# Is my Aha! bigger than yours? Investigating Individual Differences in the Experience of Insight.

Natalie M Pepping

Bachelor of Science in Psychology (Honours)

Murdoch University

*This thesis is presented in partial fulfilment of the requirements for the degree of Bachelor of Science (Honours), Murdoch University (2019).*

Murdoch
UNIVERSITY

# Declaration

*I declare that this thesis is my own account of my research and contains as its main content work which has not previously been submitted for a degree at any tertiary educational institution.*

Natalie M Pepping

# Table of Contents

**Character Count: 64, 992**

**Word Count: 9,837**

# Acknowledgements

Firstly, I would like to thank my supervisor Dr Helen Davis for her constant support and wisdom throughout the year. Dr Davis' patience, guidance and insight during every twist and turn of this project have been greatly appreciated. I would also like to thank my family, Steve, Ruby and Olivia, for their encouragement and motivation throughout this process. Steve, your unwavering faith in me, practical and technical support and calming presence have been invaluable. Finally, I would like to thank Carina, Dana and Emily. Without you and the endless stream of hilarious anecdotes, this year would have been impossible. Thank you.

# Abstract

A Neural Network Theory (NNT) account of insight processes and individual differences in cognitive ability predicts that, compared to routine problem solving, insight experiences will be associated with less involvement of control functions and will occur less frequently among people with greater fluid ability. The present study investigated the role of fluid intelligence and metacognitive control in the elicitation of Aha experiences. Seventy-six participants, predominantly university students (84% female), attempted a set of problems, including classic insight, non-insight and riddles. Subjective experiences of insight, certainty and suddenness of the solution process were measured, using a purpose-built concurrent rating apparatus and retrospective report. Participants completed Raven's Advanced Progressive Matrices (fluid intelligence) and an attention-switching task (metacognitive control). Hierarchical Generalised Linear Modelling was used to model Aha experiences as a function of item-level predictors (Level 1) and person-characteristics (Level 2). The overall odds of reporting an Aha experience were 0.42. Higher fluid intelligence, but not metacognitive control, was associated with reduced odds of reporting Aha on a problem (OR=0.88, 95% CI: 0.82,0.95), controlling for accuracy, solution suddenness, and verbal skills. Aha experiences were significantly associated with multiple theoretically meaningful retrospective and concurrent problem-solving experience ratings, with fluid intelligence moderating some associations. These findings support the NNT account of insight as a special process and fluid intelligence as a factor limiting the complexity, and accessible solution states from the initial problem representation, leading to the requirement for an alternative representation. The study demonstrates some methodological solutions to difficulties inherent in measuring insight.

# Is my Aha! bigger than yours? Investigating individual differences in the experience of insight.

A chimpanzee named Sultan, two sticks, and a hard-to-reach banana; this is the scenario from which almost a century of research investigating “Aha!” experiences emerged (Ohlsson, 1992). Early Gestalt psychologist Wolfgang Kohler observed that after repeated attempts to reach the banana, Sultan entered a failure-induced sulk. However, he became suddenly re-energised, and purposefully joined two sticks together to retrieve the banana (Kohler, 1921 cited in Ohlsson 1992). How, after numerous attempts and apparent acceptance of failure, did the solution to this intractable problem suddenly appear in Sultan's consciousness? What processes simmering beyond conscious awareness conjured a fully formed solution and planted it so abruptly into Sultan's dormant and tortured mind? "Aha!"

Aha experiences are thought to be indicative of a moment of insight and have historically been associated with exceptional creativity, scientific discovery and genius (Hill & Kemp, 2018; Metcalfe & Wiebe, 1987; Shen et al., 2016; Sternberg & Davidson, 1995). These experiences have been shown to be distinct neurophysiological phenomena (Bowden & Jung-Beeman, 2003; Kounios et al., 2006; Sandkühler & Bhattacharya, 2008; Tik et al., 2018) that facilitate memory (Danek et al., 2013), improve learning (Dominowski & Buyer, 2000), and provide motivation during difficult learning (Liljedahl, 2005).

Current definitions of insight moments describe them as the occurrence of a solution or path to a solution suddenly and unexpectedly coming to mind following a pause in active thinking when a problem-solver feels unable to make further progress (Bowden et al., 2005; Sternberg & Davidson, 1995). Insight is asserted to be a special process that is distinct from analytical problem-solving (Knoblich et al., 1999; Ohlsson, 1992; Sternberg & Davidson, 1995). Analytical problem-solving is continuous and incremental and does not engender a salient Aha experience (Schooler et al., 1993). Despite a substantial body of research seeking to demystify these processes, the specialness of insight is still the subject of much debate.

Recent research suggests that the lack of clarity is a result of the way insight is operationalised in many studies as “solving an insight problem” (Danek et al., 2016; Webb et al., 2016) without verifying that the problem-solver has experienced an Aha moment. These studies indicate that insight is not reliably evoked by these problems, suggesting the processes engaged in solving insight problems may be idiosyncratic (Danek et al., 2016; Webb et al., 2016). That is, it is possible that rather than task requirements, individual differences of the problem-solver influence whether or not insight processes are used to solve a problem and the subsequent occurence of an Aha experience.

Two individual characteristics that may be relevant to the Aha experience are fluid intelligence and metacognitive control. Fluid intelligence is defined as the ability to use controlled and deliberate mental operations to solve problems, deduce patterns, identify relations and draw inferences (McGrew, 2009). Metacognitive control is a facet of metacognition that refers explicitly to the control processes involved in regulating and directing information processing resources (Nelson & Narens, 1990). Differing levels of these abilities may influence the cognitive processes engaged during problem-solving (Dix et al., 2016). Thus, the central aim of this study is to investigate whether individual differences in fluid intelligence or metacognitive control influence whether an Aha experience occurs upon problem-solving.

The present study agrees with several others that an Aha reported by the problem-solver is verification that insight has occurred (Bowden et al., 2005; Danek et al., 2016; Webb et al., 2016). However, some researchers argue that Aha experiences occur randomly (Chuderski, 2014) or are related to post-solution affect of evaluations of the solution

(Topolinski & Reber, 2010). The second aim of the present study is to determine if Aha moments are associated with problem-solving experiences that are indicative of a special process. Due to current limitations in methodology, a new device was developed to accomplish this aim. This is described in section 2.5.

## 1.1 Operationalising Insight

Multiple studies have sought to investigate the cognitive processes that precede Aha experiences using predefined “insight problems” (for example, Ash et al., 2009; Chuderski & Jastrzebski, 2018). These problems are thought to elicit the processes theorised to precede an insight moment such as an initial, constrained problem representation, impasse, and sudden restructuring of the problem representation leading immediately to the solution (Knoblich et al., 1999; Ohlsson, 1992; Weisberg, 2015). Correctly solving these problems is used to operationalise insight, and differences in cognitive abilities between solvers and non-solvers are used to infer whether or not a special process was required to solve them (Webb et al., 2017)

Generally, these studies find insight is positively associated with cognitive abilities (for a review Chu & MacGregor, 2011). However, the assumption that incorrect solutions cannot be indicative of insight is problematic. Evidence of false insights (Danek & Wiley, 2016) suggests that while accuracy is associated with Aha (Danek & Salvi, 2018), it is not necessary for eliciting an Aha experience. Thus, if we seek to determine the processes preceding an Aha experience, analyses must be anchored on this experience, unconstrained by this assumption.

Recent studies suggest that solving insight problems does not reliably elicit Aha experiences (Danek et al., 2016; Webb et al., 2016). Webb et al. (2016) also found that Aha experiences were reported when solving problems defined as “non-insight” problems which are often used as control problems in insight studies. Non-insight problems are those that can be solved via analytical processes (Weisberg, 2015). These findings are consistent with the view of several researchers that any problem can be solved via insight or by analytical means (Bowden et al., 2005; Danek et al., 2013; Webb et al., 2016). It also indicates that the assumption that insight processes have occurred when solving insight problems is unwarranted (Webb et al., 2016). Furthermore, these studies suggest that the evocation of an insight process and subsequent Aha is not necessarily dictated by task requirements but may be influenced by problem-solver characteristics.

Consider the earlier description of Sultan and the irretrievable banana. A parallel scenario is possible. If, for example, instead of Sultan, Bea is in the enclosure. Bea tries multiple strategies to get the banana, methodically picking up tools until eventually the two sticks are in her hands and the solution becomes apparent. Bea used steady, incremental problem-solving. Comparing these two scenarios demonstrates the fundamental problem with operationalising insight as solving a predefined problem. The same solution was discovered, but the problem-solving and solution experiences were different.

There is evidence to suggest that different processes are engaged by (human) problem-solvers to solve the same problem. Fleck and Weisberg (2013) demonstrated that diverse methods are used to solve insight problems. This study used verbal protocols to delineate problem-solving processes used by participants to solve insight problems and found that different methods were reported to solve the same problems. Fleck and Weisberg (2013) reported that processes varied across individuals from analytical to insight and suggested that further investigation using a range of tasks may enhance our understanding of this variability.

The diversity of processes used to solve insight problems may explain the lack of reliability in eliciting Aha experiences using predefined tasks, and suggest the potential role of individual differences in the evocation of insight (Fleck & Weisberg, 2013). Before exploring the role of individual differences, it is necessary to seek an explanation via the theories that underpin insight research.

## 1.2 Theories of Insight

Two major perspectives have emerged in this field (Weisberg, 2015). These are referred to as the Special-Process and Nothing-Special perspectives (Bowden et al., 2005; Chuderski, 2014).

**Insight is a special process.** Special process theories describe insight during problem-solving as a result of a mental shift from an initial representation of a problem to a qualitatively different representation (Figure 1) when the configuration of the original representation fails to produce a solution. That is, the problem-solver reaches an impasse at which time they feel unable to make further progress (Ash et al., 2009). The new representation leads to the immediate and sudden emergence of a solution or understanding of the problem in the mind of the problem-solver (Ash et al., 2009), and the subsequent experience of Aha (Sternberg & Davidson, 1995; Weisberg, 2015). This process, shown in Figure 1, is thought to be spontaneous and to occur outside of conscious control (Smith, 2012).

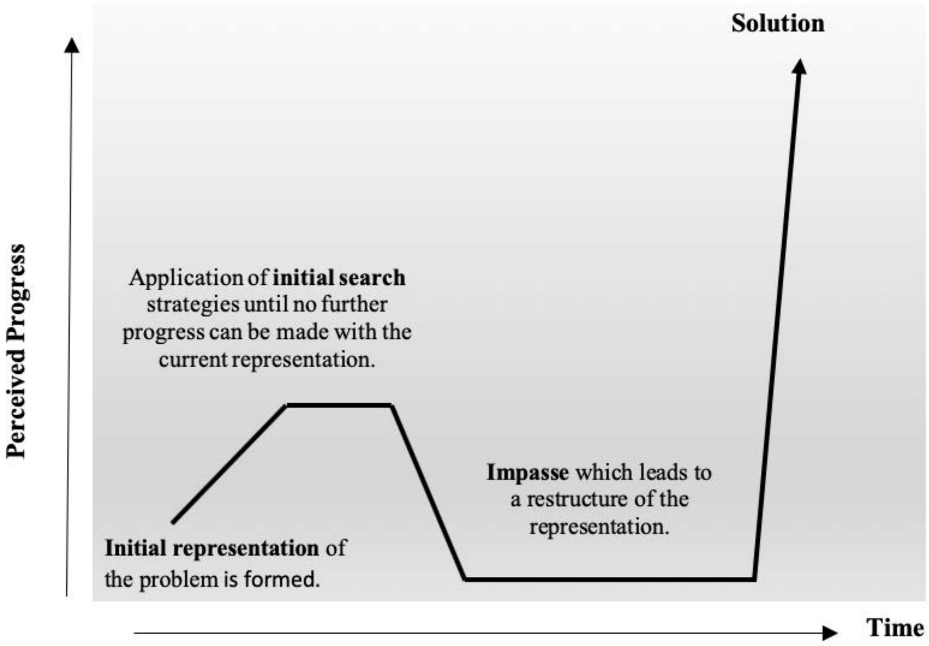


*Figure 1.* The trajectory of problem-solving with insight according to the Special-Process perspective.

***Redistribution theory.*** (RDT, Smith, 2012) RDT suggests that failure to access a solution from the initial representation of the problem causes a restructuring of the problem representation. This restructure opens the problem-space to allow a spread of activation and a previously inaccessible, new representation of the problem to form (Ash et al., 2009). The theory is an extension of Representational Change Theory described by Knoblich et al. (1999) but draws extensively from information-processing theories (Smith, 2012) emphasising the importance of the existing structure of knowledge networks and perceptual constraints inherent in problem presentation. The existing networks and perceptual constraints govern which parts or nodes of the network are triggered by the initial representation, and this defines the problem-space (Smith, 2012). Negative feedback from the application of unsuccessful analytical search strategies deactivates nodes deemed unhelpful, resulting eventually, in an "unmerited" impasse (p.4, Smith, 2012). An unmerited impasse occurs when individuals can find a solution to the problem but are prevented from doing so due to an incorrect representation of the problem. The impasse leads to a pause, during which the problem-solver feels unable to make progress, and the individual's mind feels subjectively blank (Smith, 2012). Deactivation of multiple nodes in the network leads to a redistribution of activation, resulting in novel patterns of activation, and a sudden qualitative shift in the problem representation (Smith, 2012). Special-process theories suggest that Aha experiences are a result of the suddenness of the shift in the problem representation (Sternberg & Davidson, 1995).

**Insight is Nothing-Special.** The Nothing-Special perspective suggests that differences in the experience of insight do not necessarily denote differences in mental processes underlying them (Weisberg, 2015) and that all problems can be solved via systematic processes. As shown in Figure 2, proponents of this approach suggest that the solution emerges as a result of processes that are controlled, incremental and depend upon memory and reasoning (Chuderski, 2014; MacGregor, Ormerod, & Chronicle, 2001; Weisberg, 2015).

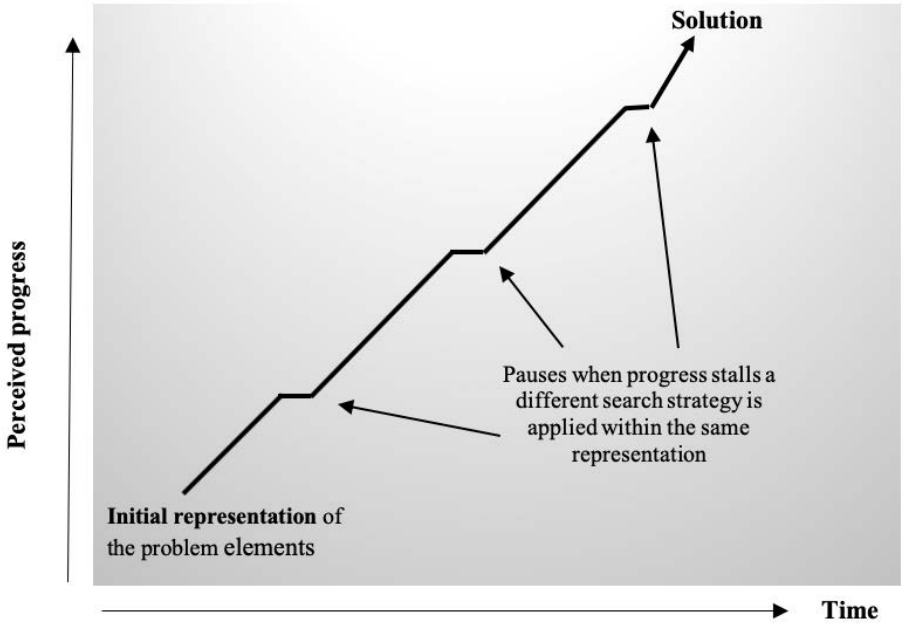


*Figure 2.* The trajectory of problem-solving with insight according to the Nothing-Special perspective.

***Criterion for Satisfactory Progress Monitoring Theory.*** (CSP, MacGregor et al., 2001) CSP describes a process of constant monitoring of progress towards a solution, based on the elements and requirements of the problem (Chu & MacGregor, 2011). MacGregor et al. (2001) describe the basic principles of this theory as follows. The problem-solver seeks to reduce the gap between the current state and a goal-state by identifying an operator or rule that fits the known variables. When the rule is applied, movement in relation to a sub-goal is monitored to ensure sufficient progress is made. Progress is measured against a predetermined criterion which ensures the goal state is still possible after each step. If progress stalls, the problem-space is expanded to enable a more appropriate, alternative rule to be applied.

As shown in Figure 3, this theory asserts that a qualitative shift in the representation of the problem is not required, rather incremental, quantitative progress towards the solution can be made within the same problem-space (Chuderski, 2014). It has been argued that Aha experiences are not systematically related to preceding processes but are related to retrospective evaluations of the solution (Topolinski & Reber, 2010), or a quirk of timing related to the simultaneous realisation of the solution and neurocognitive updating of the problem-space (Chuderski, 2014).

As illustrated in Figure 3, key theoretical differences lie in the accessibility of the solution from the initial representation, the systematicity and continuity of the problem-solving process and the requirement for cognitive restructuring to access the solution.

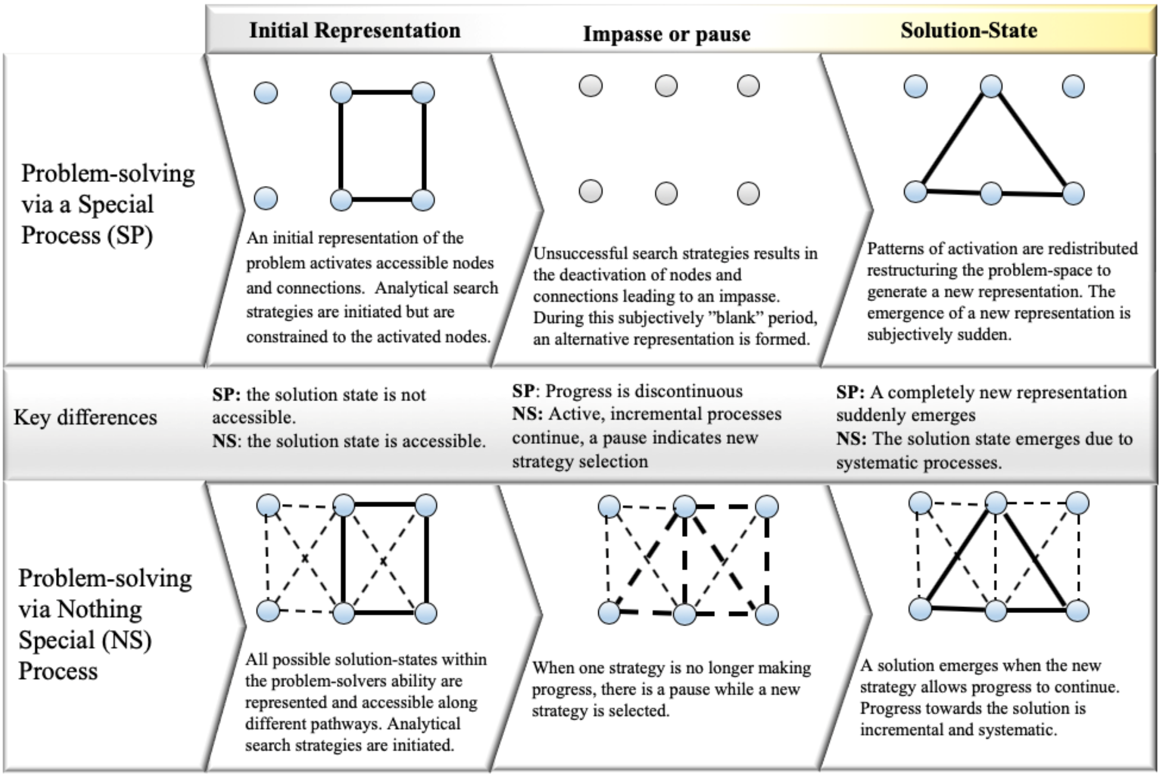


*Figure 3.* Visual representation of problem-solving processes according to the Special Process and Nothing Special perspectives.

## 1.3 The Role of Individual Differences

**Fluid Intelligence.** Fluid intelligence is described as the ability to problem-solve and adapt to novel situations (Cattell, 1963). Network Neuroscience Theory (NNT, Barbey, 2018) describes individual differences in fluid intelligence in terms of the varying extent and connectivity of underlying cognitive architecture, which is a product of knowledge and experience. This theory draws on an information-processing analogy, conceptualising these neural networks as nodes and connections (Barbey, 2018). Perceptual and cognitive processes activate accessible nodes and their connections (Barbey, 2018). Flexible patterns of node activation and more connectivity across the network is indicative of higher fluid intelligence (Barbey, 2018). Specifically, NNT suggests that access to a higher number of weak associative connections enhances fluid ability. Weak associative connections are indicative of indirect learning as opposed to strong connections, which are a result of direct experience (Barbey, 2018).

Several studies support this explanation (Jung & Haier, 2007; Preusse et al., 2011; Suprano et al., 2019). For example, an extensive review by Jung and Haier (2007) reported that higher fluid intelligence is associated with more extensive networks of activity. Preusse et al. (2011) found that higher fluid intelligence, compared to average fluid intelligence, was characterised by flexible modulation of broad regional activity. Suprano et al. (2019) found that children with higher compared to average intelligence scores had different resting-state network activation. In the context of Aha experiences, these theories may explain the possible role of fluid intelligence in the elicitation of Aha experiences.

As illustrated in Figure 4, it may be that fluid intelligence is a limiting factor that constrains the initial problem representation and the number of possible solution states that can be reached within it. For example, individuals with higher fluid intelligence, thus, more weak associative connections, would generate a more complex initial representation compared to someone with lower fluid intelligence who are limited to representing the problem predominantly via strong connections (Barbey, 2018). A more complex initial representation would increase the number of accessible solution-states, increasing the likelihood of solving the problem via analytical search strategies, negating the requirement for restructuring and new representation. To summarise, individuals with higher fluid intelligence would be more likely to solve problems via systematic processes without an Aha, while special processes and an Aha experience are more likely for individuals with lower intelligence.

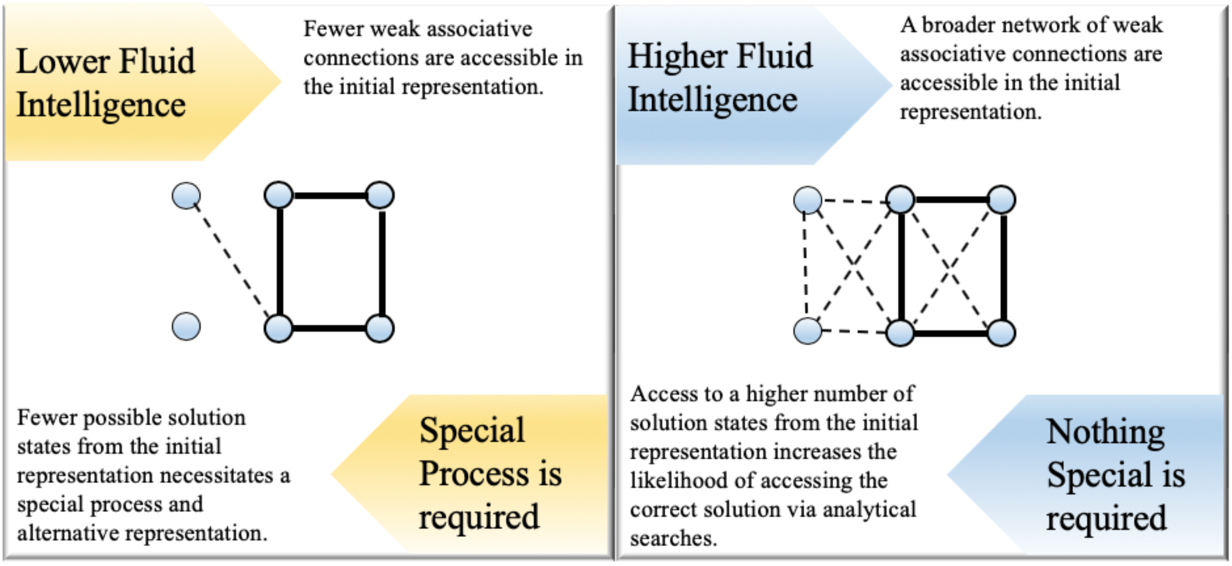
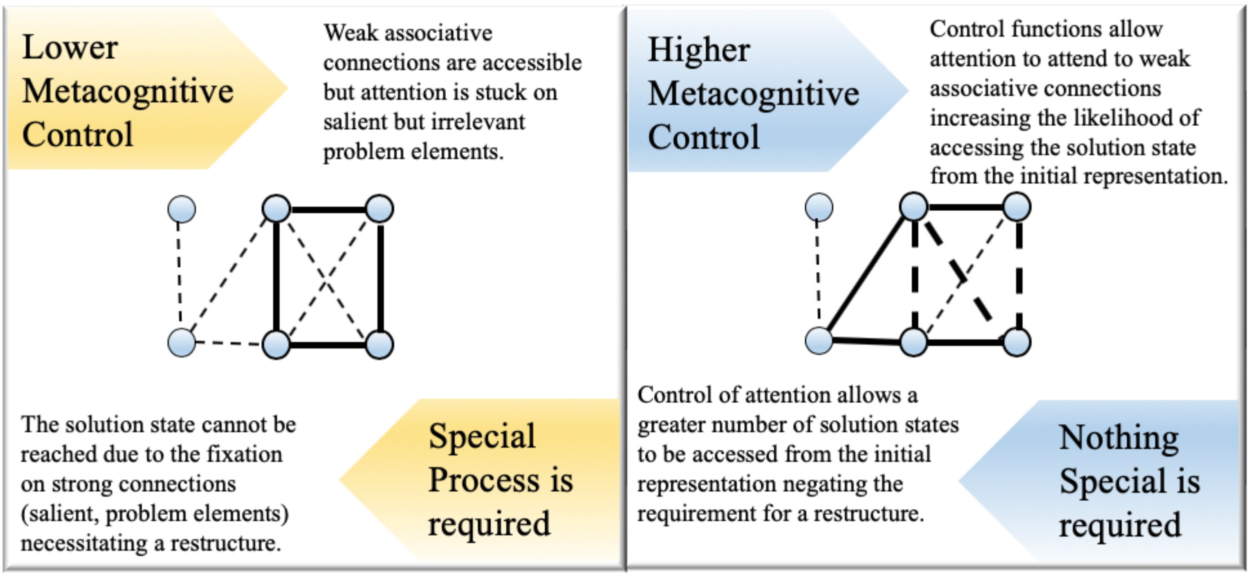


*Figure 4.* Proposed initial representation for individuals with lower and higher fluid intelligence

**Metacognitive Control.** Control processes may also influence the processes engaged to solve a problem by directing attention away from the dominant and more salient connections. Metacognitive control refers to processes that regulate and direct information processing resources (Nelson & Narens, 1990). Individual differences in control functions such as prepotent-response inhibition and goal-directed memory retrieval have been linked to creative cognition which is important for problem-solving (Beaty et al., 2019). Prepotent-response inhibition refers to the ability to prevent dominant response tendencies from interfering with divergent thinking (Benedek & Fink, 2019).

As shown in Figure 5, individuals with higher metacognitive control may be able to selectively attend to critical but non-dominant pathways (Beaty et al., 2019) to resist salient pathways and direct attention towards relevant weak associative connections. In contrast, individuals with lower metacognitive control may have less control over attention, meaning they are unable to direct attention away from salient, but irrelevant aspects of the problem

(Figure 5). This would mean that individuals with higher compared to lower metacognitive control would be less likely to experience an Aha.



Weak associative connections are accessible but attention is stuck on salient but irrelevant problem elements.

Control functions allow attention to attend to weak associative connections increasing the likelihood of accessing the solution state from the initial representation.

The solution state cannot be reached due to the fixation on strong connections (salient, problem elements) necessitating a restructure.

Control of attention allows a greater number of solution states to be accessed from the initial representation negating the requirement for a restructure.

*Figure 5.* Proposed effect of lower and higher metacognitive control on problem-solving processes

Importantly, these explanations may account for some of the variability in Aha experiences when solving insight problems. False insights may also be accounted for under this explanation as the new representation of the problem would be subjectively sudden, but may not be accurate as the cognitive architecture is still constrained by each individuals knowledge and experience (Barbey, 2018). To test whether variability in fluid intelligence and metacognitive control influence problem-solving processes, and subsequent Aha experiences, we need to measure Aha experiences and problem-solving experiences that distinguish between the special process and nothing special perspectives.

## 1.4 Measuring Insight

The second aim of the present study was to investigate whether theoretically meaningful problem-solving experiences predict Aha experiences. Key theoretical differences include the requirement for cognitive restructuring and the systematicity and continuity of problem-solving processes (Knoblich et al., 1999; MacGregor et al., 2001; Ohlsson, 1992; Weisberg, 2015). These differences may give rise to distinct problem-solving experiences (Laukkonen & Tangen, 2018), such as differing levels of systematicity and feelings of certainty throughout problem-solving. For example, solving a problem via a special process may feel less systematic and entail a higher proportion of time feeling uncertain due to the subjective experience of failed initial search strategies and impasse.

Retrospective reports of Aha and solution suddenness are often interpreted as an indication of cognitive restructuring (Salvi et al., 2016; Webb et al., 2017). However, it has been argued that retrospective ratings of suddenness and Aha may be random (Chuderski, 2014), or are not dissociable from affective responses related to retrospective evaluations of the solution itself (Topolinski & Reber, 2010; Webb et al., 2017). Webb et al. (2017) suggested that evaluations of accuracy may elevate post-solution affective ratings due to confidence in the solution and that methods need to be developed to confirm they are dissociable. Topolinski and Reber (2010) suggested that the sudden ease of processing the (now) obvious solution is responsible for the subjective experience of Aha.

Measuring problem-solving experiences in real-time and retrospectively will help to determine whether special processes precede Aha experiences and whether they are dissociable from solution-related affect. For example, if we can predict Aha experiences from reports of systematicity, suddenness and uncertainty retrospectively and throughout problem-solving, this would support the argument that special processes precede an Aha experience. However, methods for measuring these elements are limited, particularly for real-time problem-solving experiences (Laukkonen & Tangen, 2018; Webb et al., 2016).

Current methods used to capture problem-solving experiences in real-time such as onscreen pop-up feelings of warmth ratings (Hedne et al., 2016; Kizilirmak et al., 2018;

Laukkonen & Tangen, 2018) may be disruptive and often fail to provide sufficient data points for analyses (Laukkonen et al., 2018). A device recently proposed by Laukkonen et al. (2018), continuously measures embodied aspects of insight using hand-squeeze-strength as an indicator of perceived progress. This device is promising but may not be suitable for long periods. Thus, a new measurement device will be used to capture ratings of uncertainty during problem-solving. (Refer to Section 2.5)

## 1.5 The Present Study

The central aim of the present study was to investigate whether the frequency of problem-solving culminating in an Aha experience was predicted by individual differences in fluid intelligence and metacognitive control. An NNT account of insight suggests individual differences may influence when Aha experiences (Barbey, 2018). Fluid intelligence was measured using Raven's Advanced Progressive Matrices. Metacognitive control was measured using an attention-switching paradigm.

Another aim was to test if Aha experiences are associated with problem-solving experiences indicative of a special processes. Participants solved a range of problems and rated their problem-solving experiences for each item. Problem-solving experiences were measured using retrospective ratings of Aha, suddenness and systematicity, and concurrent ratings of uncertainty. Due to the nested data structure, hierarchical linear modelling (HLM) was used to analyse the relationship between item responses (Level 1) and person characteristics (Level 2). In HLM the assumption of independence of observations is relaxed, allowing for analyses of nested data (McNeish & Stapleton, 2016) while minimising the likelihood of a Type I error (Bryk & Raudenbush, 1992).

### Hypothesis 1: Aha Experiences and Special Processes

If Aha experiences are indicative of a special process, the following predictions were made:

a. Retrospective solution suddenness ratings will be higher for problems solved with an accompanying Aha experience, compared to solutions without an Aha.

b. Retrospective problem-solving processes will be rated as less systematic for solutions with an Aha experience compared to those without an Aha.

c. Concurrent ratings of uncertainty indicative of non-continuous problem-solving trajectories. This will be indicated by a higher proportion of the problem-solving time, indicating uncertainty about being able to solve the problem.

d. Concurrent ratings will be indicative of uncertainty at the point of solution-finding. This will be determined by concurrent ratings of uncertainty recorded in the 3 seconds before indicating a solution has come to mind.

### Hypothesis 2: Fluid Intelligence and Aha Experiences

If higher fluid intelligence influences the complexity and accessibility of the initial representation of a problem, it is predicted that fluid intelligence will be negatively associated with the rate of Aha experiences reported by participants.

### Hypothesis 3: Metacognitive Control and Aha Experiences

If greater metacognitive control influences the processes engaged in solving a problem, it is predicted that metacognitive control will be negatively associated with the rate of Aha experiences reported by participants.

## Method

### 2.1 Ethics Approval

This study was approved by Murdoch University's Human Research Ethics Committee before data collection (2019/051, Appendix A).

## 2.2 Participants

Volunteer participants were recruited from Murdoch University, social networking websites and word of mouth. A total of 76 participants (84% women) were recruited, of whom two were excluded due to incomplete data. The remaining participants ($N$ = 74) were all 18 years or older (18-35 yrs 67%, 36-54 yrs 30%, 55+ yrs 3%). All participants had normal or corrected-to-normal vision and met English language requirements for university entrance. Participants were allocated 1.5 course credits or a $5.00 voucher for a café at Murdoch University for their participation.

**Sample size.** The study included 76 participants (Level 2 units) with a minimum of 8 problems (Level 1 units). McNeish and Stapleton (2016) found that for hierarchical models with binary outcome variables, the number of units at level 2 (participants) was particularly important, and models with fewer than 30 level 2 units resulted in biased fixed effect estimates. When the number of level 1 units (problems) within each level 2 unit was above five, it was found to produce unbiased estimates (simulation included 100 units at level 2; (McNeish & Stapleton, 2016). Additionally, penalised quasi-likelihood estimates (PQL) provided unbiased variance components with models, including more than ten groups (Austin, 2010).

## 2.3 Design

A hierarchical nested design was used with item responses (Level 1) nested within-person (Level 2). At Level 1 variables were: Aha, retrospective ratings of solution suddenness and problem-solving systematicity, and concurrent ratings of uncertainty at the moment of solution, and the proportion of time indicating uncertainty during problem-solving.

## 2.4 Materials

**Problem-solving task.** Problems were compiled from insight literature and online sources (Appendix B). The final set consisted of thirty classic insight and non-insight problems as classified by previous research, and riddles. An example of a predefined insight problem is: *“How much earth is there in a hole 3 ft by 3 ft by 3 ft?”* An example of a non-insight problem is: *“There are 235 books in a library. On Monday, 123 books are taken out. On Tuesday, 56 books are brought back. How many books are there now?”* For the present study, no prior classification was applied.

Problems were presented via computer in black Times New Roman font (size 26) within a white box, over a grey background using the program Psychological Experiment Based Language (PEBL; Mueller & Piper, 2014).

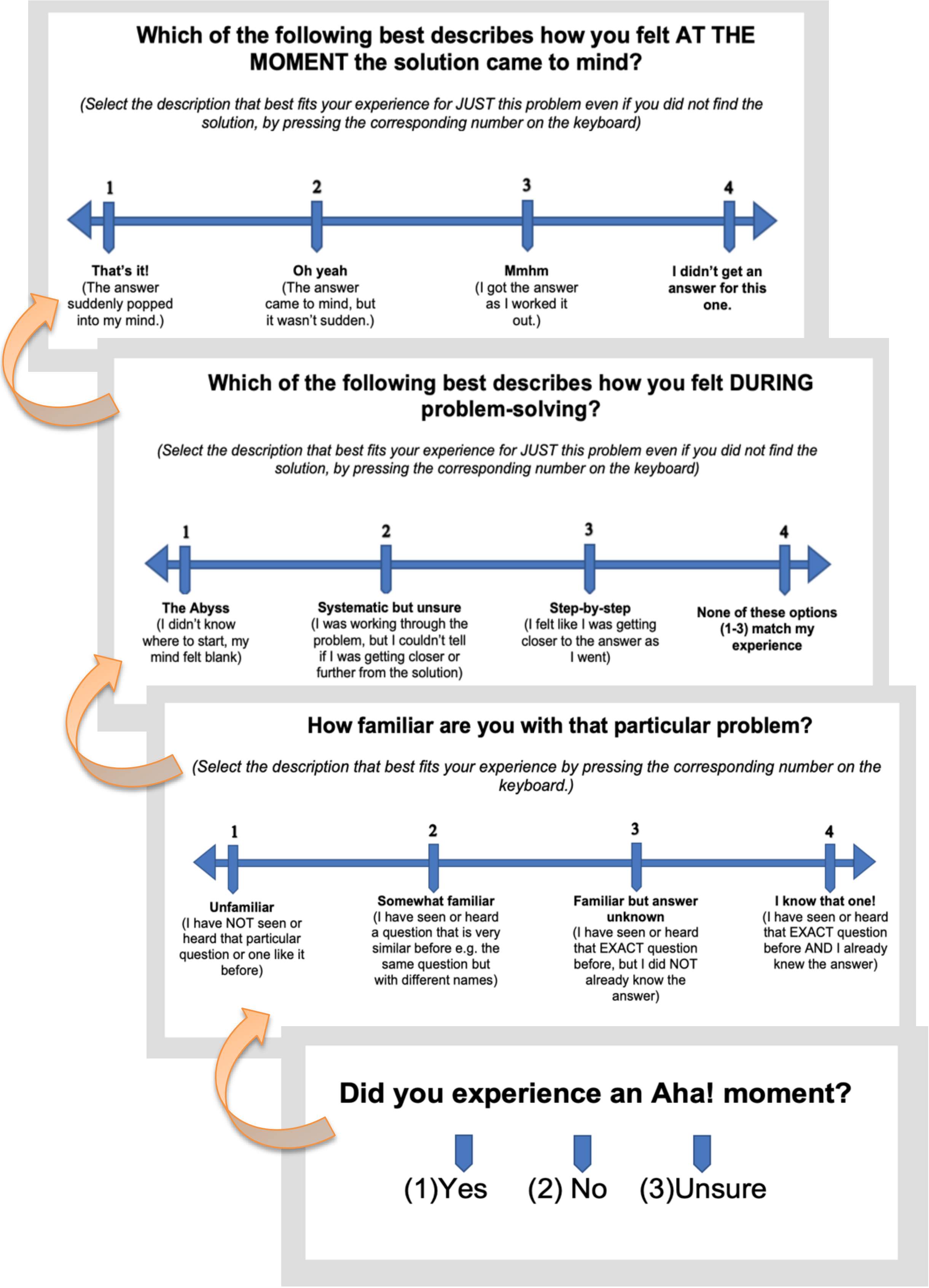
**Problem-Solving Experiences.** Problem-solving experiences were measured using retrospective and concurrent rating scales (Figure 6). Retrospective measures integrated into PEBL were Aha, Familiarity, Solution Suddenness and Systematicity. Uncertainty was measured concurrently via a new concurrent response scale (see Apparatus).

Qualitative descriptors were added for retrospective scales to capture experiences that distinguish between analytical and special processes. For example, “That’s it” would be indicative of sudden solution finding characteristic of restructuring described by the special process view (Salvi et al., 2016), whereas “Mmhm” would reflect analytical processes. For Systematicity, “The Abyss” would be indicative of unsystematic problem-solving that included periods of uncertainty or being subjectively “blank” as described by the special-process view (Smith, 2012), whereas “Step-by-step” would reflect analytical processes.

**Fluid Intelligence.** A 12-item subset of the 36-item Raven’s Advanced Progressive Matrices, Set II (RAPM, Raven et al., 1988) was used to measure fluid intelligence due to time constraints in the present study. RAPM is a widely used non-verbal test that assesses abstract reasoning ability (Raven et al., 1997). RAPM is a valid measure of fluid intelligence, is independent from cultural factors and has a high g-loading (Deary, 2004). RAPM comprises of 3 X 3 visual pattern matrices with a piece missing. Participants select the missing piece from eight options to complete the matrix.

Selection of the subset of items for the short scale was guided by Rasch item difficulty estimates (Boone, 2016) and item-total correlations (Kline, 1986) taken from data previously obtained from a sample of Murdoch University undergraduate students who completed the full 36-item Set II RAPM for another study (Appendix C). Rasch analysis is a statistical procedure that allows item difficulty to be isolated meaning item difficulty scores are sample-independent (Andrich, 1988).

Items were selected to cover a range of difficulties of approximately equal intervals to retain as much fluid ability as possible. This method is appropriate for measures that are unidimensional, such as RAPM (Waschl et al., 2016). Items with a correlation of 0.3 or more were retained (Nunally & Bernstein, 1994). Finally, scores for the subset of items were extracted from the full-scale data, and Pearson's product-moment correlation was calculated. These indicated a strong positive association between full-scale and subset scores ($r = 0.89$, $N = 107$, $p = .010$) supporting the item selection. (Appendix C). The final items included in the subset were items 4, 5, 12, 13, 14, 15, 19, 26, 31, 34, 35, and 36.



**Which of the following best describes how you felt AT THE MOMENT the solution came to mind?**

*(Select the description that best fits your experience for JUST this problem even if you did not find the solution, by pressing the corresponding number on the keyboard)*

1 **That's it!** (The answer suddenly popped into my mind.)

2 **Oh yeah** (The answer came to mind, but it wasn't sudden.)

3 **Mmhm** (I got the answer as I worked it out.)

4 **I didn't get an answer for this one.**

**Which of the following best describes how you felt DURING problem-solving?**

*(Select the description that best fits your experience for JUST this problem even if you did not find the solution, by pressing the corresponding number on the keyboard)*

1 **The Abyss** (I didn't know where to start, my mind felt blank)

2 **Systematic but unsure** (I was working through the problem, but I couldn't tell if I was getting closer or further from the solution)

3 **Step-by-step** (I felt like I was getting closer to the answer as I went)

4 **None of these options (1-3) match my experience**

**How familiar are you with that particular problem?**

*(Select the description that best fits your experience by pressing the corresponding number on the keyboard.)*

1 **Unfamiliar** (I have NOT seen or heard that particular question or one like it before)

2 **Somewhat familiar** (I have seen or heard a question that is very similar before e.g. the same question but with different names)

3 **Familiar but answer unknown** (I have seen or heard that EXACT question before, but I did NOT already know the answer)

4 **I know that one!** (I have seen or heard that EXACT question before AND I already knew the answer)

**Did you experience an Aha! moment?**

(1)Yes (2) No (3)Unsure

*Figure 6*. Retrospective rating scales in order of presentation.

**Metacognitive Control.** A computerised attention-switching paradigm was used to measure metacognitive control. Attention-switching requires the ability to attend to two cognitive tasks concurrently (Verhaeghen, Steitz, Sliwinski, & Cerella, 2003). Performance on this task is associated with control processes (Altmann & Gray, 2008) and is an indicator of metacognitive control. The task used in the present study was based on a version developed by Verhaeghen et al. (2003).

Participants were presented with a single-digit number (1-4,6-9). The number was red or yellow and was presented centre screen on a black background. Before each block began, a rule was presented (Table 1). The first two blocks of trials involved applying a single rule to categorise the number, and the third block required participants to switch between the previously learned rules, contingent on the colour of the number for each trial.

Table 1

*Switching task rule and stimulus details*

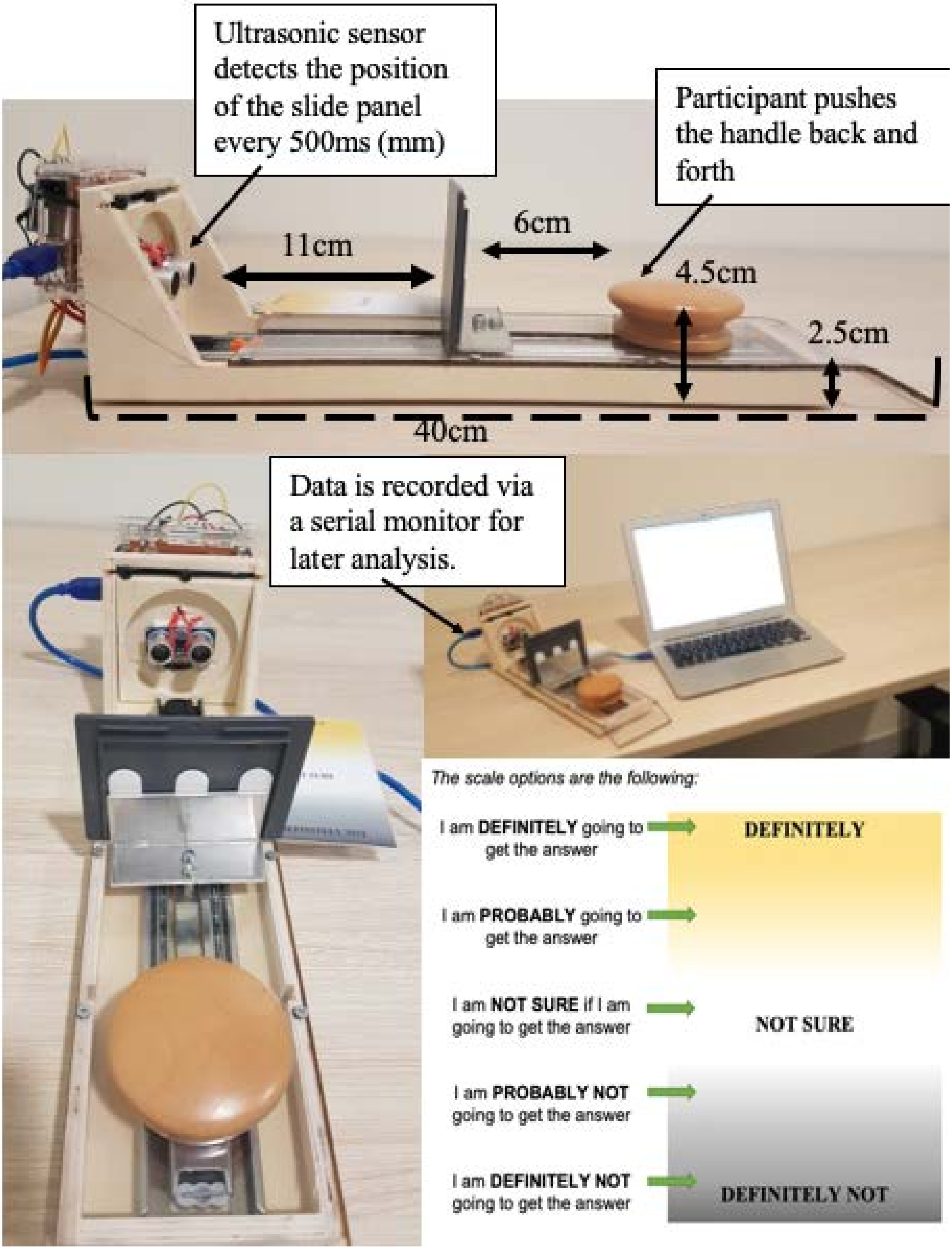
| Block | Task type | Task rule | Stimulus details | Presentation sequence |
|---|---|---|---|---|
| 1 | Single task | If the number is SMALL (less than 5) press the “Z” key, if the number is LARGE (more than 5), press the “/” key. | Yellow only single-digit number 1-4,6-9 | Random |
| 2 | Single task | If the number is odd press the “Z” key, if the number is even, press the “/” key. | Red only single-digit number 1-4,6-9 | Random |
| 3 | Dual task | If the number is red, press the “Z” key for odd, and the “/” key for even. If the number is yellow, press the “Z” key for SMALL (less than 5), and the “/” key for LARGE (more than 5). | Yellow and red single-digit number 1-4,6-9 | 4 red, 4 yellow repeating |

**Vocabulary.** The Mill Hill Vocabulary Scales (MHVS) Set B was used to measure vocabulary (Raven et al., 1997). This 34-item multiple-choice test is designed to be used alongside RAPM (Raven et al., 1997).

## 2.5 Apparatus

All computer tasks were presented on a 13.3-inch (1440 x 900mm) screen on an Apple MacBook Air.

**The Concurrent Response Scale (CRS).** The CRS includes a mechanical sliding scale, which was used by participants to continuously rate their level of uncertainty while thinking about each problem (Figure 7). Distances were recorded via a terminal emulator (Serial, 2014; accuracy of ±2mm) to correspond to a position on the uncertainty scale.


Ultrasonic sensor detects the position of the slide panel every 500ms (mm)

Participant pushes the handle back and forth

Data is recorded via a serial monitor for later analysis.

The scale options are the following:

I am **DEFINITELY** going to get the answer → **DEFINITELY**

I am **PROBABLY** going to get the answer →

I am **NOT SURE** if I am going to get the answer → **NOT SURE**

I am **PROBABLY NOT** going to get the answer →

I am **DEFINITELY NOT** going to get the answer → **DEFINITELY NOT**

*Figure 7.* Concurrent response scale labelled diagram (top), participant view (left), testing setup (centre right) and uncertainty rating scale options (bottom right).

### 2.6 Procedure

An information sheet was emailed to participants before participation and presented upon arrival. Before obtaining informed consent, questions raised by participants were answered and the purpose of the study and procedures were described. See Appendix D for forms. Participants were tested individually in a quiet, climate-controlled room. Tasks were completed in the following order: RAPM, Switching task, MHVS, Problem-solving task (including integrated demographic and CRS training). The whole session took 90 minutes.

**Psychometric tests.** RAPM and Set B of MHVS were administered as hard-copy tests according to standardised instructions (Raven et al., 1997). Participants were given 12 minutes to complete the RAPM subset by writing their answer on a response form (Appendix E). Item 1 from Set I was used as a demonstration item for this task. Scores for this test are a total of correct responses ranging from 0-12 (Raven, 2000).

**Computerised tasks.** ***Switching task.*** Instructions were presented on screen before each block. Participants were instructed to respond quickly and accurately. Each block included 8 practice trials and 64 experimental trials. If multiple errors were made in the dual task, an additional practice was offered to ensure they were able to complete the task. Tasks 1 and 2 were counterbalanced across participants.

Switch-cost was calculated as the difference in median reaction time (MRT) for correct single task trials switch and MRT for correct dual task trials (Verhaeghen et al., 2003) Medians were used in the present study instead of means to minimise the impact of outlying responses unrelated to the task. Error rates were also recorded. Higher switch-costs scores were indicative of lower metacognitive control.

***Task instructions.*** Before the problem-solving task, participants completed demographic questions about their age category and gender (Appendix F).

The following vignette about Aha experiences was provided onscreen:

*“An insight moment is an experience like an “Aha!” when an idea or answer to a problem comes to mind suddenly. When you experience an insight moment, you may not be able to describe how you came up with the answer, but you will feel confident that it is correct*”

This description is similar to that used in other studies (Bowden & Jung-Beeman, 2003). The use of the CRS was explained, and the meaning of the scale was provided (see Appendix G for instructions).

***CRS training exercise.*** Before the problem-solving task, participants completed a brief training exercise to familiarise them with using the CRS. Participants were asked to slide the CRS panel to the location corresponding statements that appeared on the screen one at a time (for example, “I am definitely not going to get the answer”), and then press “return/enter” to indicate when the panel was in the correct position. Intraclass correlations (ICC) indicated that CRS positions were located with excellent reliability across participants. The average measure ICC was .99 with a 95% confidence interval from .99 to 1.00 $F(29,2059) = 1116.34$, $p < .001$ (Appendix H).

***Problem-solving task.*** The task sequence (see Appendix I) was explained with written instructions, verbal elaboration and role-play of the task sequence to ensure full comprehension of task requirements. Problems were presented in the same order each time. No feedback was given, and participants were not given paper and pencil for working. While working on each problem, participants were asked to continuously rate their un/certainty with regard to solving the problem by sliding the handle of the CRS up (more certain) and down (less certain) the scale (Figure 7).

Participants were instructed to think about the problem until the solution came to mind, and to press "return/enter" as soon as they had the answer. They then typed the answer into a box that appeared on the screen (Appendix I). Problems timed out after 4 minutes at which time they were instructed to type “N/A” and not guess the answer. Following each problem, participants completed the retrospective rating scales in the order shown in Figure 6.

Problems were presented in blocks (blocks 1-3, 8 problems, block 4, 6 problems). Participants were instructed to continue working on the problems for 30 minutes after which the experiment timed-out (after the current problem was completed). Due to the difficulty of the task and time constraints 30 minutes was judged to be an appropriate length of time for this task.

***Unertainty-at-Solution.***The furthest distance recorded in the 3 seconds before the participant indicated they had a solution was used as an indicator of “Uncertainty at Solution”. Larger distances were indicative of less certainty.

***Uncertainty-During.*** The proportion of time spent feeling uncertain was calculated from the data points across the total problem-solving time (number of data points recorded >100 mm/total number of data points). More time spent in this zone (“Definitely not”) were indicative of discontinuous processes.

**Debrief.** After completing all tasks, participants were debriefed, and any questions were answered.

**Statistical analysis.** SPSS (IBM Corp, 2017) was used for descriptive statistics and assumption testing. Analyses to test all hypotheses were conducted using Hierarchical Linear and Nonlinear Modelling (HLM) programme, version 7.03 (Raudenbush, Bryk, & Congdon, 2013). Restricted penalised quasi-likelihood estimation was used when binary outcomes were assessed (McNeish & Stapleton, 2016), and restricted maximum likelihood was used for all other analyses.

## Results

### 3.1 Descriptive Statistics

**Level 1 Item-Response Variables.** As shown in Table 2, the mean Solution-Suddenness rating was approximately at the midpoint of the response scale ("Oh yeah/It wasn't sudden"). The mean Systematicity rating was approximately at the "Systematic but unsure of progress" descriptor. The mean Uncertainty-at-Solution sat in the "Not sure" CRS zone, and participants on average spent 16% of the time reporting Uncertainty.

Table 2

*Descriptive Statistics for Concurrent and Retrospective Ratings (Level 1)*

| Level 1 Predictor | Min | Max | *M* | *SD* |
|---|---|---|---|---|
| Solution-Suddenness | 1 | 3 | 2.04 | 0.87 |
| Systematicity | 1 | 4 | 2.14 | 0.96 |
| Uncertainty-at-Solution (mm) | 33.00 | 144.00 | 75.67 | 24.15 |
| Uncertainty-During (Proportion of time in "Definitely Not" CRS zone) | .00 | .95 | .16 | .27 |

Overall, participants answered a total of 1454 items, with an average of 19.75 per participant. Item-specific statistics are shown in Table 3. Aha was reported for all items (21.7 – 92.0%). Item "Rearrange" had the highest rate of reported Aha and the highest rate of time-outs. The mean rate of accuracy was 67% and mean rate of Aha was 57%.

Table 3

*Descriptive Statistics for Item Responses*

| Item<br>(order of presentation) | % of Time-out items not rated "Familiar" (*N* = 1364) | Descriptive statistics for valid responses only (*N* = 1061) | | |
|---|---|---|---|---|
| | | Aha (%) | Accuracy (%) | Mean time to solve (seconds) |
| Rearrange (13) | 61.8 | 92.0 | 84.0 | 96.0 |
| Professor Bumble (29) | 0.0 | 83.3 | 100.0 | 25.1 |
| Not alive (30) | 0.0 | 83.3 | 100.0 | 50.2 |
| Coffee (17) | 0.0 | 76.1 | 100.0 | 13.9 |
| Lake (19) | 3.0 | 74.2 | 74.2 | 18.9 |
| Next Letter (26) | 14.3 | 72.7 | 45.5 | 34.5 |
| Ancient Invention(5) | 1.4 | 70.6 | 92.6 | 24.8 |
| Once in a minute (22) | 17.6 | 69.2 | 92.3 | 39.6 |
| David (9) | 0.0 | 68.9 | 28.9 | 8.4 |
| Checkers (8) | 12.2 | 68.8 | 75.0 | 61.3 |
| Throw (21) | 4.2 | 68.2 | 68.2 | 45.2 |
| Hockey (27) | 0.0 | 66.7 | 100.0 | 12.9 |
| Days (11) | 10.0 | 64.5 | 59.7 | 55.9 |
| Mr Hardy (24) | 6.7 | 61.5 | 76.9 | 41.0 |
| Captain Scott (3) | 7.2 | 58.7 | 74.6 | 42.4 |
| All Day Long (14) | 1.6 | 58.3 | 73.3 | 27.4 |
| Coins (1) | 6.9 | 50.0 | 42.4 | 32.2 |
| Paul (25) | 7.1 | 50.0 | 66.7 | 29.6 |
| Ocean Liner (12) | 9.9 | 49.2 | 22.2 | 65.7 |
| Brothers (6) | 28.4 | 48.1 | 25.0 | 97.8 |
| Four Legs (18) | 26.7 | 46.9 | 37.5 | 54.1 |
| Marbles (23) | 0.0 | 46.7 | 93.3 | 12.6 |
| Chalk (7) | 0.0 | 41.8 | 89.6 | 26.0 |
| Mustard Family (28) | 0.0 | 40.0 | 80.0 | 14.3 |
| Pizza (4) | 0.0 | 36.2 | 82.6 | 36.6 |
| Library (15) | 3.2 | 35.6 | 59.3 | 48.5 |
| William (20) | 0.0 | 33.3 | 56.7 | 53.8 |
| Hole (2) | 5.8 | 32.8 | 43.8 | 32.9 |
| Socks (16) | 8.5 | 28.3 | 24.5 | 81.0 |
| Pyramid (10) | 36.5 | 21.7 | 43.5 | 94.6 |

**Level 2 Person Characteristics.** As shown in Table 4, all participants reported experiencing at least one Aha during problem-solving. All participants answer more than ten problems.

Table 4

*Descriptive Statistics for Person Characteristics (Level 2)*

| | *N* | Min | Max | *M* | *SD* |
|---|---|---|---|---|---|
| Fluid Intelligence (RAPM score) | 72 | 0 | 10 | 5.31 | 2.48 |
| Metacognitive Control (Switch-Cost (ms)) | 52 | 36.00 | 1154.00 | 406.96 | 227.59 |
| Vocabulary (MHVS score) | 72 | 19 | 40 | 28.11 | 3.89 |
| | | | Problem-solving task | | |
| Number of items attempted per person | 72 | 11 | 30 | 19.75 | 5.19 |
| Number of Correct Responses | 72 | 2 | 24 | 9.84 | 4.52 |
| Number of Aha Responses | 72 | 1 | 23 | 8.39 | 4.66 |

The Level 2 predictors, RAPM and Switch-Cost, were not significantly correlated (Table 5). The number of items attempted (those for which a response was given) correlated significantly with the percentage of reported Aha experiences.

Table 5

*Correlations for Person Characteristics (Level 2)*

| Variable | 1. | 2. | 3. | 4. | 5. |
|---|---|---|---|---|---|
| 1. RAPM score | - | | | | |
| 2. Switch-Cost (*N*=73) | .09 | | | | |
| 3. MHVS score | .32** | .10 | | | |
| 4. Items attempted during problem-solving task | .21 | -.13 | .38** | | |
| 5. % Accuracy | .28* | -.08 | .09 | .04 | |
| 6. % Aha Reported | .04 | -.16 | .24* | .65** | .16 |

*Note. N=74, *Significant at the .05 level (2-tailed), **Significant at the .01 level (2-tailed)*

### 3.2 Hierarchical Linear Modelling

**Assumptions.** HGLM assumes the outcome variable is binary, continuous predictor variables are linearly related to the log odds of the outcome variables, observations are independent at Level 1, and predictor variables at Level 2 are not highly correlated. The logit linearity of continuous Level 1 predictors was assessed via the Wald test statistic. All were found to be non-significant, thus this assumption was satisfied. Correlations of Level 2 variables were analysed and satisfied this assumption. (Refer to Appendix K).

For continuous outcome variables, HLM was used. HLM requires the the assumption of homogeneity of variance and normally distributed residuals to be satisfied. This assumption was violated in anlyses when the outcome variable was Uncertainty-at-Solution. However, as suggested by Raudenbush and Bryk (2002), when level 2 coefficients are the focus of the primary hypotheses, violations of this assumption are not serious. To account for normality, robust standard error estimates were interpreted. It was observed that the results for robust and non-robust standard error were similar, indicating this assumption was not violated.

**Data preparation**. Before conducting analyses, 393 invalid item responses were removed from the data leaving 1061 items. Invalid items were those rated as Familiar-Known (90), items with no response within 4 minutes (157), and those that participants were unsure if they had experienced an Aha (146). Familiar items were removed as participants already knew the solution, thus problem-solving experiences for these are not relevant.

To enable modelling of item-level effects on the rate of Aha, all Level 1 item response ratings were person-centred by subtracting ratings for each item from the mean across all items. To allow tests of the person-level effect on the rate of Aha, Level 2 person variables were grand-mean centred by subtracting each person score (RAPM or Switch-cost) from the mean of all people. The number of participants in each analysis varies due to missing data or an inadequate number of items per cell.

**Modelling strategy.** To model the relationships between Level 1 and 2 predictors and the rate of reporting an Aha experience, the following steps were taken:

1. ***Unconditional Models.*** A model was estimated with no predictors to gauge the magnitude of variability between individuals in the rate of reporting an Aha, Uncertainty-at-Solution, and Uncertainty-During. Significant variance in the intercepts suggests there is sufficient variability to warrant HGLM (Raudenbush & Bryk, 2002).
2. ***Level 1: Conditional Models of Concurrent and Retrospective Ratings.*** To model the relationship between retrospective and concurrent item response ratings at Level 1, Uncertainty-at-Solution was modelled as a function of retrospective Solution-Suddenness and Uncertainty-During problem-solving was modelled as a function of retrospective Systematicity ratings. Significant positive slopes indicate congruence between ratings.

3. ***Level 1: Conditional Models predicting Aha.*** To test the effect of each Level 1 predictor on the rate of reported Aha experiences, each Level 1 predictor was estimated separately. Significant positive slopes are indicative of support for hypothesis 1.
4. ***Level 2: Variance components analyses.*** The variance components from the hierarchical models from Step 3 were analysed. Substantial unexplained variance, denoted by significant variance components warrants adding Level 2 predictors.
5. ***Level 2: Conditional Models.*** To test the effect of Fluid Intelligence on the rate of Aha reported, RAPM scores were added to Step 3 models estimating the effect of Solution-Suddenness and Uncertainty-at-Solution, as a Level 2 predictor of Level 1 Aha intercept and slope. The same steps were taken to test the effect of Metacognitive Control with the inclusion of Switch-Cost. Accuracy and Vocabulary were included as control variables.

### 3.3 Level 1 Predictors of Aha (Steps 1-3)

**Step 1: Unconditional Models.** The unconditional model for Aha had an intercept of 0.42. The Aha outcome variable represents the expected log-odds of an Aha being reported when all variables are allowed to vary randomly (Moineddin, Matheson, & Glazier, 2007). This intercept value can be transformed into a probability of .40, 95% CI = (.20,.64) of reporting an Aha for the average person in the sample. The variance in intercept between persons was significant, $\chi^2(72) = 221.18$, $p < .001$.

The unconditional model for Uncertainty-at-Solution was significant, $\chi^2(71) = 221.18$, $p < .001$, with 81% of the variance related to items at Level 1 and 19% related to individual differences (ICC = .19). The unconditional model for Uncertainty-During was significant $\chi^2(71) = 320.79$, $p < .001$, with 87% of the variance related to items at Level 1 (ICC = .13.) (Appendix L details all analyses.)

**Step 2. Level 1 Conditional Models of Concurrent and Retrospective Ratings.**

***Uncertainty-at-Solution and Solution-Suddenness.*** As shown in Figure 8, higher retrospective ratings of suddenness of solution-finding were associated with uncertainty reported at the point of solution. The positive, significant slope indicates that the retrospective and concurrent reports were significantly associated (refer to Table 6).

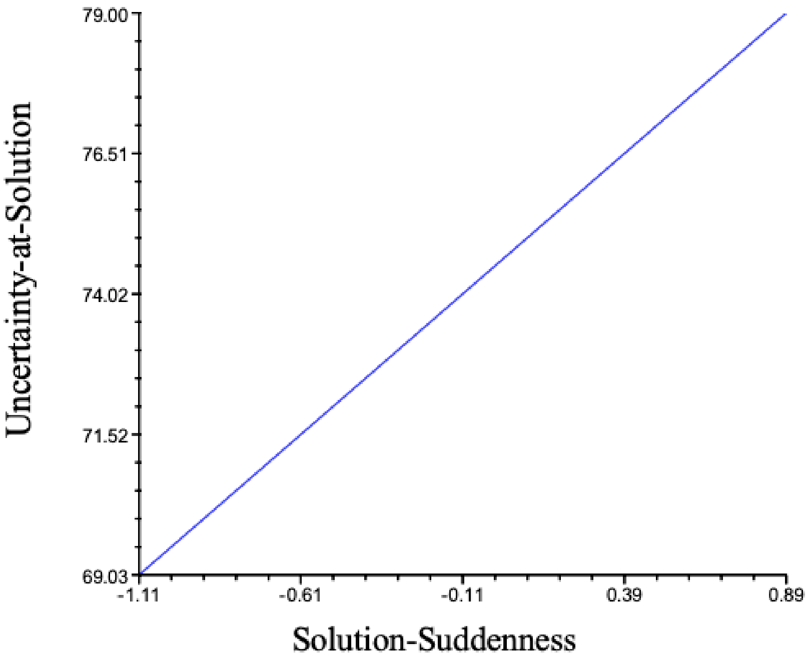


*Figure 8.* Solution Suddenness as a predictor of Uncertainty-at-Solution.

***Uncertainty-During and Systematicity.*** Similarly, retrospective Systematicity ratings were significantly associated with greater Uncertainty-During. As shown in Figure 9, when participants rated their problem-solving processes as less systematic (represented by higher values) they also spent a higher proportion of total problem-solving time indicating they were “Definitely not going to solve the problem”. This indicates that the concurrent and retrospective ratings were congruent.

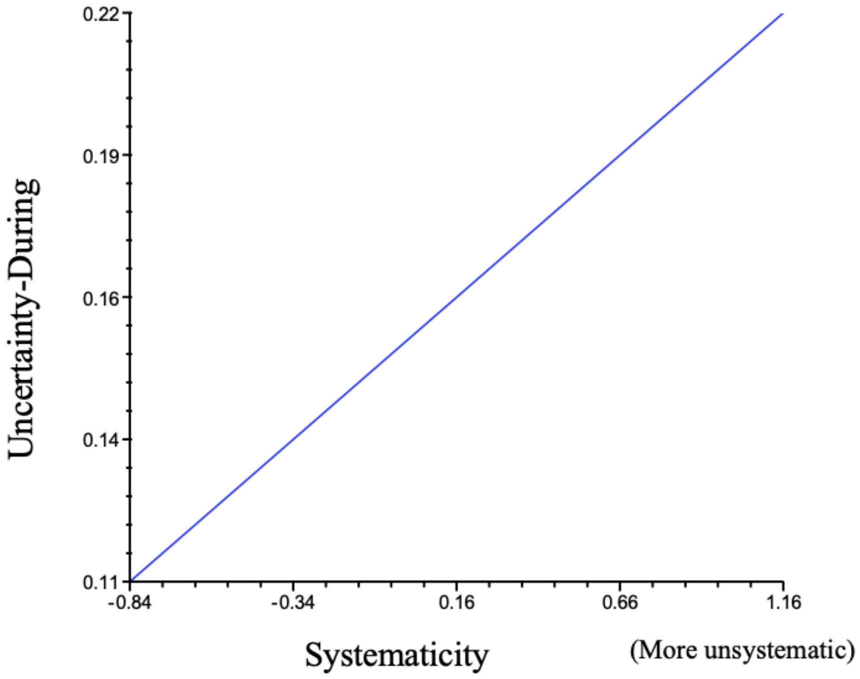


*Figure 9.* Systematicity as a predictor of Uncertainty-During problem-solving.

Table 6

*Final Estimation of Fixed Effects for Models of Level 1 Concurrent Item Ratings Predictors of Retrospective Item Ratings (with robust standard errors)*

| | Coefficient | *SE* | *t*-ratio | Approx. *d.f.* | *p*-value |
|---|---|---|---|---|---|
| Uncertainty-at-Solution Intercept | 74.55 | 1.41 | 52.98 | 71 | <.001 |
| Solution-Suddenness Slope | 4.99 | 1.03 | 4.84 | 71 | <.001 |
| Uncertainty-During Intercept | 0.15 | 0.01 | 11.15 | 71 | <.001 |
| Systematicity Slope | 0.05 | 0.12 | 4.26 | 71 | <.001 |

**Step 3. Level 1 Conditional Models**. ***Accuracy.*** As shown in Figure 10, 74% of item responses with an Aha were correct, whereas only 47% of items solved without an Aha were

correct. Aha experiences were a significant, positive predictor of correct responses. These analyses indicate that problems solved with a reported Aha were 3.34 times more likely to be correct compared to those solved without an Aha experience (Refer to Table 7, Panel A). This finding supports hypothesis 1.

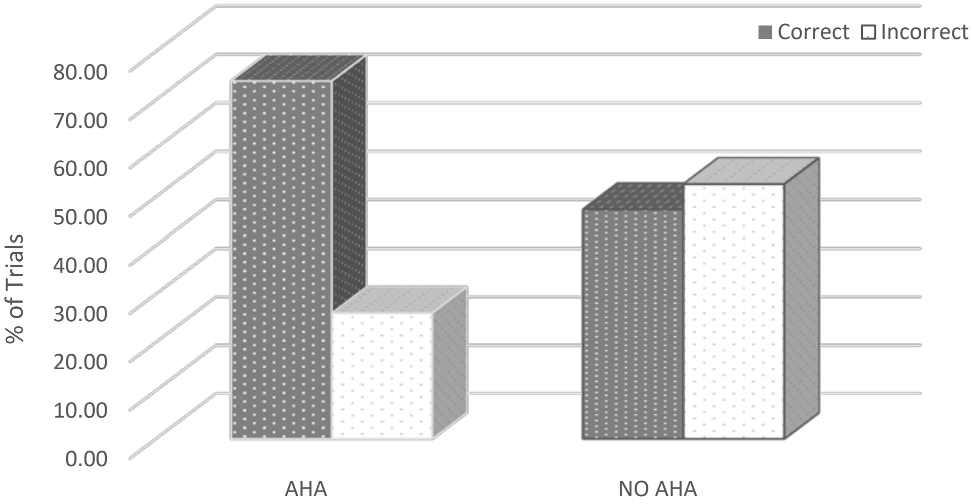


*Figure 10*. Accuracy (%) for trials reported as Aha or No Aha

***Retrospective ratings of Solution-Suddenness.*** As illustrated in Figure 11, 67% of Aha trials were retrospectively reported as “Popped into mind” compared to 12.5% of trials without an Aha. The opposite pattern was reported for “I worked it out”. Consistent with hypothesis 1, statistical analyses indicated that Solution-Suddenness was a significant, positive predictor of Aha. Specifically, a one-unit increase in Solution-Suddenness increases the odds of reporting an Aha by 366%. (Refer to Table 7, Panel B.) A one-unit increase is equivalent to one standard deviation (*SD* = 0.87).

Table 7

*Final Estimation of Fixed Effects for Models of Level 1 Predictors of Aha (with robust standard errors) and Odds Ratio*

| Panel | | Coefficient | *SE* | *t*-ratio | Approx. *d.f.* | *p*-value | Odds Ratio [95% CI] |
|---|---|---|---|---|---|---|---|
| A | Accuracy Intercept | 0.60 | 0.09 | 7.02 | 72 | <.001 | 1.83<br>[1.54, 2.17] |
| | Aha Slope | 1.21 | 0.12 | 10.48 | 72 | <.001 | 3.34<br>[2.66, 4.21] |
| B | Aha Intercept | 0.54 | 0.12 | 4.37 | 72 | <.001 | 1.72<br>[1.34, 2.21] |
| | Solution-Suddenness Slope | 1.54 | 0.11 | 13.62 | 72 | <.001 | 4.66<br>[3.72, 5.85] |
| C | Aha Intercept | 0.43 | 0.11 | 3.86 | 72 | <.001 | 1.53<br>[1.23, 1.91] |
| | Systematicity Slope | 0.23 | 0.09 | 2.67 | 72 | .009 | 1.26<br>[1.06, 1.50] |
| D | Aha Intercept | 0.42 | 0.11 | 3.72 | 71 | <.001 | 1.52<br>[1.21.1.90] |
| | Uncertainty-at-Solution Slope | 0.01 | 0.00 | 3.49 | 71 | <.001 | 1.01<br>[1.01, 1.02] |
| E | Aha Intercept | 0.41 | 0.11 | 3.70 | 71 | <.001 | 1.51<br>[1.21, 1.88] |
| | Uncertainty-During Slope | 0.14 | 0.28 | 0.48 | 71 | .632 | 1.15<br>[0.65, 2.02] |

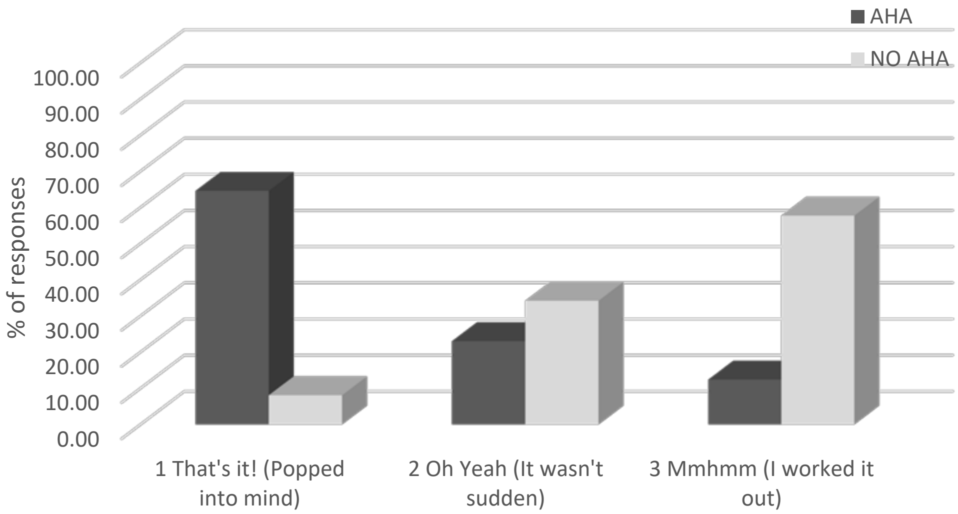


*Figure 11.* Percentage of items in each Solution Suddenness response category for Aha and No Aha trials

***Retrospective ratings of Systematicity.*** Across all items, systematic problem-solving (Step-by-Step) was reported most frequently, regardless of whether an Aha was experienced (Figure 12). However, a higher percentage of no Aha trials were reported as Step-by-Step. The reverse was true for all other options which were indicative of decreasing systematicity or unexplainable processes.

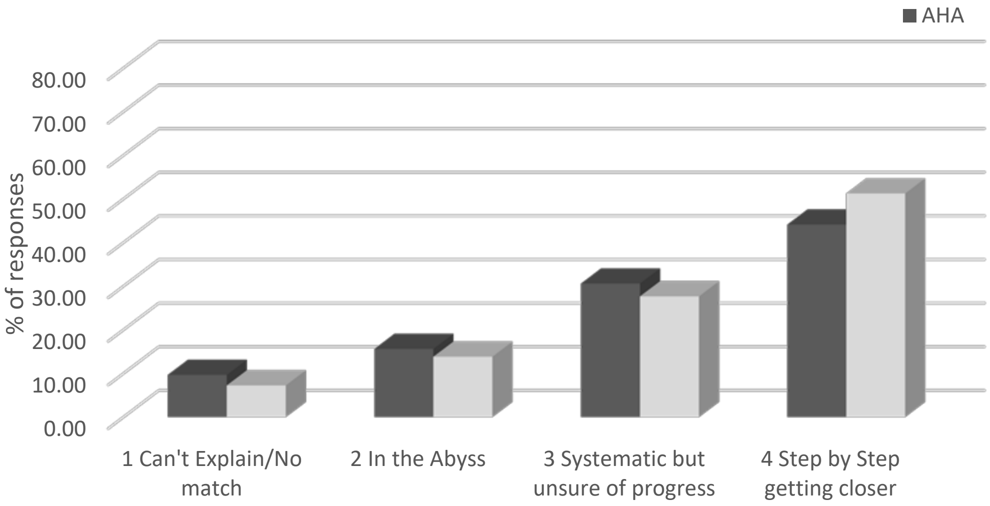


*Figure 12.* Percentage of items in each Systematicity response category for Aha and No Aha trials

Low Systematicity was a significant predictor of Aha. Specifically, when all other variables are held constant, a one-unit decrease in Systematicity increases the odds of reporting an Aha by 26%. (Refer to Table 7, Panel C). A one-unit decrease is equivalent to one standard deviation (*SD* = 0.96). This finding is consistent with predictions under hypothesis 1.

***Concurrent ratings of Uncertainty-at-Solution.*** Participants' ratings of Uncertainty-at-Solution were different for trials with, compared to without, a reported Aha. As illustrated in Figure 13, at the point of solution, a higher percentage of participants who were indicating they would "Definitely Not" solve the problem, reported an Aha experience (17% compared to 12%). Conversely, a higher percentage of participants indicating they were certain about solving the problem at the point when they found the solution did not subsequently report an Aha (45.3% compared to 49.8%).

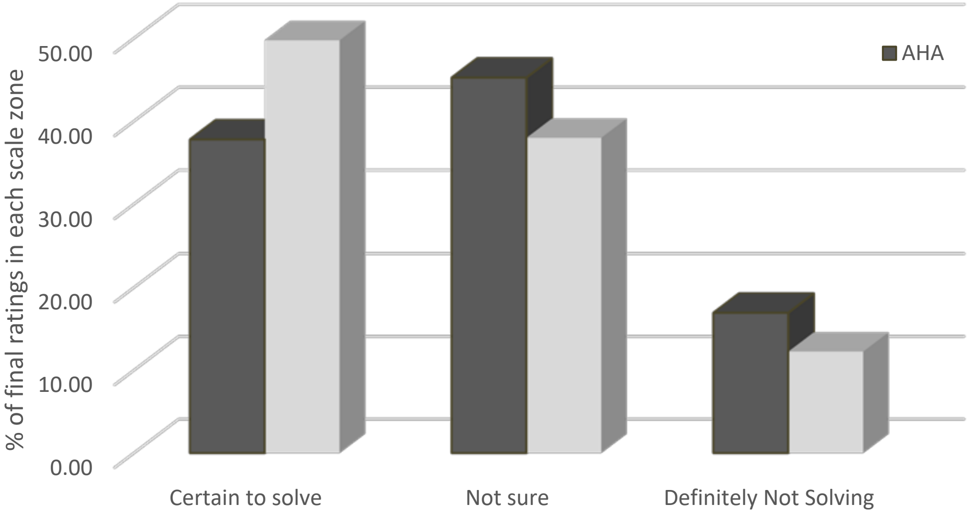


*Figure 13.* Percentage of items in each Uncertainty-at-Solution rating (CRS Zone) for Aha and No Aha trials when the solution came to mind.

Uncertainty-at-Solution was a significant predictor of Aha. Specifically, a one-unit increase was associated with an increased odds of Aha by 1%. (Refer to Table 7, Panel D.) A one-unit increase is equivalent to a shift of one standard deviation (*SD* = 24.15mm) along the CRS scale. This finding is consistent with predictions under hypothesis 1.

***Concurrent ratings of Uncertainty-During problem-solving.*** As illustrated in Figure 14, the overall proportion of problem-solving time participants spent indicating the were definitely going to solve the problem (Certain to solve) was higher for trials where no Aha was reported (.27) compared to trials where an Aha was reported (.21). However, the same proportion of time was spent in the “Definitely Not” CRS zone for trials where an Aha was reported and No Aha trials (.16).

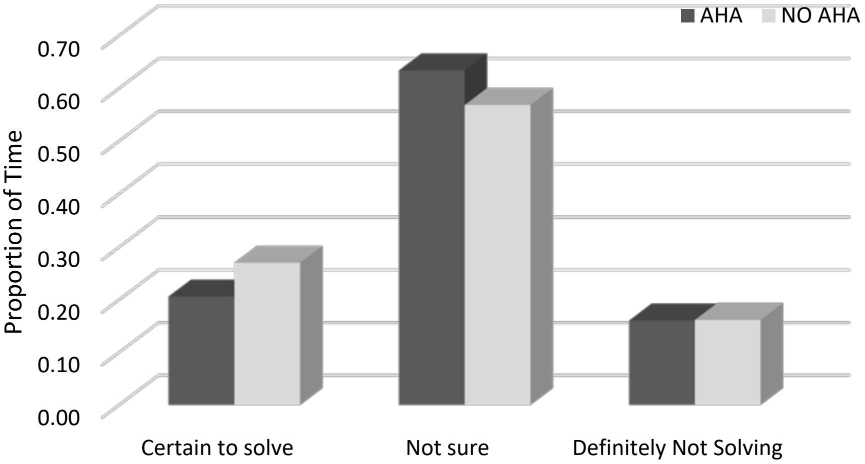


*Figure 14.* Proportion of problem-solving time spent in each CRS zone for Aha and No Aha trials. (Uncertainty-During)

Statistical analyses indicated that uncertainty during problem-solving did not significantly predict Aha. (Refer to Table 7, Panel E).

### 3.4 Level 2 Predictors of Aha

**Step 4: Level 2 Variance components analyses.** Prior to testing the effect of Level 2 predictors RAPM scores and Switch-Cost, preliminary tests were conducted to ensure there was significant variance in the Aha intercept. Variance components from the Level 1 analyses (above) were analysed for Solution-Suddenness and Unertainty-at-Solution (Table 8, Panel A

& B). The analyses were repeated for the subsample of participants with acceptable levels of performance on the Metacognitive Control task (Table 8, Panel C & D). Significant variance existed in the Aha intercepts and Level 1 predictor slopes across all analyses, allowing us to proceed to step 5.

**Step 5: Level 2 Conditional Models.** The following steps were taken to test hypothesis 2 and 3. If at any step the model failed to reveal significance, not further analyses were conducted. Before conducting analyses for hypothesis 3 with Metacognitive control as a Level 2 predictor, participants with an error rate of more than 5% on the switching task were removed (*n* = 20)**.** High error rates may indicate that partipants had not learned each single task adequately meaning performance on the dual task is not interpretable as switch-cost. Outlier analysis for the switching task was conducted using SPSS (IBM Corp, 2017) no extreme outliers were identified (Appendix M).

a. Beginning with the model from Step 3 with Solution Suddenness as Level 1 predictor, a Level 2 predictor was added (RAPM or Switch-Cost).
b. Control variables Accuracy at Level 1 and MHVS scores at Level 2 were then added. This ensured that any apparent effects of RAPM (or Switch-Cost) were not confounded with individual differences in accuracy or vocabulary. Parallel analyses were conducted with Uncertainty-at-Solution as the Level 1 predictor. Slopes were allowed to vary randomly.

Statistically significant effects on Aha intercepts would indicate that Level 2 predictors explained significant variance in the log-odds of reporting an Aha experience during problem-solving. Full details of these analyses can be found in Appendix K. The most complex models are reported below.

Table 8

*Final Estimation of Variance Components for Models of Level 1 Predictors of Aha*

| Panel | Random Effect | *SD* | Variance Component | *df* | $\chi^2$ | *p*-value |
|---|---|---|---|---|---|---|
| | | | Fluid Intelligence Sample | | | |
| A | Aha Intercept | 1.39 | 1.93 | 72 | 235.98 | <.001 |
| | Solution-Suddenness Slope | 1.08 | 1.17 | 72 | 133.11 | <.001 |
| B | Aha Intercept | 0.89 | 0.79 | 71 | 223.61 | <.001 |
| | Uncertainty-at-Solution Slope | 0.02 | 0.0003 | 71 | 95.84 | .026 |
| | | | Metacognitive Control Sample | | | |
| C | Aha Intercept | 1.41 | 2.00 | 50 | 235.98 | <.001 |
| | Solution-Suddenness Slope | 0.96 | 0.91 | 50 | 79.25 | .005 |
| D | Aha Intercept | 0.89 | 0.79 | 50 | 163.14 | <.001 |
| | Uncertainty-at-Solution Slope | 0.02 | 0.0005 | 50 | 81.40 | .004 |

*Note.* The degrees of freedom are different for A and B because CRS data was not available for one participant.

***Fluid intelligence.*** Analyses at all steps outlined above continued to reveal significant effects. The final model included Level 1 predictor Solution-Suddenness, Level 2 predictor RAPM scores, and control variables MHVS scores and Accuracy**.** As shown in Table 9 (Panel A), RAPM scores had a statistically significant negative effect on the Aha intercept in this model. Solution-Suddenness slopes were not different across fluid intelligence groups (Panel C). This effect was retained when controlling for MHVS scores and Accuracy. Accuracy also had a significant effect on the Aha intercecpt.

As demonstrated in Figure 15, higher RAPM scores were associated with a reduction in the log-odds of reporting an Aha. Specifically, a one standard deviation increase in RAPM scores were associated with a 12% reduction in the odds of reporting Aha. These results are consistent with hypothesis 2.

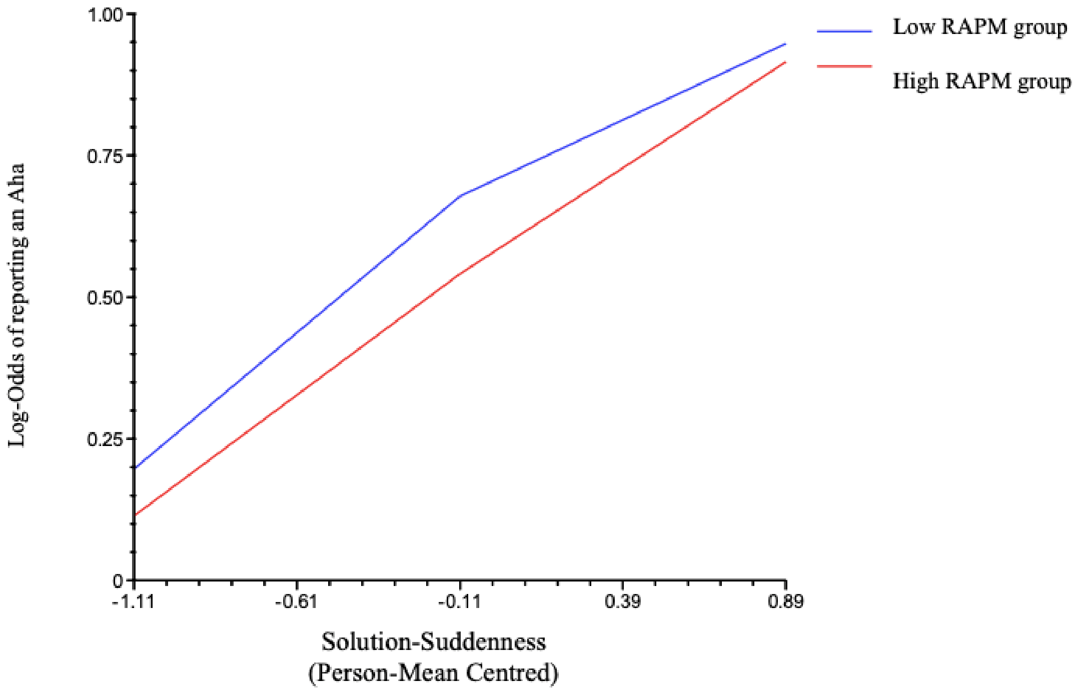


*Figure 15*. The effect of RAPM scores on the slope representing the relationship between Solution-Suddenness and the log-odds of reporting Aha

Table 9

*Final Estimation of Fixed Effects for Models of Level 1 Aha Predictor Solution-Suddenness and Level 2 Predictor RAPM Scores Controlling for Accuracy and MHVS Scores (with robust standard errors) and Odds Ratio*

| Panel | | Coefficient | *SE* | *t*-ratio | Approx. *d.f.* | *p*-value | Odd Ratio [95 % CI] |
|---|---|---|---|---|---|---|---|
| A | Level 1 (item response) | | | | | | |
| | Aha intercept | 0.52 | 0.11 | 4.71 | 70 | <.001 | 1.68 [1.35,2.10] |
| | Level 2 (person) RAPM Intercept | -0.13 | 0.04 | -3.49 | 70 | <.001 | 0.88 [0.82,0.95] |
| | MHVS Intercept | 0.003 | 0.04 | 0.09 | 70 | .931 | 1.00 [0.93, 1.08] |
| B | Level 1 (item response) | | | | | | |
| | Accuracy Slope | 0.66 | 0.09 | 7.13 | 70 | <.001 | 1.93 [1.60, 2.31] |
| | Level 2 (person) | | | | | | |
| | Accuracy x RAPM Slope | -0.13 | 0.04 | -0.36 | 70 | .722 | 0.99 [0.92, 1.06] |
| | Accuracy x MHVS Slope | -0.007 | 0.03 | -0.28 | 70 | .783 | 0.99 [0.94, 1.08] |
| C | Level 1 (item response) | | | | | | |
| | Solution-Suddenness Slope | 1.39 | 0.10 | 13.67 | 70 | <.001 | 4.02 [3.28,4.93] |
| | Level 2 (person) | | | | | | |
| | Solution-Suddenness xRAP Slope | 0.02 | 0.03 | 0.71 | 70 | .483 | 1.02 [0.96,1.09] |
| | Solution-Suddenness xMHVS Slope | -0.003 | 0.03 | -0.12 | 70 | .906 | 1.00 [0.95, 1.05] |

*Uncertainty-at-Solution.* Analyses at all steps outlined above continued to reveal significant effects. The final model included Level 1 predictor Uncertainty-at-Solution, Level 2 predictor RAPM scores, and control variables MHVS scores and Accuracy**.** RAPM scores had a statistically significant negative effect on the Aha intercept in this model (Table 10, Panel A). This effect was retained when controlling for MHVS scores and Accuracy. As demonstrated in Figure 16, a one standard deviation increase in RAPM scores were associated with an 11% reduction in the odds of reporting Aha.

Significant slopes were also observed in this model. As shown in Figure 16, individuals with lower fluid intelligence had similar odds of reporting an Aha regardless of uncertainty at the point of solution, whereas individuals with higher fluid intelligence had lower odds of reporting an Aha when reporting uncertainty at the point of solution compared to when reporting certainty.

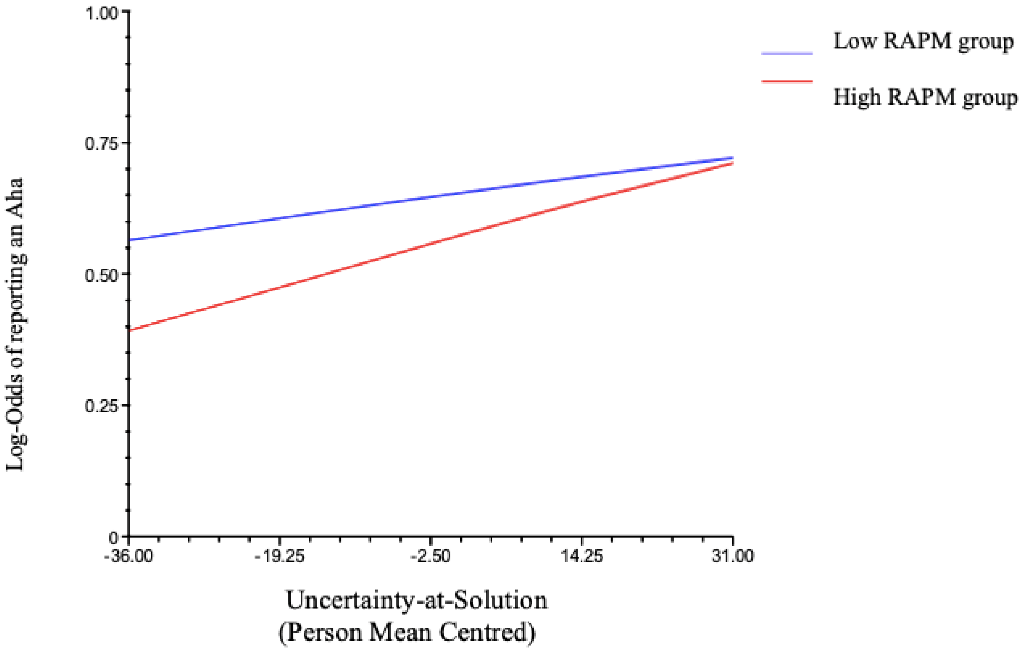


*Figure 16.* The effect of RAPM scores on the slope representing the relationship between Uncertainty-at-Solution and the log-odds of reporting Aha

Table 10

*Final Estimation of Fixed Effects for Models of Level 1 Aha Predictor Uncertainty-at-Solution and Level 2 Predictor RAPM Scores Controlling for Accuracy and MHVS Scores (with robust standard errors) and Odds Ratio*

| Panel | | Coefficient | *SE* | *t*-ratio | Approx. *d.f.* | *p*-value | Odd Ratio [95 % CI] |
|---|---|---|---|---|---|---|---|
| | Level 1 (item response) | | | | | | |
| | Aha intercept | 0.45 | 0.11 | 4.12 | 69 | <.001 | 1.56 [1.26,1.94] |
| A | Level 2 (person) RAPM Intercept | -0.11 | 0.90 | -3.06 | 69 | .003 | 0.89 [0.83,0.96] |
| | MHVS Intercept | 0.002 | 0.04 | 0.04 | 69 | .965 | 1.00 [0.93, 1.08] |
| | Level 1 (item response) | | | | | | |
| | Accuracy Slope | 1.07 | 0.11 | 9.86 | 69 | <.001 | 2.90 [2.34, 3.60] |
| B | Level 2 (person) | | | | | | |
| | Accuracy x RAPM Slope | -0.004 | 0.04 | -0.10 | 69 | .919 | 1.00 [0.92, 1.08] |
| | Accuracy x MHVS Slope | -0.02 | 0.03 | -0.62 | 69 | .536 | 0.98 [0.93, 1.04] |
| | Level 1 (item response) | | | | | | |
| | Certainty-at-Solution Slope | 0.01 | 0.00 | 3.66 | 69 | <.001 | 1.01 [1.01, 1.02] |
| C | Level 2 (person) | | | | | | |
| | Uncertainty-at-Solution x RAPM Slope | 0.003 | 0.001 | 2.15 | 69 | .035 | 1.00 [1.00,1.01] |
| | Uncertainty-at-Solution x MHVS Slope | 0.00 | 0.001 | 0.26 | 69 | .798 | 1.00 [1.00, 1.00] |

***Metacognitive Control.*** The initial models including Level 1 predictors Solution Suddenness and Uncertainty-at-Solution (Step 5a) and Level 2 predictor, Switch-Cost failed to achieve statictical significance. Thus, hypothesis 3 was not supported and no further steps were undertaken (Table 11).

Table 11

*Final Estimation of Fixed Effects for Models of Level 1 Predictors Solution-Suddenness and Uncertainty-at-Solution and Level 2 Predictor Switch-Cost (with robust standard errors) and Odds Ratio*

| Panel | | Coefficient | *SE* | *t*-ratio | Approx. *d.f.* | *p*-value | Odd Ratio [95 % CI] |
|---|---|---|---|---|---|---|---|
| A | Level 1 (item response) Aha/ intercept | 0.41 | 0.15 | 2.8 | 50 | .007 | 1.51 [1.13,2.03] |
| | Level 2 (person) Intercept | -0.00 | 0.00 | -0.19 | 50 | .852 | 1.00 [1.00,1.00] |
| | Solution-Suddenness Slope | 1.59 | 0.12 | 12.953 | 50 | <.001 | 4.93 [3.85, 6.31] |
| | Solution-Suddenness x Switch-Cost Slope | -0.00 | 0.00 | -1.22 | 50 | .227 | 1.00 [1.00,1.00] |
| B | Level 1 (item response) Aha intercept | 0.31 | 0.13 | 2.37 | 49 | .022 | 1.36 [1.05,1.77] |
| | Level 2 (person) Intercept | -0.00 | 0.00 | -0.31 | 49 | .759 | 1.00 [1.00,1.00] |
| | Uncertainty-at-Solution Slope | 0.01 | 0.00 | 3.23 | 49 | .002 | 1.02 [1.01,1.02] |
| | Uncertainty-at-Solution x Switch-Cost Slope | -0.00 | 0.00 | -0.44 | 49 | .661 | 1.00 [1.00,1.00] |

# Discussion

The central aim of the present study was to investigate the relationship between reported Aha experiences, fluid intelligence and metacognitive control. It was proposed that fluid intelligence and metacognitive control would be negatively associated with the odds of reporting an Aha experience due to differences in underlying cognitive architecture and control processes described by Network Neuroscience Theory (Barbey, 2018). Consistent with predictions, fluid intelligence negatively predicted the probability of reporting an Aha. Metacognitive control was not significantly associated with Aha experiences. Thus, hypothesis 3 was not supported.

A secondary aim was to explore the relationship between Aha experiences and problem-solving experiences. It was proposed that Aha experiences are non-random and are predicted by problem-solving experiences suggestive of a special process; such as feelings of suddenness, systematicity and uncertainty. Overall, the results of the present study supported this proposal.

To gain a complete picture of Aha experiences and their relationship with preceding problem-solving experiences, analyses were not constrained by the assumption that insight occurs only for correct solutions. Thus, all problems with a solution were included in all analyses, except problems of which participants had prior knowledge as solutions may be indicative of memory rather than problem-solving processes. Removing this assumption is a strength of this study, as many previous studies have assumed only correct solutions are indicative of insight (Chuderski & Jastrzebski, 2018; Gilhooly & Fioratou, 2009). However, as described by Danek and Wiley (2016), false insights occur, and this was also found in the present study. By including all responses, we were better able to determine what differentiates Aha experiences from problems solved without an Aha.

## 4.1 The Aha Experience

The present results indicate that Aha experiences are systematically associated with self-reported problem-solving experiences. Specifically, the odds of reporting an Aha were higher when participants reported uncertainty at the time of solution (concurrent rating) when retrospective ratings of suddenness were high, and problem-solving processes were reported as unsystematic. Importantly, retrospective ratings of suddenness and systematicity were congruent with concurrent ratings of uncertainty. However, the proportion of time spent feeling uncertainty during problem-solving was not significantly associated with Aha experiences.  The strongest effect was found for retrospective ratings of solution suddenness, which consistent with other research (Webb et al., 2017), and is unsurprising as participants reporting an Aha are likely to answer retrospective ratings that align with this.  Overall, these findings are consistent with previous research (Danek et al., 2014; Webb et al., 2017), but also constitute new knowledge.

A small but significant association was found between concurrent reports of uncertainty at the point of solution and retrospective reports of higher solution suddenness. Uncertainty at the moment of solution was also associated with higher odds of reporting an Aha. This indicates that Aha experiences and solution suddenness are detectable in the moments before the solution emerged, suggesting they are dissociable from subsequent evaluations of accuracy, confidence or relief. This result is consistent with recent work by Laukkonen and Tangen (2018) who also found distinct patterns of low progress ratings preceding Aha experiences using a dynamometer. These findings offer support for the argument that Aha experiences are indicative of a special process (Salvi et al., 2016) rather than post-solution affect (Topolinski & Reber, 2010).

The present study also found that Aha experiences are positively associated with retrospective reports of unsystematic problem-solving processes. Process systematicity was also associated with the proportion of time participants spent feeling unable to solve the problem. Specifically, retrospective reports of unsystematic processes predicted uncertainty during problem-solving. The congruence between these ratings suggests that feeling uncertain for a higher proportion of time is experienced by participants as unsystematic.

Concurrent reports of uncertainty during problem-solving were not a significant predictor of Aha experiences, indicating that low levels of certainty averaged throughout problem-solving are not necessarily associated with Aha experiences. However, characteristics of the rating scale may have unintentionally influenced how some participants reported their certainty levels during problem-solving. For example, one participant suggested they would never use the "Definitely Not" end of the scale because they "don't ever give up".

Perceived systematicity of problem-solving processes has not been measured in previous insight research. In doing so, we sought to differentiate processes preceding Aha experiences from those without an Aha. Proponents of the nothing special view suggest that problem-solving relies on controlled, systematic process such as memory, attention and reasoning (Chuderski & Jastrzebski, 2018). The special process perspective asserts that processes leading to an Aha experience are beyond conscious awareness and thus cannot be controlled or explained by problem-solvers (Davidson & Sternberg, 1984; Knoblich et al., 1999; Smith, 2012). The present findings offer support for the latter.

It was proposed that when an Aha was reported, problem-solving was discontinuous (indicated by more uncertainty) and would be reported by problem-solvers as unsystematic. It was argued that this would be most likely when the direction of thinking preceding the impasse was disconnected from the solution state. That is, when unsuccessful lines of thought led to an impasse, which resulted in the immediate emergence of an alternative representation and the solution state, problem-solvers would report unsystematic processes. Some support was found for this proposition. However, inconsistency in the findings suggests refinement of the concurrent measurement may be efficacious to understand this relationship further.

### 4.2 Fluid Intelligence and Aha Experiences

A key finding of the study is a consistent, small, but significant negative effect of fluid intelligence on the odds of reporting an Aha. This negative relationship was maintained when controlling for both accuracy and verbal ability. No significant difference in the number of items answered was found depending on fluid intelligence. Thus, this result is unlikely to be due to people of higher fluid intelligence, giving a higher number of item responses than those of lower fluid intelligence.

These results offer support for the proposition outlined in section 1.3 that fluid intelligence is influential in dictating the ensuing processes, and subsequent solution experience due to more extensively connected underlying cognitive architecture (Barbey, 2018; Jung & Haier, 2007). Specifically, it was proposed that weak associative connections are indicative of fluid intelligence level (Barbey, 2018), and these connections dictate what information is accessible from the initial representation (Smith, 2012). This accessibility influences whether the solution can be found via analytical processes or whether a special process is required, thus systematically influencing the subsequent problem-solving and solution experience.

The relationship between solution suddenness and Aha experiences was not significantly different depending on fluid intelligence. This suggests that Aha experiences are characterised by the same subjective experience of suddenness, irrespective of fluid intelligence level, but occur less frequently for individuals with higher fluid intelligence. Interestingly, lower fluid intelligence was associated with reporting significantly more certainty at the moment the solution came to mind regardless of whether an Aha was reported.

The nested design of this study allowed for more fine-grained analyses of the effect of item-specific variables and person variables. This result shows that at item-level, uncertainty upon solution-finding is associated with Aha experiences, but when accounting for individual differences, the magnitude of this relationship varies. This finding may suggest differences in metacognitive monitoring (Veenman & Beishuizen, 2004) preceding an Aha experience that could be examined in future research.

A particularly noteworthy result is the paradoxical relationship between Aha, accuracy and fluid intelligence. Although accuracy was positively associated with both Aha experiences and fluid intelligence; fluid intelligence and Aha were negatively associated. This suggests that individuals with higher fluid intelligence reported a higher number of correct solutions, and correct responses were more likely to elicit an Aha. However, the odds of individuals with higher fluid intelligence reporting an Aha were lower compared to those with lower fluid intelligence. Overall, this offers further support for the proposed negative relationship between fluid intelligence and Aha.

Differences in metacognitive monitoring, which is closely related to fluid intelligence (Rozencwajg, 2003) may explain this result. For example, overconfidence in the solution due to less effective monitoring (Dunlosky & Rawson, 2012) may mean that individuals with lower fluid intelligence are more likely to report false insights. In contrast, individuals with higher fluid intelligence may be more likely to evaluate a salient but incorrect solution systematically.  It may be efficacious for future research to investigate this possibility.

### 4.3 Metacognitive Control and Aha Experiences

In contrast to predictions, metacognitive control was not significantly associated with reduced odds of reporting an Aha experience. However, this result may reflect issues with the switching task used to measure metacognitive control, which resulted in a third of participants being excluded from analyses due to a high error rate. The number of single-task trials may not have been sufficient for participants to learn the rules before undertaking the dual-task, compromising the validity of this task. Other studies have used more trials in the single task than used here (Verhaeghen et al., 2003). The data from this task did not correlate with fluid intelligence, which is inconsistent with other studies (Veenman & Beishuizen, 2004) and strengthens concerns about the validity of this task.

### 4.4 Evaluation of the Concurrent Response Scale

A considerable strength of this study is the implementation of a new device to capture participant ratings in real-time. The lack of tools to measure theoretically meaningful information problem-solving trajectories and experiences is an oft-noted limitation in insight research (Bowden et al., 2005; Laukkonen & Tangen, 2018). The CRS enabled us to capture participants’ ratings and analyse these in a way that enhanced our understanding of Aha experiences and their relationship to theoretically distinct problem-solving experiences. Furthermore, although simple data analyses were used in this preliminary study, a large number of data points recorded by the CRS may enable more sophisticated measures of problem-solving trajectories in future studies.

This device represents a substantial improvement over currently available methods, which may disrupt theoretically important processes and may not be suitable for problem-solving over more extended periods. The CRS also has potential application as a broad, continuous measure of subjective state in other research contexts. However, limitations related to the CRS must be noted. Although there were prompts throughout the problem-solving task, and a researcher monitored its use throughout, a handful of participants reported that they forgot to use the CRS for some problems. However, this may have caused an observer-expectancy effect (Yantz & McCaffrey, 2005).

## 4.5 Limitations of the study

Further limitations of the study must also be considered. The retrospective rating scales may not have captured the problem-solving experiences of all participants. Although extensive scripted instructions and explanations of the scales were provided verbatim for all participants, and several participants noted that the qualitative anchors were an accurate representation of their experiences, this may not have been so for all participants. Fatigue and task difficulty may have altered participant affect towards the end of the problem-solving task. However, the data suggests Aha reports did not vary greatly by item order thus, this is unlikely have influenced results.

Once invalid responses were removed, the number of items for some participants may have been below the minimum suggested (five items) in simulation studies, which may limit the statistical power of the study. The sample was predominantly female, and a large proportion were university students, who may be more familiar with the types of problems used, and testing scenarios compared to the general population. However, the range of accuracy and relatively low number of problems rated as familiar suggests this is unlikely to limit generalisability of these findings substantially.

## 4.6 Conclusion and Future Directions

Overall, the present study strengthens the support for measures of perceived progress and suddenness as indicators of cognitive restructuring during problem-solving described by the special process perspective (Knoblich et al., 1999; Smith, 2012). New insights into the relationship between perceived systematicity and Aha were uncovered. These findings challenge the assertion by proponents of the nothing special perspective that Aha experiences are random events unrelated to preceding mental activity (Chuderski, 2014).

Evidence of the contingent nature of special processes during problem-solving supports the proposition that underlying cognitive architecture may dictate the processes activated to solve a problem, and the subsequent solution experience. Future research may look to explore the influence of task difficulty on the relationship between fluid intelligence and Aha experiences. For example, there may be a threshold level of difficulty beyond which fluid intelligence is no longer a negative predictor of Aha.

Increased difficulty may increase the likelihood of individuals with higher fluid intelligence being unable to find a solution within the initial representation, thus necessitating a restructure. Whereas, increased difficulty would reduce the likelihood of solution-finding (Preusse et al., 2011) for those with lower fluid intelligence, reducing the number of Aha experiences. The results of this study may be of practical importance in the context of metacognitive monitoring and accuracy. For example, in an educational environment, training in metacognition may encourage the adoption of more evaluative processes during problem-solving (Miller & Geraci, 2011).

To our knowledge, this is the first study to assess the relationship between individual differences in fluid intelligence and metacognitive control, Aha experiences and theoretically meaningful problem-solving experiences that precede them. The present research adds to the body of empirical literature in theoretical, practical and methodological terms and provides directions for future research. The introduction of a new device will enhance future studies.

This device may also be applied more broadly across other domains where continuous, real-time measurement is desirable.

# References

Altmann, E. M., Gray, W. D. (2008). An integrated model of cognitive control in task switching. *Psychological Review*, *115*(3), 602–639. http://dx.doi.org.libproxy.murdoch.edu.au/10.1037/0033-295X.115.3.602

Andrich, D. (1988). *Rasch Models for Measurement*. https://doi.org/10.4135/9781412985598

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). https://doi.org/10.7771/1932-6246.1056

Austin, P. C. (2010). Estimating multilevel logistic regression models when the number of clusters is low: a comparison of different statistical software procedures. *The International Journal of Biostatistics*, *6*(1). https://doi.org/10.2202/1557-4679.1195

Barbey, A. K. (2018). Network neuroscience theory of human intelligence. *Trends in Cognitive Sciences*, *22*(1), 8–20. https://doi.org/10.1016/j.tics.2017.10.001

Beaty, R. E., Seli, P., & Schacter, D. L. (2019). Network neuroscience of creative cognition: Mapping cognitive mechanisms and individual differences in the creative brain. *Current Opinion in Behavioral Sciences*, *27*, 22–30. https://doi.org/10.1016/j.cobeha.2018.08.013

Benedek, M., & Fink, A. (2019). Toward a neurocognitive framework of creative cognition: The role of memory, attention, and cognitive control. *Current Opinion in Behavioral Sciences*, *27*, 116–122. https://doi.org/10.1016/j.cobeha.2018.11.002

Boone, W. J. (2016). Rasch analysis for instrument development: why, when, and how? *CBE Life Sciences Education*, *15*(4). https://doi.org/10.1187/cbe.16-04-0148

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

Bowden, E. M., Jung-Beeman, M., Fleck, J., & Kounios, J. (2005). New approaches to demystifying insight. *Trends in Cognitive Sciences*, *9*(7), 322–328. https://doi.org/10.1016/j.tics.2005.05.012

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical Linear Models: Applications and Data Analysis Methods*. Newbury Park, USA: SAGE Publications, Inc.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, *54*(1), 1–22. http://dx.doi.org.libproxy.murdoch.edu.au/10.1037/h0046743

Chu, Y., & MacGregor, J. (2011). Human performance on insight problem solving: A review. *The Journal of Problem Solving*, *3*(2). https://doi.org/10.7771/1932-6246.1094

Chuderski, A., & Jastrzebski, J. (2018). Much ado about aha!: Insight problem solving is strongly related to working memory capacity and reasoning ability. *Journal of Experimental Psychology: General*, *147*(2), 257–281. https://doi.org/10.1037/xge0000378

Chuderski, Adam. (2014). How well can storage capacity, executive control, and fluid reasoning explain insight problem solving. *Intelligence*, *46*, 258–270. https://doi.org/10.1016/j.intell.2014.07.010

Danek, A. H., Fraps, T., Müller, A., Grothe, B., & Öllinger, M. (2014). It’s a kind of magic – what self-reports can reveal about the phenomenology of insight problem solving. *Frontiers in Psychology*, *5*, 1-11. https://doi.org/10.3389/fpsyg.2014.01408

Danek, A.H., Wiley, J., & Öllinger, M. (2016). Solving classical insight problems without aha! Experience: 9 Dot, 8 Coin, and matchstick arithmetic problems. *Journal of Problem Solving*, *9*(1), 47–57. https://doi.org/10.7771/1932-6246.1183

Danek, A. H., Fraps, T., von Müller, A., Grothe, B., & Öllinger, M. (2013). Aha! Experiences leave a mark: Facilitated recall of insight solutions. *Psychological Research*, *77*(5), 659–669. https://doi.org/10.1007/s00426-012-0454-8

Danek, A. H., & Salvi, C. (2018). Moment of Truth: Why Aha! Experiences are Correct. *The Journal of Creative Behavior*, *0*(0), 1-3. https://doi.org/10.1002/jocb.380

Danek, A. H., & Wiley, J. (2016). What about false insights? Deconstructing the Aha! experience along its multiple dimensions for correct and incorrect solutions separately. *Frontiers in Psychology*, *7*(2077), 1-14. https://doi.org/10.3389/fpsyg.2016.02077

Davidson, J. E., & Sternberg, R. J. (1984). The role of insight in intellectual giftedness. *Gifted Child Quarterly*, *28*(2), 58–64. https://doi.org/10.1177/001698628402800203

Deary, I. (2004). An “instantaneous” estimate of a lifetime’s cognitive change. *Intelligence*, *32*(2), 113–119. https://doi.org/10.1016/j.intell.2003.06.001

Dix, A., Wartenburger, I., & van der Meer, E. (2016). The role of fluid intelligence and learning in analogical reasoning: How to become neurally efficient? *Neurobiology of Learning and Memory*, *134*, 236–247. https://doi.org/10.1016/j.nlm.2016.07.019

Dominowski, R. L., & Buyer, L. S. (2000). Retention of problem solutions: the re-solution effect. *The American Journal of Psychology*, *113*(2), 249–274. https://doi.org/10.2307/1423730

Dunlosky, J., & Rawson, K. A. (2012). Overconfidence produces underachievement: Inaccurate self evaluations undermine students’ learning and retention. *Learning and Instruction*, *22*(4), 271–280. https://doi.org/10.1016/j.learninstruc.2011.08.003

Fleck, J. I., & Weisberg, R. W. (2013). Insight versus analysis: Evidence for diverse methods in problem solving. *Journal of Cognitive Psychology*, *25*(4), 436–463. https://doi.org/10.1080/20445911.2013.779248

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. https://doi.org/10.1080/13546780903178615

Hedne, M. R., Norman, E., & Metcalfe, J. (2016). Intuitive feelings of warmth and confidence in insight and noninsight problem solving of magic tricks. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01314

Hill, G., & Kemp, S. M. (2018). Uh-oh! What have we missed? A qualitative investigation into everyday insight experience. *The Journal of Creative Behavior, 52*(3), 201–211. https://doi.org/10.1002/jocb.142

IBM Corp. (2017). IBM SPSS Statistics for Macintosh (Version 25.0). Armonk, NY: IBM Corp.

Jung, R. E., & Haier, R. J. (2007). The Parieto-Frontal Integration Theory (P-FIT) of intelligence: Converging neuroimaging evidence. *Behavioral and Brain Sciences*, *30*(2), 135–154. https://doi.org/10.1017/S0140525X07001185

Kizilirmak, J. M., Serger, V., Kehl, J., Ollinger, M., Folta-Schoofs, K., & Richardson-Klavehn, A. (2018). Feelings-of-warmth increase more abruptly for verbal riddles solved with in contrast to without aha! experience. *Frontiers in Psychology, 9*,(1404), 1-11. https://doi.org/10.3389/fpsyg.2018.01404

Kline, P. (1986). *A handbook of test construction: Introduction to psychometric design.* London, UNITED KINGDOM: Methuen.

Knoblich, G., Ohlsson, O., 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. http://dx.doi.org.libproxy.murdoch.edu.au/10.1037/0278-7393.25.6.1534

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. https://doi.org/10.1111/j.1467-9280.2006.01798.x

Laukkonen, R. E., & Tangen, J. M. (2018). How to detect insight moments in problem solving experiments. *Frontiers in Psychology*, *9*(MAR). https://doi.org/10.3389/fpsyg.2018.00282

Laukkonen, R., Ingledew, D., & Tangen, J. (2018). Getting a grip on insight: An embodied measure of Aha! and metacognition during problem solving. *PsyArXiv*. https://doi.org/10.31234/osf.io/fyhwb

Liljedahl, P. G. (2005). Mathematical discovery and affect: The effect of AHA! experiences on undergraduate mathematics students. *International Journal of Mathematical Education in Science and Technology, 36*(2–3), 219–234. https://doi.org/10.1080/00207390412331316997

MacGregor, J., Ormerod, T. C., & Chronicle, E. P. (2001). Information processing and insight: A process model of performance on the nine-dot and related problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(1), 176–201.

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence, 37*(1), 1–10. https://doi.org/10.1016/j.intell.2008.08.004

McNeish, D. M., & Stapleton, L. M. (2016). The effect of small sample size on two-level model estimates: A review and illustration. *Educational Psychology Review*, *28*(2), 295–314. https://doi.org/10.1007/s10648-014-9287-x

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

Miller, T. M., & Geraci, L. (2011). Training metacognition in the classroom: The influence of incentives and feedback on exam predictions. *Metacognition and Learning; New York*, *6*(3), 303–314. http://dx.doi.org.libproxy.murdoch.edu.au/10.1007/s11409-011-9083-7

Mueller, S. T., & Piper, B. J. (2014). The psychology experiment building language (PEBL) and PEBL test battery. *Journal of Neuroscience Methods*, *222*, 250. https://doi.org/10.1016/j.jneumeth.2013.10.024

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In *In G. Bower (Ed.) The psychology of learning and motivation* (Vol. 26, pp. 125–140). New York: Academic Press.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill

Ohlsson, S. (1992). Informaton-processing explanations of insight and related phenomena. In M. Keane & K. J. Gilhooly (Eds.). In *Advances in the psychlogy of thinking* (Vol. 1, pp. 1–44). London: Harvester-Wheatsheaf.

Preusse, F., Meer, E. eVan D., Deshpande, G., Krueger, F., & Wartenburger, I. (2011). Fluid intelligence allows flexible recruitment of the parieto-frontal network in analogical reasoning. *Frontiers in Human Neuroscience, 5*. https://doi.org/10.3389/fnhum.2011.00022

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical liner models: Applications and data analysis methods* (Vol. 1.) Sage.

Raudenbush, S., Bryk, T., & Congdon, R. (2013). Hierarchical linear and nonlinear modeling (Version 7.03). Retrieved from www.ssicentral.com

Raven, J. (2000). The raven’s progressive matrices: Change and stability over culture and time. *Cognitive Psychology*, *41*(1), 1–48. https://doi.org/10.1006/cogp.1999.0735

Raven, J. C., Court, J. H., & Raven, J. (1997). *Manual for Raven’s Progressive Matrices and Vocabulary Scale. Section 5: Mill HIll Vocabulary Scales.* Oxford, UNITED KINGDOM: Oxford Psychologists Press Ltd.

Rozencwajg, P. (2003). Metacognitive factors in scientific problem-solving strategies. *European Journal of Psychology of Education; Dordrecht*, *18*(3), 281–294. http://dx.doi.org.libproxy.murdoch.edu.au/10.1007/BF03173249

Salvi, C., Bricolo, E., Kounios, J., Bowden, E., & Beeman, M. (2016). Insight solutions are correct more often than analytic solutions. *Thinking & Reasoning*, *22*(4), 443–460. https://doi.org/10.1080/13546783.2016.1141798

Sandkühler, S., & Bhattacharya, J. (2008). Deconstructing insight: EEG Correlates of insightful problem solving. *PLoS One*, *3*(1). https://doi.org/10.1371/journal.pone.0001459

Schooler, J., Ohlsson, S., & Brooks, K. (1993). Thoughts beyond words: When language overshadows insight - ProQuest. *Journal of Experimentl Psychology*, *122*(2), 166–183. http://dx.doi.org.libproxy.murdoch.edu.au/10.1037/0096-3445.122.2.166

Shen, W., Yuan, Y., Liu, C., & Luo, J. (2016). In search of the “Aha!’’ experience: Elucidating the emotionality of insight problem-solving.” *British Journal of Psychology*, *107*(2), 281–298. https://doi.org/10.1111/bjop.12142

Smith, C. L. (2012). Stellan Ohlsson: Deep learning: How the mind overrides experience. *Science & Education*, *21*(9), 1381–1392. https://doi.org/10.1007/s11191-012-9449-5

Sternberg, R. J., & Davidson, J. E. (1995). *The Nature of Insight*. Cambridge: MIT Press.

Suprano, I., Delon-Martin, C., Kocevar, G., Stamile, C., Hannoun, S., Achard, S., … Sappey-Marinier, D. (2019). Topological modification of brain networks organization in children with high intelligence quotient: A resting-state fMRI study. *Frontiers in Human Neuroscience*, *13*, 1–12. https://doi.org/10.3389/fnhum.2019.00241

Tik, M., Sladky, R., Luft, C. D. B., Willinger, D., Hoffmann, A., Banissy, M. J., … Windischberger, C. (2018). Ultra-high-field fMRI insights on insight: Neural correlates of the Aha!-moment. *Human Brain Mapping*, *39*(8), 3241–3252. https://doi.org/10.1002/hbm.24073

Topolinski, S., & Reber, R. (2010). Gaining Insight Into the "Aha" Experience. *Current Directions in Psychological Science*, *19*(6), 402–405. https://doi.org/10.1177/0963721410388803

Veenman, M. V. J., & Beishuizen, J. J. (2004). Intellectual and metacognitive skills of novices while studying texts under conditions of text difficulty and time constraint. *Learning and Instruction*, *14*(6), 621–640. https://doi.org/10.1016/j.learninstruc.2004.09.004

Verhaeghen, P., Steitz, D. W., Sliwinski, M. J., & Cerella, J. (2003). Aging and dual-task performance: A meta-analysis. *Psychology and Aging*, *18*(3), 443–460. http://dx.doi.org.libproxy.murdoch.edu.au/10.1037/0882-7974.18.3.443

Waschl, N., Nettelbeck, T., Jackson, S., & Burns, N. (2016). Seeing reason: The dimensionality of the Advanced Raven's Progressive Matrices. *Personality and Individual Differences*, *101*, 525. https://doi.org/10.1016/j.paid.2016.05.344

Webb, M. E., Little, D. R., & Cropper, S. J. (2016). Insight is not in the problem: Investigating insight in problem solving across task types. *Frontiers in Psychology*, *7*(SEP). https://doi.org/10.3389/fpsyg.2016.01424

Webb, M. E., Little, D. R., & Cropper, S. J. (2017). Once more with feeling: Normative data for the aha experience in insight and noninsight problems. *Behavior Research Methods*, 1–22. https://doi.org/10.3758/s13428-017-0972-9

Webb, M. E., Little, D. R., Cropper, S. J., & Roze, K. (2017). The contributions of convergent thinking, divergent thinking, and schizotypy to solving insight and non-insight problems. *Thinking & Reasoning*, *23*(3), 235–258. https://doi.org/10.1080/13546783.2017.1295105

Weisberg, R. W. (2015). Toward an integrated theory of insight in problem solving. *Thinking & Reasoning*, *21*(1), 5–39. https://doi.org/10.1080/13546783.2014.886625

Yantz, C. L., & McCaffrey, R. J. (2005). Effects of a Supervisor's Observation on Memory Test Performance of the Examinee. *Journal of Forensic Neuropsychology*, *4*(2), 27–38. https://doi.org/10.1300/J151v04n02_03