate

Altogether, the results of Study 3 are unusual. Two conditions from Study 2 were replicated in Study 3 (the guided and fixed eye movements with diagrams conditions), but the

previously found embodied effect of guiding eye movements improving solution rates did not replicate. It is possible that there are individual differences in the samples from Study 2 and Study 3 that may explain why this effect seemed to disappear in Study 3. In particular, Study 2 and Study 3's data were collected during different semesters (Spring and Summer, respectively). The following chapter examines these individual differences between both studies and examine whether they were influential in predicting problem solving success rates.

Spatial Ability

There were several notable effects of spatial ability on the problem solving process in Study 3. There was a general benefit of spatial ability for performance on Red Adair, but there was also a significant moderating effect of spatial ability and diagrams on this same problem. The results suggested that in the absence of diagrams on scratch paper, spatial ability was a significant predictor of success on Red Adair. However, when diagrams were present on participants' scratch paper, their spatial ability was not relevant for performance. Recall that scratch paper was available for all three problems, but the diagrams vs. blank scratch paper manipulation only occurred for the Radiation Problem (all subsequent scratch papers were blank). So, this effect is potentially a latent benefit of diagrams aiding in convergence schema formation. The nature of this effect is likely that diagrams even the playing field between low and high spatials, helping them to visualize the spatial aspects of the convergence solution that aid in schema construction.

On the other hand, those who used blank scratch paper had to rely on their spatial ability to be successful on later problems, as they had no diagrams to aid in visualizing the convergence schema. Those in the blank scratch paper condition who had higher spatial ability scores were

most successful on the first transfer problem. This could highlight a suppressing effect of diagrams aiding in convergence solution schema formation for high spatials. In other words, those higher in spatial ability may be better able to construct mental models of spatial solutions when they are not provided with a diagram as a starting point.

Analogical Problem Solving

Study 3 again failed to support that any of the experimental manipulations impacted the analogical problem solving process. Prior solution rates, eye movement guidance, and diagram presence all failed to predict meaningful differences in transfer problems. Because the embodied effect found in Study 2 did not replicate in Study 3, these data are limited in that an embodied effect on analogical transfer cannot be eliminated. However, taken together with the results of Study 2, there is not strong evidence to support that eye movement guidance helps problem solvers embody the convergence solution as identified through analogical transfer.

Similarly, there were no significant predictors of participants noticing the convergence schema across all three problems. Compared to Study 2, there were a substantially fewer number of participants who did notice the common schema across all three problems, and it is possible there simply was not enough statistical power to make predictions about recognizing the convergence schema in Study 3. However, as mentioned in the previous section, spatial ability may have impacted analogical transfer from problem 1 to problem 2, but it is unclear whether spatial ability impacts analogical transfer for the first transfer problem, or if that problem in particular benefits from spatial ability.

Influence of Diagrams on Creativity and Problem Solving Success

In Study 3, there was a consistent significant effect of diagrams that emerged with both solution rates on the radiation problem, as well as creativity scores on written solutions. Previous research supports that diagrams are useful during problem solving (e.g., Beveridge & Parkins, 1987; Gick & Holyoak, 1983; Pedone et al., 2001), however, the diagram in the present study was very simple and did not contain any information about the solution to the problem. In these previously cited studies, each diagram contained some perceptual information (e.g., directional arrows, shading, etc.) that helped convey how the tumor was to be x-rayed. The present study's results support that the diagrams facilitated creativity, helping problem solvers generate a greater variety of solutions, as well as more original solutions.

The post-hoc analyses suggest that the beneficial effects of diagrams on problem solving may be better explained by diagrams improving creativity. It is reasonable to think that scratch paper with diagrams might give participants additional information (compared against blank scratch paper) that could help guide their responses toward the convergence solution. However, the post-hoc mediation analyses support that diagrams were partially more likely to aid in convergence solution formation *because* they primed more creative solutions to the radiation problem. Flexibility, Elaboration, and Total Solutions were all greater among those who solved the radiation problem, and Elaboration and Total Solutions appeared to be stronger predictors of success than diagrams because their addition into the logistic regression models weakened the predictive strength of diagrams. Flexibility and Total solutions were also greater in the diagrams conditions, further highlighting the lack of independence of these effects.

It is also possible that these diagrams serve as a visual aid that helps problem solvers construct a mental representation of the problem. Previous research indicates that more detailed mental representations are helpful for problem solving (Chi et al., 1989; Chi & Glaser, 1985), so this is one possible explanation for how the diagrams aid in problem solving. It could also be possible that detailed mental representations allow problem solvers to generate more solutions, which could be another way to make sense of the mediating effects described in the previous sections.

CHAPTER SIX: COMPARING STUDY 2 AGAINST STUDY 3

Methodology for Group Comparisons (Study 2 vs. Study 3)

In both Study 2 and Study 3, there were two conditions that saw the radiation problem, underwent a guided or fixed eye movements digit-tracking task, and then attempted to solve the radiation problem using scratch paper with diagrams. These conditions in Study 2 showed a significant embodied effect replicating effects found in previous literature. However, these same two groups showed essentially zero difference in Study 3. There are several possibilities as to why this might be the case. First, I examine variables that were different between those who solved the radiation problem, and those who did not. Next, I examine demographic differences between the two studies, as well as individual difference measures and performance measures. In the following sections, I highlight which variables influenced solution rates, which demographic variables were different between the two studies, and then assess whether they differentially affected solution rates in Study 2 and Study 3. Any variable identified as influential for problem solving success that is also significantly different between the two groups could hold an explanation underlying the replication issue. Finally, I examine posterior probabilities for these effects based on re-analyses of previous studies' available data (from Thomas & Lleras, 2007; 2009).

Variables Related to Problem Solving Success

Several *t*-tests were completed to examine variables that were different in those who solved the problem compared to those who did not solve the problem. It is possible that a variable that was influential in problem solving was differently distributed between the two studies. These *t*-tests compare the following variables: Creativity-related variables (Originality,

Flexibility, Elaboration, Total Solutions, Redundant Solutions), Demographic variables (Age, Gender, Class Rank, STEM Majors, Medical-Related Majors), Cognitive performance measures (Verbal ability from Word Endings, Word Beginnings, and Vocabulary Tests, and Spatial Ability from the PFT and CRT measures), Self-reports of liking to solve puzzles, whether they used scratch paper for the radiation problem, and whether they solved the radiation problem. Examining these variables could aid in understanding whether any meaningful variables for problem solving were different in each study, and whether this difference impacted solution rates. There were five variables significantly different between those who solved the radiation problem and those who did not: Verbal ability (which was addressed in the previous section), Elaboration, Self-reports of liking to solve puzzles, Vocabulary Test Score, PFT score, and whether or not participants used the scratch paper. The results of these *t*-tests are displayed in Table 21.

Table 21. T-Tests for Differences in Demographic Variables Between Those Who Solved the Radiation Problem and Those Who Did Not

	Solved Problem?	*M*	*SD*	*t*(149)	*d*
Originality	No (n = 75) Yes (n = 76)	0.89 0.76	0.96 0.88	0.88	0.15
Flexibility	No Yes	3.31 3.38	1.31 1.29	-0.31	0.05
Elaboration	No Yes	446.02 533.20	253.27 250.37	-2.10*	0.35
Total Solutions	No Yes	3.92 4.11	1.55 1.61	-0.73	0.12
Redundant Solutions	No Yes	0.60 0.73	0.91 1.23	-0.71	0.12
Age	No Yes	19.46 19.92	2.27 4.00	-0.88	0.14
Gender	No Yes	1.49 1.59	0.50 0.50	-1.25	0.21
Class Rank	No Yes	2.59 2.71	0.99 1.06	-0.75	0.12
"I Like Solving Puzzles"	No Yes	3.23 3.70	0.95 0.94	-2.98**	0.49
STEM Major	No Yes	0.66 0.76	0.48 0.43	-1.27[a]	0.21
MED Related Major	No Yes	0.33 0.32	0.47 0.47	0.13	0.02
Word Beginnings Score	No Yes	13.77 15.11	4.82 5.02	-1.65	0.27
Word Endings Score	No Yes	16.04 16.85	4.93 5.41	-0.96	0.16
Vocabulary Score	No Yes	6.47 7.79	3.56 3.29	-2.32*	0.38
CRT Score	No Yes	47.83 49.41	18.92 16.91	-0.53	0.09
PFT Score	No Yes	9.04 10.51	4.00 3.81	-2.27*	0.38
Problem 1 Scratch Paper Used	No Yes	0.89 0.98	0.31 0.12	-2.49*	0.39
Problem 1 Solved	No Yes	464.44 494.46	133.35 114.39	-1.45	0.24

Note. *p < .05, [a]adjusted *df* were used as Levene's test indicated a violation of the equality of variance assumption. *d* = Cohen's *d* effect size (provided as absolute values). Significance of statistical tests was determined based on Bonferroni corrections for familywise error rate in each family of tested variables, where appropriate (i.e., creativity variables, verbal ability variables, spatial ability variables).

Variables Different by Study

Next, several independent samples *t*-tests were performed to check whether the participants in Study 2 were different from those in Study 3. These *t*-tests mirror the previous tests, except that the groups are compared by study instead of problem solving success. Means, standard deviations, and the results of these group comparison tests are provided in Table 22.

Table 22. T-Tests for Differences in Demographic Variables Between Study 2 and Study 3's Replicated Guided and Fixed Conditions

	Study	*M*	*SD*	*t*(149)	*d*
Originality	2 (*n* = 79) 3 (*n* = 72)	0.70 0.97	0.91 0.92	-1.85	0.30
Flexibility	2 3	3.38 3.32	1.25 1.33	0.29	0.05
Elaboration	2 3	517.92 451.08	266.38 238.15	1.62	0.26
Total Solutions	2 3	4.09 3.93	1.53 1.61	0.62	0.10
Redundant Solutions	2 3	0.71 0.61	1.23 0.83	0.57	0.09
Age	2 3	19.83 19.46	2.12 3.96	0.72	0.12
Gender	2 3	1.44 1.63	0.50 0.49	-2.37*	0.39
Class Rank	2 3	2.78 2.46	1.05 0.97	1.94	0.32
"I Like Solving Puzzles"	2 3	3.56 3.32	1.02 0.91	1.47	0.24
STEM Major	2 3	0.68 0.71	0.47 0.46	-0.33	0.05
MED Related Major	2 3	0.25 0.39	0.44 0.49	-1.73[a]	0.30
Word Beginnings Score	2 3	15.03 13.78	5.14 4.76	1.54	0.25
Word Endings Score	2 3	16.85 16.00	4.97 5.30	1.01	0.17
Vocabulary Score	2 3	7.76 6.31	3.21 3.68	2.59**	0.42
CRT Score	2 3	49.67 47.21	18.24 17.55	0.84	0.14
PFT Score	2 3	9.90 9.65	4.27 3.79	0.38	0.06
Problem 1 Scratch Paper Used	2 3	0.94 0.93	0.25 0.26	0.11	0.02
Problem 1 Solved	2 3	0.42 0.47	0.50 0.50	-0.69	0.11

Note. $*p < .05$, $**p < .01$ [a]adjusted *df* were used as Levene's test indicated a violation of the equality of variance assumption. *d* = Cohen's *d* effect size (provided as absolute values). Significance of statistical tests was determined based on Bonferroni corrections for familywise error rate in each family of tested variables, where appropriate (i.e., creativity variables, verbal ability variables, spatial ability variables).

The results show a significant difference in Gender and Vocabulary Test score between Studies 2 and 3. The sample in Study 2 had more males than females, but this distribution was reversed in Study 3. Since gender is also known to be an influence for solving the radiation problem (favoring males; Antonietti & Gioletta, 1995), it could have been a factor in why the

effect did not replicate. The sample in Study 2 also had a higher Vocabulary Test score compared to Study 3. It is possible that verbal ability may relate to a greater likelihood of problem solving success.

To test whether the group differences in these variables could account for the solution rate differences between Study 2 and Study 3, two logistic regressions were performed in two blocks: The first block included the variable of interest (either Vocabulary Test score or Gender) along with the Study variable (Study 2 or Study 3), and the second block included the product term of these two variables. This logistic regression model permits testing whether the group differences differentially predicted success on the radiation problem as a function of Study.

For the first model with Vocabulary Test score, the first block was statistically significant $\chi^2(2, N = 149) = 6.80$, $p = .033$, Nagelkerke $R^2 = .06$), with only Vocabulary Test score predicting success on the radiation problem. The product term of both variables was added to the second block, but did not significantly increase variance explained $\Delta\chi^2(1, N = 149) = 1.20$, $p = .27$, ΔNagelkerke $R^2 = .01$), yielding an overall model that was statistically significant $\chi^2(3, N = 149) = 8.00$, $p = .046$, Nagelkerke $R^2 = .07$), but with no individually significant variables (all ps > .28). Regression coefficients are provided in Table 23.

Table 23. Logistic Regression Model for Vocabulary Test Score and Study Predicting Solution Rates for the Radiation Problem

	Variable	B	S.E.	Wald	Exp(B)
Block 1	Constant	-2.16	1.03	4.43*	0.15
	Vocabulary Test Score	0.13	0.05	6.00*	1.13
	Study	0.42	0.35	1.46	1.52
Block 2	Constant	-0.07	2.15	0.00	0.94
	Vocabulary Test Score	-0.16	0.27	0.37	0.85
	Study	-0.39	0.82	0.23	0.68
	Study*Score	0.11	0.10	1.19	1.12

Note. $*p < .05$; $**p < .01$.

This same model was tested using Gender. Although there was a significant effect of gender predicting problem solving success before entering it into the logistic regression model ($p = .05$), the first block with Gender and Study was not statistically significant $\chi^2(2, N = 148) = 1.87$, $p = .39$, Nagelkerke $R^2 = .02$), with neither Gender nor Study predicting success on the radiation problem. The product term of both variables was added to the second block and did not significantly increase variance explained $\Delta\chi^2(1, N = 148) = 0.01$, $p = .94$, ΔNagelkerke $R^2 = .00$), yielding an overall model that also was not statistically significant $\chi^2(3, N = 148) = 1.87$, $p = .60$, Nagelkerke $R^2 = .017$). In essence, although the Gender distribution was significantly different between Study 2 and Study 3, this difference did not significantly impact the problem solving success rate differences between Study 2 and Study 3 (see Table 24).

Table 24. Logistic Regression Model for Gender and Study Predicting Solution Rates for the Radiation Problem

	Variable	B	S.E.	Wald	Exp(B)
Block 1	Constant	-1.6	0.92	1.87	0.28
	Gender	0.38	0.34	1.26	1.47
	Study	0.19	0.34	0.30	1.20
Block 2	Constant	-1.46	2.74	0.28	0.23
	Gender	0.51	1.72	0.09	1.66
	Study	0.26	1.11	0.06	1.30
	Study*Gender	-0.05	0.68	0.01	0.95

Note. $*p < .05$; $**p < .01$.

Taken together, these results indicate that there was no differential effect for Verbal Ability or Gender on problem solving success, suggesting that even though these variables were significantly different between Study 2 and Study 3, this difference did not impact solution rates.

Examining Posterior Probabilities

In efforts to better understand what the lack of replication between Studies 2 and 3 implies for future research, posterior probabilities were examined. This analysis assumes a conservative equal likelihood for null and alternative hypotheses prior to the results from Thomas and Lleras (2007). This assumption is conservative, considering that Grant and Spivey (2003) found evidence supporting Thomas and Lleras (2007)'s claims for an embodied effect of eye movements; however, Grant and Spivey's (2003) conditions were not directly comparable to Thomas and Lleras (2007)'s, hence employing a conservative prior null hypothesis likelihood estimate. Using logistic regression, comparing their closely related guided condition (n = 18) and fixed condition (n = 18) solution rates yields a near-significant effect of eye movement guidance on solution rates ($p = .08$, Observed Power = .88). Employing Bayes' Theorem, the

posterior belief in the null hypothesis given this result produces a posterior probability of p = .083 that the null hypothesis is true.

This posterior probability (p = .083) was input into a new calculation integrating the results from Thomas and Lleras (2009) with Bayes' Theorem. Their closely related guided and fixed conditions (both ns = 23) yield a significant effect of eye movement guidance on solution rates (p = .04, Observed Power = .70). This produces a posterior probability of p = .005 that the null hypothesis is true, given the data.

These calculations were performed in order to inform the prior probabilities for examining Studies 2 and 3. When examining posterior probabilities for both present studies, both studies independently suggest a low likelihood of a null effect given the data (posterior probabilities p = .00009 for Study 2, and p = .047 for Study 3). Assuming the effects from previous research are genuine, the results from the present work do not lend support to a null effect of eye movement guidance on problem solving.

Comparing Solution Rates to Previous Research

One possible reason why the results of the present studies could have been different in previous research would be if the solution rates were higher overall. If the studies in the present samples were simply better at solving the radiation problem, the beneficial effects of guiding eye movements might not be detectable. With regard to overall solution rates, Studies 2 (41%) and 3 (37%) were not out of the range of solution rates seen across all conditions in Grant and Spivey (2003; 44%), Thomas and Lleras (2007; 36%), and Thomas and Lleras (2009; 26%). However, it is worth noting that these solution rates are higher than the traditional 10% solution rate seen in early studies (Gick & Holyoak, 1980; 1983). However, there are two reasons why the solution

rates in the present and relevant previous work should be higher. In the early studies, verbalizing or writing were the only means of providing solutions. In the present studies, written and drawn solutions were required, and in Grant and Spivey and Thomas and Lleras's studies, drawn solutions were required. This may explain a portion of the difference in that drawing a solution is driven by fundamentally different processes (e.g., visuospatial) than writing or verbalizing a solution. Additionally, incubation was not present in the early studies, which may have lowered solution rates. In sum, the solution rates in the present work do not appear unreasonable compared to relevant previous research.

Discussion

The results highlighted do not provide a clear answer as to why the embodied effect did not replicate from Study 2 to Study 3. Gender and a measure of verbal ability were significantly different between the two studies, but these variables did not differentially predict success in solving the radiation problem for each study. It should be noted that there were very few statistical differences in the samples between each study which could indicate that they were not drastically different samples, or that some unmeasured variable drove the difference in effects. An alternative possibility is that the embodied effect from previous literature that was replicated in Study 2 is temperamental in nature or simply a small effect that does not replicate easily. There is reason to believe that this may be the case, as several other studies have failed to find embodied effects that aid in solving the radiation problem (e.g., Cooperrider & Goldin-Meadow, 2014; Hostetter et al., 2016; Trowbridge, 2016). However, it should be noted that the analysis of posterior probabilities that the null hypothesis was true (given the data) did not lend much support to this claim. Instead, it would appear that a failure to replicate an embodied effect in

Study 3 still left little overall support for a null effect of eye movement guidance on problem solving. With that being considered, Study 2 should probably carry more weight in terms of evidence for embodied effects on problem solving, but the mediating effect of creative thinking should not be ignored.

CHAPTER SEVEN: GENERAL DISCUSSION

Implications for Analogical Problem Solving

The proposed research expands the literature on analogical problem solving by incorporating eye-movement priming studies into an analogical problem solving paradigm to assess whether the benefits of guided eye movements transfer to later problem solving. The data from Studies 2 and 3 do not support that guiding eye movements in a pattern embodying a problem's solution aids in analogical transfer. However, in Study 2, these effects may have been eliminated due to the provision of hints. Also, because of the replication issue in Study 3, it may be that because the embodied effect never replicated in the first problem, it could not be detected through analogical transfer in the transfer problems. Although the data do not support an embodied benefit to analogical problem solving, they do not rule out this possibility. More research is needed to determine if embodied priming can aid analogical transfer.

Implications for Embodied Cognition

It is worth noting that the results from Studies 2 and 3 conflict with one another. Two of the groups in Study 2 (Perceptual Simulation Group and Simple Incubation Group) and two of the groups in Study 3 (Guided + Diagrams and Fixed + Diagrams) were identical in terms of methodology. However, the difference in solution rates for those groups in Study 2 were significant, in favor of the perceptual simulation group, and in line with previous research, whereas the groups in Study 3 were statistically equivalent. A more in-depth analysis of the contrasting individual differences in the groups between these studies did reveal significant differences in verbal ability. Verbal ability was higher in the Study 2 sample but did not differentially predict solution rates in either sample. Considering that there were essentially no

impactful differences in the samples of Study 2 and Study 3, it is possible that the embodied effect of guiding eye movements is not as robust as would be expected based on Grant and Spivey (2003) and Thomas and Lleras's (2007; 2009) claims. Another possibility is that there is some additional unmeasured variable responsible for the stark difference in the results of the present Studies 2 and 3.

Study 3's results fail to support an initial embodied effect, and they also fail to support any latent embodied effects that could be expected based on the expected effect of a perceptual simulation improving one's mental representation of a problem that aids in analogical transfer. However, it is worth noting that the failure to elicit an initial embodied effect may have been the reason that no latent embodied effects emerged in the data. Future research is needed to determine whether the benefits conferred by guided eye movements are truly embodied in nature, to the extent that they facilitate a perceptual simulation of the convergence solution.

Taking posterior probabilities into consideration, the most relevant results would be those found in Study 2. For embodied cognition, Study 2 supports that the benefits of guiding eye movements for solving the radiation problem are due to an embodied priming of creative thinking – not necessarily because guiding eye movements fosters a perceptual simulation of the convergence solution. It is possible that there is unique variance in both predictors of solution rates (guiding eye movements and creative thinking), but the current results support that guiding eye movements resulted in an increase in creative thinking, which led to an increase in solution rates. These effects mirror those of previous attentional priming of creativity studies (e.g., Friedman et al., 2003; Wegbreit et al. 2012), which argue a general effect of visual attention on creativity. Because creativity is a key element in success with the radiation problem (Dow & Mayer, 2004; Gick & Holyoak, 1980), it is more plausible to argue that guiding eye movements

influences creative thinking which increases solution rates. Without knowing the impacts of guiding eye movements on creative thinking in previous studies, it is less reasonable to argue that guiding eye movements in the pattern of the convergence solution fosters a perceptual simulation of that solution.

Implications for Creativity

In all three studies, creativity effects were measured based on the written responses that participants had provided. In Study 1, no meaningful creativity effects emerged due to the manipulation. In Study 2, there appeared to be a benefit to creativity due to incubation as well as eye movement guidance. In Study 3, broad creativity benefits emerged due to the manipulation of diagrams, and it appears they mediated the association between diagrams and problem solving success. It is possible that these creativity effects are concomitant with any benefits due to guiding eye movements. However, it is interesting to note that these effects did not emerge until participants were required to draw their solutions. It is possible that the act of drawing itself facilitates creativity in problem solving.

At the time of this writing, the present work is the first to examine creative thinking using solutions to the radiation problem. Both Study 2 and Study 3's experimental manipulations yielded differences in solution rates that were partially explained by a concomitant increase in measures of creative thinking. One possibility is that the experimental manipulations from previous studies (e.g., relevant analogies, diagrams, animations, gestures) may also have benefited problem solving through creativity. Future research should explore whether the traditional manipulations that benefit analogical problem solving may be intertwined with creativity.

It may appear as if creative thinking effects are more pronounced in Study 3 than in Study 2, but it should be recognized that the reason for this difference is likely due to the manipulation of the scratch paper variable. All groups had diagrams on their scratch paper in Study 2, whereas half of the groups had diagrams and the other half had blank scratch paper in Study 3. Thus, any differences in creativity in Study 3 could be due to blank scratch paper stifling creativity.

One concern about creativity is that it is a difficult construct to measure (e.g., Piffer, 2012). The present study employed a divergent thinking style approach to measuring creative thinking by coding and analyzing participants' written responses to problem solving, based on Guilford's (1967) theories of divergent thinking. This approach is argued to be one of the more feasible measures of creativity (Piffer, 2012), but Piffer critiques that it only measures a relatively transitory state known as creative thinking rather than a trait of creativity. Based on previous research that examined creative performance after short visual attention tasks (e.g., Friedman et al., 2003; Wegbreit et al. 2012), it is likely that the effects of creative thinking that were observed in this study were due to the experimental manipulations, rather than simply trait-based creativity.

Lastly, it is worth mentioning that there were a small number of participants who might be more closely drawn to majors requiring creativity to be successful (e.g., fine arts, performing arts, decorative arts, applied arts, or arts education; Silvia & Nusbaum, 2012). Silvia and Nusbaum supported that those who majored in the previously described areas had more creative accomplishments and were more creative in their day-to-day lives. It is likely that people who choose to concentrate their studies in these areas exhibit creative traits, and participants in the present studies holding these majors might also perform differently on creative thinking tasks,

compared to non-arts majors. Participants' major was further explored to better understand whether those who majored in arts differed on the creative thinking measures, compared to those who did not major in the arts.

Unfortunately, the present studies had only a total of 31 arts majors across all three studies (Study 1 n = 10; Study 2 n = 12; Study 3 n = 9), and none of these participants were significantly higher on any of the creative measures when compared to those concentrating in non-arts majors (all *p*s > .15). Although the number of arts majors in each study was small, their performance on measures of creative thinking in problem solving were not significantly better than non-arts majors.

Disambiguating Embodied Priming and Creativity Priming

In Study 2, there appeared to be a significant effect of guiding eye movements, such that those whose eye movements were guided in a pattern congruent with the convergence solution were more likely to solve the radiation problem. Eye movement guidance also appeared to generate significant differences on several creativity measures – Elaboration, Total Solutions, and Redundant Solutions. Because the Attentional Priming hypothesis posited that the problem solving benefits of guiding eye movements are actually due to a boost in creativity, a mediation analysis was performed to disambiguate these effects. Elaboration was found to significantly mediate the benefits of guiding eye movements. This lends support to eye movement guidance facilitating increased creativity, rather than embodying the solution, which leads to better problem solving success rates. It should be noted that this was the only creativity variable that mediated this relationship, so it is possible there may be some unique effects of eye movement guidance that fosters embodiment. Future research should seek to further disambiguate these

concomitant effects by examining alternative manipulations of the tracking task (e.g., without a diagram, in a different pattern) or similarly difficult insight problems that have a disparate spatial solution (i.e., one not mirrored by the tracking task).

In Study 3, it is important to note that the differences in creativity scores tended to mediate the relationship between the scratch paper conditions (diagrams vs. blank) and solution rates on the radiation problem. The nature of these relationships was such that those in the scratch paper with diagrams conditions showed higher measures of creativity than those in the blank scratch paper conditions, and this difference in creativity explained the difference in problem solving success rates between the two conditions. Although at first glance it appeared that the scratch paper with diagrams conditions performed better on the radiation problem, it was ultimately better explained by a concomitant increase in creativity. Qualitatively, these mediating effects were much more consistent in nature than those from Study 2, and more strongly suggest that scratch paper with diagrams improve solution rates through an enhancement of creativity.

Implications for Spatial Ability

Across all three studies, there were not any consistent effects for spatial ability. Considering the spatial nature of the radiation problem and its analogues, it is surprising that no major effects emerged. However, it is worth noting that spatial ability appeared to influence solution rates on the Red Adair in Study 3. Since participants did not receive solutions to problems in Study 3, it is possible that this effect illuminates a benefit for the analogical reasoning process of structure mapping. Those with higher spatial ability should theoretically be

better able to visualize and link the structures between analogical problems, and this appeared to improve solution rates – particularly for those who used blank scratch paper.

Spatial ability never emerged as a predictor of success on the Radiation Problem, but it appears that spatial ability was a significant predictor of success on Red Adair in both studies. This could suggest that those with higher spatial ability are more likely to spontaneously transfer the solution from the Radiation Problem to Red Adair. Or more simply, spatial ability could have been particularly influential for Red Adair. The present study did not counterbalance the presentation order of the two transfer problems, as Corkill and Fager (1995) demonstrated that the presentation order of transfer problems did not affect analogical transfer. If spatial ability is more influential for Red Adair than The General, future research would see this variable as a significant predictor of performance, no matter how far removed it was from the source problem. However, if spatial ability is a meaningful predictor of analogical transfer, it would be a significant predictor of performance on whatever transfer problem followed the source problem.

Lastly, it is worth recognizing that spatial ability was measured during the last portion of the study, after any experimental manipulation had taken place. This position in the procedure was chosen because of concerns that spatial tests could have primed spatial thinking, which could have affected any of the problem solving tasks. Conversely, there is a similar possibility that any differences observed in spatial ability were due to the manipulation occurring prior to spatial ability measurement.

Liking to Solve Puzzles

In each study, participants answered a scaled Likert-type question to indicate the degree to which they like solving puzzles. In Study 2, this variable was significantly different between

those who solved the problem and those who did not (see Table 21). This variable uniquely predicted variance in solving the radiation problem – it did not covary with any experimental manipulations. Furthermore, this variable was not statistically different across the samples in each study. This variable was intended as a simple variable to assess whether "liking" the task explained all of the benefits seen by experimental manipulations. It is worth noting that liking to solve puzzles could represent other potentially relevant factors that could have impacted problem solving success. Need for Cognition (e.g., Cacioppo, Petty, & Kao, 1983) is a construct that assesses the degree to which one enjoys engaging in thinking. People with higher Need for Cognition may enjoy problem solving tasks which increases their likelihood of successful problem solving. Liking to solve puzzles may also represent a motivational factor that increased people's likelihood of problem solving success. In general, this variable may represent affect-related aspects of the individual that could impact learning and problem solving.

Practical Applications

Although the evidence is not strong for an embodied effect of improving problem solving, there appears to be evidence for increased solution generation in problem solving as a function of guiding eye movements and incubation. In Study 2, guiding eye movements increased the number of solutions generated, like the attentional priming of creativity effects in previous literature (e.g., Friedman et al., 2003). Additionally, there was a benefit for Incubation in terms of solution elaboration, such that people who saw their problem before undergoing their digit-tracking task, then resumed attempting to solve the problem were more likely to generate more detailed solutions than those who only saw the problem after doing their digit-tracking task. Practically, these results echo the traditional understanding of incubation effects (e.g.,

Segal, 2004), that temporarily removing oneself from a problem can benefit solution generation upon revisiting the problem.

There were also significant effects of diagrams on creativity in Study 3. Although diagrams did help improve solution rates, it was because they increased creativity which increased the likelihood of solving the radiation problem. These results suggest that diagrams can improve problem solving success by enriching the attempted solutions during the problem solving process. A possible cognitive mechanism for this effect could be that diagrams aid in the formation of mental representations of problems, which facilitate greater flexibility in problem solving. A practical application of these results is that learners and problem solvers could ease their learning and/or problem solving processes by drawing out diagrams of the problems they are trying to solve. This effect relates to Vygotsky's theoretical concept of scaffolding (e.g., Wood, Bruner, & Ross, 1976) during learning, such that a diagram facilitates mental model formation that aids in problem solving. This may also represent a traditional benefit of multimedia learning (e.g., Mayer, 2002), such that diagrams combined with other material (e.g., problem text) enhance understanding.

With regard to the interaction between spatial ability and diagrams in Study 2 (see Figure 28), diagrams seemed to equalize performance differences among those with various levels of spatial ability; however, those with the high levels of spatial ability were most likely to be successful when they did not receive diagrams on their scratch paper during the radiation problem. This may represent an expertise reversal effect (e.g., Kalyuga, 2009), such that the scaffolding benefits of diagrams are too much information for those with high spatial ability, who should be able to construct mental representations of the problem more easily on their own.

This may explain why those with high spatial ability had lower performance when a diagram was provided for them compared to when they did not receive a diagram.

In retrospect, it is possible that all of the aids to problem solving seen in previous research (e.g., diagrams, animations, gesture interaction) that improved solution rates only did so through an improvement in creativity. In the present studies, guiding eye movements improved creative thinking which improved solution rates. Using diagrams during problem solving also improved creative thinking which improved solution rates. At the time of this writing, the present work is the only research to have examined creativity from a divergent thinking perspective, and in two studies, creative thinking explained some or all of the problem solving benefits of the experimental manipulations. Future research should consider the effects of creative thinking during problem solving to better understand the ways in which aids to problem solving benefit cognition. Furthermore, it may be valuable to revisit earlier factors that influence problem solving success to better understand whether they facilitate creative thinking, or whether there are unique aspects of analogies, diagrams, animations, and enacted gestures that contribute to problem solving success.

Limitations and Future Directions

As previously mentioned, the embodied effect from Study 2 and previous literature did not replicate in Study 3. The post-hoc analyses implicated a difference in verbal ability between each study's sample which was influential in problem solving, but the relationship between verbal ability and problem solving was not significantly different between the Study 2 and Study 3 samples. It is possible that the disparate effects were driven by a variable that was not measured in this study. One potential variable that may be relevant to this area of research is

Need for Cognition (e.g., Cacioppo et al., 1983). Self-report measures of "I like to Solve Puzzles" were significantly related to problem solving rates, but they did not mediate any relationships between experimental variables and problem solving success rates. This variable is likely related to Need for Cognition, which may encompass similar effects that aid in problem solving. Another potential area of interest may be in motivation. Those who rated themselves higher on liking to solve problems may have been more motivated to solve each problem. This aspect may not be captured by Need for Cognition, so an alternative motivation variable (e.g., success motivation) may be a useful variable for future work.

Again, a significant limitation is that the embodied effect did not replicate from Study 2 to Study 3. This limits the conclusions that can be drawn about the existence of this embodied effect and whether it is a consistent, replicable effect. It should be mentioned that the present work provides unfavorable evidence for the existence of an embodied effect. The only time the effect successfully replicated (Study 2), it appeared to be mediated by creativity. Although there were methodological reasons the effect may not have replicated in Study 1 (nearly all participants chose not to draw their solutions), there is no identifiable reason why it should not have replicated in Study 3. Taken together, the results of this study lend more support to the attentional priming hypothesis – that guiding eye movements improves creativity – rather than fostering a perceptual simulation of the problem. Future research is needed to continue examining this effect, perhaps with a new set of analogical problems with different spatial solutions or a new pattern of eye movement guidance that follows. Furthermore, future researchers should not ignore creativity variables in their studies of problem solving, as it may help researchers understand benefits to problem solving more comprehensively.

It did not appear that there were any benefits for analogical transfer during problem solving as a result of guiding eye movements. It is possible that this effect is not long lasting and may need to be refreshed prior to solving other problems (e.g., re-run the tracking task before attempting each problem). This could mean that perceptual simulations that aid in mental representation construction are also short-lived. Future research should examine whether this effect endures over time or if it is a transitory effect. It is also possible that any eye movement guidance may need to be made more explicit in order for it to be helpful for problem solving and analogical transfer.

Conclusion

The concept of guiding people's eyes to influence the way they think is a potentially potent implicit mechanism that could have its uses (e.g., learning, problem solving) or abuses (marketing). This idea was espoused in previous studies (Grant & Spivey, 2003; Thomas & Lleras, 2007; 2009), but the present work found evidence against its existence through a mediating effect as well as a failure to replicate. Across all three studies, the present work only replicated the effect in one study, and it may have been dependent upon external factors rather than experimentally replicated factors. If this embodied priming effect is legitimate, it is possible that it is inconsistent or potentially not robust enough to emerge with slight methodological changes. Regardless, this effect is worthy of future investigation while evidence for embodied cognitive effects are still emerging.

APPENDIX A: IRB APPROVAL LETTER

University of Central Florida Institutional Review Board
Office of Research & Commercialization
12201 Research Parkway, Suite 501
Orlando, Florida 32826-3246
Telephone: 407-823-2901 or 407-882-2276
www.research.ucf.edu/compliance/irb.html

Approval of Human Research

From: **UCF Institutional Review Board #1**
FWA00000351, IRB00001138

To: **Bradford Schroeder**

Date: **October 13, 2017**

Dear Researcher:

On 10/13/2017 the IRB approved the following human participant research until 10/12/2018 inclusive:

Type of Review:	UCF Initial Review Submission Form Expedited Review
Project Title:	Thinking and Problem Solving
Investigator:	Bradford Schroeder
IRB Number:	SBE-17-13231
Funding Agency:	
Grant Title:	
Research ID:	N/A

The scientific merit of the research was considered during the IRB review. The Continuing Review Application must be submitted 30days prior to the expiration date for studies that were previously expedited, and 60 days prior to the expiration date for research that was previously reviewed at a convened meeting. Do not make changes to the study (i.e., protocol, methodology, consent form, personnel, site, etc.) before obtaining IRB approval. A Modification Form **cannot** be used to extend the approval period of a study. All forms may be completed and submitted online at https://iris.research.ucf.edu .

If continuing review approval is not granted before the expiration date of 10/12/2018, approval of this research expires on that date. When you have completed your research, please submit a Study Closure request in iRIS so that IRB records will be accurate.

Use of the approved, stamped consent document(s) is required. The new form supersedes all previous versions, which are now invalid for further use. Only approved investigators (or other approved key study personnel) may solicit consent for research participation. Participants or their representatives must receive a copy of the consent form(s).

All data, including signed consent forms if applicable, must be retained and secured per protocol for a minimum of five years (six if HIPAA applies) past the completion of this research. Any links to the identification of participants should be maintained and secured per protocol. Additional requirements may be imposed by your funding agency, your department, or other entities. Access to data is limited to authorized individuals listed as key study personnel.

In the conduct of this research, you are responsible to follow the requirements of the Investigator Manual.

On behalf of Sophia Dziegielewski, Ph.D., L.C.S.W., UCF IRB Chair, this letter is signed by:

Kamille Chap

Signature applied by Kamille Chaparro on 10/13/2017 04:33:43 PM EDT

IRB Coordinator

APPENDIX B: CONVERGENCE SOLUTION CODING SCHEME

Coding the Convergence Solution

Adapted from Gick and Holyoak (1980)

A solution is considered a convergence solution for the radiation problem if it has:

1. Rays applied to tumor from different directions
2. At low intensity
3. Simultaneously

This coding scheme can be modified to be more abstract and suitable to analogous problems:

A solution is considered a convergence solution for the radiation problem analogies if it has:

1. Forces applied to central target from different directions
2. At low intensity
3. Simultaneously

APPENDIX C: REUSE AND REPRINT PERMISSIONS

Order Completed

Thank you for your order.

This Agreement between Mr. Bradford Schroeder ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

Your confirmation email will contain your order number for future reference.

printable details

License Number	4442070541016
License date	Oct 04, 2018
Licensed Content Publisher	Elsevier
Licensed Content Publication	Cognitive Psychology
Licensed Content Title	Schema induction and analogical transfer
Licensed Content Author	Mary L. Gick,Keith J. Holyoak
Licensed Content Date	Jan 1, 1983
Licensed Content Volume	15
Licensed Content Issue	1
Licensed Content Pages	38
Type of Use	reuse in a thesis/dissertation
Portion	excerpt
Number of excerpts	2
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Title of your thesis/dissertation	EYE MOVEMENTS AND SPATIAL ABILITY: INFLUENCES ON THINKING DURING ANALOGICAL PROBLEM SOLVING
Expected completion date	Nov 2018
Estimated size (number of pages)	215
Requestor Location	Mr. Bradford Schroeder ORLANDO, FL 32817 United States Attn: Mr. Bradford Schroeder
Publisher Tax ID	98-0397604
Total	0.00 USD

ORDER MORE CLOSE WINDOW

Order Completed

Thank you for your order.

This Agreement between Mr. Bradford Schroeder ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

Your confirmation email will contain your order number for future reference.

printable details

License Number	4442071190639
License date	Oct 04, 2018
Licensed Content Publisher	Elsevier
Licensed Content Publication	Cognitive Psychology
Licensed Content Title	Schema induction and analogical transfer
Licensed Content Author	Mary L. Gick,Keith J. Holyoak
Licensed Content Date	Jan 1, 1983
Licensed Content Volume	15
Licensed Content Issue	1
Licensed Content Pages	38
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Original figure numbers	Figure 1
Title of your thesis/dissertation	EYE MOVEMENTS AND SPATIAL ABILITY: INFLUENCES ON THINKING DURING ANALOGICAL PROBLEM SOLVING
Expected completion date	Nov 2018
Estimated size (number of pages)	215
Requestor Location	Mr. Bradford Schroeder ORLANDO, FL 32817 United States Attn: Mr. Bradford Schroeder
Publisher Tax ID	98-0397604
Total	0.00 USD

ORDER MORE CLOSE WINDOW

Order Completed

Thank you for your order.

This Agreement between Mr. Bradford Schroeder ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

Your confirmation email will contain your order number for future reference.

printable details

License Number	4442430495590
License date	Oct 05, 2018
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Memory & Cognition
Licensed Content Title	Visual representation in analogical problem solving
Licensed Content Author	M. Beveridge, E. Parkins
Licensed Content Date	Jan 1, 1987
Licensed Content Volume	15
Licensed Content Issue	3
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	2
Will you be translating?	no
Circulation/distribution	<501
Author of this Springer Nature content	no
Title	EYE MOVEMENTS AND SPATIAL ABILITY: INFLUENCES ON THINKING DURING ANALOGICAL PROBLEM SOLVING
Instructor name	n/a
Institution name	n/a
Expected presentation date	Nov 2018
Portions	Figures 2 and 3
Requestor Location	Mr. Bradford Schroeder ORLANDO, FL 32817 United States Attn: Mr. Bradford Schroeder
Billing Type	Invoice
Billing address	Mr. Bradford Schroeder ORLANDO, FL 32817 United States Attn: Mr. Bradford Schroeder
Total	0.00 USD

ORDER MORE CLOSE WINDOW

Order Completed

Thank you for your order.

This Agreement between Mr. Bradford Schroeder ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

Your confirmation email will contain your order number for future reference.

printable details

License Number	4442070990643
License date	Oct 04, 2018
Licensed Content Publisher	Elsevier
Licensed Content Publication	Cognition
Licensed Content Title	Covert shifts of attention function as an implicit aid to insight
Licensed Content Author	Laura E. Thomas,Alejandro Lleras
Licensed Content Date	May 1, 2009
Licensed Content Volume	111
Licensed Content Issue	2
Licensed Content Pages	7
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Original figure numbers	Figure 1
Title of your thesis/dissertation	EYE MOVEMENTS AND SPATIAL ABILITY: INFLUENCES ON THINKING DURING ANALOGICAL PROBLEM SOLVING
Expected completion date	Nov 2018
Estimated size (number of pages)	215
Requestor Location	Mr. Bradford Schroeder ORLANDO, FL 32817 United States Attn: Mr. Bradford Schroeder
Publisher Tax ID	98-0397604
Total	0.00 USD

ORDER MORE CLOSE WINDOW

Order Completed

Thank you for your order.

This Agreement between Mr. Bradford Schroeder ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

Your confirmation email will contain your order number for future reference.

printable details

License Number	4442450249444
License date	Oct 05, 2018
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Memory & Cognition
Licensed Content Title	The use of diagrams in analogical problem solving
Licensed Content Author	Roberto Pedone, John E. Hummel, Keith J. Holyoak
Licensed Content Date	Jan 1, 2001
Licensed Content Volume	29
Licensed Content Issue	2
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Will you be translating?	no
Circulation/distribution	501 to 1000
Author of this Springer Nature content	no
Title	EYE MOVEMENTS AND SPATIAL ABILITY: INFLUENCES ON THINKING DURING ANALOGICAL PROBLEM SOLVING
Instructor name	n/a
Institution name	n/a
Expected presentation date	Nov 2018
Portions	Figure 1
Requestor Location	Mr. Bradford Schroeder ORLANDO, FL 32817 United States Attn: Mr. Bradford Schroeder
Billing Type	Invoice
Billing address	Mr. Bradford Schroeder ORLANDO, FL 32817 United States Attn: Mr. Bradford Schroeder
Total	0.00 USD

ORDER MORE | CLOSE WINDOW

Order Completed

Thank you for your order.

This Agreement between Mr. Bradford Schroeder ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

Your confirmation email will contain your order number for future reference.

printable details

License Number	4441521304398
License date	Oct 03, 2018
Licensed Content Publisher	Elsevier
Licensed Content Publication	Cognitive Psychology
Licensed Content Title	Analogical problem solving
Licensed Content Author	Mary L. Gick,Keith J. Holyoak
Licensed Content Date	Jul 1, 1980
Licensed Content Volume	12
Licensed Content Issue	3
Licensed Content Pages	50
Type of Use	reuse in a thesis/dissertation
Portion	excerpt
Number of excerpts	1
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Title of your thesis/dissertation	EYE MOVEMENTS AND SPATIAL ABILITY: INFLUENCES ON THINKING DURING ANALOGICAL PROBLEM SOLVING
Expected completion date	Nov 2018
Estimated size (number of pages)	215
Requestor Location	Mr. Bradford Schroeder ORLANDO, FL 32817 United States Attn: Mr. Bradford Schroeder
Publisher Tax ID	98-0397604
Total	0.00 USD

ORDER MORE CLOSE WINDOW

APPENDIX D: RESEARCH QUESTIONS AND HYPOTHESES FOR EACH STUDY

Study 1 Research Questions and Hypotheses	Supported?
In Study 1, the following research questions were preliminarily tested:	
What duration tracking task is appropriate to yield differences in solution rates?	No effect found
Does spatial ability predict greater success in solving the radiation problem?	
Hypothesis 3: Those with higher spatial ability will perform better on the radiation problem.	No effect found
Hypothesis 3a: Those with higher spatial ability will perform better on the radiation problem, independent of the effects of the experimental manipulations.	No effect found
Hypothesis 3b: The benefits of the experimental manipulations will be more pronounced in those with lower spatial ability, whereas they will not be different among those with higher spatial ability.	No effect found
Does a brief period of incubation aid in the problem solving process?	
Hypothesis 4 (incubation): Participants who view a problem before undergoing a digit-tracking or free-viewing task will perform better on the radiation problem than those who simply go straight into attempting to solve it	No effect found

Study 2 Research Questions and Hypotheses	Supported?
In Study 2, the following research questions were tested:	
Does guiding eye movements foster a perceptual simulation of the radiation problem that leads to improved success rates?	
Hypothesis 1 (Perceptual Simulation): Guiding eye movements will lead to higher solution rates on the radiation problem, independent of any attentional priming effects on creativity.	Partial support, but effect was mediated by one creativity variable
Does guiding eye movements prime creativity that incidentally results in an improvement in problem solving success rates?	
Hypothesis 2 (Attentional Priming): guiding eye movements only primes creativity (not a perceptual simulation of the problem), improving problem solving success rates.	Full support from Elaboration, Partial support from other creativity variables Total Solutions and Redundant Solutions
Hypothesis 2a (Flexibility Hypothesis): guiding eye movements prime divergent thinking, which results in a greater quantity of solutions generated, leading to a greater likelihood of solving the problem	Partial support - guiding eye movements did increase solutions generated, but this increase did not lead to improved problem solving performance
Does spatial ability predict greater success in solving the radiation problem?	
Hypothesis 3: Those with higher spatial ability will perform better on the radiation problem and analogical problems.	Partial support - PFT was significant predictor of success on Red Adair and The General, but not radiation problem
Hypothesis 3a: Those with higher spatial ability will perform better on the radiation problem and analogical problems, independent of the effects of the experimental manipulations.	No effect found
Hypothesis 3b: The benefits of the experimental manipulations will be more pronounced in those with lower spatial ability, whereas they will not be different among those with higher spatial ability.	No effect found
Does a brief period of incubation aid in the problem solving process?	
Hypothesis 4 (incubation): Participants who view a problem before undergoing a digit-tracking or free-viewing task will perform better on the radiation problem than those who simply go straight into attempting to solve it.	Partial support - those who saw their problem before their digit-tracking task generated significantly more solutions to the radiation problem than those who saw it for the first time when they attempted to solve it.

Study 3 Research Questions and Hypotheses	Supported?
In Study 3, the following research questions were tested:	
Does guiding eye movements foster a perceptual simulation of the radiation problem that leads to improved success rates?	
Hypothesis 1 (Perceptual Simulation): Guiding eye movements will lead to higher solution rates on the radiation problem, independent of any attentional priming effects on creativity	No effect found
Hypothesis 1a (Analogical Boost): Those whose eye movements are guided will be more likely to draw a convergence solution to the subsequent analogical problems.	No effect found
Does guiding eye movements prime creativity that incidentally results in an improvement in problem solving success rates?	
Hypothesis 2 (Attentional Priming): guiding eye movements only primes creativity (not a perceptual simulation of the problem), improving problem solving success rates.	No effect found
Hypothesis 2a (Flexibility Hypothesis): guiding eye movements prime divergent thinking, which results in a greater quantity of solutions generated, leading to a greater likelihood of solving the problem	No effect found
Does spatial ability predict greater success in solving the radiation problem?	
Hypothesis 3: Those with higher spatial ability will perform better on the radiation problem and analogical problems	Partial support - CRT was associated with Red Adair Performance
Hypothesis 3a: Those with higher spatial ability will perform better on the radiation problem and analogical problems, independent of the effects of the experimental manipulations	No effect found
Hypothesis 3b: The benefits of the experimental manipulations will be more pronounced in those with lower spatial ability, whereas they will not be different among those with higher spatial ability	Partial support using the PFT on Red Adair - High spatials performed better in the condition without diagrams, but those in the diagram condition were roughly equal, regardless of spatial ability
How does a diagram influence solution rates?	
Hypothesis 5 (Demand Characteristic): Those who attempt to solve the problem on scratch paper with the diagram of the problem will be more likely to draw the convergence solution to the radiation problem, but only for those whose eye movements are guided	No effect found
Hypothesis 6 (Visual Aid): The mere presence of a diagram on scratch paper will boost solution rates, compared to those who receive blank scratch paper.	Full support, but this effect was partially mediated by creative thinking variables.

APPENDIX E: MULTIVARIATE ANALYSES OF VARIANCE FOR CREATIVITY VARIABLES

For each study, a multivariate analysis of variance was performed on the following dependent variables of creativity: Flexibility, Originality, and Elaboration. Two other creativity variables were not used – Total Number of Solutions and Redundant Solutions. Both of these were created out of the Flexibility variable, and due to multicollinearity issues, they were not included in the MANOVA analysis. These variables were assessed to better understand research questions about creative thinking, and are discussed in their respective study's analysis sections.

For Study 1, a MANOVA was used to compare whether any condition's creative thinking scores differed as a function of media type (guided, fixed, free-viewing) or duration (4 minutes, or 7 minutes). There were no significant differences in creative thinking scores as a function of media type (p = .17), duration (p = .47), or their interaction (p = .88).

For Study 2, a MANOVA was used to compare whether any condition's creative thinking scores differed as a function of guidance type (guided, fixed) or problem context (problem before tracking, skip to tracking). There was a significant multivariate effect for guidance type, $F(3, 167) = 4.34$, $p = .006$, Wilks' $\Lambda = .93$, $\eta^2_p = .07$, and a significant multivariate effect for problem context, $F(3, 167) = 3.69$, $p = .013$, Wilks' $\Lambda = .94$, $\eta^2_p = .06$. The interaction between both variables was not statistically significant (p = .31). Upon further examining which creativity variables drove these multivariate effects, eye movement guidance only significantly impacted Elaboration scores, $F(3, 169) = 8.23$, $p = .003$, $\eta^2_p = .05$. The nature of this effect was such that those in the guided eye movements conditions had significantly higher elaboration scores than those in the fixed eye movements conditions. For problem context, there were significant differences identified in Flexibility, $F(3, 169) = 4.66$, $p = .032$, $\eta^2_p = .03$ and Elaboration $F(3, 169) = 7.36$, $p = .007$, $\eta^2_p = .04$. In both cases, those who saw the problem before undergoing

their digit tracking task had higher Flexibility and Elaboration scores than those who only saw the problem after their tracking task.

For Study 3, a MANOVA was used to compare whether any condition's creative thinking scores differed as a function of guidance type (guided, fixed) or diagrams on scratch paper (diagrams present, blank paper). The multivariate effect for guidance type was not statistically significant (p = .34), but there was a significant multivariate effect for diagrams, F(3, 137) = 4.59, p = .004, Wilks' Λ = .91, η^2_p = .09. The interaction between both variables was not statistically significant (p = .61). Upon further examining which creativity variables drove these multivariate effects, diagrams significantly impacted Flexibility scores, F(3, 139) = 6.97, p = .009, η^2_p = .05 and Originality scores, F(3, 139) = 12.02, p = .001, η^2_p = .08. The nature of this effect was such that those in the scratch paper with diagrams conditions had significantly higher Flexibility and Originality scores than those in the blank scratch paper conditions.

REFERENCES

Alibali, M. W., Bassok, M., Olseth, K. L., Syc, S. E., & Goldin-Meadow, S. (1995). Gestures reveal mental models of discrete and continuous change. In P*roceedings of the Seventeenth Annual Conference of the Cognitive Science Society , 17,* 391-396.

Antonietti, A., & Gioletta, M. A. (1995). Individual differences in analogical problem solving. *Personality and Individual Differences, 18*(5), 611-619.

Barsalou, L. W. (1999). Perceptions of perceptual symbols. *Behavioral and Brain Sciences, 22*(4), 637-660.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59*, 617-645.

Bassok, M. (2003). Analogical transfer in problem solving. In J.E. Davidson & R. J. Sternberg (Eds.)*The psychology of problem solving* (pp. 343-369). New York: Cambridge University Press.

Beveridge, M., & Parkins, E. (1987). Visual representation in analogical problem solving. *Memory & Cognition*, *15*(3), 230-237.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis.* Chichester, England: Wiley

Carroll, J. B. (1993). Abilities in the domain of visual perception. In J. B. Carrol (Ed.), *Human cognitive abilities: A survey of factor-analytic studies*. (pp. 304-363). Cambridge, UK: Cambridge University Press.

Catrambone, R., Craig, D. L., & Nersessian, N. J. (2006). The role of perceptually represented structure in analogical problem solving. *Memory & Cognition*, *34*(5), 1126-1132.

Chi, M. T. & Glaser, R. (1985). Problem solving ability. In R. Sternberg (Ed.), *Human abilities: An information processing approach* (pp.227-250). San Francisco, CA: Freeman.

Chi, M. T., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, *5*(2), 121-152.

Chi, M. T., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, *13*(2), 145-182.

Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis, 58*, 7-19.

Cooperrider, K., & Goldin-Meadow, S. (2014, January). The role of gesture in analogical problem solving. In *Proceedings of the Cognitive Science Society*, *36*, 2068-2072.

Corkill, A. J., & Fager, J. J. (1995). Individual differences in transfer via analogy. *Learning and Individual Differences*, *7*(3), 163-187.

Craig, D. L., Nersessian, N. J., & Catrambone, R. (2002). Perceptual simulation in analogical problem solving. In L. Magnani & N. J. Nersessian (Eds.) *Model-Based Reasoning: Science, Technology, & Values* (pp.167-191). New York: Kluwer Academic/Plenum Publishers.

Davidson, J. E., & Sternberg, R. J. (1998). Smart problem solving: How metacognition helps. *Metacognition in educational theory and practice*, 47-68.

Dow, G. T., & Mayer, R. E. (2004). Teaching students to solve insight problems: Evidence for domain specificity in creativity training. *Creativity Research Journal*, *16*(4), 389-398.

Duncker, K. (1945). On problem-solving (L. S. Lees, Trans.). *Psychological Monographs, 58*(5), i-113.

Ekstrom, R., French, J., Harman, H., & Dermen, D. (1976). *Kit of factor-referenced cognitive tests.* Princeton, NJ: Educational Testing Services.

Eliot, J., & Smith, I. M. (1983). *An international directory of spatial tests*. Cengage Learning Emea.

Farah, M. J., Hammond, K. M., Levine, D. N., & Calvanio, R. (1988). Visual and spatial mental imagery: Dissociable systems of representation. *Cognitive Psychology*, *20*(4), 439-462.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175-191.

Fremont, M. (2018, March 1). *Grey's Anatomy recap: The new girl in town.* Retrieved from http://www.vulture.com/2018/03/greys-anatomy-recap-season-14-episode-13.html

French, J. W. (1951). The description of aptitude and achievement tests in terms of rotated factors. *Psychometric Monographs, 5*.

Friedman, R. S., Fishbach, A., Förster, J., & Werth, L. (2003). Attentional priming effects on creativity. *Creativity Research Journal*, *15*(2-3), 277-286.

Gick, M. L. (1985). The effect of a diagram retrieval cue on spontaneous analogical transfer. *Canadian Journal of Psychology, 39*(3), 460.

Gick, M. L. (1989). Two functions of diagrams in problem solving by analogy. In H. Mandl & J. R. Levin (Eds.), *Advances in psychology, 58. Knowledge acquisition from text and pictures* (pp. 215-231). North Holland, Amsterdam: Elsevier.

Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology, 12*(3), 306-355.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15*(1), 1-38.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, *7*(2), 155-170.

Gentner, D., Brem, S., Ferguson, R., Wolff, P., Markman, A. B., & Forbus, K. D. (1997). Analogy and creativity in the works of Johannes Kepler. In T. B. Ward, S.M. Smith & J. Vaid (Eds.) *Creative thought: An investigation of conceptual structures and processes* (pp. 403-459). Washington, DC: American Psychological Association

Grant, E. R., & Spivey, M. J. (2003). Eye movements and problem solving: Guiding attention guides thought. *Psychological Science, 14*(5), 462-466.

Guilford, J. P. (1967). The nature of human intelligence. New York: McGraw Hill

Hayes, A. F. (2012). PROCESS: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling [White paper].

Hegarty, M., & Kozhevnikov, M. (1999). Types of visual–spatial representations and mathematical problem solving. *Journal of Educational Psychology*, *91*(4), 684-689.

Hegarty, M., & Sims, V. K. (1994). Individual differences in mental animation during mechanical reasoning. *Memory & Cognition*, *22*(4), 411-430.

Holyoak, K. J. (1995). Problem solving. In E. E. Smith and D. N. Osherson (Eds.) *Thinking: An invitation to cognitive science* (pp. 267-296). Cambridge, MA: MIT Press

Holyoak, K. J., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory & Cognition, 15*(4), 332-340.

Hostetter, A. B., & Alibali, M. W. (2008). Visible embodiment: Gestures as simulated action. *Psychonomic Bulletin & Review, 15*(3), 495-514.

Hostetter, A. B., Wieth, M. B., Foster, K. D., Moreno, K., & Washington, J. (2016) Effects of gesture on analogical problem solving: When the hands lead you astray. In *Proceedings of the Thirty eighth Annual Conference of the Cognitive Science Society, 38*, 1685-1690.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological review*, *87*(4), 329.

Just, M. A., & Carpenter, P. A. (1985). Cognitive coordinate systems: Accounts of mental rotation and individual differences in spatial ability. *Psychological review*, *92*(2), 137.

Kalyuga, S. (2009). The expertise reversal effect. In *Managing Cognitive Load in Adaptive Multimedia Learning* (pp. 58-80). IGI Global.

Kasof, J. (1997). Creativity and breadth of attention. *Creativity Research Journal*, *10*(4), 303-315.

Kirsh, D. (2010). Thinking with external representations. *Ai & Society, 25*(4), 441-454.

Knoblich, G., Ohlsson, S., Haider, H., & Rhenius, D. (1999). Constraint relaxation and chunk decomposition in insight problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(6), 1534-1555.

Knoblich, G., Ohlsson, S., & Raney, G. E. (2001). An eye movement study of insight problem solving. *Memory & Cognition, 29*(7), 1000-1009.

Kosslyn, S. (1990). Mental imagery. In D. Osherson, S. Kosslyn, & J. Hollerbach (Eds.), *Visual cognition and action* (pp. 73-98). Cambridge, MA: MIT Press.

Kozhevnikov, M., Hegarty, M., & Mayer, R. E. (2002). Revising the visualizer-verbalizer dimension: Evidence for two types of visualizers. *Cognition and Instruction, 20*(1), 47-77.

Kubricht, J. R., Lu, H., & Holyoak, K. J. (2017). Individual differences in spontaneous analogical transfer. *Memory & Cognition, 45*(4), 576-588.

Logie, R. H. (1995). *Visuo-spatial working memory.* Hove, UK: Lawrence Erlbaum.

Lohman, D. F. (1996). Spatial ability and g. In I. Dennis & P. Tapsfield (Eds.), *Human abilities: Their nature and measurement* (pp. 97-116). Mahwah, NJ: Erlbaum.

Lucas, É. (1882). *Récréations mathématiques* (Vol. 1). Gauthier-Villars.

Maier, N. R. F. (1930). Reasoning in humans. I. On direction. *Journal of Comparative Psychology, 10*(2), 115-143.

Maier, N. R. F. (1931). Reasoning in humans. II. The solution of a problem and its appearance in consciousness. *Journal of Comparative Psychology, 12*(2), 181-194.

Mandler, J. M., & Cánovas, C. P. (2014). On defining image schemas. *Language and Cognition*, *6*(4), 510-532.

Mayer, R. E., & Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology*, *86*(3), 389-401.
McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, *22*(3), 276-282.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Norman, D. A. (1986). Reflections on cognition and parallel distributed processing. *Parallel distributed processing: Explorations in the microstructure of cognition*, *2*, 531-546.

Novick, L. R., & Sherman, S. J. (2003). On the nature of insight solutions: Evidence from skill differences in anagram solution. *The Quarterly Journal of Experimental Psychology: Section A*, *56*(2), 351-382.

Ormerod, T. C., MacGregor, J. N., & Chronicle, E. P. (2002). Dynamics and constraints in insight problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(4), 791-799.

Pedone, R., Hummel, J. E., & Holyoak, K. J. (2001). The use of diagrams in analogical problem solving. *Memory & Cognition*, *29*(2), 214-221.

Pellegrino, J. W., Alderton, D. L., & Shute, V. J. (1984). Understanding spatial ability. *Educational Psychologist*, *19*(4), 239-253.

Piffer, D. (2012). Can creativity be measured? An attempt to clarify the notion of creativity and general directions for future research. *Thinking Skills and Creativity*, *7*(3), 258-264.

Raven, J. (2000). *Manual for Raven's Progressive Matrices and Vocabulary Scales: Research Supplement No. 3, 2000 Ed*. Oxford Psychologists Press.

Resnick, L. B., & Glaser, R. (1976). Problem Solving and Intelligence. In L. B. Resnick (Ed.) *The nature of intelligence* (pp. 205-230). Hillsdale, NJ: Erlbaum.

Rozenblit, L., Spivey, M.J., & Wojslawowicz, J. (2002). Mechanical reasoning about gear-and-belt systems: Do eye movements predict performance? In M. Anderson, B. Meyer, & P. Olivier (Eds.), *Diagrammatic representation and reasoning* (pp. 223– 240). Berlin: Springer-Verlag.

Salthouse, T. A., Babcock, R. L., Mitchell, D. R. D., Palmon, R., Skovronek, E. (1990). Sources of individual differences in spatial visualization ability. *Intelligence*, *14*, 187–230.

Schooler J. W., Melcher J. (1995). The ineffability of insight. In S. M. Smith, T. B. Ward, R. A. Finke (Eds.) *The Creative Cognition Approach* (pp.97-133). Cambridge, MA: MIT Press.

Segal, E. (2004). Incubation in insight problem solving. *Creativity Research Journal*, *16*(1), 141-148.

Shepard, R. N. (1978). Externalization of mental images and the act of creation. In B. S. Randhawa and W. E. Coffman (Eds.), *Visual learning, thinking, and communication,* (pp. 135-189). New York: Academic Press.

Silvia, P. J., & Nusbaum, E. C. (2012). What's your major? College majors as markers of creativity. *The International Journal of Creativity and Problem Solving*, *22*(2), 31-44.

Sio, U. N., & Ormerod, T. C. (2009). Does incubation enhance problem solving? A meta-analytic review. *Psychological Bulletin*, *135*(1), 94-120.

Smith, S. M., & Blankenship, S. E. (1991). Incubation and the persistence of fixation in problem solving. *The American journal of psychology*, *104*, 61-87.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of personality and social psychology*, *69*(5), 797.

Sternberg, R. J. (1977). Component processes in analogical reasoning. *Psychological Review*, *84*(4), 353-378.

Sternberg, R. J., & Gardner, M. K. (1983). Unities in inductive reasoning. *Journal of Experimental Psychology: General*, *112*(1), 80.

Thibaut, J. P., & French, R. M. (2016). Analogical reasoning, control and executive functions: a developmental investigation with eye-tracking. *Cognitive Development*, *38*, 10-26.

Thomas, L. E., & Lleras, A. (2007). Moving eyes and moving thought: On the spatial compatibility between eye movements and cognition. *Psychonomic Bulletin & Review*, *14*(4), 663-668.

Thomas, L. E., & Lleras, A. (2009). Covert shifts of attention function as an implicit aid to insight. *Cognition*, *111*(2), 168-174.

Trowbridge, F. A. (2016). *Action and analogical problem solving*. (Master's Thesis). North Carolina State University, Raleigh, NC.

Vartanian, O. (2009). Variable attention facilitates creative problem solving. *Psychology of Aesthetics, Creativity, and the Arts*, *3*(1), 57.

Vartanian, O., Martindale, C., & Kwiatkowski, J. (2007). Creative potential, attention, and speed of information processing. *Personality and Individual Differences*, *43*(6), 1470-1480.

Vincent, A. S., Decker, B. P., & Mumford, M. D. (2002). Divergent thinking, intelligence, and expertise: A test of alternative models. *Creativity Research Journal*, *14*(2), 163-178. doi:10.1207/S15326934CRJ1402_4

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: a meta-analysis and consideration of critical variables. *Psychological Bulletin*, *117*, 250-270.

Wegbreit, E., Suzuki, S., Grabowecky, M., Kounios, J., & Beeman, M. (2012). Visual attention modulates insight versus analytic solving of verbal problems. *The Journal of Problem Solving*, *4*(2), 94.

Weisberg, R. W., & Alba, J. W. (1981). An examination of the alleged role of" fixation" in the solution of several" insight" problems. *Journal of Experimental Psychology: General*, *110*(2), 169.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, *9*(4), 625-636.

Wood, D., Bruner, J., & Ross, G. (1976). The role of tutoring in problem solving. Journal of Child Psychology and Psychiatry, 17(2), 89-100.

Zabelina, D., Saporta, A., & Beeman, M. (2016). Flexible or leaky attention in creative people? Distinct patterns of attention for different types of creative thinking. *Memory & Cognition*, *44*(3), 488-498.