Old Dominion University
ODU Digital Commons

Psychology Theses & Dissertations | Psychology

Fall 2015

Support for Gestalt Versus Business-As-Usual Theories of Insight Depends on Operational Definition of Insight

Kimberly D. Lee
Old Dominion University

Follow this and additional works at: https://digitalcommons.odu.edu/psychology_etds

Part of the Cognitive Psychology Commons

Recommended Citation
Lee, Kimberly D.. "Support for Gestalt Versus Business-As-Usual Theories of Insight Depends on Operational Definition of Insight" (2015). Master of Science (MS), thesis, Psychology, Old Dominion University, DOI: 10.25777/h2b3-mv17
https://digitalcommons.odu.edu/psychology_etds/17

This Thesis is brought to you for free and open access by the Psychology at ODU Digital Commons. It has been accepted for inclusion in Psychology Theses & Dissertations by an authorized administrator of ODU Digital Commons. For more information, please contact digitalcommons@odu.edu.

SUPPORT FOR GESTALT VERSUS BUSINESS-AS-USUAL THEORIES OF INSIGHT DEPENDS ON OPERATIONAL DEFINITION OF INSIGHT

by

Kimberly Diane Lee
B.A. May 2002, Christopher Newport University
B.S. May 2012, Old Dominion University

A Thesis Submitted to the Faculty of
Old Dominion University in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

PSYCHOLOGY

OLD DOMINION UNIVERSITY
December 2015

Approved by:

Ivan K. Ash (Director)

Elaine M. Justice (Member)

James P. Bliss (Member)

ABSTRACT

SUPPORT FOR GESTALT VERSUS BUSINESS-AS-USUAL THEORIES OF INSIGHT DEPENDS ON OPERATIONAL DEFINITION OF INSIGHT

Kimberly Diane Lee
Old Dominion University, 2015
Director: Dr. Ivan K. Ash

Some theories propose that insight involves automatic processes that are responsible for restructuring. Other theories postulate that the mechanisms surrounding restructuring are controlled and effortful. The current study tested these theories by comparing different methodology and operational definitions that have been used in previous research to investigate the nature of "Aha!" experiences and impasse in insightful problem solving. One hundred two undergraduate psychology students from Old Dominion University completed working memory tasks, six classic insight problems, and gave initial problem representation ratings for the insight problems before solution attempt. Using a think-aloud protocol, we assessed the occurrence of impasse during the problem solving phase. After solving each problem, participants completed self-reported, measures of the Aha! experience—solution confidence, how sudden a solution appeared, and the effort required. Results demonstrated distinctly different response patterns between self-reported ratings of insight and the empirically coded measure of impasse when compared with all other variables of interest. This suggests that the Aha! ratings lack construct validity as an assessment of insight. Further, we replicated contradictory working memory correlations found in previous research with the self-report ratings and impasse coding, suggesting that discrepancies in the literature were the result of how insight was assessed. These findings call into question previous research utilizing self-report Aha! ratings.

Keywords: insight, problem solving, working memory, impasse, Gestalt psychology

© 2015 Kimberly Diane Lee. All Rights Reserved.

hoc opus dedicatur ad familiam meam et cervisiam quae perduxit me illuc

ACKNOWLEDGMENTS

Foremost, I would like to thank my advisor, Dr. Ivan Ash, for sharing his expertise and guiding me through this process, particularly in how to better communicate my work both in writing and presentation. I am also grateful for the impact his proficiency and dedication to rigorous experimental design has had on my own personal convictions regarding the science of psychology.

I would also like to extend my thanks to Dr. Elaine Justice and Dr. James Bliss for their service on my thesis committee. Their encouragement, insightful discussion, and discerning comments have greatly improved the quality of this manuscript as well as expanded my perspective regarding this research.

I would also like to acknowledge the hard work and dedication of my research assistants, Aekta Javia, Alonzo Anderson, Holly Fitzgerald, Jacqueline Guzman, and Kimberly Sears. Even the long hours of coding video data did not dampen their enthusiasm for this project. It was a joy to see them get excited over new discoveries.

And finally, words cannot express my gratitude to my family—my 'ohana—for their continued love and support.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Everyone has experienced sudden insight—that flash of understanding that seems to come out of nowhere, from a place where no progress was being made; the *Aha!* moment. These experiences feel qualitatively different from non-insightful problem solving, e.g., solving an arithmetic equation, and researchers have long investigated whether these two types of problem solving involve qualitatively different cognitive processes. And if so, whether the feelings associated with the Aha! experience are truly indicative of the underlying cognitive mechanisms.

Theories of insight have evolved into two competing classes: theories that stem from Gestalt psychologists who proposed special, automatic processes that are specific to insightful solutions (Ash & Wiley, 2006; Bowden & Jung-Beeman, 2003; Fleck, 2008; Stellan Ohlsson, 2011) and theories that propose no differences between insightful and non-insightful problem solving (Chein & Weisberg, 2013; Davidson, 1995; Gilhooly & Murphy, 2005; Kaplan & Simon, 1990). As detailed in Ash, Cushen, and Wiley (2009), difficulty in researching the insight phenomenon has resulted in various methodologies and operational definitions of insight. We propose that these differences are responsible for some of the conflicting evidence produced in the literature. The current study aimed to reconcile these differences by testing predictions based on competing theory while directly comparing several operational definitions of insight.

Historical Perspective

As an emerging scientific field, psychology was rooted in Associationist views that became increasingly stringent over time (Ash, Jee, & Wiley, 2012). Culminating with Behaviorism, this approach was confined to observable behaviors and their contingent causes from the external world. Internal mental processes that might affect learning and behavior were dismissed as unknowable or, at the most extreme, nonexistent. Behaviorists attempted to explain all learning as a gradual, passive process in which co-occurring stimuli become associated with a desired outcome. Thorndike's (1911) famous "puzzle box" experiments exemplify this type of learning. For these experiments, Thorndike created special puzzle boxes in which a cat was required to activate some type of mechanism, e.g., a wire pull or a foot treadle, to gain freedom. To begin, the cat would try many ineffectual means to escape the box. Eventually, it would randomly release the correct mechanism and was rewarded with escape from confinement. When returned to the box, the cat would again engage in ineffectual activities until randomly alighting on the correct mechanism. Over time, the random activities decreased as the cat slowly learned to associate the correct mechanism with the door release.

It was from this constrained backdrop of Associationist learning theory that the Gestalt movement was born. Gestalt theorists were dissatisfied with the notion that human behavior could be wholly explained by passive, external connections. They proposed that the internal representation of a problem or situation was paramount to the external environment. Koffka (1935) spoke of two types of environments that influence behavior: the *geographical* and the *behavioral*. The geographical environment encompasses the physical world and all the specific elements therein, e.g., the color, texture, and weight of surrounding objects. The behavioral environment is how a person *interprets* this information to form a mental representation which can depend on many things including what they choose to focus on in the environment, the goals that are motivating them, and their knowledge and previous experience with the environment or its objects.

It is the changes that take place in the behavioral environment that were of particular interest to the Gestalt psychologists. They believed that to solve certain types of problems, a reorganization of the initial mental representation must occur, a process known as *restructuring* (Duncker & Lees, 1945; Maier, 1931; Wertheimer, 1954/1959). They also proposed that *insight*—the sudden realization of a problem's solution—follows restructuring. Köhler (1948) was the first to demonstrate what appeared to be this type of insightful learning. In one of his best-known experiments, Köhler provided a chimpanzee, Sultan, with two hollow rods and bananas that were just out of reach from his cage. The bananas could not be obtained by using either rod individually, but could be reached by inserting the end of one rod into the other to create one long rod. Sultan unsuccessfully attempted to obtain the bananas by using one of the short rods, a strategy that had previously worked. He continued in this vein for an hour before giving up, having reached *impasse*, a point in which active problem solving ceases and the solver is unsure how to proceed. While sitting, he examined the rods and happened to hold them in a straight line. Upon seeing the line, he inserted the end of one rod into the other and immediately went to retrieve the bananas. In Gestalt parlance, the unsuccessful, single-rod technique was part of the initial, inappropriate problem representation which was later, after impasse, restructured to include the successful, double-rod strategy. Unlike Thorndike's cats, which required numerous trials to gradually associate the stimulus with the desired outcome, Sultan appeared to experience one-trial learning as he immediately connected the poles in future attempts. One-trial learning is at odds with Behaviorist learning theory and Köhler's experiments lent credence to the idea of insightful learning.

Gestalt psychologists turned their attention to insightful problem solving in humans, which required the development of laboratory techniques that would allow for insight to occur. Insight problems are a particular type in which the problem itself typically induces an inappropriate mental problem representation. The creation of this faulty representation goes beyond the scope of the actual parameters presented in the problem. Prior problem solving experience leads a solver to make assumptions about the

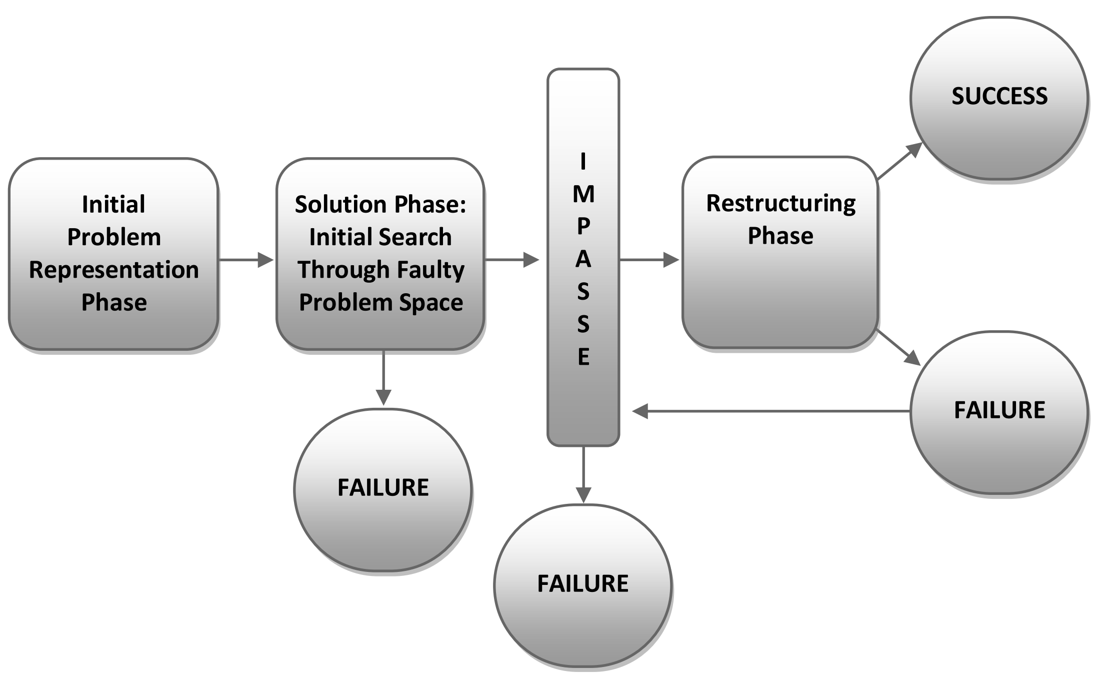


Figure 1. The Gestalt theory of insight problem solving presented in an information processing framework (Ash & Wiley, 2006).

constraints of the problem space as well as additional operations or rules that must be followed. The correct solution is not obtainable from this inappropriate representation and the solver may reach a point in which they are no longer making progress towards the goal state. After reaching this impasse, the problem representation must be restructured in order to realize the correct solution (see Figure 1 for a diagram of the Gestalt view of insight problem solving presented within an information processing framework).

Duncker's (1945) candle problem experiment illustrates this type of insight inducing problem. In this experiment, participants were given candles, a matchbox with matches, and some tacks. They were asked to affix the candle to the wall without allowing wax to drip. In one condition, participants were given the matchbox with the matches inside while the other group received an empty matchbox with the matches separate. He found that participants were much more likely to solve the puzzle—tacking the matchbox to the wall for use as a shelf for the candle—if they received an empty matchbox. The full matchbox reinforced the function of the matchbox as a container, an inappropriate mental representation for this problem reinforced by previous learning, and fixation on this function made it more difficult for the participants to restructure their problem representation to find the correct solution. Early researchers continued to make contributions supporting the Gestalt concept of insight (Maier, 1931; Wertheimer, 1954/1959), but were not without detractors anchored in Associationist theory. This conflict is echoed in modern insight problem solving research.

The Current State of Insight Problem Solving Research

Currently, there is a division in the literature regarding the processes that underlie insight problem solving, specifically, restructuring. Restructuring is typically unnecessary in analytic, non-insight problems as the goal state can be reached through strategic means from the initial, correct problem representation (e.g., arithmetic, Tower of Hanoi task). As such, it is key to investigating processes that may be specific to insight. Some researchers have proposed that automatic processes, such as spreading activation in semantic memory (Ohlsson, 1992; 2011), chunk decomposition and constraint relaxation (Knoblich, Ohlsson, Haider, & Rhenius, 1999), and switching between fine- and course-grain processing (Jung-Beeman et al., 2004), are responsible for restructuring (Ash & Wiley, 2006; Fleck, 2008; Schooler, Ohlsson, & Brooks, 1993). Others subscribe to the *Business-As-Usual* theory which posits that insightful problem solving is no different from analytic problem solving and the mechanisms involved in restructuring are controlled and effortful (Chein & Weisberg, 2013; Davidson, 1995; Gilhooly & Murphy, 2005; Kaplan & Simon, 1990). To facilitate discussion, Gilhooly and Murphy (2005) provided a useful framework for this debate: a dual process model.

Dual process models have been established as viable approaches to conceptualize thinking (Evans, 2003, 2011; Kahneman, 2003; Stanovich & West, 2000). These models distinguish between two systems of processing. System 1 is associated with automatic, fast processing which is not under conscious control or subject to the limitations of working memory. System 2 is associated with deliberate, sequential processing and *is* constrained by the demands of working memory. Evans (2008) advocated a semantic change that relates this division as *types* of processing as opposed to systems of processing. He reasoned that though the processes associated with System 2 are relatively stable across the literature, the processes labeled as System 1 are widely varied and better conceptualized as types of processes that fall under the umbrella of the attributes associated with it.

Investigations specific to the debate surrounding the Type 1 or Type 2 processes involved in insight problem solving often depend on working memory performance. Working memory is a system that allows for temporary storage and management of information (Engle, Tuholski, Laughlin, & Conway, 1999). It is constrained by a very limited capacity and is typically measured by performance on a primary task while concurrently managing a secondary task. This methodology is used to measure an individual's capacity for controlled processes, such as focusing attention. By correlating working memory, the hallmark of Type 2 processing, to problem solving performance or to certain stages of the problem solving process, researchers can infer whether Type 1 or Type 2 processes are dominant.

Working Memory and Restructuring

Ash and Wiley (2006) recognized that Type 2 processes are involved in the insight problem solving process even within the scope of automatic restructuring theories. They argued that insightful solutions begin with the same strategic, analytic procedures used in non-insightful problem solving which are employed within the initial search phase through the faulty problem space. These procedures utilize Type 2 processes as the solver directs attentional resources towards accomplishing their goal. Accordingly, both automatic restructuring theories and Type 2 restructuring theories allow for the impact of individual differences in working memory on insight problem solving ability when considering the *entire* process, though only Type 2 restructuring necessitates it. To investigate restructuring within this line of reasoning, Ash and Wiley isolated the restructuring phase with an experimental approach. They designed two sets of matched insight problems. Matching problems shared the same instructions, but more importantly, shared the same representational change that must occur during restructuring to find the correct solution, i.e., the problems were solved in the same manner. The difference between the two sets was that one set was simplified so that the initial faulty problem space was small and exhausted almost immediately. The other set allowed for many possible moves within the initial faulty problem space so that it took more time before the solver exhausted the problem space and came to impasse. By effectively eliminating the pre-impasse stage from the simplified set, Ash and Wiley could make inferences about the restructuring phase by comparing solving ability from the two sets with working memory capacity. They found that working memory correlated to solution rates for the large initial problem space set, with higher working memory span participants solving more insight problems, but did not correlate with the small initial problem space set. As Type 2 restructuring would predict a relationship between working memory and both problem sets, these results were interpreted as relating Type 2 processes to the initial, faulty search phase, but not with the restructuring phase, supporting the Gestalt theory of automatic restructuring.

Ricks, Turley-Ames, and Wiley (2007) explored the relationship between working memory and compound remote associates (CRA) problems. These problems were originally developed for the Remote Associates Test by Mednick (1968) as a way to test creative problem solving. As this type of problem can be experienced as either an insight or non-insight problem depending on individual differences, they have become a popular means of studying insight (e.g., Bowden & Jung-Beeman, 2003; Cunningham, MacGregor, Gibb, & Haar, 2009; Jung-Beeman et al., 2004; Kounios et al., 2006). In a CRA problem, the solver is presented with a set of three seemingly unrelated words, e.g., *aid, rubber, wagon*. They are then tasked with finding a fourth word that will pair with each of the three cue words. In this example, *band* is the correct solution: *band-aid, rubber band, bandwagon*. Successful CRA solutions require the solver to efficiently search long term memory stores to produce solution attempts, reject inappropriate solutions, and keep the inappropriate solutions from interfering with subsequent searches. As these processes are managed by Type 2 mechanisms, Ricks et al. argued that individual differences in working memory would predict CRA performance. But they were particularly interested in the idea that increased working memory capacity could sometimes handicap the CRA problem solving process. They posited that the increased attentional focus afforded to participants with higher working memory spans could amplify inappropriate response fixation.

To test this hypothesis, Ricks et al. used two sets of CRA problems: a neutral set and a set that was designed to mislead the solver into selecting an inappropriate, knowledge domain-oriented response. They then explored the effect of working memory in solvers who had differential levels of domain knowledge on these two sets. They found that working memory did predict solution success in the neutral set as well as with solvers with low domain knowledge in the domain misleading set. But solution success was significantly hampered in the domain misleading set for participants with high working memory spans and high levels of domain specific knowledge, who solved even fewer CRA problems than participants with average working memory. These results confirmed that Type 2 processes are involved in CRA problem solving, as well as indicated that the working memory sub-processes responsible for individual differences are linked to attentional focus. However, it is important to note that this study examined CRA problems in general and did not attempt to differentiate between insightful and non-insightful solutions.

Fleck (2008) compared insight and analytic problems using measures of fluid intelligence, verbal and spatial short term memory, and working memory. Restructuring, defined here as representational change from the initial problem representation to the final solution, was assessed through a verbal protocol in which solvers verbalized their problem solving process and through retrospective reports. To begin, Fleck confirmed that restructuring took place at a high frequency in the insight problems, but low frequency in the analytic problems. She also found that working memory strongly predicted solution success in analytic problems, but not in insight problems. However, insight solutions were predicted by verbal short term memory, a component of working memory, which was not predictive of analytic success. These results provided evidence of distinct processing between analytic and insightful solutions and were interpreted as support for automatic restructuring theories.

Gilhooly and Fioratou (2009) also found support for process differences between insightful and analytic problem solving. Using a large problem set—much larger than previously used in this line of research—they investigated the relationship between problem type and verbal working memory, spatial working memory, inhibition, and switching. They found that both verbal and spatial working memory predicted solution success in insight problems, but the executive functions of inhibition and switching did not. For analytic problems, they found that working memory as well as switching contributed to problem solving. These processes differences were interpreted, in part, as support for automatic restructuring. However, it is important to note that Gilhooly and Fioratou relied on previous research to define insightful versus analytic problems and did not attempt to assess the actual occurrence of restructuring.

Chein and Weisberg (2013) used a self-report method to identify the occurrence of insight in CRA problems. After solving a problem, they asked participants to rate their problem solving experience on a 4-point Aha! rating scale anchored at *Strategy* and *Insight*. They explained that a problem solved by strategy is one in which the solver comes upon the correct solution unawares, and only through an effortful, strategic process, e.g., a trial-and-error process, can the solver confirm the accuracy of the solution. Sudden insight "means that as soon as you thought of the word, you knew that it was the answer. The solution word came with a feeling that it was correct ('It popped into my head'; 'Of course!'; 'I had an Aha!')" (Appendix B, para. 1-3). Responses were then used to categorize problems as analytic or insightful. They found that working memory positively correlated to solution success in problems classified as insightful, a finding they argued indicated the presence of Type 2 processes throughout the entire problem solving process. They also found that solution times were significantly faster for problems identified as involving insight. This is at odds with the Gestalt view which predicts slower solving times for insightful solutions since a solver must first exhaust a faulty problem space before obtaining the solution. These results were interpreted as support for Type 2 restructuring and the Business-As-Usual theory.

Assessing Restructuring

Insight problem solutions are often distinguished by the Aha! experience. Unlike analytic problems in which incremental progress is felt, insight solutions seem to occur unexpectedly with little to no prior feeling of progress being made towards the goal state (Metcalfe & Wiebe, 1987). This sudden solution realization can be accompanied by a strong feeling that is typically described as an Aha! experience or a *Eureka!* moment. Because of the challenges inherent in experimentally isolating stages of the problem solving process or even in identification of the insight phenomenon, this subjective Aha! experience is often used to determine the occurrence of insight, and more specifically restructuring.

Bowden and Jung-Beeman (2003) used these feelings to establish correlations with neurological activity during CRA problem solving, developing the first instance of the Aha! rating scale (also, Jung-Beeman et al., 2004). After solving each problem, they asked solvers to rate their problem solving experience on a single continuum in terms of strategy versus insight. They found sudden bursts of brain activity just prior to solution realization in problems that were identified as occurring with an Aha! experience, but not with problems that were solved strategically. As the activity preceded the solution, these bursts were interpreted as an automatic process occurring beyond the solvers control.

As previously described, Chein and Weisberg (2013) used a streamlined version of this measure in their work. But this Aha! rating scale has the potential to create confounded results. The anchors do not reflect a single construct leaving room for subjective interpretations and response confusion that would not be available in a scale anchored between extremes, e.g., *Not* and *Very*. Further, several constructs are inherent in the instructions: solution confidence, how sudden a solution manifests, and the amount of effort involved in the process. If these constructs are not highly related, results would be influenced by which component felt most pertinent at the time of the rating. For example, a participant who happens to solve a problem quickly and easily, without impasse or restructuring, would likely choose an insight rating because they did not feel the process was effortful or that they actively employed a strategy. Or perhaps they interpreted suddenness in terms of how quickly they obtained the answer. Conversely, a solver who realizes the solution after impasse and restructuring, may choose a strategic rating because they expended a lot of mental resources and tried many strategies before coming to the solution.

At no point in previous research has the validity been assessed for this scale. Even if the three constructs inherent in the scale consistently relate with one another when assessing insight, the subjective feelings associated with the Aha! experience may not be reliable indicators of restructuring. Ash et al. (2009) argued against conflating the experience of an Aha! moment with the process of restructuring. Gleaning evidence from previous research, they highlighted vulnerabilities implicit when solvers are asked to assess their own metacognitive processes. Further, no information processing theory of automatic restructuring includes *feelings* as mechanisms of the insight problem solving process. Rather, these Aha! feelings are an epiphenomenon that be explained by the theories.

However, all automatic restructuring theories denote the importance of impasse in activating the restructuring process that must occur in insightful solutions. As such, Ash et al. (2012) chose to operationally define the occurrence of insight as solution after impasse. In this study, impasse was empirically assessed using a think-aloud protocol. Using both insight problems and arithmetic problems, they found that this method of coding the restructuring phase confirmed predictions based on the Gestalt theory of insight. They assessed initial problem representation through importance ratings administered prior to the solving phase and found that problems solved without impasse began with a more appropriate initial problem representation than problems solved with impasse, confirming the assumption that insightful solutions begin with a faulty problem representation. They also found that problems involving impasse took markedly longer to solve than problems solved without impasse. This result supported Gestalt theory, but was directly contradicted by Chein and Weisberg's (2013) finding that problems involving an Aha! experience were solved much faster than those categorized as analytical. Lastly, Ash et al. found that none of these observations were echoed in the data from the arithmetic problems indicating a qualitative distinction between analytical and insightful problem solving.

Current Study

We propose that conflicting evidence in the literature regarding working memory and the insight process is a product of how insightful solutions have been assessed. Ash and Wiley (2006) isolated the restructuring phase by manipulating the initial problem space, but did not attempt to confirm the occurrence of restructuring. Ricks et al. (2007) did not attempt to assess restructuring at all. Fleck (2008) attempted to assess restructuring, but chose representational change as the operational definition instead of solution after impasse. Gilhooly and Fioratou (2009) relied on previous research to differentiate between insightful and analytic problems and did not attempt to confirm the occurrence of restructuring. Chein and Weisberg (2013) attempted to assess restructuring, but used a self-report method that has not been validated, finding results that contradicted previous research.

We also propose that the self-report Aha! rating scale is specifically responsible for the contradictory findings that Chein and Weisberg (2013) reported. This scale had not been tested for its psychometric soundness and there was no evidence that it displayed construct validity or that it was even related to impasse and restructuring. Further, we propose that the subjective, self-reported feelings associated with insight are a questionable means for coding restructuring. As all theories of insight indicate that restructuring occurs after impasse, we argue that empirically coding for impasse is a more theoretically valid method of assessing the occurrence of restructuring.

The two major goals of this study were to 1) test the validity of the Aha! rating scale and 2) to test predictions based on automatic restructuring and Type 2 restructuring theories using different operational definitions of insight.

To test the validity of the Aha! rating, we separated the scale into the three dimensions of insight articulated within the scale's instructions: confidence, suddenness, and effort. We further reduced the opportunity for response confusion by anchoring each scale with extremes, *Not* and *Very*. If the Aha! rating scale possesses strong construct validity, we expected that these three ratings would be highly correlated for correctly solved problems. If the Aha! rating scale demonstrates convergent validity, we expected that these three ratings would be highly correlated with solution after impasse. To further confirm convergent validity, we expected that the Aha! ratings would closely mirror impasse results for all remaining hypothesis testing.

All theories of insight assume that insightful solutions begin from an inaccurate problem representation. Using initial problem representation ratings, we expected to confirm this assumption with all four operational definitions of insight: high confidence, high suddenness, low effort, and solution after impasse.

Gestalt theory predicts longer solving times for insightful solutions as they require the solver to exhaust an inappropriate problem space before entering impasse. Ash et al. (2012) provided evidence to support this prediction. Using the Aha! rating scale, Chein and Weisberg (2013) found the opposite: that insightful solutions were associated with faster solving times than non-insightful solutions. We attempted to replicate these findings using the four different definitions of insight.

Finally, the Gestalt theory of insightful problem solving proposes that restructuring is an automatic process that does not involve attention. This predicts a weaker correlation between solution success and working memory for problems solved with restructuring than for problems solved without restructuring. The Business-As-Usual theory proposes that restructuring involves Type 2 processes that are controlled and effortful. This predicts that correlations between working memory and solution success should be equally high whether the solution is insightful or not. We assessed these predictions by comparing solution success with working memory for all four definitions of insight.

CHAPTER 2

METHOD

Participants

One hundred two Old Dominion University undergraduate psychology students volunteered to participate in this study (79.4% women; M_{age} = 21.15, *SD* = 5.49, age range: 18–53; 93.1% native English speakers; 74.5% had college algebra within five years; 28.4% psychology majors; see Tables 1 and 2 for further demographic information). A power analysis was conducted using G*Power 3.1 (Faul, Erdfelder, Buchner, & Lang, 2009; Faul, Erdfelder, Lang, & Buchner, 2007) to determine the necessary sample size of 72 to detect an effect size of .28 with 80% power using a bivariate correlation and alpha set at .05. The effect size used in this analysis was based on results from Ash and Wiley (2006) and chosen as a more conservative estimate of the relationship between working memory and insight problem solving ability than other reported effect sizes in the literature. Participants were required to be at least 18 years of age and have normal or normal-to-corrected vision. Participants were recruited through the Department of Psychology's online research participation system and received one and a half research credits for their participation. This study was approved by Old Dominion University's Institutional Review Board (IRB) and ethical guidelines set forth by the American Psychological Association were followed. All participants signed an IRB approved Informed Consent Form that explained their rights as a volunteer (Appendix A).

Table 1

Participant Demographic Frequency Data

	n	%
Sex		
Female	81	79.4
Male	21	20.6
Total	102	100.0
English		
English has ALWAYS been my primary language.	95	93.1
English has been my primary language for MORE than 10 years.	2	2.0
English has been my primary language for LESS than 10 years.	0	0.0
English has been my primary language for LESS than 5 years.	2	2.0
English is NOT my primary language.	3	2.9
Total	102	100.0
Algebra		
I've had a college level algebra class in the past 5 years.	76	74.5
I've had a college level algebra class in the past 10 years.	5	4.9
I've had a college level algebra class more than 10 years ago.	1	1.0
I've NEVER had a college level algebra class.	20	19.6
Total	102	100.0

Table 2

Participant Major Frequency Data

	n	%
Accounting	1	1.0
Biology	10	9.8
Business Management	1	1.0
Business Management / Psychology	1	1.0
Chemistry	1	1.0
Civil Engineering	1	1.0
Criminal Justice	8	7.8
Criminal Justice / Sociology	1	1.0
Dental Hygiene	4	3.9
Electrical Engineering Technology	1	1.0
Engineering	1	1.0
Exercise Science	6	5.9
Health Science	1	1.0
Human Services	5	4.9
Industrial Technology	2	2.0
Interdisciplinary Studies	1	1.0
Marine Biology / Applied Mathematics	1	1.0

Table 2 (*continued*)

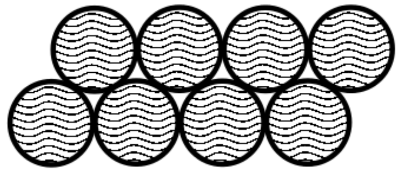
	n	%
Mechanical Engineering	1	1.0
Nuclear Medicine Technology	1	1.0
Nursing	14	13.7
Occupational and Technical Studies Training Specialist	1	1.0
Parks and Recreation Management	1	1.0
Parks, Recreation, and Tourism	1	1.0
Physical Education	1	1.0
Pre Medical Biology	1	1.0
Psychology	29	28.4
Total	102	100.0

Materials

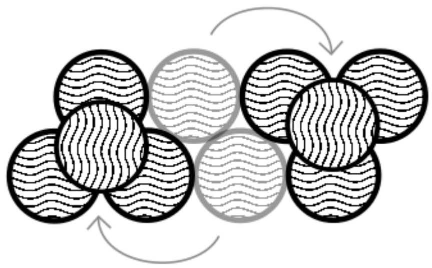
All sections of this study, with the exception of the problem solving phase, were administered on a Dell Optiplex 780 computer running Windows Vista with a monitor set to a resolution of 1280 × 1024 pixels. The tasks were programmed using E-Prime 2.0 software (Schneider, Eschman, & Zuccolotto, 2012a, 2012b). During the problem solving phase, participants completed the insight problems with paper and pencil while the Aha! ratings were administered on the computer.

Insight problems. This study used problems selected from Ash and Wiley (2006; 2008; see Appendix B). These problems have been classified as insight problems in previous research because they tend to elicit an incorrect initial problem representation in which a final solution cannot be found unless restructuring of the problem occurs. To be able to assess initial problem representation, a final criterion in selection was that each problem could be broken down into discrete components that were easily coded for importance in the final solution.

Initial problem representation ratings. Using a method developed by Ash & Wiley (2008), we assessed initial problem representation by having participants rate each *a)* Coins Problem: In this problem, there are 8 coins. Move **2 coins** so that each coin touches **exactly** 3 other coins. The coins will need to be separated into **two groups**.



b) Coins problem solution.



c) Component importance coding.

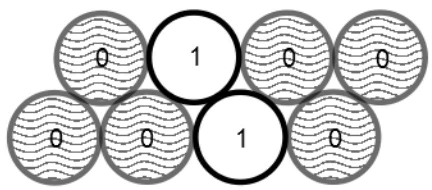


Figure 2. The coins problem, its solution, and how key and distractor components (each coin) are coded for importance in the final solution. Participants are asked to rate the importance of each coin on a scale of (1) *Not Important* to (7) *Very Important*. Scores for the key components are averaged together and scores for the distractor components are reverse coded, then averaged together. A final representation score is created by averaging these two scores together.

component of the insight problems for their importance in finding the solution. For example, the insight problem illustrated in Figure 2 shows a collection of coins with the instructions, *In this problem, there are 8 coins. Move 2 coins so that each coin touches exactly 3 other coins. The coins will need to be separated into two groups.* In this problem, there are only two coins that can be moved to create the final, correct solution.

These two coins are key components and are coded as being used in the solution (1), while all other coins are distractor components and are coded as not being used in the solution (0). Participants were presented with an image of the problem with the component they were rating colored red. They were asked to rate *How important will the coin in red be in finding the solution?* on a scale of (1) *Not Important* to (7) *Very Important*.

Before entering the rating phase, they were instructed to give immediate impressions and to quickly rate each component without attempting to solve. To accommodate for the varying number of components in each problem, a weighted mean was used to create a final representation rating. Ratings for key components were averaged together and ratings for distractor components were reverse coded, then averaged together. These two scores were averaged together to create an overall initial problem representation rating in which higher scores indicated better initial problem representation.

Aha! ratings. To investigate the subjective feelings associated with insight problem solving, participants were asked to rate their problem solving experience on three dimensions: confidence, suddenness, and effort. Ratings were made on a 4-point scale upon completion of each problem. Participants were asked, *How confident are you in your solution?*, anchored at (1) *Not Confident* and (4) *Very Confident*; *How sudden did the solution come to you?*, anchored at (1) *Not Sudden* and (4) *Very Sudden*; and *How much effort was required to find the solution?*, anchored at (1) *Not Effortful* and (4) *Very Effortful*. Instructions for these ratings included detailed descriptions for each dimension (Appendix C). To replicate the dichotomous insight coding used by Chein and Weisberg

(2013), scores of one or two were collapsed into a single non-insightful category and scores of three or four were similarly collapsed as insightful for both confidence and suddenness ratings. Effort was reverse coded as non-insightful for scores of three or four and insightful for scores of one or two.

Impasse coding and think-aloud protocol. In this study, we empirically assessed impasse utilizing a think-aloud protocol. During the problem solving phase, participants were asked to continuously describe their mental processes. The think-aloud protocol for this phase was video recorded for later data coding. Before beginning the problem solving phase, the experimenter read the following:

> Using the packet provided, you will have four minutes to solve each problem. There is one problem per page. You must show ALL YOUR WORK for each problem and CIRCLE the final answer. You may explain your solution in writing, arrows, or other diagrams as necessary as long as the solution is clear. ALL problems have a solution and answers such as "impossible" or "not solvable" will be counted as incorrect. So that we can understand what you are doing while you solve each problem, you will talk through your problem solving process out loud. This section will be video recorded for later data coding.
>
> This ‘think aloud’ process should be like a stream of consciousness: whatever is going on in your mind is what you should verbalize. If you are reading the problem, read it out loud. If you are writing on the paper, verbalize what you are doing. Verbalize anything you are thinking and be sure to keep talking through the entire process.

> You will not be given any feedback regarding how close you are to solution or the accuracy of your solution, but you will be reminded to keep talking if you stop verbalizing during the process. Please remember to speak loud enough for the video recording.
>
> After completing each problem, you will be asked three rating questions (on the computer) that describe your problem solving process. You will enter your response by using the mouse to click on your choice. After completing the rating questions, you will wait for the experimenter before continuing to solve the next problem.
>
> We will now do a practice problem so you can get used to the think aloud protocol and see how the rating questions work. When you are ready to begin, you will turn the practice page over and read the instructions out loud. After you read the instructions, I will ask if you are ready to begin. I will then begin the four minute timer. When you have the solution, RAISE YOUR HAND and I will stop the timer. You will then answer the rating questions on the computer using the mouse. REMEMBER TO KEEP THINKING OUT LOUD. Do you have any questions?

The videos were later coded by four independent raters looking for the occurrence of impasse. Following the impasse coding procedure outlined in Ash et al. (2012), the operational definition of impasse was taken from Knoblich, Ohlsson, Haider, and Rhenius (1999) who define impasse as the "cessation of overt problem-solving behavior" which is "accompanied by a subjective feeling of not knowing what to do" (p. 1534). Raters looked for instances that supported this definition: periods of silence, especially

Table 3

Inter-rater Reliability for Impasse Coding

	Cohen's kappa	*SE*	*p*	95% CI Lower	95% CI Upper
Coder 1 x Coder 2	.82	0.07	< .001	.68	.95
Coder 1 x Coder 3	.84	0.06	< .001	.71	.96
Coder 1 x Coder 4	.84	0.06	< .001	.72	.96
Coder 2 x Coder 3	.80	0.07	< .001	.66	.94
Coder 2 x Coder 4	.76	0.08	< .001	.61	.91
Coder 3 x Coder 4	.78	0.07	< .001	.63	.92

Note. *n* = 93.

after reminders to continue the verbal protocol, verbal indications of impasse such as "I don't know what to do" or "I'm lost", physical stillness in which the participant ceased to actively write or point to the problem, verbal expressions of frustration or frustrated body language, and mindless repetition of the instructions. All raters reviewed a subset of 93 problems in which Cohen's kappa was used to assess agreement between raters. Using Altman's (1991) standards of agreement, agreement between raters ranged from good, $\kappa = .76$, $p = < .001$, 95% CI [.61, .91], to very good, $\kappa = .84$, $p = < .001$, 95% CI [.72, .96] (Table 3).

Video for the think-aloud protocol was captured with a Panasonic PV-GS150 digital video camcorder on a Macintosh computer running Mac OS X using SecuritySpy software. The tripod mounted camcorder was positioned approximately four and a half feet in front of the seated participant (Figure 3).

Working memory tasks. Working memory capacity was measured using replications of Kane et al.'s (2004) reading span task (RSPAN) and operation span task (OSPAN). These are dual-process tasks that require participants to maintain information

Figure 3. Laboratory set-up for this experiment. Participants were seated at the computer to the left and the experimenter was seated at the computer to the right. The camera was positioned approximately four and a half feet in front of the participant.

while concurrently executing a secondary task. Working memory capacity is then assessed by performance on information recall. For analysis, a single working memory score was created by averaging the standardized composite scores from both the RSPAN and OSPAN.

Reading span task. The RSPAN task requires participants to retain, then later recall a set of letters from their short-term memory. While attempting to retain these letters, a secondary reading comprehension task must be completed. In this task, participants are presented with a sentence that they must read aloud. All sentences are grammatically correct, but in half of the sentences a noun has been replaced with an irrelevant noun rendering the sentence nonsensical. The participant must state whether the sentence makes sense or not by indicating "yes" or "no". The sentences are followed by a letter that the participant must read aloud and attempt to retain. For example, if the participant is presented with the following, *Andy was stopped by the policeman because he crossed the yellow heaven. ? R*, they should say: “Andy was stopped by the policeman because he crossed the yellow heaven... no… r.” After the participant reads the letter aloud, the experimenter immediately advances to the next stimulus screen. After a certain number of these sentences and letters, the participant is cued to write down all the letters they can recall from that particular set in the order they occurred. The sets range from two to five letters and each set size is repeated three times with new stimuli for a total of 12 sets (see Appendix D for complete stimuli list). Each recalled set is graded on the total number of letters recalled divided by the total number of letters presented in that set. This is an all or nothing grading in which credit is only given if all letters in the set are reproduced in the correct order. A final composite RSPAN score is created by summing the weighted set scores.

Operation span task. The OSPAN task follows the same procedure as the RSPAN with two differences: the primary memory task asks participants to recall a set of words instead of letters and the secondary task has been replaced with a simple arithmetic equation and solution in which the participant must decide if the solution presented is correct (e.g., *IS (9 ÷ 3) + 2 = 2 ? AUNT*; see Appendix E for complete stimuli list).

Reliability and validity. Both working memory tasks have been adapted from Engle et al. (1999) in which short term memory tasks, working memory tasks, general fluid intelligence, and both the verbal and math portions of the Scholastic Aptitude Test (SAT) were examined for divergent and convergent construct relationships. Using factor analysis and structural equation modeling, their results supported that these are distinct, though related constructs. Also using structural equation modeling techniques, Kane et al.

(2004) investigated the relationship between verbal working memory tasks (RSPAN and OPSAN) and spatial working memory tasks. They found that the two categories share 70%–80% of their variance indicating that working memory tasks, regardless of domain specification, largely measure a domain-general construct. They further reported that the RSPAN and OSPAN have good reliability with Cronbach's alphas of .78 and .80 respectively. In this study, the RSPAN and OSPAN together were found to be reliable, producing a Cronbach's alpha of .84.

Procedure

Participants were recruited via the Psychology Department's online research participation system. They attended their individual, one and a half hour session at a laboratory located on the campus of Old Dominion University. Upon arrival, they were seated at a computer station and given time to read and sign the Informed Consent Form. After the participants signed the form, they proceeded through the experimental phases (Figure 4). A balanced Latin square approach was used to produce six order conditions for the insight problems. Participants were randomly assigned to one of the six order

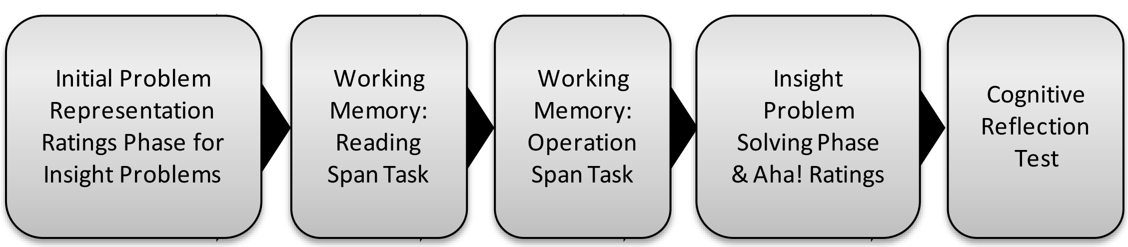


Figure 4. Experiment procedural flow. A balanced Latin square approach was used to produce six order conditions for the insight problems. Participants were randomly assigned to one of the six order conditions. This order condition was maintained for the initial problem representation phase as well as for the problem solving phase. During the problem solving phase, each insight problem was followed by three Aha! ratings.

conditions. This order condition was maintained for the initial problem representation phase as well as for the problem solving phase. All directions were read aloud by the experimenter, but were also represented on the computer screen so that the participant could read along with them. Each section contained a practice phase.

Participants began with the initial problem representation phase in which each insight problem was presented and the participant was asked about their familiarity with the problem as well as to rate the importance of the problem components. They then completed the working memory phase, first the RSPAN and then the OSPAN, and continued with the problem-solving phase and think-aloud protocol. After completion of each insight problem, the participant gave their Aha! ratings on the computer before continuing to the next problem. Following the problem solving phase, participants took the Cognitive Reflection Test. This section was a pilot test for future research and will not be discussed at this time. At the conclusion of the experiment, the participant was debriefed, allowed to ask questions, and asked to refrain from discussing the experiment with others in the research pool.

CHAPTER 3

RESULTS AND DISCUSSION

Data Screening

Thirteen participants were removed from all analyses for a final sample size of 89. Of these 13, six participants did not finish the experiment and were removed due to incomplete data. We reduced the potential for confounded results caused by language difficulties by removing seven participants who indicated that English was not their primary language (including two previously mentioned as missing data). Using criteria set forth by Engle et al. (1999), participants must have maintained at least 85% accuracy on the secondary working memory tasks to be retained for analyses. This reduces the potential for biased or inaccurate working memory scores (e.g., participants who are attempting to employ rehearsal or other mnemonic strategies to improve their recall performance or participants who may have language difficulties). Five participants failed to meet this minimum requirement (including three previously mentioned for removal due to missing data or language).

Overall Analysis Strategy

Not all solvers experience insight while completing classic insight problems. To this end, it is necessary to analyze the data at the level of observation. This allows for participants with differential insight and solution rates to contribute to the different categories of observation, e.g., solution success, impasse, and confidence. To accommodate the differing number of observations that a single participant can contribute, information from each of the six insight problems were considered as individual data points and analyzed at the level of observation where indicated.

Descriptive Statistics

Table 4 presents the descriptive statistics for mean initial representation scores by problem and by participant, working memory span, number of correctly solved problems, and mean solving time for correctly solved problems. Table 5 presents the solution rates for the insight problems. Solution rates were lower than expected based on previous data from Ash and Wiley (2006) and Ash et al. (2012). From a total of 534 problems, there was a 24.2% success rate, while 22.3% of problems were completed with incorrect solutions and 53.6% were left unsolved. Of correctly solved problems, the circles problem garnered the highest solution rate (27.1%), whereas the coins problem had the lowest solution rate (3.1%). Table 6 presents the frequency data for the four measures of insight. Of note is the low variability within confidence for correctly solved problems.

Table 4

Descriptive Statistics for Experiment Measures

				95% CI			
	N	*M*	*SD*	Lower	Upper	Skew	Kurtosis
Initial Problem Representation							
By Problem	534	4.27	1.13	4.17	4.36	0.33	0.24
By Participant	89	4.27	0.49	4.16	4.37	0.44	0.83
Working Memory	89	0.00[a]	0.93	-0.20	0.20	0.25	-0.01
Total Correctly	89	1.45	1.22	1.19	1.71	0.82	0.62
Mean Solving Time[b]	67[c]	123.57	63.16	108.17	138.98	0.24	-1.14

Note. *N* = 534; a = working memory span scores were standardized; b = for correctly solved problems; c = twenty-two participants (24.7%) did not solve any problems.

Table 5

Solution Rates By Problem Type

	Correct		Incorrect		Did Not Solve	
All Problems, *N* = 534						
Matchsticks #1	17	3.2%	21	3.9%	51	9.6%
Matchsticks #2	22	4.1%	8	1.5%	59	11.0%
Glasses	32	6.0%	24	4.5%	33	6.2%
Circles	35	6.6%	16	3.0%	38	7.1%
Coins	4	0.7%	40	7.5%	45	8.4%
Squares	19	3.6%	10	1.9%	60	11.2%
Total	129	24.2%	119	22.3%	286	53.6%
Correctly Solved Problems, *N* = 129						
Matchsticks #1	17	13.2%				
Matchsticks #2	22	17.1%				
Glasses	32	24.8%				
Circles	35	27.1%				
Coins	4	3.1%				
Squares	19	14.7%				
Total	129	100.0%				

Table 6

Frequency Data for Measures of Insight

	Solved with Insight		Solved without		DNS/Incorrect	
All Problems, *N* = 534						
Confidence	107	20.0%	22	4.1%	405	75.8%
Suddenness	68	12.7%	61	11.4%	405	75.8%
Effort	60	11.2%	69	12.9%	405	75.8%
Impasse	80	15.0%	49	9.2%	405	75.8%
Correctly Solved Problems, *N* =						
Confidence	107	82.9%	22	17.1%		
Suddenness	68	52.7%	61	47.3%		
Effort	60	46.5%	69	53.5%		
Impasse	80	62.0%	49	38.0%		

Note. DNS = did not solve.

Participants felt confident about their solutions 82.9% of the time for correctly solved problems.

Relationships Between Measures of Insight

Self-report ratings of the Aha! experience used in previous research have been presented on a single continuum scale that encompasses several dimensions: confidence, suddenness, and effort. Within the context of this scale, the occurrence of an insightful solution is operationalized as the solver being highly confident in their solution (high confidence), that the solution seemed to come suddenly (high suddenness), and that little effort was necessary to obtain the solution (low effort). As such, it is expected that parsing these dimensions into three unique ratings would produce strong levels of agreement when sorting problems solved with insight and without. Further, if self-report Aha! ratings are truly reflective of post-impasse restructuring, it is expected that they would strongly relate to solutions categorized by the Gestalt defined occurrence of insight, solution after impasse. To examine these relationships, a series of chi square tests of independence were performed at the level of observation on correctly solved problems (*N* = 129; see Tables 7-12 for frequency data). Chi square goodness of fit tests were then used to isolate differences within significant relationships.

Table 7

Observed Frequencies for Confidence × Suddenness

	Suddenness		
	High (Insight)	Low (No Insight)	Total
Confidence			
High (Insight)	61 (47.3%)	46 (35.7%)	107 (82.9%)
Low (No Insight)	7 (5.4%)	15 (11.6%)	22 (17.1%)
Total	68 (52.7%)	61 (47.3%)	129 (100.0%)

Table 8

Observed Frequencies for Suddenness × Effort

	Effort				Total	
	Low (Insight)		High (No Insight)			
Suddenness						
High (Insight)	41	(31.8%)	27	(20.9%)	68	(52.7%)
Low (No Insight)	19	(14.7%)	42	(32.6%)	61	(47.3%)
Total	60	(46.5%)	69	(53.5%)	129	(100.0%)

Table 9

Observed Frequencies for Effort × Confidence

	Confidence				Total	
	High (Insight)		Low (No Insight)			
Effort						
Low (Insight)	53	(41.1%)	7	(5.4%)	60	(46.5%)
High (No Insight)	54	(41.9%)	15	(11.6%)	69	(53.5%)
Total	107	(82.9%)	22	(17.1%)	129	(100.0%)

Table 10

Observed Frequencies for Impasse × Confidence

	Confidence				Total	
	High (Insight)		Low (No Insight)			
Impasse						
Yes (Insight)	64	(49.6%)	16	(12.4%)	80	(62.0%)
No (No Insight)	43	(33.3%)	6	(4.7%)	49	(38.0%)
Total	107	(82.9%)	22	(17.1%)	129	(100.0%)

Table 11

Observed Frequencies for Impasse × Suddenness

	Suddenness				Total	
	High (Insight)		Low (No Insight)			
Impasse						
Yes (Insight)	30	(23.3%)	50	(38.8%)	80	(62.0%)
No (No Insight)	38	(29.5%)	11	(8.5%)	49	(38.0%)
Total	68	(52.7%)	61	(47.3%)	129	(100.0%)

Table 12

Observed Frequencies for Impasse × Effort

	Effort		
	Low (Insight)	High (No Insight)	Total
Impasse			
Yes (Insight)	22 (17.1%)	58 (45.0%)	80 (62.0%)
No (No Insight)	38 (29.5%)	11 (8.5%)	49 (38.0%)
Total	60 (46.5%)	69 (53.5%)	129 (100.0%)

Confidence × Suddenness. The data revealed a significant relationship between confidence and suddenness, $\chi^2(1, N = 129) = 4.65$, $p = .031$, $\varphi = .19$. However, follow up analyses showed that this relationship was not diagnostic of the occurrence of insight (Figure 5). Though participants who believed their answer came to them suddenly were likely to feel very confident about their solution, $\chi^2(1, n = 68) = 42.88$, $p = < .001$, participants who did not experience feelings of suddenness were still very confident in their responses, $\chi^2(1, n = 61) = 15.75$, $p = < .001$. Participants who were highly confident in their solution, $\chi^2(1, n = 107) = 2.10$, $p = .147$, or had little confidence in their solution, $\chi^2(1, n = 22) = 2.91$, $p = .088$, showed no differences between feelings of suddenness. These results suggest that the significant relationship between confidence and suddenness is driven by a general level of confidence in all solutions and is not related to solution suddenness.

Suddenness × Effort. There was a significant relationship between suddenness and effort, $\chi^2(1, N = 129) = 10.98$, $p = .001$, $\varphi = .29$ (Figure 6). Participants who felt their solution came to them suddenly showed no differences between feelings of effort, $\chi^2(1, n = 68) = 2.88$, $p = .090$, but participants who did not feel their solution was sudden reported significantly more effort was involved, $\chi^2(1, n = 61) = 8.67$, $p = .003$.

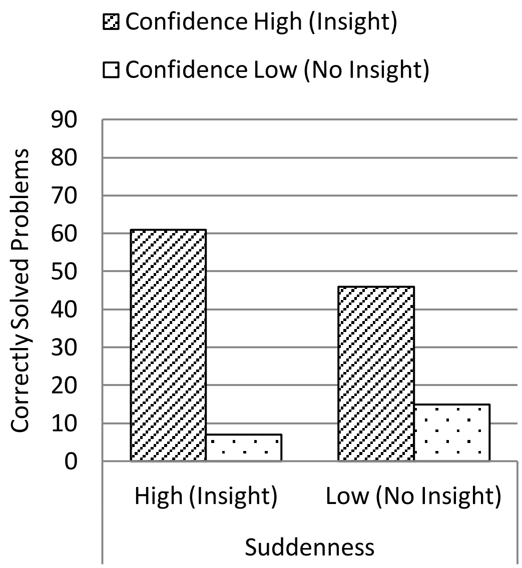


Figure 5. Relationship between confidence and suddenness for correctly solved problems.

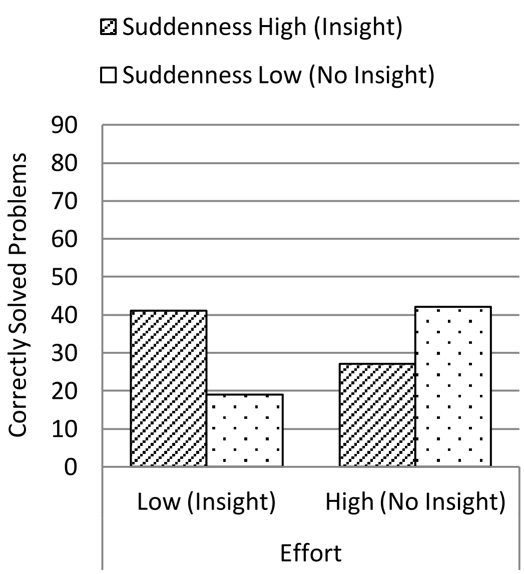


Figure 6. Relationship between suddenness and effort for correctly solved problems.

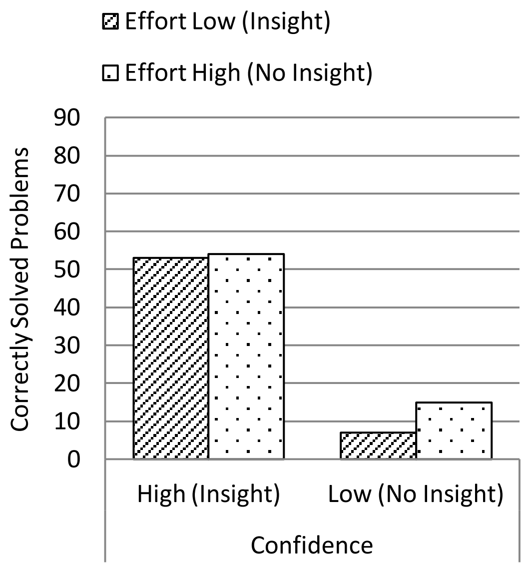


Figure 7. Relationship between effort and confidence for correctly solved problems.

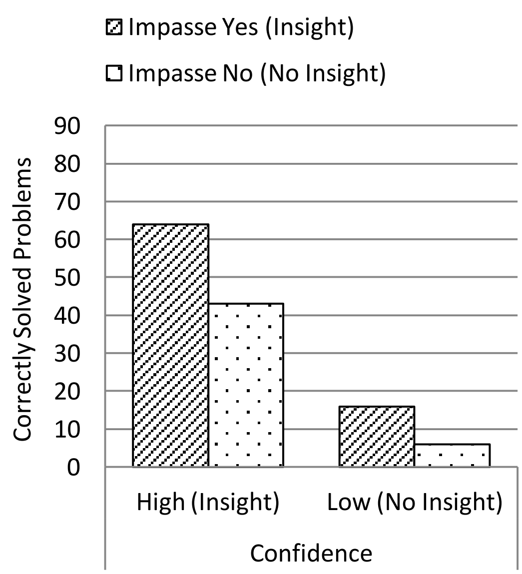


Figure 8. Relationship between impasse and confidence for correctly solved problems.

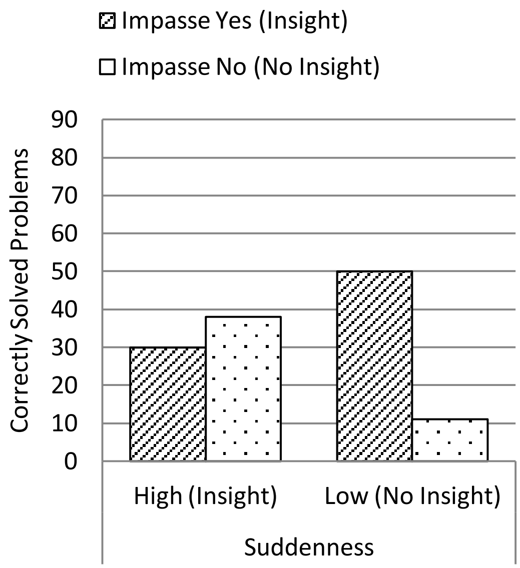


Figure 9. Relationship between impasse and suddenness for correctly solved problems.

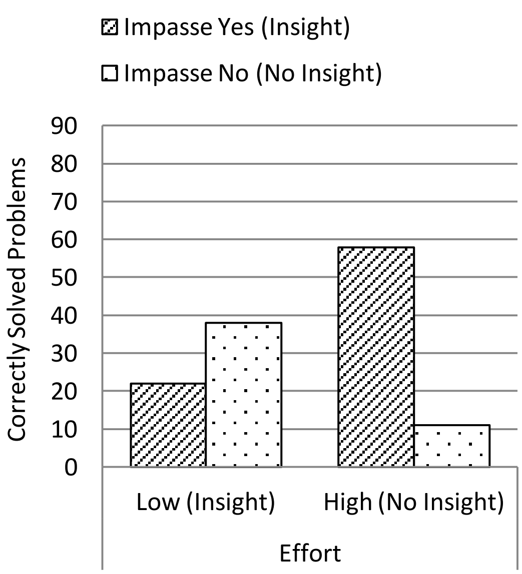


Figure 10. Relationship between impasse and effort for correctly solved problems.

Feelings of little effort were more often accompanied by feelings of suddenness, $\chi^2(1, n = 60)$ 8.07, $p = .005$. However, high effort was not significantly associated with suddenness, $\chi^2(1, n = 69) = 3.26$, $p = .071$. Though weakly associated, these constructs did show evidence of validity when assessing Aha! experiences.

Effort × Confidence. There was no significant relationship between effort and confidence, $\chi^2(1, N = 129) = 2.30$, $p = .129$, $\varphi = .13$. As Figure 7 illustrates, these constructs are unrelated aspects of the Aha! experience.

Impasse × Confidence. The data revealed no significant relationship between impasse and confidence, $\chi^2(1, N = 129) = 1.29$, $p = .256$, $\varphi = .10$ (Figure 8), indicating that solvers were confident in their solutions regardless of whether they came to impasse.

Impasse × Suddenness. As suddenness and effort appear to be the only related constructs within the Aha! rating scale, it is of particular interest how they relate to problems solved with and without impasse. The relationship between impasse and suddenness was significant, $\chi^2(1, N = 129) = 19.56$, $p = < .001$, $\varphi = .39$, but follow up analyses revealed an interesting pattern to the data (Figure 9). Participants who came to impasse were less likely to indicate that their solution was sudden, $\chi^2(1, n = 80)$ 5.00, $p = .025$. Participants who did not come to impasse reported high feelings of suddenness, $\chi^2(1, n = 49) = 14.88$, $p = < .001$. There were no significant differences between coming to impasse in solutions rated as sudden, $\chi^2(1, n = 68) = 0.94$, $p = .332$. However, participants who did not feel that their solution was sudden were very likely to experience impasse, $\chi^2(1, n = 61)$ 24.93, $p = < .001$. This data pattern illustrates the *opposite* relationship expected if the Aha! rating correlated with the Gestalt concept of restructuring.

Impasse × Effort. A similar and even stronger pattern emerged from the significant relationship between impasse and effort, $\chi^2(1, N = 129) = 30.60$, $p = < .001$, $\varphi = .49$. Participants who came to impasse were more likely to rate their solutions as very effortful, $\chi^2(1, n = 80)$ 16.20, $p = < .001$, whereas participants who did not experience impasse were more likely to feel that their response was not effortful, $\chi^2(1, n = 49)$ 14.88, $p = < .001$. Participants who did not feel their solution took a lot of effort were less likely to experience impasse, $\chi^2(1, n = 60)$ 4.27, $p = .039$. Participants who felt their solution was effortful were very likely to have experienced impasse, $\chi^2(1, n = 69)$ 32.01, $p = < .001$. This relationship is the opposite of what is expected if impasse and effort are convergent measures of insightful problem solving (Figure 10).

In summary, the first goal of these analyses was to determine whether the three constructs inherent in previously used self-report Aha! rating scales showed evidence of construct validity when separated into three distinct ratings. Our findings illustrate that these three constructs are not strongly related when assessing the occurrence of insight. Only the constructs of suddenness and effort appeared to be associated.

The second goal of these analyses was to determine the relationships between the self-report Aha! ratings and solution after impasse, an operational definition designed to reflect the Gestalt view of restructuring. Though suddenness and effort appear to be related constructs when assessing insight, they are both inversely related to solution after impasse, a pattern that is opposite of predications based on any theory of insight which acknowledges that restructuring occurs in solutions attained after impasse. Therefore, previous studies utilizing the Aha! rating scale have inappropriately categorized some solutions obtained *without* impasse as insightful experiences based on feelings of suddenness and low effort.

Insight and Initial Problem Representation

Integral to the Gestalt theory of insight is the idea that the process to an insightful solution begins with an inaccurate problem representation. This assumption was explored by examining the relationships between initial problem representation and the four different methods of operationalizing insight (high suddenness, high confidence, low effort, or solution after impasse). This was an exploratory investigation for the three self-report measures and an attempt to replicate results from Ash et al. (2012) for impasse. To do this, a series of mixed model analyses of variances, which compared insightful solutions, solutions made without insight, and problems that were not solved or had incorrect solutions, were performed on mean initial representation scores at the level of observation (N = 534; see Table 13 for descriptive statistics). In these analyses, participant was included as a random effect variable. Results did not reveal differences in initial problem representation and whether the problem solving process was experienced as insightful for the three self-report measures: confidence, $F(2, 98.93) = 0.20$, $p = .817$, partial $\eta^2 = .004$, suddenness, $F(2, 99.02) = 0.99$, $p = .376$, partial $\eta^2 = .02$, and effort, $F(2, 98.68) = 1.88$, $p = .159$, partial $\eta^2 = .04$. Only impasse produced significant differences, $F(2, 105.28) = 4.29$, $p = .016$, partial $\eta^2 = .08$. Participants who correctly solved problems without experiencing impasse were more likely to begin with a better initial problem representation (M = 4.94, SE = 0.18) than solvers who came to impasse (M = 4.11, SE = 0.13), $F(1, 20.29) = 6.51$, $p = .019$, partial $\eta^2 = .24$. This supports the

Table 13

Descriptive Statistics for Initial Representation Ratings as a Function of Insight

				95% CI	
	n	*M*	*SE*	Lower	Upper
Confidence					
High (Insight)	107	4.32	0.12	4.09	4.56
Low (No Insight)	22	4.48	0.24	4.02	4.95
Did Not Solve/Incorrect	405	4.25	0.06	4.13	4.36
Suddenness					
High (Insight)	68	4.61	0.15	4.32	4.91
Low (No Insight)	61	4.20	0.15	3.90	4.50
Did Not Solve/Incorrect	405	4.25	0.06	4.13	4.36
Effort					
Low (Insight)	60	4.67	0.16	4.36	4.97
High (No Insight)	69	4.14	0.14	3.87	4.41
Did Not Solve/Incorrect	405	4.25	0.06	4.13	4.36
Impasse					
Yes (Insight)	80	4.11	0.13	3.86	4.37
No (No Insight)	49	4.94	0.18	4.59	5.28
Did Not Solve/Incorrect	405	4.25	0.06	4.13	4.36

Note. *N* = 534.

Gestalt theory of insightful problems solving and replicated findings from Ash et al. (2012).

These results challenge the appropriateness of using the self-report Aha! ratings as operational definitions of insight. If they captured the Gestalt concept of restructuring, we would expect the patterns of data to mirror that of the impasse data. Though not significant, Figure 11 illustrates that both suddenness and effort actually appear to have the opposite pattern from what is expected: insightful experiences, i.e., problems rated with high suddenness and low effort began from a better initial problem representation. This is directly at odds with the assumption that insightful problem solving must begin with an inaccurate problem representation.

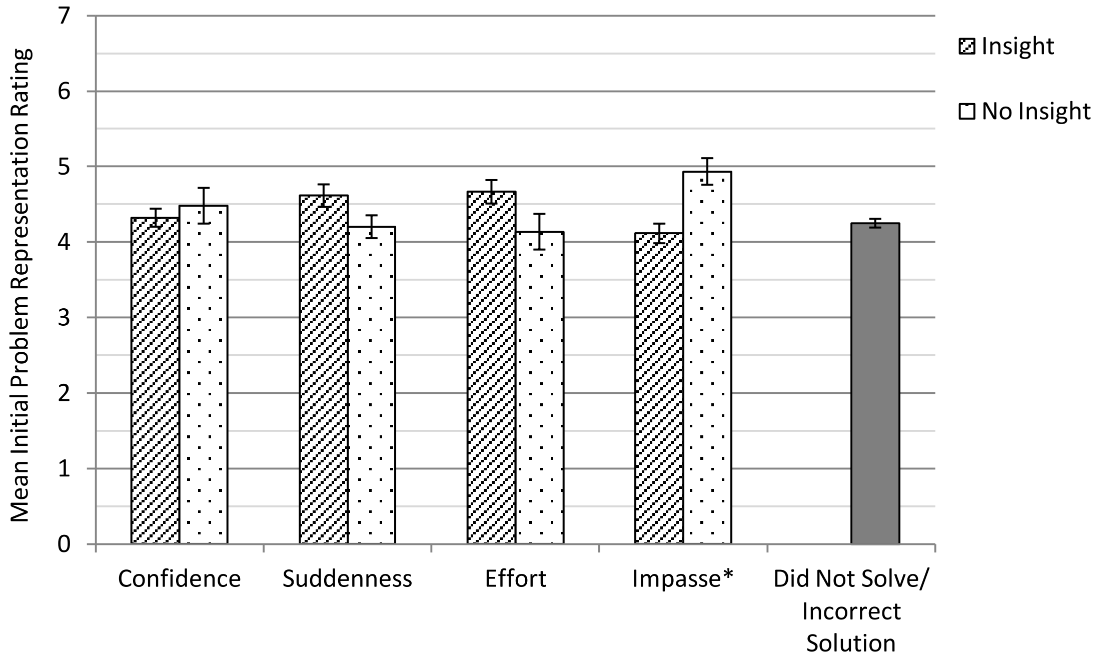


Figure 11. Mean initial representation rating scores as a function of insight in confidence, suddenness, effort, and impasse. Note that asterisks indicate significant differences.

Insight and Solving Time

Two specific studies found contradictory evidence regarding solution speed for problems solved with insight. Ash et al. (2012) demonstrated that solutions obtained after impasse came significantly slower than non-insightful solutions. Chein and Weisberg (2013) found that insightful solutions were produced much quicker than non-insightful solutions. One primary difference between these two studies was the operational definition of insight. Ash et al. defined insight as solution after impasse coded by impartial raters, while Chein and Weisberg utilized the self-report Aha! rating scale. The current study allowed us to compare solving times using both methodologies within the same problem set.

A series of mixed model analyses of variances (problems solved with insight versus problems solved without insight), in which participant was included as a random effect variable, were performed on mean solving times at the level of observation for correctly solved problems (*N* = 129; see Table 14 for descriptive statistics). No relationship was detected between confidence and solving time, $F(1, 10.22) = 0.13$, $p = .731$, partial $\eta^2 = .01$. Using suddenness as an indicator of insight revealed a trend towards significance, $F(1, 14.51) = 3.72$, $p = .073$, partial $\eta^2 = .20$, with solutions rated as sudden associated with faster solving times (*M* = 97.63, *SE* = 7.28) than solutions rated as not very sudden (*M* = 149.65, *SE* = 7.38). The same effect was observed when effort was used as the indicator for insight, $F(1, 17.66) = 27.78$, $p = < .001$, partial $\eta^2 = .61$. Low effort solutions were significantly faster (*M* = 69.60, *SE* = 6.25) than high effort solutions (*M* = 154.88, *SE* = 5.55). Overall, this replicated Chein and Weisberg's results who found

Table 14

Descriptive Statistics for Solving Time as a Function of Insight

				95% CI	
	n	*M*	*SE*	Lower	Upper
Confidence					
High (Insight)	107	122.33	6.36	109.56	135.10
Low (No Insight)	22	126.48	12.58	101.22	151.74
Suddenness					
High (Insight)	68	97.63	7.28	82.99	112.27
Low (No Insight)	61	149.65	7.38	134.81	164.49
Effort					
Low (Insight)	60	69.60	6.25	57.00	82.20
High (No Insight)	69	154.88	5.55	143.70	166.06
Impasse					
Yes (Insight)	80	154.32	5.84	142.53	166.12
No (No Insight)	49	48.99	7.87	33.11	64.87

Note. *N* = 534.

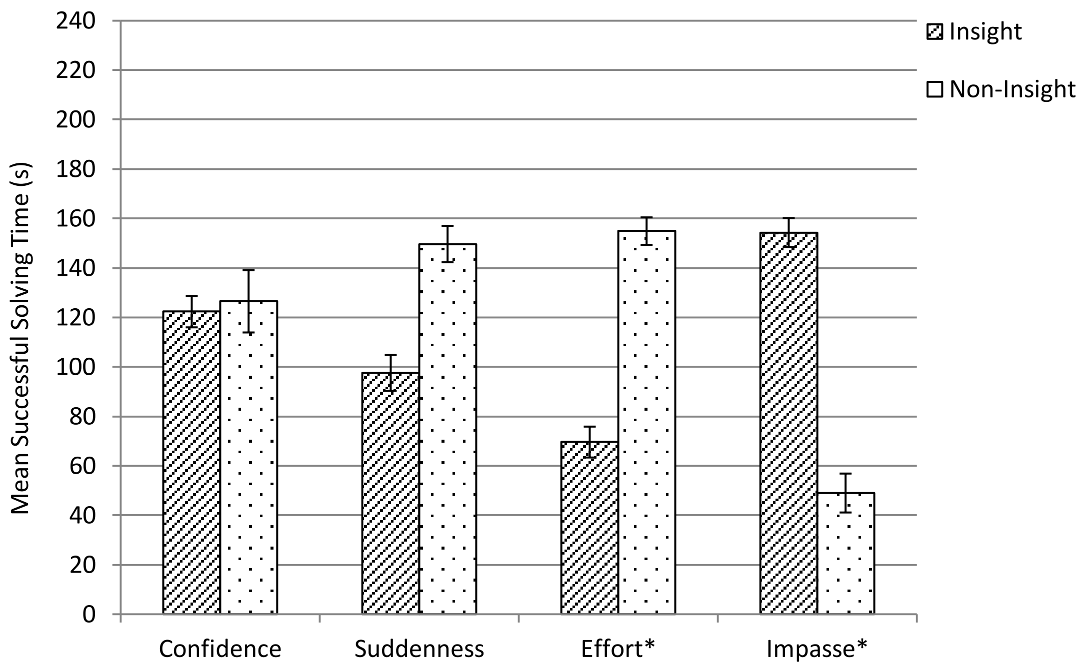


Figure 12. Mean solving time as a function of insight for correctly solved problems in confidence, suddenness, effort, and impasse. Asterisks indicate significant differences.

that problems reported as insightful, i.e., high confidence, high suddenness, and low effort, were solved faster than problems reported with low Aha! ratings. However, when insight was operationalized as problems solved after impasse, the opposite pattern was found, $F(1, 20.49) = 36.79$, $p = < .001$, partial $\eta^2 = .64$. Solutions after impasse were associated with significantly slower solving times ($M = 154.32$, $SE = 5.84$) than problems solved without impasse ($M = 48.99$, $SE = 7.87$), a replication of Ash et al. (2012). As Figure 12 illustrates, these analyses clearly demonstrate that the contradictory results from previous literature regarding solving time and insight are a function of how insight was operationally defined.

Insight and Working Memory

Working memory has been a prominent feature in research investigating the insight problem solving process, being used to support both the Gestalt and the Business-As-Usual theories of insight. We continued this line of research by comparing the four measures of insight with working memory. These analyses were conducted at the participant level (N = 89). Consistent with previous research, overall problem solving ability was significantly related to working memory, $r(87) = .213$, $p = .045$. We replicated procedures used in Chein and Weisberg (2013) and created two sets of proportions for each definition of insight: 1) number of problems solved with insight over total number of problems; 2) number of problems solved without insight over total number of problems; 3) number of problems solved with insight over total number of problems less problems solved without insight; and 4) number of problems solved without insight over the total number of problems less problems solved with insight (see Table 15 for descriptive statistics). The last two proportions were included as more precise measures of individual differences in success at insightful or non-insightful problem solving. After calculating these proportions for each of the four operational definitions of insight—high confidence, high suddenness, low effort, and impasse after solution—we then correlated these proportions with working memory scores. The two different methods for calculating the proportions did not produce major differences between the correlations with working memory (Table 16). Therefore, we discuss the results in terms of the proportions calculated using the total number of problems.

For the self-report ratings assessed as insightful, the proportions of problems

Table 15

Descriptive Statistics for Proportions of Problems Solved

	M	*SD*	95% CI Lower	95% CI Upper	Skew	Kurt
Proportion Solved / Total # Problems	0.24	0.20	0.20	0.28	0.82	0.62
Confidence						
Proportion Solved with Insight / Total # Problems	0.20	0.21	0.16	0.24	1.16	1.21
Proportion Solved with No Insight / Total # Problems	0.04	0.08	0.02	0.06	1.79	2.45
Proportion Solved with Insight / Total # Problems Less No Insight	0.20	0.21	0.16	0.24	1.15	1.18
Proportion Solved with No Insight / Total # Problems Less Insight	0.05	0.09	0.03	0.07	1.79	2.37
Suddenness						
Proportion Solved with Insight / Total # Problems	0.13	0.18	0.09	0.17	1.72	3.04
Proportion Solved with No Insight / Total # Problems	0.11	0.15	0.08	0.15	1.27	1.38
Proportion Solved with Insight / Total # Problems Less No Insight	0.13	0.18	0.09	0.17	1.72	3.03
Proportion Solved with No Insight / Total # Problems Less Insight	0.13	0.16	0.09	0.16	1.09	0.36
Effort						
Proportion Solved with Insight / Total # Problems	0.11	0.17	0.08	0.15	1.74	2.72
Proportion Solved with No Insight / Total # Problems	0.13	0.14	0.10	0.16	0.82	0.01
Proportion Solved with Insight / Total # Problems Less No Insight	0.11	0.18	0.08	0.15	1.75	2.79
Proportion Solved with No Insight / Total # Problems Less Insight	0.15	0.16	0.12	0.18	0.70	-0.53
Impasse						
Proportion Solved with Insight / Total # Problems	0.15	0.14	0.12	0.18	0.74	-0.10
Proportion Solved with No Insight / Total # Problems	0.09	0.16	0.06	0.13	2.46	6.71
Proportion Solved with Insight / Total # Problems Less No Insight	0.15	0.15	0.12	0.18	0.70	-0.17
Proportion Solved with No Insight / Total # Problems Less Insight	0.11	0.19	0.07	0.14	2.32	5.72

Note. N = 89.

Table 16

Relationships Between Working Memory Span and Insight

	r(WM, Insight/Total)	*r*(WM, No Insight/Total)	*r*(WM, Insight/[Total-No Insight])	*r*(WM, No Insight/[Total-Insight])
Confidence	.254*	-.114	.254*	-.096
Suddenness	.203	.047	.204	.103
Effort	.198	.066	.199	.132
Impasse	.050	.220*	.056	.226*

Note. * < .05.

solved with high confidence were significantly correlated with working memory, $r(87) = .254$, $p = .016$. High suddenness, $r(87) = .203$, $p = .056$, and low effort, $r(87) = .198$, $p = .063$, also showed the same positive relationships with working memory. No significant relationships were found between working memory and the proportions of problems solved without insight, i.e., low confidence, low suddenness, and high effort. Overall, this replicated results reported by Chein and Weisberg (2013) which were interpreted as evidence that Type 2 processes are involved with restructuring.

However, when insight was operationalized as solution after impasse, there was no evidence of a relationship between the proportion of problems solved with insight and working memory, $r(87) = .050$, $p = .644$. Further, the proportion of problems solved without impasse was significantly related to working memory, $r(87) = .220$, $p = .038$. This replicates previous research that has shown no relationship between working memory and restructuring in insightful problem solving (Ash & Wiley, 2006; Fleck, 2008; Gilhooly & Fioratou, 2009). Therefore, when insight is operationally defined as solution after impasse, our results match predictions made by automatic restructuring theories. However, when insight is operationally defined in terms of self-report Aha! experiences, we find evidence predicted by the Business-As-Usual theory of insight.

CHAPTER 4

CONCLUSIONS

Many studies have used a self-report Aha! rating scale to determine the occurrence of insight (e.g., Bowden & Jung-Beeman, 2003; Chein & Weisberg, 2013; Jung-Beeman et al., 2004). In this study, we investigated whether this was a valid way to assess insightful problem solving experiences. The anchors of this scale—*strategic* and *insight*, constructs that are not extremes on the same continuum—are described in terms of solution confidence, how sudden a solution manifests, and the effort required during the solving process. Of concern is that these multi-barreled instructions create the potential for response confusion. Our data clearly suggests that this is the case. Separating this scale into its three unique subcomponents reveals little in the way to support construct validity. Confidence is not related to either suddenness or effort and although suddenness and effort do appear to be related, it is a weaker association that does not justify their combination into a single measurement. When presenting these three fairly independent constructs as a single scale, participants must choose their response given which element is most salient to them at the time, inadvertently creating what is essentially three different dependent variables.

Our data further demonstrate that the Aha! rating scale is wholly unrelated to the Gestalt defined concept of restructuring, the key to all investigations regarding the insight problem solving process. The Gestalt theory of insight is based on the assumptions that insightful solutions begin from an inaccurate initial problem representation and that restructuring, the process of overcoming this faulty problem space and formulating the correct problem representation, occurs only after a period of impasse. When operationally defining insight based on the criteria of *solution after impasse*, we found no evidence of convergent validity with any of the self-report definitions of insight: high confidence, high suddenness, or low effort. Confidence was unrelated to impasse after solution. High suddenness and low effort were actually associated with solutions obtained *without* impasse.

We also found that participants solving after impasse began with a less appropriate initial problem representation than those who solved without impasse, a pattern predicted by Gestalt theory. However, the three self-report ratings were not significantly associated with initial problem representation. This anomalous pattern was repeated when assessing solving time. Solutions after impasse were significantly associated with slower solving times, a result that is predicted by Gestalt theory. If a solver must exhaust a faulty problem representation before coming to impasse, it stands to reason that it would take longer to acquire a solution than if the solver began with an accurate problem representation. Of the self-report ratings, confidence was unrelated to solving time and suddenness and effort were, again, predictive of the *opposite* pattern than expected: solutions rated with high suddenness and low effort were produced at a faster rate.

In total, the data present no evidence that the self-report Aha! ratings are at all appropriate in assessing Gestalt defined restructuring. Further, our results indicate that participants interpreted suddenness and effort as reflecting solutions that came quickly and easily, i.e., did not involve impasse and started from a more accurate problem representation. This suggests that previous studies using the Aha! rating scale have, at least in part, erroneously classified solutions as insightful with no instance of impasse or restructuring. This has resulted in data patterns that may be contrary to what is expected from the Gestalt view of insightful problem solving, explaining some of the discrepancies found in the literature.

Our final set of analyses corroborates previous research supporting competing theories of insightful problem solving, providing further evidence that different operational definitions of insight are driving contradictory findings. Analyzing the relationship between insight and working memory, we demonstrated that assessing insight with the Aha! ratings resulted in support for the Business-As-Usual theory, while assessing insight as solution after impasse supported the Gestalt theory of insight. These differences can be attributed to the miscategorization of quick and easy solutions as involving restructuring, when they actually were neither misrepresented nor involved the occurrence of impasse.

Our results clearly illustrate that the self-report Aha! rating scale lacks validity. The profound implication of this statement is that over a decade of research using this scale as the sole means of assessing restructuring has produced questionable results. Further, our data imply that the feelings traditionally associated with an Aha! experience are not necessarily indicative of the underlying processes. Many researchers have overemphasized the importance of these feelings and presented the Aha! experience as synonymous with restructuring with little evidence that these constructs are reliably related (see Ash et al., 2009, for review of this issue). While our results indicated that the subjective Aha! experience is not an appropriate method for assessing restructuring, they clearly support that focusing on how a person overcomes impasse is. By this definition, our work consistently demonstrated support for the Gestalt theory that restructuring is a component specific to the insight problem solving process with the deeper implication that modern information processing theories grounded in this view accurately assume that this process is largely automatic.

REFERENCES

Altman, D. G. (1991). *Practical statistics for medical research.* New York: Chapman and Hall.

Ash, I. K., Cushen, P. J., & Wiley, J. (2009). Obstacles in Investigating the Role of Restructuring in Insightful Problem Solving. *The Journal of Problem Solving, 2*(2), Ariticle 3. doi: http://dx.doi.org/10.7771/1932-6246.1056

Ash, I. K., Jee, B. D., & Wiley, J. (2012). Investigating Insight as Sudden Learning. *The Journal of Problem Solving, 4*(2), 1-27.

Ash, I. K., & Wiley, J. (2006). The nature of restructuring in insight: An individual-differences approach. *Psychonomic Bulletin & Review, 13*(1), 66-73. doi: 10.3758/BF03193814

Ash, I. K., & Wiley, J. (2008). Hindsight bias in insight and mathematical problem solving: Evidence of different reconstruction mechanisms for metacognitive versus situational judgments. *Memory & Cognition, 36*(4), 822-837. doi: 10.3758/MC.36.4.822

Bowden, E. M., & Jung-Beeman, M. (2003). Aha! Insight experience correlates with solution activation in the right hemisphere. *Psychonomic Bulletin & Review, 10*(3), 730-737. doi: 10.3758/BF03196539

Chein, J. M., & Weisberg, R. W. (2013). Working memory and insight in verbal problems: Analysis of compound remote associates. *Memory & Cognition*, No Pagination Specified. doi: 10.3758/s13421-013-0343-4

Cunningham, J. B., MacGregor, J. N., Gibb, J., & Haar, J. (2009). Categories of insight and their correlates: An exploration of relationships among classic-type insight problems, rebus puzzles, remote associates and esoteric analogies. *The Journal of Creative Behavior, 43*(4), 262-280. doi: 10.1002/j.2162-6057.2009.tb01318.x

Davidson, J. E. (1995). The suddenness of insight. In R. J. S. J. E. Davidson (Ed.), *The nature of insight* (pp. 125-155). Cambridge, MA, US: The MIT Press.

Duncker, K., & Lees, L. S. (1945). On problem-solving. *Psychological Monographs, 58*(5), i-113. doi: 10.1037/h0093599

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, *128*(3), 309-331. doi: 10.1037/0096-3445.128.3.309

Evans, J. S. B. T. (2003). In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences, 7*(10), 454-459. doi: 10.1016/j.tics.2003.08.012

Evans, J. S. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, *59*, 255-278. doi: 10.1146/annurev.psych.59.103006.093629

Evans, J. S. B. T. (2011). Dual-process theories of reasoning: Contemporary issues and developmental applications. *Developmental Review, 31*(2-3), 86-102. doi: 10.1016/j.dr.2011.07.007

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149-1160.

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*, 175-191.

Fleck, J. I. (2008). Working memory demands in insight versus analytic problem solving. *European Journal of Cognitive Psychology, 20*(1), 1-38. doi: 10.1080/09541440601016954

Gilhooly, K. J., & Fioratou, E. (2009). Executive functions in insight versus non-insight problem solving: An individual differences approach. *Thinking & Reasoning, 15*(4), 355-376. doi: 10.1080/13546780903178615

Gilhooly, K. J., & Murphy, P. (2005). Differentiating insight from non-insight problems. *Thinking & Reasoning, 11*(3), 279-302. doi: 10.1080/13546780442000187

Jung-Beeman, M., Bowden, E. M., Haberman, J., Frymiare, J. L., Arambel-Liu, S., Greenblatt, R., . . . Kounios, J. (2004). Neural activity when people solve verbal problems with insight. *Public Library of Science Biology, 4*, 1-23.

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist, 58*(9), 697-720. doi: 10.1037/0003-066X.58.9.697

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The Generality of Working Memory Capacity: A Latent-Variable Approach to Verbal and Visuospatial Memory Span and Reasoning. *Journal of Experimental Psychology: General, 133*(2), 189-217. doi: 10.1037/0096-3445.133.2.189

Kaplan, C. A., & Simon, H. A. (1990). In search of insight. *Cognitive Psychology, 22*(3), 374-419. doi: 10.1016/0010-0285(90)90008-R

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. doi: 10.1037/0278-7393.25.6.1534

Koffka, K. (1935). *Principles of Gestalt psychology*. Oxford, England: Harcourt, Brace.

Köhler, W. (1948). The mentality of apes (1917). *Readings in the history of psychology* (pp. 497-505). East Norwalk, CT, US: Appleton-Century-Crofts.

Kounios, J., Frymiare, J. L., Bowden, E. M., Fleck, J. I., Subramaniam, K., Parrish, T. B., & Jung-Beeman, M. (2006). The prepared mind: Neural activity prior to problem presentation predicts subsequent solution by sudden insight. *Psychological Science, 17*(10), 882-890. doi: 10.1111/j.1467-9280.2006.01798.x

Maier, N. R. F. (1931). Reasoning in humans. II. The solution of a problem and its appearance in consciousness. *Journal of Comparative Psychology, 12*, 181-194. doi: 10.1037/h0071361

Mednick, S. A. (1968). The Remote Associates Test. *The Journal of Creative Behavior, 2*(3), 213-214.

Metcalfe, J., & Wiebe, D. (1987). Intuition in insight and noninsight problem solving. *Memory & Cognition, 15*(3), 238-246. doi: 10.3758/BF03197722

Ohlsson, S. (1992). Information processing explanations of insight and related phenomena. In M. T. Keane & K. J. Gilhooly (Eds.), *Advances in the psychology of thinking*. London: Harbester-Wheatsheaf.

Ohlsson, S. (2011). *Deep learning: How the mind overrides experience*. New York, NY, US: Cambridge University Press.

Ricks, T., Turley-Ames, K., & Wiley, J. (2007). Effects of working memory capacity on mental set due to domain knowledge. *Memory & Cognition, 35*(6), 1456-1462. doi: 10.3758/BF03193615

Schneider, W., Eschman, A., & Zuccolotto, A. (2012a). *E-Prime Reference Guide*. Pittsburgh: Psychology Software Tools, Inc.

Schneider, W., Eschman, A., & Zuccolotto, A. (2012b). *E-Prime User's Guide*. Pittsburgh: Psychology Software Tools, Inc.

Schooler, J. W., Ohlsson, S., & Brooks, K. (1993). Thoughts beyond words: When language overshadows insight. *Journal of Experimental Psychology: General, 122*(2), 166-183. doi: 10.1037/0096-3445.122.2.166

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences, 23*(5), 645-665. doi: 10.1017/S0140525X00003435

Thorndike, E. L. (1911). *Animal intelligence: Experimental studies.* Lewiston, NY, US: Macmillan Press.

Wertheimer, M. (1954/1959). *Productive thinking.* New York: Harper & Brothers.

APPENDIX A

INFORMED CONSENT DOCUMENT

INFORMED CONSENT DOCUMENT
OLD DOMINION UNIVERSITY

PROJECT TITLE: Working Memory and Problem Solving

INTRODUCTION
The purposes of this form are to give you information that may affect your decision whether to say YES or NO to participation in this research, and to record the consent of those who say YES to participation in this study.

RESEARCHERS
Responsible Project Investigator:
Ivan K. Ash, Ph.D.
Assistant Professor
College of Sciences
Department of Psychology
Old Dominion University
iash@odu.edu
757-683-4446

Investigator:
Kimberly D. Lee
Master's Candidate
College of Sciences
Department of Psychology
Old Dominion University
kleex027@odu.edu
757-683-4439

DESCRIPTION OF RESEARCH STUDY
This experiment investigates how people solve cognitive problems. Pending your voluntary consent, you will participate in a one-hour (approximately) experiment. You will be asked to solve a series of word problems, puzzle problems, memory tasks, as well as answer questions about your opinion on different aspects of these problems. Afterwards, you will be debriefed by the experimenter and will have an opportunity to ask any questions that you may have about this experiment before leaving. Approximately 300 volunteers will be participating in this study.

Video Recording:

One section of this experiment will be video recorded for later data collection. During this section, you will be asked to solve problems while verbalizing your mental problem solving processes. Your anonymity and confidentially is of utmost importance. The video will be recorded on a closed computer that is not connected to the internet nor can be otherwise accessed by any other computer. In the video, you will only be identified by your participant number and that number cannot be linked to your name or SONA I.D. number. After we have collected the data from the video, the video will be deleted. At no time will any portion of the video footage or images from the video footage be used for presentation or publishing purposes.

EXCLUSIONARY CRITERIA
You must be at least 18 years of age and have normal or corrected-to-normal vision.

RISKS AND BENEFITS
RISKS: There are no substantial risks for participants in this study. However, as with any research, there is always the possibility that you may be subject to risks that have not yet been identified. If at any point during the course of the experiment you feel uncomfortable, remember that your participation is voluntary and you may end your participation at any time without penalty.

BENEFITS: There are no direct benefits from participating in this study.

COSTS AND PAYMENTS
All participants will receive 1.5 Psychology Research Participation (SONA) credit for participation in this study.

NEW INFORMATION
If the researchers find new information during this study that would reasonably change your decision about participating, then they will give it to you.

CONFIDENTIALITY
All information obtained about you in this study is strictly confidential unless disclosure is required by law. The results of this study may be used in reports, presentations and publications. All results will be reported in the aggregate, and the researcher will not identify you. Although your name and email were used to make your appointment and will be used to assign research credit, you will be assigned a participant number which cannot be connected to this information. This number will be used to organize all your responses. Therefore, your identity can never be associated with your questionnaire responses or performance data. Your responses will be held in strict confidentiality, in accordance and observation with ethical guidelines established by the American Psychological Association (APA).

WITHDRAWAL PRIVILEGE
It is OK for you to say NO. Even if you say YES now, you are free to say NO later, and walk away or withdraw from the study at any time. Your decision will not affect your relationship with Old Dominion University, or otherwise cause a loss of benefits to which you might otherwise be entitled. Also, the investigators reserve the right to withdraw your participation at any time throughout this investigation.

COMPENSATION FOR ILLNESS AND INJURY
If you say YES, then your voluntary consent in this document does not waive any of your legal rights. It is highly improbable and unlikely that any illness or injury will result from your participation with this research project. However, in the event of any harm arising from this study, neither Old Dominion University nor the researchers are able to give you any money, insurance coverage, free medical care, or any other compensation for such injury. In the event that you suffer harm as a result of participation in this research project, you may contact the Responsible Project Investigator, Dr. Ivan Ash at 757-683-4446, the current IRB chair, Dr. George Maihafer at 757-683-4520, or the Old Dominion University Office of Research at 757-683-3460 who will be glad to review the matter with you.

VOLUNTARY CONSENT
By signing this form, you are saying several things. You are saying that you have read this form or have had it read to you, that you are satisfied that you understand this form, the research study, and its risks and benefits. The researchers should have answered any questions you may have had about the research. If you have any questions later on, please contact the Responsible Project Investigator, Dr. Ivan Ash, at 757-683-4446.

If at any time you feel pressured to participate, or if you have any questions about your rights or this form, then you should call Dr. George Maihafer, the current IRB chair, at 757-683-4520, or the Old Dominion University Office of Research, at 757-683-3460.

And importantly, by signing below, you are telling the researcher YES, that you agree to participate in this study. The researcher should give you a copy of this form for your records.

Participant's Printed Name	**Participant's Signature**	**Date**

INVESTIGATOR'S STATEMENT

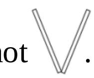
I certify that I have explained to this subject the nature and purpose of this research, including benefits, risks, costs, and any experimental procedures. I have described the rights and protections afforded to human subjects and have done nothing to pressure, coerce, or falsely entice this subject into participating. I am aware of my obligations under state and federal laws, and promise compliance. I have answered the subject's questions and have encouraged him/her to ask additional questions at any time during the course of this study. I have witnessed the above signature(s) on this consent form.

Investigator's Printed Name	**Investigator's Signature**	**Date**

APPENDIX B

INSIGHT PROBLEMS

In the problem below, matchsticks form Roman numerals in an equation. Notice that both sides of the equation are **not** equal. Make these matchsticks into a correct arithmetic equation by moving **only a single matchstick**. The specific rules are:

A) Only one matchstick can be moved.
B) A matchstick cannot be discarded; that is, it can only be moved from one position in the equation to another.
C) An upright matchstick cannot count as a slanted stick, so  is not .
D) The result must be a correct arithmetic equation.

Move one matchstick to make the following into a correct equation.

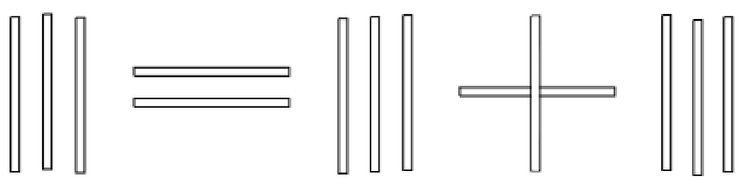


Figure B1. Matchsticks operator problem.

In the problem below, matchsticks form Roman numerals in an equation. Notice that both sides of the equation are **not** equal. Make these matchsticks into a correct arithmetic equation by moving **only a single matchstick**. The specific rules are:

A) Only one matchstick can be moved.
B) A matchstick cannot be discarded; that is, it can only be moved from one position in the equation to another.
C) An upright matchstick cannot count as a slanted stick, so  is not .
D) The result must be a correct arithmetic equation.

Move one matchstick to make the following into a correct equation.

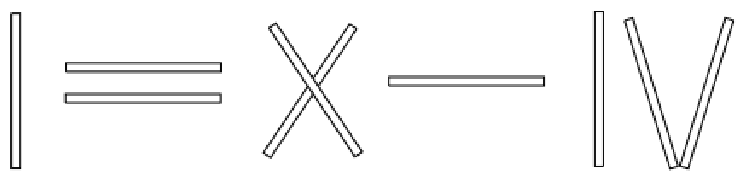


Figure B2. Matchsticks numeral problem.

This problem consists of 6 glasses. The first 3 glasses contain liquid. Describe how you could make it so no 2 glasses containing liquid are next to each other and no 2 empty glasses are next to each other, while keeping 3 of the 6 glasses full. To do this, you are only allowed to move **1 glass.**

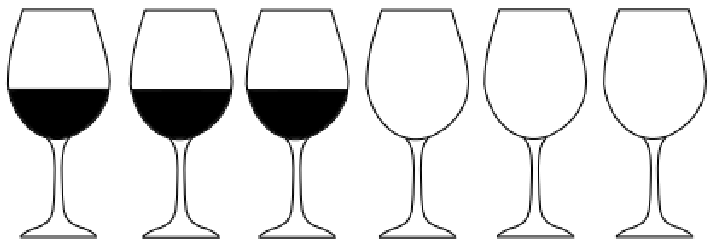


Figure B3. Glasses problem.

In this problem, individual circles form a triangle that points to the top of the page. Move **3 circles only** to get the triangle to point to the bottom of the page.

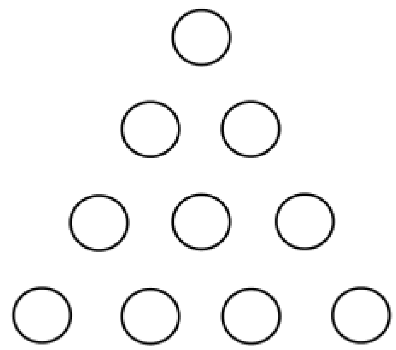


Figure B4. Circles problem.

In this problem, there are 8 coins. Move **2 coins** so that each coin touches **exactly** 3 other coins. The coins must be separated into **two groups**.

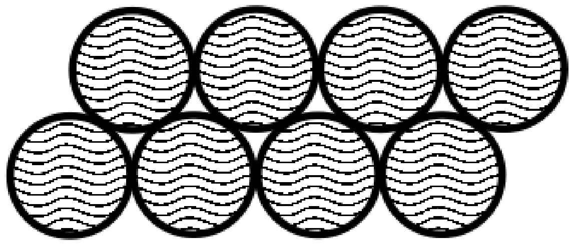


Figure B5. Coins problem.

Move 3 of the sticks to make 5 squares.

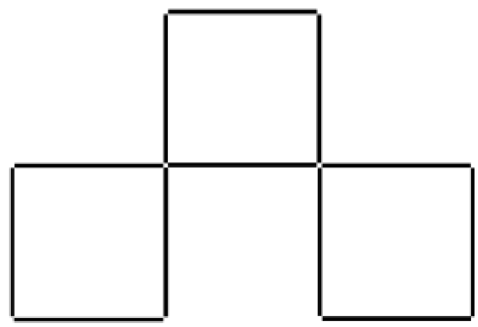


Figure B6. Squares problem.

APPENDIX C

AHA! RATINGS

Sometimes when we solve a problem, we feel unsure that our answer is the correct solution (*Not Confident*). Other times, we feel absolutely certain that the answer we have come up with is correct (*Very Confident*).

How confident are you in your solution? *Not Confident* (1) to *Very Confident* (4)

Sometimes when we solve a problem, we feel like we consistently make progress towards the solution; that we take incremental steps that lead to the final solution and the solution didn't just come "out of the blue" (*Not Sudden*). Other times, we can work on a problem and not feel as though we are making any progress towards a solution, but the solution will just "pop into our head" (*Very Sudden*).

How sudden did the solution come to you? *Not Sudden* (1) to *Very Sudden* (4)

Sometimes when we solve a problem, the answer comes to us easily, without having to use a lot of mental effort (*Not Effortful*). Other times, we have to think very hard and it feels like it takes a lot of mental work to come up with a solution (*Very Effortful*).

How much effort was required to find the solution? *Not Effortful* (1) to *Very Effortful* (4)

APPENDIX D

READING SPAN STIMULI

No matter how much we talk to him, he is never going to change.	?	J
The prosecutor's dish was lost because it was not based on fact.	?	M
Every now and then I catch myself swimming blankly at the wall.	?	F
We were fifty lawns out at sea before we lost sight of land.	?	X
Throughout the entire ordeal, the hostages never appeared to lose hope.	?	L
Paul is afraid of heights and refuses to fly on a plane.	?	R
The young pencil kept his eyes closed until he was told to look.	?	B
Most people who laugh are concerned about controlling their weight.	?	Q
When Lori shops she always looks for the lowest flood.	?	H
When I get up in the morning, the first thing I do is feed my dog.	?	M
After yelling at the game, I knew I would have a tall voice.	?	X
Mary was asked to stop at the new mall to pick up several items.	?	L
When it is cold, my mother always makes me wear a cap on my head.	?	Q
All parents hope their list will grow up to be intelligent.	?	H
When John and Amy moved to Canada, their wish had a huge garage sale.	?	B
In the fall, my gift and I love to work together in the yard.	?	F
At church yesterday morning, Jim's daughter made a terrible plum.	?	R
Unaware of the hunter, the deer wandered into his shotgun range.	?	J
Since it was the last game, it was hard to cope with the loss.	?	J
Because she gets to knife early, Amy usually gets a good parking spot.	?	B
The only furniture Steve had in his first bowl was his waterbed.	?	R
Last year, Mike was given detention for running in the hall.	?	Q
The huge clouds covered the morning slide and the rain began to fall.	?	X
After one date I knew that Linda's sister simply was not my type.	?	M
Jason broke his arm when he fell from the tree onto the ground.	?	H
Most people agree that Monday is the worst stick of the week.	?	L
On warm sunny afternoons, I like to take a walk in the park.	?	F
With intense determination he overcame all obstacles and won the race.	?	B
A person should never be discriminated against based on his race.	?	M
My mother has always told me that it is not polite to shine.	?	L
The lemonade players decided to play two out of three sets.	?	F

Raising children requires a lot of dust and the ability to be firm.	?	H
The gathering crowd turned to look when they heard the gun shot.	?	J
As soon as I get done taking this envy I am going to go home.	?	X
Sue opened her purse and found she did not have any money.	?	Q
Jill wanted a garden in her backyard, but the soil was mostly clay.	?	R
Stacey stopped dating the light when she found out he had a wife.	?	F
I told the class that they would get a surprise if they were orange.	?	R
Jim was so tired of studying, he could not read another page.	?	Q
Although Joe is sarcastic at times, he can also be very sweet.	?	X
Carol will ask her sneaker how much the flight to Mexico will cost.	?	L
The sugar could not believe he was being offered such a great deal.	?	H

APPENDIX E

OPERATION SPAN STIMULI

IS (10 ÷ 2) – 3 = 2 ? sea
IS (10 ÷ 10) – 1 = 2 ? class
IS (7 ÷ 1) + 2 = 7 ? paint

IS (3 ÷ 1) – 2 = 3 ? cloud
IS (2 x 1) – 1 = 1 ? pipe
IS (10 ÷ 1) + 3 = 13 ? ear
IS (9 x 2) + 1 = 18 ? flame
IS (9 ÷ 1) – 7 = 4 ? bike

IS (8 x 4) – 2 = 32 ? bean
IS (9 x 3) – 3 = 24 ? arm
IS (4 ÷ 1) + 1 = 4 ? ground

IS (10 ÷ 1) – 1 = 9 ? hole
IS (8 x 4) + 2 = 34 ? dad

IS (6 x 3) + 2 = 17 ? kid
IS (6 ÷ 3) + 2 = 5 ? fork
IS (6 x 2) – 3 = 10 ? jail
IS (8 ÷ 2) + 4 = 2 ? hat
IS (8 ÷ 2) – 1 = 3 ? lamp

IS (9 ÷ 1) – 5 = 4 ? cave
IS (6 ÷ 2) – 2 = 2 ? back
IS (7 x 2) – 1 = 14 ? hall
IS (6 x 2) – 2 = 10 ? fern

IS (2 x 2) + 1 = 4 ? man
IS (7 x 1) + 6 = 13 ? world

IS (10 ÷ 1) + 1 = 10 ? calf
IS (4 x 4) + 1 = 17 ? fish
IS (3 x 3) – 1 = 8 ? cheek

IS (3 x 1) + 2 = 2 ? bread
IS (4 ÷ 2) + 1 = 6 ? germ
IS (5 ÷ 5) + 1 = 2 ? dock

IS (2 x 3) + 1 = 4 ? game
IS (9 ÷ 3) – 2 = 1 ? nerve
IS (10 ÷ 2) – 4 = 3 ? wax
IS (5 ÷ 1) + 4 = 9 ? tin
IS (10 x 2) + 3 = 23 ? church

IS (7 ÷ 1) + 6 = 12 ? beach
IS (3 x 2) + 1 = 6 ? card

IS (6 x 4) + 1 = 25 ? job
IS (9 ÷ 3) – 1 = 2 ? cone
IS (8 ÷ 1) – 6 = 4 ? brass
IS (9 x 1) + 9 = 1 ? street

VITA

KIMBERLY DIANE LEE

Department of Psychology
250 Mills Godwin Life Sciences Bldg
Norfolk, VA 23529
757-683-4439 (office)
kleex027@odu.edu

Education

Bachelor of Arts, Music Theatre Performance — May 2002
Christopher Newport University, Newport News, VA

Bachelor of Science, Psychology — May 2012
Old Dominion University, Norfolk, VA

Master of Science, Experimental Psychology — December 2015
Old Dominion University, Norfolk, VA

Conference Presentations & Posters

Ash, I. K., May, R. W., & Lee, K. D. (2012, May). *Hindsight bias on confidence and outcome likelihood judgments due to different mechanisms.* Poster presentation at the 24th Annual Convention of the Association of Psychological Sciences: Chicago, IL.

Lee, K. D. , May, R. W., & Ash, I. K. (2012, April). *Not all hindsight biases are the same: Effects of expectation and surprise on memory for prior judgments.* Paper presentation at the Virginia Psychological Association 2012 Spring Conference: Norfolk, VA.