**Accident and Agency: A Mixed Methods Study Contrasting Luck and Interactivity in Problem Solving**

Wendy Ross

London Metropolitan University

and

Frédéric Vallée-Tourangeau,

Kingston University

Address correspondence to Wendy Ross, Psychology Department, London Metropolitan University, 166-220 Holloway Road, N7 8DB, UNITED KINGDOM, [w.ross@londonmet.ac.uk](mailto:w.ross@londonmet.ac.uk). We thank Andrea Marin for her time watching the video corpus alongside us and Paul March, Linden Ball and three anonymous reviewers for helpful comments on a previous version of this manuscript.

**Abstract**

Problem solving in a materially rich environment requires interacting with chance. Sixty-four participants were invited to solve 5-letter anagrams presented as movable tiles: half of the conditions allowed the participants to move the tiles as they wished, the other half only allowed random shuffling (without rearranging the tiles post shuffling) thus contrasting pure luck with an interactive model. We hypothesised that shuffling would break unhelpful mental sets and introduce beneficial unplanned problem-solving trajectories. However, participants performed significantly worse when shuffling, which suggests luck plays less of a role than has been previously suggested. Granular analysis of seven critical cases revealed arbitrary path dependency across both conditions and moments of missed luck. It also questions current models of non-agentic luck and the ability to separate agent and luck. This research has implications for fostering better problem solving in an uncertain and fluid world.

Keywords: Serendipity, Interactivity, Chance Discovery, Luck

**Accident and Agency: A Mixed Methods Study Contrasting Luck and Interactivity**

The material world is in a constant state of change. This flux offers various opportunities for action, which come under the broad label of chance (Bardone & Magnani, 2013). Traditional models of problem solving tend to consider these moments of chance as irrelevant to an understanding of the underlying process of solution discovery. However, anecdotal evidence from scientific discoveries (Copeland, 2019; Sand, 2020) and interview data (Csikszentmihalyi, 1996) alongside informal and formal observations of experimental participants (Steffensen et al., 2016) suggests that chance—often in the form of accidental unplanned movements—plays a non-trivial role in problem solving and creativity.

Incorporating chance into models of problem solving requires embracing environmental fluctuations and a wider perspective on what constitutes cognition. A systemic perspective (F. Vallée-Tourangeau & Vallée-Tourangeau, 2020; G. Vallée-Tourangeau & Vallée-Tourangeau, 2017) posits that the boundary between the brain and world is permeable and that cognition is better conceived of as an interaction between internal processes and cognitive objects, that is as a coupling of the mental and the material. By cognitive objects here we mean those artifacts that are designed to scaffold and support cognition and become cognitive through their use (Heersmink, 2013). If the extended mind is constituted through interactivity in this way, then problem solving is best considered an emergent product of a cognitive ecosystem (Steffensen, 2017) and this system must necessarily incorporate chance. We suggest that this chance aspect of environmental engagement has been under explored even in those cognitive models and theories which propose a more embedded view of cognition. It may however be fruitful to consider this aspect, especially in relation to non-linear insight problem solving. Research from the area of chance discovery suggests this bears a close resemblance to insight problem solving in requiring both sudden shifts and gradual realisations (Terai & Miwa, 2013) along with distinctive phenomenological markers (Makri & Blandford, 2012). It is plausible that there is an overlap between the two research domains.

**Chance, Luck and Serendipity**

Chance refers to the continual flow of environmental fluctuations (Griffith, 2010). It becomes manifest to the human agent through accidents which can either be positive or negative (Ross, forthcoming a) and once that chance becomes significant to a person then it is categorised as either luck or serendipity (Rescher, 1995) depending on the level of agential involvement. There are two main positions which define the class of events to be called luck: the lack of control (Coffman, 2009) and the modal (Pritchard, 2014) accounts. The lack of control account stipulates that the event in question occurs when there is no agentic involvement from the person experiencing that luck either in the lead up to the event, or crucially, after. For example, winning the lottery is dependent on the correct numbers being drawn; this is something over which the person has no control (Coffman, 2009). In turn, the modal account relies on counterfactual thinking: a lucky event is one that could not have happened in the nearest of possible worlds. While the modal account does not require a lack of agency, as Carter and Peterson (2017) argue, the lack of agency is implicit because an agent’s actions would be likely to have the similar results in another possible world. The only thing that would generate different results is something beyond the agent’s control. Therefore, a model which takes luck into account reverses traditional cognitive models, which locate epistemic credit in the person; rather epistemic credit is environmentally situated.

Serendipity, on the other hand, refers to a mixture of ‘accident and sagacity’, that is, it emerges at the intersection of a person and luck (Copeland, 2019). Serendipity can be triangulated in terms of an environmental, personal and temporal dimensions (Makri & Blandford, 2012; Rubin et al., 2011). In other words, serendipity involves the right person in the right place at the right time. Material and social worlds are dynamically entwined and produce a rich array of possible perceptions and action possibilities. The fortuitous and felicitous qualities of some of these possibilities when enacted by an intentional agent is the main focus of serendipity (Ross, 2020). On this view epistemic credit for discovery is distributed across both agent and environment and the clear distinctions between the two are collapsed.

In a more concrete manner, luck in the case of problem solving describes those things beyond the control of the problem solver, but which cause a material change in her environment that becomes significant to her. While this may be characterised as external volatility (Ohlsson, 2018), the modal account requires it to be on a smaller scale than hypothesised by Ohlsson. Since larger environmental changes are likely to traverse several possible worlds, luck is more clearly seen in minor ‘accidents’. What is important for the current study is that luck maintains a distinction between the problem solver and her environment: “something happens to a person, it is not something that he does” (Griffith, 2010, p. 46). In contrast, serendipity presupposes an active agent who both generates and responds to environmental contingencies. In other words, accidents arise all the time, it is the role of the agent in taking advantage of them which is important and requires further attention.

**Luck and Serendipity in Problem Solving**

There has been little research explicitly into the role of chance in the form of either luck or serendipity in problem solving but its role has been anecdotally observed. In these informal observations it is often recast as an accident, although the lack of systematic consideration means it is unclear if it is accident generated by a stochastic process (luck) or accident enacted through action (serendipity). Accident here is taken to be something which occurs in the external environment caused by unintentional actions over objects (Ross, forthcoming a).

To date, randomness—a form of non-agent directed luck—has been seen to improve performance in a word production task over and above direct interaction with movable letter tiles (Kirsh, 2014, although see Ross & Vallée-Tourangeau, 2021). In Kirsh (2014) participants were invited to generate as many words as they could either with a static display, a display which could be moved and in a random shuffle condition. More words were produced in this shuffle condition where the aleatoric process generates a wider range of candidates to be quickly rejected. Word verification carries a lower cognitive cost than generation; randomness is really a series of hints activating liminal words. Indeed, Kirsh (2014) argues that randomness in this case broadens the search space and is functionally equivalent to an adding additional member to a team. In a similar task, Gavurin (1967) also found a significant boost for a random condition over a static display in anagram solving and further hypothesised that a typical anagram solver may be held back by a strategic fixation with pursuing solutions based on her knowledge of English lexical structure. The addition of unplanned and unexpected hints should break an impasse wrought by this solution pathway. These hints driven by randomness are similar to non-agentic luck and imply that a problem re-representation can be better facilitated by unintended moves than by a problem solver doggedly reifying unhelpful ideas.

In addition, across the literature there has been some informal consideration of the way problem solving is facilitated by unplanned changes while interacting with movable representations. However, it is hard to fully extract the nature of the interaction posited, whether it is an “accident” or the results of purposeful interactions. Fleck and Weisberg (2013) refer to this environmentally driven cognition as ‘data driven restructuring’: “Data-driven restructuring included instances when the individual changed his or her representation of the problem in response to something he or she saw from the physical configuration of the problem […]. Observations occurred as the participant was attempting to construct or implement another heuristic-based solution” (Fleck & Weisberg, 2013, p. 452). In this instance it is unclear whether these observations were driven by luck or serendipity (the description implies unintentional action) but their nature deserves further study.

Other studies reported observations of accidental success, but it is not always clear whether that accident is an enacted one. Fioratou and Cowley (2009) describe a version of the cheap necklace problem[[1]](#footnote-1) presented with actual link-chains and observe that 6 of the 21 solvers (almost a third) solved the problem through the exploitation of an “accident”. This same observation occurs in Chuderski et al. (2020, p. 18) who suggest that “in the matchstick algebra problem, it is arguably easier to arrive at the solution by accident or trial and error, for instance by realizing as a result of a random movement of a stick that it could act as a negative sign”. Here accident is presumably luck as the agent does not move intentionally or with a plan in mind and it is contrasted with more active trial and error. Most clearly, in one of the few qualitative approaches to problem solving, Steffensen et al. (2016) undertake a finely grained analysis of a participant solving the 17-animal problem[[2]](#footnote-2). The analysis attributes the realisation that the problem requires overlapping sets to have been triggered by “serendipity” (p. 92) or an accident. Here agency and epistemic credit appear to be distributed across person and environment. The extent of agential involvement becomes important to categorise if we wish to explore this aspect further.

This distinction between luck and serendipity is important when considering the extent to which helpful hints can support problem solving (Kirsh, 2009). Neth and Payne (2011) compared the performance of participants in a yoked coin addition task. The factorial design compared performance in four conditions: when allowed to move the coins and given a random array, prevented from moving coins and given a random array, moving and given the array of a previous successful participant and no moving and given a successful array. To some extent being given an already sorted and helpful array could be seen as luck – if it were to happen in the “wild” it would be a helpful hint. However, participants made fewer errors when allowed to interact with an unhelpful array than when given the static already sorted and helpful array, suggesting that the actions may be important to fully leverage the advantages of a helpful array and that pure luck is less important than proposed by Kirsh and Gavurin (as outlined above). Instead, it may be the interaction with the environment which at once generates accidents and drives the response to them which is important. It is this tension we wish to explore.

The aim of this study was to make a first attempt at a systematic analysis of the role of luck in problem solving and to what extent it can replace a model based on active engagement (serendipity). We move accident and chance from something that is informally observed and commented on to something that is expressly manipulated. By forefronting the role of environmental chance in this way, we hope to increase our understanding of the nature of our interaction with cognitive objects beyond considering them as simple working memory extensions (e.g., Risko & Gilbert, 2016)

**Laboratory Research on Interactivity**

Moving away from a research programme which focuses on internal cognitive processes to one that accounts for interaction with cognitive objects in a controlled manner is not an easy task and many of the studies in this area have taken place under the theoretical framework of interactivity or systemic cognition (F. Vallée-Tourangeau & Vallée-Tourangeau, 2020). Interactivity describes the coordination of a problem solver with her environment (Steffensen, 2017). It is commonly examined by allowing an experimental participant access to movable problem representations and assesses whether this leads to increased performance. Broadly, an interactive problem-solving experimental condition encourages the problem solver to make full use of the cognitive scaffolds and extends the cognitive ecosystem to reflect more accurately embedded and situated problem solving (Steffensen & Vallée-Tourangeau, 2018). A low interactivity condition focuses on the performance of a problem solver while decoupled from this system and restricted in her movements so as to better isolate internal cognitive processes. The two conditions are then compared to establish whether there is an augmentative effect in interacting with the environment.

Expanding the problem space in the lab to incorporate engagement with problem solving materials tends to increase success in solving insight problems, suggesting that such materials can themselves scaffold or even trigger moments of insight (Henok et al., 2020; Steffensen et al., 2016; Vallée‐Tourangeau, 2014). However, solution rates with interactivity vary as a function of the type of problem and the material provided to the participants (Chuderski et al., 2020), which suggests that in these nonlinear tasks, interactivity leads to complex and unpredictable outcomes. Therefore, we are arguing for increased attention to be paid to the mechanisms through which interactivity can support or otherwise affect cognition with the understanding that it may not be a straightforwardly augmentative process (Ross & F. Vallée-Tourangeau, 2021; F. Vallée-Tourangeau et al., 2020)

There are two main possible explanations for the augmentative effect of interactivity. Both are likely to play a role in problem solving activity. First, the material world can act as a cognitive offload, which supports and augments the problem solver’s internal resources (Clark & Chalmers, 1998; Risko & Gilbert, 2016; G. Vallée-Tourangeau & Vallée-Tourangeau, 2017), reflecting planned and purposeful moves by the problem solver. Second, these interactive environments (whether explicitly conceived as thus or not) allow the environment to play a constitutive role in the cognitive ecosystem through unplanned and unexpected changes. The unpredictable nature of these change may account for the inconsistent results in this area. It is this second explanation that the current study was designed to investigate by contrasting a passive interaction with a changing display with a more interactive one which we tentatively suggest can reflect pure luck or serendipity.

Furthermore, it is plausible that the role of accident may be a more important explanatory factor for performance on problems solved through insight than for analytical problems. Insight problems are those that are not solved by linear step by step processes but rather require an abductive leap and therefore the role of working memory and cognitive offloading is necessarily smaller (Gilhooly & Webb, 2018). The strong evidence for the benefits of interactivity in analytical problem solving is not matched by similar evidence in insight problem solving (see Chuderski et al., 2020; Vallée‐Tourangeau, 2014). If interactivity is most beneficial as a working memory scaffold, it makes sense that the clear augmentative benefits are somewhat dissipated in insight problem solving. However, theories of insight problem solving require a restructuring of an initial misleading representation. If insight is reliant on taking a path which has not been planned and having a sudden shift in representation, changes in environmental information may as act as trigger for restructuring (Ohlsson, 2018). Therefore, it seems likely that the role of luck as an external trigger of restructuring might be more important in non-linear tasks.

The traditional model of insight problem solving also requires an impasse. Indeed, Ohlsson writes, ‘Without the impasse, there is no insight, only smooth progress’ (1992, p.4) Participants labour away on a problem until they realise that their initial problem representation is incorrect and then they experience impasse before there is a helpful or unhelpful restructuring and so the cycle continues. However, the requirement for impasse is no longer central to theories of insight; for example, Danek (2018) labels it as optional in her model of non-monotonic problem solving. Furthermore, Fleck and Weisberg (2013) report low levels of impasse using verbal protocols during problem solving with traditional insight problems. It seems more likely that luck will be important when participants have exhausted the search space and require inspiration from outside of the internally generated search space and therefore, we would expect it to be more important to those participants who follow a classic insight path of impasse and restructuring. In this case the restructuring would be wrought by external random changes rather than an internal reprocessing.

**Anagrams as an Insight Task**

Insight is methodologically problematic to evoke but is often accompanied by a sudden restructuring of the problem space and clear phenomenological markers: “Aha” moments (F. Vallée-Tourangeau, 2018). Such moments are recognised when they occur but do so unreliably in a natural environment (Friedlander & Fine, 2018) and so those interested in the cognitive mechanisms of insight have developed a battery of tasks designed to elicit this feeling in the laboratory. These tasks range from traditional riddles such as the antique coin problem[[3]](#footnote-3) to tasks such as Compound Remote Associates (CRA; Bowden & Jung-Beeman, 2003)) where participants are required to find a link word between three seemingly unrelated words. Although these do not always elicit pure ‘insight’ they allow for a large number of repeated trials with a reasonable amount of control over the difficulty level and the predicted outcome (Webb et al., 2016). More recently, the level of insight required to solve the problem has not been assumed but has rather been measured by self-report (Bowden et al., 2005; Danek et al., 2014).

The anagram task has a long history both in recreational problem solving and as an experimental task. A true anagram is a full word that can be rearranged using all of the letters only once to create another word, such as EATERS which can be rearranged to spell EASTER. More usually though, the letters are presented in a scrambled non-word form such as STEAER and the problem solver is invited to uncover the target word from this letter string. Theoretically, anagrams can be solved by incremental and step by step processing and, in the past, have been often characterised as non-insight problems (Gilhooly & Murphy, 2005). However, it seems likely that they can also be solved in an insightful way: recent normative research indicates that the anagrams elicit reliably high ratings of insight (Webb et al., 2018) and anagrams are now often used as insight problem solving stimuli (Ellis et al., 2011; Kounios et al., 2009; Valueva et al., 2016). It is likely that they are solved through a mixture of insightful and analytical processes.

However, the current study did not primarily aim to answer questions about the nature of anagram solutions per se: Anagrams provide a convenient platform for the study of luck in problem solving using an interactive procedure with letter tiles that can be moved. First, shuffling the letter tiles that configure an anagram is a meaningful action that could yield the problem solution or at least steps toward it in a way that randomly reconfiguring the words of a riddle are unlikely to. Second, the difficulty of an anagram depends to some extent on distractor bigrams, so provoking random changes may be crucial for breaking mental set and making unplanned leaps through the problem space (Maglio et al., 1999). Additionally, the letters form easily traceable and identifiable problem components so the problem-solving process can be traced, physically, through these behavioural markers and changes in the physical composition of the problem; these changes permit the detailed coding of the transformation of the problem over time. As participants interact with the problem, the researcher can monitor the changes in the problem-solving environment, chart their path, and unveil solution trajectories as well as unproductive dead ends.

In sum, these rich and concrete data offer a strong opportunity to assess the role of luck and serendipity in problem solving. We designed two conditions – interactive and shuffle. Similar to the word production tasks in Kirsh (2014) and Ross and F. Vallée-Tourangeau (2021), the interactive condition allowed participants to move and interact with the letter tiles however they saw fit and in the shuffle condition they were able to pick up the tiles and shake them to shuffle them but were then allowed no further interaction. Thus, we aimed to test whether luck in the form of uncontrolled changes to the problem array (shuffle) aided problem solving more than a dynamic coupling (serendipity).

**The Current Study**

The current study was designed to test whether a model based on luck or one based on serendipity captures better the way problems are solved. We employed a mixed methods design. Mixed method research is particularly useful to pick up the role of chance in an experimental situation using video data and close analysis of participant behaviours. While it is possible to manipulate experimental variables to assess the overall effect of luck, the idiosyncratic nature of chance means that such broad analysis may miss these moments. On the other hand, the claim in the literature that luck is responsible for solutions should be tested systematically. Additionally, traditional research on luck and serendipity relies on self-report and thus is filtered through a lens of human agency (Copeland, 2019). A detailed, highly granular examination of the nature of small moments of microserendipity (Ross & F.Vallée-Tourangeau, 2020, 2021) will add to our overall understanding of these complex phenomena.

Much like Gavurin (1967), we hypothesised that the benefits of random suggestion and the usefulness of new bigrams would lead to more anagrams being solved (and indeed selected an anagram as a promising platform to increase the likelihood of luck playing a key role). We further hypothesised in line with Kirsh (2014) that randomness would be particularly useful when the participants had encountered impasse and were ‘out of ideas’ by generating novel ideas. For this reason, we asked participants to indicate when they felt stuck.

We were also interested in the extent to which the participants chose to interact with the tiles. To this end we had only two conditions (an interactive and a shuffle condition) and omitted a low interactivity condition—where participants are prevented entirely from moving the tiles–which acts as a typical control condition in other experiments in interactivity. This was to ensure that the participants did not feel obliged to move or interact with the tiles and because we were mainly interested in the contrast between luck and serendipity–a low interactivity condition where participants are not able to move the tiles affords neither and may prime the participants to opt to move unnaturally in the other conditions. Rather we preregistered the creation of a third post hoc condition of non-movers if enough participants opted to not move the tiles. This allowed us to empirically test the hypothesis that participants would interact with the environment even in the absence of a condition which has artificially constrained their movements.

We therefore preregistered the following quantitative hypotheses:

H1: At impasse, random moves (shuffling) will lead to a greater proportion of solved anagrams in (a) the 3 seconds post impasse and (b) overall.

H2: Moving tiles will lead to lower reported rates of impasse overall.

H3: Latency to impasse will be longer in the free interactivity condition.

H4: When people are stuck, they move the tiles (whether shuffling or freely).

The qualitative research questions were generated after the quantitative research phase.

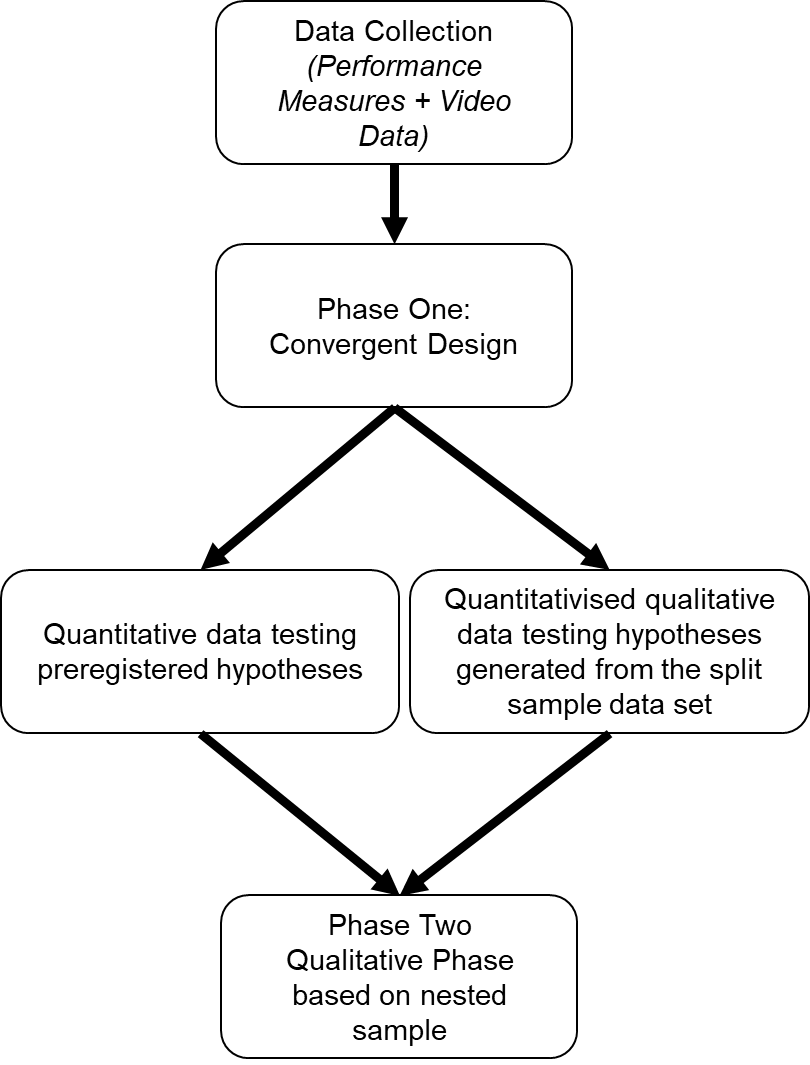
**Method**

The current study was pre-registered prior to data collection and the preregistration can be found here: <https://osf.io/we3ru/?view_only=730b84dacac24f1cafc315c20dd4fc2b>.

**Mixed Methods Design**

The current study combines three forms of data. First, the aggregated mean performance (proportion of correct anagrams and latency to solution) across the participants and measures of self-report (feelings of insight) data. Second, behavioural data which were produced by observation and then quantitativised using the methods of interaction analysis (Bakeman & Gottman, 2009; Bakeman & Quera, 2011). In this instance, the coding schemes were pre-specified after watching a randomly selected and then discarded group of participants (a split sample – see Figure 2). Finally, we present evidence from video based finely grained analysis to elaborate on the causal mechanisms for the results reported in phases 1 and 2 of the data analysis. Therefore, a fixed quantitative section and a fixed qualitative section reliant on quantitavisation of qualitative data (convergent subsection) are combined with an emergent qualitative section which is based on coding of the video data collected in parallel with initial codes developed convergently but refined after the quantitative analysis (QUAN 🡪 qual; Creswell, & Plano Clark, 2018)..

**Figure 1**

*The Modified Convergent Design*

As outlined in Figure 1, this makes it a modified version of an explanatory sequential study (Creswell & Plano Clark, 2018). The mixed methods sequential design typically consists of two distinct phases of data collection whereas in this study, data collection for both the QUAN and the qual phase happened simultaneously, which is similar to a convergent design. However, we follow Creswell and Plano Clark (2018) in suggesting that the timing is less relevant that the underlying intent. The quantitative data and the quantitativised qualitative data form the first phase of the research while the finely grained, narrative analysis of the video data was emergent and developed to answer questions raised by the quantitative data. Such a design is easier to implement with the global richness of video data, which can be re-oriented to answer underlying questions (Saldaña, 2016).

Thus, while the current study was quantitative dominant (Onwuegbuzie & Combs, 2011), it also aimed to establish the benefits of introducing qualitative data to understand not what the effect of the experimental manipulation was but how that effect was manifest (see Ball & Ormerod [2017] and Steffensen [2016] for a discussion on qualitative research in cognitive psychology). This technique has been used by the authors in previous research into lab based problem-solving experiments (Ross & F. Vallée-Tourangeau, 2021; F. Vallée-Tourangeau et al., 2020; see also Ross & F. Vallée-Tourangeau, in press, for a longer discussion of the methodology).

**Participants**

Eighty participants, representing a mixture of undergraduate and post graduate psychology students (67 female) ranging in age from 19 to 55 (*M* =25.73, *SD* = 8.84) were recruited in return for course credits. This participant sample was divided into two sections: an exploratory sample to generate observational hypotheses and a confirmatory sample which was used to test these and the preregistered hypotheses. The exploratory sample was not used to test any hypotheses and was discarded after hypothesis generation. A nested sample was selected from the confirmatory sample for the qualitative phase of analysis. This is illustrated in Figure 2.

***Confirmatory Sample Participant Characteristics***

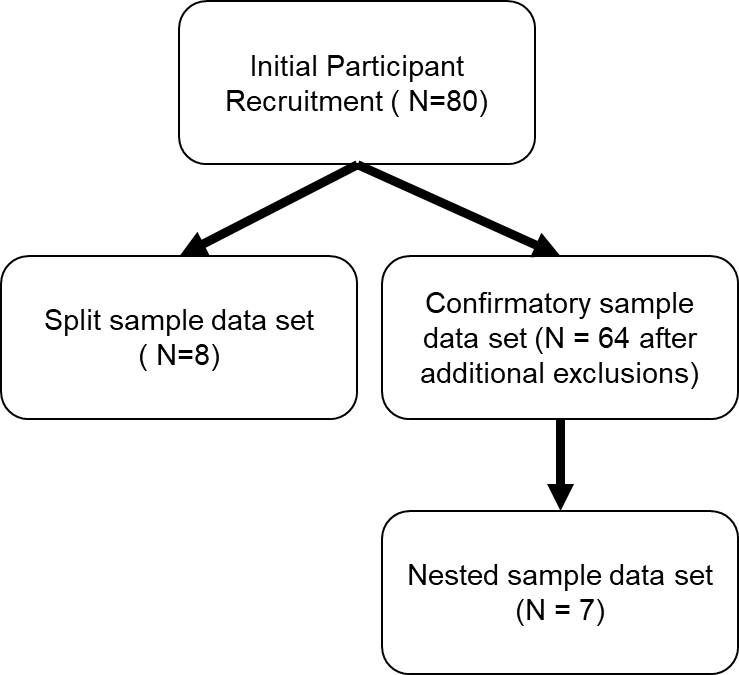
After exclusions (outlined below) the final quantitative sample consisted of 64 participants (55 female) with ages ranging from 19 to 55 (*M* = 25.19, *SD* = 8.68). All participants had performance data collected and were analysed at a coarse-grained behavioural level (e.g., time spent moving the letter tiles).

***Exploratory Sample Participant Characteristics***

To our knowledge an observational study of problem solving in this task has not be carried out before so to generate behavioural quantitative hypotheses 10% (8) of the dataset (7 female, *Mage* = 30.38, SD = 9.50, Range = 19 - 43) was selected at random[[4]](#footnote-4) and constituted the exploratory sample to support principled hypothesis generation (Anderson & Magruder, 2017). These were then excluded from the final analysis.

**Figure 2**

*The Different Samples*

****

***Further Exclusions***

Eight further participants (5 female, *Mag*e = 25.16, *SD* = 9.41, Range = 19- 44) were excluded from the confirmatory sample: five for data loss (three video loss; two total data loss whose demographic details are not reported here or above), one for consistent failure to follow instructions, one who became visibly upset during the verbal fluency task (the task is further explained below) and one who did not adhere to the instructions requiring English as a first language.

**Excluded Trials.** Six trials were excluded either because of experimenter error or because the participant was not following instructions. Commensurate with the preregistration, a further 20 trials which involved participants solving the anagram in under 3 seconds were excluded. Trials were also excluded from the final data set if the participants did not move the tiles. This is more fully outlined in the first section of the results. None of the trial exclusions resulted in a participant being fully excluded.

***Nested Sample***

After the quantitative analysis phase was concluded a nested sample of participants was selected from the confirmatory sample (See Qualitative Results for demographic details).

**Design**

The participants were invited to solve eight five letter anagrams: four in a shuffle condition where movement was restricted to picking up the tiles, shaking them in their hands and then randomly relaying them in a straight line and four in an interactive condition where they were able to interact with the tiles in any way they chose. The interactivity level (interactive, shuffle) was a repeated-measures factor. The anagrams were presented one at a time with the conditions alternating every two anagrams (leading to four blocks of two anagrams). The order of the initial condition was counterbalanced across participants.

**Materials**

The anagrams were selected from the set of hard five letter anagrams presented in Novick and Sherman (2008). To ensure that the target words were well known among the participant population, all 20 target words were presented to a sample of 25 undergraduate students from the same participant pool (who did not go on to take part in the full study). The participants were asked to rate their familiarity with the words on a Likert-type scale of 1 - 5 where 1 was “not familiar at all” and 5 was “very familiar”. These scores were then averaged to create a familiarity index. This led to the selection of eight most familiar target words which were presented in the same initial array to each participant, see Table 1. All were considered to be familiar to the target population. The order of the target words was counterbalanced across conditions.

**Table 1**

*The Target Word and the Presented Array*



The participants were also profiled along two dimensions of individual differences, anagram expertise and verbal fluency. Their level of anagram expertise was assessed using the questions in Novick and Sherman (2003), which require them to state how often they solve anagrams (every day, a few times a week, once a week, a few times a month, once a month, a few times a year, never) and how they would rate their anagram expertise in relation to other students (on a scale of -5 to +5 with 0 being average). Following, Novick and Sherman, the frequency was converted to a 1 -7 scale and two items collapsed to create a composite experience (Cronbach’s α = .75). The participants were also profiled on their phonemic verbal fluency using the F-A-S test and their semantic verbal fluency using the ANIMALS test (Whiteside et al., 2016). Both required participants to say aloud over the course of a minute words that fit the given category. The ANIMALS test required them to name animals, the F-A-S test words which begin with the letters F, A and S. The total number of words from all four trials were added together to yield a global verbal fluency score.

Additionally, if a participant solved an anagram successfully, they were asked to answer a series of insight questions (taken from Webb et al., 2018). This required them to rate using a Likert-type scale ranging from 1-100 where 1 = low and 100 = high : (1) the confidence that the given response was correct (very unsure to very sure), (2) the strength of the aha experience (very weak to very strong), (3) the pleasantness of the insight experience (very unpleasant to very pleasant), (4) the surprising nature of the insight experience (not surprising at all to very surprising), and (5) the feeling of impasse before the insight experience (no impasse at all to very stuck). This informed exploratory analyses on the potential processes required in that particular problem.

**Procedure**

Data were collected in a purpose-built observation lab with three cameras (see Figure 3) and answers were recorded on a Qualtrics interface (www.qualtrics.com). First, participants were invited to complete demographic information, which included the measures of anagram expertise. Before being invited to solve the anagrams, participants were asked to solve two practice anagrams: one interactive and one shuffle. If they did not opt to shuffle in the practice condition, they were invited to practice shuffling as a stand-alone procedure. The anagrams were presented in the form of movable tiles (2.7cm x 2.4cm) taken from the word game ApplettersTM . They were presented on a small tray behind a screen and latency was recorded from the time the screen was removed.

**Figure 3**

*The Experimental Set-up*



Participants were given 90 seconds to solve each anagram. If the participants did not solve the anagram in 90 seconds, they were told the answer. Before each trial they were reminded of the experimental condition and the requirement to announce impasse (adapted from Fedor et al., 2015) using the following wording:

**Interactive**

You will have 5 letter tiles in front of you which can be rearranged to make a common English word. When you know the word then please call it out. You do not have to move the tokens around but it may help you to do so. You may know the answer straight away but you may also get stuck, we would therefore like you to indicate when you are stuck, that is when you have exhausted all the solutions that initially come to mind. Please do this by saying, ‘I’m stuck’. You can do this as many times as you like. You may then carry on playing with the tokens to solve the problem.

**Shuffle**

You will have 5 letter tiles in front of you which can be rearranged to make a common English word. When you know the word then please call it out. You do not have to move the tokens but it may help you to do so. If you are going to rearrange them you must pick them all up in one go and shake them in your hands like a dice. You may know the answer straight away but you may also get stuck, we would therefore like you to indicate when you are stuck, that is when you have exhausted all the solutions that initially come to mind. Please do this by saying, ‘I’m stuck’. You can do this as many times as you like. You may then carry on shuffling the tokens to solve the problem.

If they solved the anagram, they were invited to rate their feelings of insight. Finally, participants were invited to do the verbal fluency tests: The F-A-S was administered first followed by the ANIMALS.

**Quantitative Results**

**Non-moving Trials**

In order to elicit natural behaviour, participants were told that they were under no obligation to move the letter tiles. However, a trial where the tiles were not moved does not reflect the experimental manipulation so therefore, an additional 128 trials were excluded because the participants did not move the tiles. No single participant was excluded because each participant moved the tiles in at least one trial. Eighty-six of these trials were from the shuffle condition and 42 were from the interactive condition. Of these trials, the overwhelming majority were correctly solved: all of the trials in the interactive condition were solved and 82 of the 86 (95.3%) shuffle trials were solved. These trials were also solved quickly: In the interactive condition, the average latency to a correct solution was 18.01s (*SD* = 16.09) while in the shuffle condition it was 19.16s (*SD* = 15.70). That no one participant opted to not move the letter tiles at all indicates that participants would opt to move when they needed to and when given the opportunity. Ceiling rates of success and low latency to solution would suggest that these were trials where the participants did not require external support to scaffold their answer.

**Deviations from Preregistered Analyses**

Several changes were made to the preregistered analyses. First, the high proportion of excluded trials and the possible effect of the individual anagrams on the overall effect of condition led to a change in analysis plan from paired *t* tests to a mixed models analysis (we thank a reviewer of an earlier version of this manuscript for this suggestion). To retain the largest numbers of trials but to account for the repeated measures design and the counterbalancing of the anagram stimuli, a mixed model was constructed for the different hypotheses with participant and anagrams as crossed random effects (see Barr et al., 2013 on the use of maximal structures for confirmatory hypothesis testing).

The preregistration suggested that we would only analyse those participants who reported impasse. While we followed this analysis plan (as reported below) during the experiment it became clear that the reporting of impasse varied between participants and did not match observed behavioural markers nor post task reports. Therefore, before any statistical analyses were run, the decision was made to analyse the whole data set. These inconsistencies are examined in more detail in the qualitative section of the results.

Additionally, H1(a) could not be tested because no participants successfully solved the anagrams 3 seconds after declaring impasse. H4 also could not be tested because only 19 trials involved participants who did not move before declaring impasse.

**Analysis**

All analyses were conducted in jamovi (The jamovi project, 2020). We used a mixed-model approach to analyse problem solving performance using the GAMLj package (Gallucci, 2019). Parameter estimates for all models are displayed in Tables 3 and 4. For the dichotomous variables (correct, impasse experienced) we used a binomial logit-link function in a generalised mixed model and for the linear variables (latency, insight, time to first interaction, proportion of time spent moving) we constructed linear mixed models. All models followed a maximal structure and had the participants and the anagram stimuli as crossed random factors with random intercepts and slopes.

**Overall Performance**

After exclusions 355 trials remained across the 64 participants and all 8 anagrams. There were more trials in the interactive condition (205) than in the shuffle condition (150). A greater proportion of trials (70.73%) were successfully solved in the interactive condition than in the shuffle condition where under half the trials were successfully solved (49.33%). To test our hypothesis that there would be a difference between the conditions, we specified an

**Table 3**

*Generalised Mixed Models for Binary Outcomes*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Model | Effect | Estimate | SE | exp(B) | 95% CI Lower | 95% CI Higher | z | p |
| Condition 🡪 Outcome | (intercept) | .36 | .15 | 1.44 | 1.06 | 1.94 | 2.35 | .019 |
|  | Shuffle-Interactive | -.66 | .18 | 0.52 | 0.37 | 0.73 | -3.73 | <.001 |
|  |  |  |  |  |  |  |  |  |
| Condition + Time to Interaction 🡪 | (intercept) | .75 | .30 | 2.11 | 1.17 | 3.82 | 2.40 | .014 |
| Outcome | Shuffle- Interactive | -.98 | .36 | 0.38 | 0.19 | 0.77 | 2.68 | .007 |
|  | Time to Interaction | -.03 | .02 | 0.97 | 0.94 | 1.00 | -1.91 | .056 |
|  |  |  |  |  |  |  |  |  |
| Condition🡪 Impasse | (intercept) | .48 | .27 | 0.62 | 0.37 | 1.04 | -1.79 | .073 |
|  | Shuffle-Interactive | .55 | .32 | 1.75 | 0.91 | 3.29 | 1.69 | .092 |
|  |  |  |  |  |  |  |  |  |
| (Impasse Set) Condition 🡪 Outcome | (intercept) | -.62 | .18 | 0.54 | 0.38 | .77 | -3.43 | <.001 |
|  | Shuffle -Interactive | -.54 | .29 | 0.58 | 0.33 | 1.03 | -1.85 | .064 |
| (Interactive Set) Time spent moving | (intercept) | .66 | .18 | 1.94 | 1.37 | 2.74 | 3.72 | <.001 |
|  | Time moving | -.47 | .53 | 0.63 | 0.22 | 1.77 | -0.88 | .38 |
|  |  |  |  |  |  |  |  |  |

**Table 4**

*Mixed Models for Continuous Outcomes*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Model | Effect | Estimate | SE | 95% CI Lower | 95% CI Higher | df | t | p |
| Condition 🡪 Latency to Correct | (intercept) | 37.60 | 2.04 | 33.64 | 41.60 | 8.87 | 18.46 | <.001 |
| Solution | Shuffle-Interactive | 11.00 | 3.73 | 3.73 | 18.40 | 15.48 | 2.96 | .010 |
|  |  |  |  |  |  |  |  |  |
| Condition 🡪 Time to Impasse | (intercept) | 18.54 | 2.13 | 14.36 | 22.70 | 11.70 | 8.69 | <.001 |
|  | Shuffle- Interactive | 4.62 | 3.00 | -1.26 | 10.50 | 14.30 | 1.54 | .146 |
|  |  |  |  |  |  |  |  |  |
| (Impasse Set) Condition 🡪 Latency to | (intercept) | -0.62 | 0.18 | 0.38 | 0.77 |  | -3.43 | <.001 |
| Correct Solution | Shuffle -Interactive | -0.54 | 0.29 | 0.33 | 1.03 |  | -1.85 | .064 |
|  |  |  |  |  |  |  |  |  |
| **Insight Ratings by Condition** |  |  |  |  |  |  |  |  |
| Stuck | (intercept) | 43.80 | 3.09 | 37.77 | 49.90 | 15.45 | 14.18 | <.001 |
|  | Shuffle- Interactive | 12.30 | 4.12 | 4.18 | 20.30 | 4.93 | 2.98 | .032 |
|  |  |  |  |  |  |  |  |  |
| Surprise | (intercept) | 57.30 | 3.10 | 51.23 | 63.40 | 18.00 | 18.51 | <.001 |
|  | Shuffle- Interactive | 5.06 | 3.79 | -2.37 | 12.5 | 12.4 | 1.33 | .206 |
|  |  |  |  |  |  |  |  |  |
| Aha | (intercept) | 66.77 | 3.32 | 60.25 | 73.28 | 19.3 | 20.08 | <.001 |
|  | Shuffle- Interactive | 1.84 | 39.3 | -5.86 | 9.53 | 14.7 | 0.47 | .647 |
|  |  |  |  |  |  |  |  |  |
| Correct | (intercept) | 84.59 | 2.43 | 79.82 | 89.35 | 18.45 | 34.79 | <.001 |
|  | Shuffle- Interactive | -4.42 | 2.83 | -9.97 | 1.13 | 5.29 | -1.56 | .176 |
|  |  |  |  |  |  |  |  |  |
| Pleasant | (intercept) | 83.02 | 1.58 | 79.92 | 86.11 | 53.48 | 52.61 | <.001 |
|  | Shuffle- Interactive | 0.18 | 1.70 | -3.14 | 3.50 | 7.27 | 0.11 | .92 |

additional model which included experimental condition as a fixed factor. This indicated that the condition had significant effect on the probability of getting the anagram correct, χ² (1, N = 355) = 14.1, *p* < .001.

Latency to solution can only be measured in successful trials. Two hundred and nineteen trials (145 interactive, 74 shuffle) remained across 60 participants and all 8 anagrams which allowed the comparison of the latency to solution in correct trials. The average latency to solution was faster in the interactive condition (*M* = 32.07, *SD* = 22.61) than the shuffle condition (*M* = 43.22, *SD* = 23.61). A linear model was constructed to test the effect of condition on latency using condition as a fixed effect and participant and anagram as crossed random effects: Condition was a significant predictor of solution latencies, *F*(1, 15.5) = 8.75, *p* = .010

**Impasse**

Overall, 36% of trials in the interactive condition resulted in the participants declaring impasse while 48% of trials in the shuffle condition had the participants declaring impasse which supports H1. When a model was fitted with condition as a fixed factor and anagram and participants as random crossed factors, condition was not a significant predictor of declaring impasse,χ² (1, N = 355) = 2.84, *p* = .092.

Contrary to our hypothesis, participants declared impasse faster in the interactive trials (*M* = 16.47 seconds, *SD* = 25.86) than in the shuffle trials (*M*=21.56 seconds, *SD* = 28.24), however when a model was fitted with condition as a fixed factor and anagram and participants as random crossed factors, condition was not a significant predictor of time to impasse, *F*(1, 14.3) = 2.37, *p* =. 146.

**Performance post impasse**

We were initially particularly interested in what happens when a participant declares impasse. Seventy-four trials in the interactive condition and 72 trials in the shuffle condition across 55 participants and all 8 anagrams involved a participant declaring impasse. There was a greater proportion of correct trials in the interactive condition post-impasse with 41.9% of trials resulting in the correct answer while only 22.2% of trials were correct in the shuffle condition. However, a general linear model with condition as a fixed effect and participant and anagram as crossed random effects suggests that condition was not a significant predictor of the probability of a correct answer in this case, *χ*² = 3.01, *p* = .083.

An analysis of latency to solution when a participant declared impasse suggested that trials were solved faster in the interactive condition (*M*= 57.02, *SD* = 24. 82) than the shuffle condition (*M* = 61.21, *SD* = 27.13). Only 47 trials could be included in this analysis which prevented the inclusion of participant as a random grouping effect. A linear mixed model was constructed with condition as the fixed effect and anagram for the random effect which demonstrated that condition was not a significant predictor of the time to solution, *F* < 1.

**Insight**

We further conducted exploratory analyses on the 214 successful trials for differences in reported insight. These are illustrated in Figure 4. Linear modelling with condition as a fixed effect and participant and anagram as a random effect suggested that condition was only a significant predictor in the case of being stuck *F*(1, 4.93) = 8.854, *p* =.032. Across all other dimensions condition was not a significant predictor (smallest *p* = .176 for “correct”). Only 31 trials were solved correctly in the interactive condition after impasse and 16 in the shuffle conditions precluding sensible statistical analysis. We discuss the low levels of success in those participants below.

**Figure 4**

*Insight Ratings Across the Five Dimensions for both the Interactive and Shuffle Conditions (a) Over the Whole Data Set (b) When Impasse Was Not Declared (c) When Impasse Was Declared. Error Bars Represent Standard Errors of the Mean.*



**Quantitative Behavioural Analysis**

Two further quantitative hypotheses were generated from reviewing the split sample.

H5: The time to first interaction (whether movement or shuffle) would predict performance.

H6: The proportion of time spent interacting with the tiles in the interactive condition would predict performance.

The average time to first interaction with the tiles was 9.48 seconds (*SD* = 8.49) in the trials in the interactive condition and 15.04 seconds (*SD* = 14.64) in the shuffle condition. A linear mixed model with condition as a fixed effect and anagram and participant as crossed random effects indicates that condition was a significant predictor of the time to first interaction, *F*(1,55.8)= 19.7, *p* <.001. We were interested in whether adding time to first interaction improved the model for success outlined above, the generalised mixed models suggest some relationship, but this was not significant one, *χ*² = 3.66, *p* =.056. However, it is worth noting that the borderline nature of the results here may suggest that the time to first engagement with the tiles would be worth investigating further.

The proportion of time spent interacting with the tiles was added as a fixed effect to a model with anagram and participants as crossed random effects to assess if this was a significant predictor or performance; the model suggests that it was not, *χ*² = .996, *p* =.318.

**Summary of Quantitative Results**

Our pre-registered hypotheses were not upheld. There was no significant benefit to shuffling with the tiles after impasse on either performance or latency to solution. Rather the opposite was found: being allowed to interact freely with the tiles predicted more success and also a faster solution rate. This was not modified by any behaviour that we tested. Although participants were faster to interact with the tiles in the interactive condition this was not a significant factor when included in the model. The time spent actually rearranging rather than contemplating the tiles was also not a significant predictor of success suggesting that a brute measure of interaction time could not explain the difference in performance.

In the case of the whole data set, the direction of the effect was reversed from what we expected with more trials being solved correctly and faster in the interactive condition. This reversal of effect informed the subsequent qualitative analyses.

**Qualitative Results**

Although the qualitative data were collected alongside the quantitative data and initial in vivo analysis and first round coding was carried out in parallel, the qualitative research questions were informed by the statistical analysis and carried out subsequent to those in line with the sequential explanatory design.

We formed two main qualitative exploratory research questions:

1. Why were participant less successful in the shuffle condition?
2. Why were there such low levels of declared impasse?

**Qualitative Analysis of Audio-Video Data**

Bakeman and Quera (2011) distinguish between two types of analysis of video data, narrative qualitative and systematic. The behavioural coding outlined above reflected systematic video analysis, the analysis in this section reflects a more narrative analysis. Such an analysis of video data differs from interpretative analysis of text-based data because it is informed by a mimetic assumption – that is, that the video data in some respect represent what occurred rather than an interpretation of events (Knoblauch, 2013). The role of video allows the researcher to move to a more purely observational role while still acknowledging the interpretivist nature of the analysis.

The analysis carried out in this section drew on grounded theory method (Charmaz, 2014; Glaser & Strauss, 1967; Urquhart et al., 2009). The first author was the original researcher and so the initial coding process happened in vivo through traditional experimental notebooks, before happening again during the initial sampling (the convergent phase) and then being further refined through the selection of critical cases (the explanatory phase). Constant comparisons were made between the different videos and participants. Theoretical sampling emerged first through the selection of critical cases and later through the reanalysis and revisiting of the video data.

The behavioural coding reported in the quantitative section required the first author to watch each participant in the quantitative data set in close detail. This preliminary review (Heath et al., 2010) resulted in the generation of several initial areas of interest which seemed salient to the whole data set. These aspects covered participant behaviour and also the role of chance, that is the continual flow of environmental fluctuations. This first cycle of coding was descriptive and exploratory supported by analytical memos and yielded three main observations. First, it was noted that participants often missed environmental changes that would be helpful (missed chance); second, that problem solving in this instance advanced incrementally through a combination of heuristics and trial and error (interactivity); and third, that the problem solution was often dependent on an arbitrary choice made early on in the problem-solving trajectory (path dependency).

These observations were then assessed in greater depth through the lens of seven cases. These cases were not selected randomly (see Seawright & Gerring, 2008 for a discussion on selection of cases) but rather the first author kept a series of analytical memos during the first cycle coding phase which informed the opportunistic selection of cases to illustrate the observations. The cases were deemed typical in terms of the observed behaviours both by the first author and another member of the research team who also watched the entire video corpus. The characteristics of the participants of this phase are described in Table 5.

**Table 5**

*Characteristics of the Nested Sample*

**

*a Calculated as a total of the phonemic and semantic fluency. b Assessed using the scale from the Novick and Sherman (2003). c Calculated as SD from the mean of the experimental population.*

This analysis was expanded on through the selected cases and additional observations were generated. There was a recursive nature to the data analysis from initial in vivo observations to behavioural coding to conceptual construction leveraging the method of theoretical sampling. That is, those sections of video who were most relevant in that situation had their video data reanalysed in the light of the emerging theories and different sections were sampled. Video data sets are necessarily rich and so researchers in this field will employ the principle of sequentiality; the whole data set is not analysed as a totality, only key scenes are interpreted and those are done so frame by frame (Knoblauch, 2013). This is closely related to the technique of focused observation in which a particular aspect is identified as the candidate for intense observation (Willig, 2013). Video data often elicits this type of analysis because the very rich level of the data requires the researcher to selectively code in the open coding phase and revisit the data later to resample in light of conceptual developments (Heath et al., 2010). This required the videos to be watched in detail both at normal speed and at a slower (70%) speed. Additionally, at key times, frame by frame analysis was required to fully reveal the micro moments in the problem-solving trajectory (Schubert, 2013). This continued until saturation, that is, until repeated watching yielded no further observations.

The aim of the qualitative phase had always been to elucidate the cognitive processes in more finely grained and interpretative detail but the strong effect in the opposite direction of our hypotheses made such an analysis more urgent. The underlying theoretical assumptions that had driven the hypotheses required examination. The following analysis aims to shed light on the luck in problem solving and the process of insight problem solving more generally.

**The Difficulty of Distinguishing Between Luck and Serendipity**

***Path Dependency***

Problem-solving, particularly insight problem-solving, is a path dependent process. By this we mean the nature of the path taken determines the likelihood of a correct solution. The initial path selected in an insight problem is necessarily done with no vision of the end process (for were the end point clearly visible the problem would be solved analytically). Indeed, it has been suggested that people do not plan even one move ahead in some cases (Ormerod et al., 2013). The idea of path dependency echoes the ideas put forward in criterion of satisfactory progress theory (CSPT; Ormerod et al., 2002). At each stage, the problem space is restricted, and moves are selected which maximise progress towards a goal. What we would like to emphasise is that an initial choice will strongly direct the journey taken.

This seems particularly pertinent to problem solving with anagrams. For example, it is impossible to quantify the difficulty or otherwise of an anagram (Knight & Muncer, 2011) but there are a few things that seem to make it harder. One is the similarity to a different, non - target word. Thus, if the problem solver sees a collection of letters which trigger a recognisable word (e.g., **LA** in **MYLAD**; answer **MADLY**), CSPT would predict that she uses that as a starting point on the problem-solving path (e.g., exploring **LAD**). If this bigram or trigram were unhelpful, the problem solver would be less likely to solve the problem and more likely to explore unproductive paths.

The close data analysis suggested that this initial problem-solving move can happen arbitrarily. This can happen in the shuffle conditions where the change in array is likely to force the problem solver in a direction decided purely randomly: A fortunate tumble of the letter tiles would qualify as luck, that is as random chance which directs the problem solver in the correct direction, that which points the problem solver down the wrong path would be unlucky. Take for example, P32: her third shuffle trial required her to unscramble *VANIE* to find **NAÏVE**. The first shuffle yields **EIVAN** which is close to **EVIAN** (a proper noun hence disallowed), her second **VNIAE** which is not dissimilar to the original layout and unlikely to be of help with no new bigrams, the final shuffle again gives her an unhelpful array - **INEAV** - where no letters are in the correct position. Contrast this with P59 who was aiming to unscramble **MYKOS** and yielded **SKMOY** on the second shuffle – the **S** and the **Y** (and also the helpful bigram of **MO**)anchoring the answer of **SMOKY** which was indeed announced at the same time as the last tile fell indicating that it provided the answer for the participant. Similarly, when looking to unscramble **MENGO**, the same participant yielded **GNEOM** again on the second shuffle and again announced the answer at the same time as the last tile even after announcing impasse earlier. Thus, while neither shuffle produced the full word, it did produce an array which made the step to the full solution easy to see. However, it is notable that there were few examples of this ‘lucky’ shuffle which means that the quantitative results may be driven by a particularly unlucky set of shuffles.

However, this path dependency—the importance and yet arbitrary nature of that initial choice—can also be seen in the interactive conditions. P10 for example (solving the anagram **GNOME**) creates the four letters **MONE** and is left with the **G**. Observation of other problems solvers (e.g., P32) suggests that the **GM** bigram can be a pivot which leads to the solution of the problem (the **M** and the **N** presumably being orthographically similar). Here P10 makes the decision to move the **G** to the end of the array creating **MONEG** rather than **GMONE**. This decision opens up a path away from the solution; there is a vexing contingency to this moment. Selecting the correct path from a multitude of options suggests a level of unexpected omniscience.

With the same word, P36 forms the letters **GNME** across the tray and moves the letter **O** last. She places it between the **G** and the **N**, breaking a helpful bigram and frustrating her solution. Similarly, P62 is attempting to solve the anagram **PROBE**. She has formed the word **ROBE** and is left with the letter **P** which she places in the centre of the array, had it been placed at the start of the four letters the word would have become obvious. Take for example the same participant when faced with the word **MILKY** as illustrated by Table 6. Prior to this, the participant has spread the tiles out and is now reforming them in a line. She hovers between **K** and **I** and decides on the **I** which dragged down to the **M** now has **MIL** across the bottom line, that is the line of interest. She now has to decide where to place the **K**, if she places it at the end of the **MIL** trigram it will spell **MILK** and most likely trigger a solution. Instead, she places it at the inviting gap between the **I** and the **L** – knocking the **L** out of place in the process, a move which she reinforces by tidying it further away. That decision means that she is now looking at a tempting but unhelpful trigram of **MIK** rather than the more useful **MILK**.

**Table 6**

*P62: A Demonstration of the Contingency of Solution Choices*.

|  |  |  |
| --- | --- | --- |
| Time (s:ms from start of problem) | Description | Resulting Array |
|  |  |  |
| 13.597 | **M** moved to the bottom left |  |
|  |  |  |
| 14. 132 | **I** moved down to join the **M** |  |
|  |  |  |
| 15.379 | **K** moved around the outside of the array and is moved into the space knocking the **L** |  |
|  |  |  |
| 16.872 | **L** is moved to one side |  |
|  |  |  |
| 17.234 | **Y** moved down to fill the space |  |

It is important to note that these decisions are made quickly, and do not appear to be mediated by strategic thinking. The series of movements illustrated in Table 6 took place over the course of only 4 seconds. Equally, the effect of either luckiness or unluckiness is only able to be assessed by an observer who already knows the outcome. From the vantage point of the omniscient observer the moves are clumsily close to an answer[[5]](#footnote-5). However, the arbitrary nature of that choice first supports theories that suggest a lack of look ahead in problem solving and second, questions models that attribute all epistemic credit to the person’s deliberate strategy when the moves can themselves be contingent on nothing more than a lucky hunch. It is easy to trace moments when unplanned decisions lead to unfortunate pathways, it is less easy to distinguish between equally frivolous decisions which lead to a successful outcome and those which are the result of sober planning.

This suggests that our initial clear division between luck and serendipity was misguided and placed too much emphasis on an intentional agent who moved the tiles strategically in the interactive condition. Rather we can see here that the role of luck is dissipated, and that the agent is supported in performance by arbitrary decisions that later shore up good or bad fortune for her.

***Microserendipity.*** There were several moments of microserendipity, that is when moves by the problem solver trigger unexpected configurations in the array which are then noticed and used to direct the problem-solving trajectory. For example, P10 when solving **MILKY** first spells out the word **MILK** and once this is formed, she notices the left-over **Y** and uses it to create the word **MILKY**. Her voice rises in pitch in excitement and the move which attaches the Y is faster than the previous slightly laboured spelling out. The solution is triggered by a change in the array rather than a preformed plan. P32 has a similar moment when the solution is triggered by an unplanned change in the array as demonstrated by Table 7. Here the participant is solving the anagram **GNOME**. She moves the **G** and the **E** out of the line (moving the middle tiles is a common strategy) with the **G** going above the array and the **E** going below, she then moves to close the space between the **M** and the **O** which means the **M** is directly lined under the **G**. This **GM** bigram appears to lead to the rapid solution, tested first and announced 6 and a half seconds later.

**Table 7**

*P32: A Moment of Microserendipity*

|  |  |  |
| --- | --- | --- |
| Time (s:ms from start of problem) | Description | Resulting Array |
|  |  |  |
| 46.661 | Letters lined up again |  |
|  |  |  |
| 49.308 | **G** and **E** moved to either side |  |
|  |  |  |
| 49.912 | Participant tidies up the array. Considers this with hands away from the tiles |  |
|  |  |  |
| 51.888 | **N** is now moved around from top right to bottom left. |  |
|  |  |  |
| 53.088 | **G** and **N** are placed together |  |
|  |  |  |
| 54.688 | **O** placed in the correct space |  |
|  |  |  |
| 56.208 | Word formed and announced |  |

While we could outline further examples, the most striking thing revealed by our granular analysis was that the participants often missed the chance movements created by the shifting environment (see Figure 5 for exemplars). In the words of P71: ‘I feel like I can see it but I can’t’. This aspect of participant behaviour has already been noticed in passing in other qualitative analysis of problem solving: Steffensen et al. (2016) narrate how the participant in their task solves the problem after an accidental overlap, yet this overlap had accidentally occurred twice previously but was ignored and dismantled. We also noticed this missed chance in a previous study in which participants were asked to make as many words as they could from a series of seven letters (Ross & Vallée-Tourangeau, 2021). This lack of noticing becomes more salient in a task such an anagram task in which there is only one answer and the cost of ignoring the correct answer is greater than the benefit of quick moves in generating many choices.

**Figure 5**

*Moments of Missed Serendipity*

**

**Impasse**

Impasse is currently measured in inconsistent ways: through participant report (Fedor et al., 2015) or through assumed behaviours (Öllinger et al., 2014). In our study we asked people to report their moments of impasse, however the behaviours do not mirror the self -reports. Take P10 being asked to find the word **MINUS** from the initial arrangement of **SINUM**. For the first 8 seconds of the trial, she speaks to herself and rearranges the tiles in an aimless fashion. After 13 seconds, she announces that the answer is **MUSIC** and despite being told that this is incorrect, she forms the word **MUSI** and is disappointed to see there is no **C**. She continues to play with the tiles in an aimless fashion, making the start of **MUSI** several times, muttering until she declares impasse after 74 seconds. There is a tension here in how we understand impasse.

By any definition, the behaviours displayed by the participant here indicate that she is encountering impasse – she offers the wrong answer, she continues to return to an unhelpful representation, her behaviour indicates that she has no coherent plan–but she does not declare impasse as we would understand it. It is also noticeable that impasse does not appear to elicit further problem-solving attempts in all the participants and in fact appears to be a proxy for giving up. For example, P62 after 44 seconds attempting to solve the anagram **SMOKY** announces “I’m stuck on this one” and physically pushes the tray away. Soon after announcing impasse, P10 picks the tray up and tilts it so that the tiles fall randomly indicating that she has completely given up. This indicates that the current way that impasse is theorised which is as an epistemic feeling that the participant wishes to relieve by solving the problem might be missing those cases when the participants alleviate that discomfort by simply giving up. This incongruity can only be revealed through the granular analysis we have conducted here.

**Material Traces**

When problem solving takes place in an environment low in interactivity, the cognitive processes can only be theorised from proxy measures such as response time (although response times are not always recorded, e.g., Chuderski et al, 2020). Problem solving in an interactive environment allows us to trace the problem-solving act as it happens over time and the use of lettered tiles in this instance makes this process clear. Similar to Fleck and Weisberg (2013), we found diverse methods of solving the anagrams suggesting that even relatively similar stimuli cannot be relied upon to elicit the same response. Some of the trials were solved directly–a pop out solution as described by Novick and Sherman (2003) – other after consideration and yet others solved piecemeal with moments of microserendipity crossing with purposeful moves.

Take for example P71 solving the anagram **POWER**. The 10 seconds before a correct solution was announced are described in detail in Table 7.

**Table 7**

*The Process of Solving the Anagram POWER*

|  |  |  |
| --- | --- | --- |
| Time (s:ms from start of problem) | Description | Resulting Array |
|  |  |  |
| 15.241 | Starts to move **P** |  |
|  |  |  |
| 15.741 | **P** dragged to below the main array |  |
|  |  |  |
| 16.375 | **R** moved out of way and placed next to **E** |  |
|  |  |  |
| 17.310 | **O** moved down next to **P** |  |
|  |  |  |
| 18.411 | **P** and **O** grouped together |  |
| 18.411-19.178 | Array is considered, head moved back, fingers off tiles |  |
| 19.178 | **W** is selected for the next move |  |
| 19.779 | **W** is moved next to **P** and **O**. the move knocks the **R** tile and pulls it with the **W** tile in line with **P** and **O** but tilted on its side |  |
|  |  |  |
| 20.646 | P71 tidies the array so that **R** is straight |  |
|  |  |  |
| 21.780 | P71 moves the **E** down so the tiles are in a straight line. |  |
| 21.780 -23.180 | P71 considers the array |  |
| 23.180 | P71 traces her fingers over the **R** and the **E** |  |
| 24.450 | P71 forms the word **POWER** |  |
| 25.404 | P71 announces the correct answer |  |

As can be clearly seen, the answer unfolds through a reciprocal moulding of the array; P71’s gradual realisation of the answer directs the moves she makes which further support her understanding of the answer. While the pivotal event (at 19.779) seems to be the moving of the **W** tile which knocks into the **R** (a further example of microserendipity) and pulls it with it to create the useful trigger this is not the only important action. Instead, we can see a gradual unfolding of understanding over the course of the 10 seconds facilitated by purposeful moves, moments of consideration and fiddling in coordination with the accidental move of the **R**. Even the understanding that the word is **POWER** requires the participant to check before she makes the announcement, the mental representation of the word not secure until it is also represented externally, the realisation in the mind contingent on the realisation in the world.

**Discussion**

The research reported here aimed initially to establish the role of luck and serendipity in an anagram task. It was hypothesised that accidental rearrangements of the array would break impasse by splitting unhelpful bigrams and suggesting different and unconsidered paths. Rather we found the opposite: The quantitative results suggest that random shuffling is not enough to scaffold problem solving success. Solution rates were significantly higher in the interactive condition and participants solved the problem faster. Additionally, granular narrative qualitative analyses suggest that luck was dissipated across both experimental conditions.

Real world problem solving is a complex mix of internal factors, external intrusions and conscious and non-conscious coupling of the two. Selecting which aspects of the dynamic interplay of people and things to investigate in the laboratory is not an easy task. Problem solving will always be a richly contextual activity and adding the variable of random environmental fluctuations further complicates matters. However, taking the material and external world into consideration is important if we wish to scale up the findings in this field from laboratory findings to those which occur outside that environment.

First, while the extent to which interactivity can augment cognitive performance and the implications for the ontological locus of cognition are not uninteresting or unimportant questions, our behavioural analysis points to something more fundamental: Every participant opted to make use of the material world in at least one trial. Whether cognition does or does not *require* the external world is less important than noting that it frequently involves it (see also Clark, 2010). For example, work with expert designers suggests that aggregate measures of design output are the same when blindfolded or allowed to sketch (Bilda et al., 2006) but it is clear designers do sketch and draw in their everyday life and they certainly do not undertake their design process blindfolded. Taking aggregate measures of what *can* be done is possibly not as useful as assessing what *is* done if we wish to understand situated problem solving. The evidence that every participant made use of the scaffolding afforded by the lettered tiles on at least one trial suggests that whether or not it has an augmentative effect, problem solving that ignores the role of the environment is a poor shadow of how people think. In addition, chance and unintended action constitutes a large part of that environment and, as a consequence, it invites a reconceptualization of agency in problem solving and creativity.

However, the efficiency with which participants interacted with the environment varied. For one, there was no relationship between a correct solution and the time spent moving the tiles. This suggests that environmental engagement is not a panacea and nor is it consistently and linearly augmentative in a one-solution task of this type. Work from the first wave of research in interactivity and cognitive offloading suggested that interaction with external representations would necessarily augment problem solving performance as the processing costs were distributed across a large computational space, but the analysis we have drawn out in the qualitative section suggests reasons for the inconclusive data that are sometimes reported. The solution pathways in non-analytical tasks which are initially useful may be reinforced and steer a problem solver off track (Ross & F.Vallée-Tourangeau, forthcoming; F. Vallée-Tourangeau et al., 2020). It may be that interactivity functions as a form of affective and motivational support which is why the feelings of impasse were significantly lower in the interactive than the shuffle condition while the declarations were not. This was an unexpected disconnect and requires further investigation. Furthermore, there was a curious lack of attention to helpful changes in the array. The notion that the environment can yield a solution and that this is unproblematically recognised and taken up would appear to be naïve. The act of noticing appears to be a crucial component of the interaction with chance and deserves further investigation.

**Disentangling Luck and Serendipity**

Given the anecdotal evidence about the role of luck and accident in problem solving, the current study was designed to test how important random material changes were to the array. This was motivated by inconsistencies in the literature around how an accident is produced and also the role of chance in “data-driven restructuring” (Fleck & Weisberg, 2013). The quantitative phase of the current study tested the hypothesis that luck is enough to drive problem solving and support restructuring. It was hypothesised that a task such as an anagram would be particularly likely to be supported by a random generation of hints through blind shuffling of the letter tiles. The data here suggest that the opposite is true: Interacting with the tiles led to a higher overall solution rate and faster solutions than luck and mental effort alone. This means that claims such as those of Chuderski et al. (2020) that problem solutions may be arrived at by accident need to be more carefully moderated.

There is an ongoing debate in the field of luck and serendipity over the extent to which epistemic credit should be awarded for things attributable to accident (Arfini et al., 2018; Sand, 2020). The overall evidence from the experimental manipulation was that luck is not enough to spark problem solving, rather an active interaction and implementation of that luck is required: Chance has to be enacted. This has implications for the growing field of chance discovery. However, it is not clear to what extent we have done more than laid the foundations for a more systematic study of the role of luck. The nature of the experimental design meant that we could not control the anagram presentation and it may be that we encountered a particularly unlucky set of participants whose shuffles did not do as theorised. This however underlines the relational nature of luck.

Additionally, it is worth noting that at the highly granular level there seems to be an unclear division between luck and serendipity in this task. No one person threw the correct answer in a shuffle. Indeed, the number of missed serendipity moments suggest that even if the world throws an answer it is not guaranteed that this will be noticed by the problem solver and become significant. We characterise this as a form of epistemic blindness that shares a similarity with inattentional blindness (Simon & Chabris, 1999). Therefore, by mapping the moments when chance is not enacted, our data suggest it is unclear to what extent pure, non-agentic luck is even possible. If it is necessary for a concept of luck to bear some significance to the problem solver, it may be that the act of bestowing significance on an event involves a form of agency. Further theoretical and empirical refinements become necessary.

The highly granular analysis provided additional evidence for idiosyncratic problem-solving methods both between and within subjects. This observation supports previous evidence from other more detailed analysis (Bilalić et al., 2019; Cushen & Wiley, 2012; Steffensen et al., 2016). Participants solved some anagrams quickly with little obvious effort, some slowly and incrementally and some not at all. This suggests that simply aggregating data across participants may hide important patterns. The growing evidence for diverse methods in problem solving is further complicated in insight research by the number of different categories of problems (Batchelder & Alexander, 2012). It may be that large scale studies which both assume the problem-solving process from the problem and collapse across diverse problem types may only provide a limited view of the complexity of insight problem solving. We propose further micro analytical studies to assess process as much as outcome and so aid theoretical development.

**Impasse**

Ohlsson defines impasse as “a mental state in which all problem solving has come to a halt” and goes on to suggest ‘behaviourally, impasses are characterised by the cessation of problem-solving activity (1992, p. 4). Initially, impasse seemed vital to our model of luck driven problem solving and it was this which led to our framing of the preregistration as we did. After all, it is not implausible that when a participant exhausts the problem space they will benefit most from a random hint. However, our results were inconclusive, and the qualitative analysis led us to lose faith in the mid task self-report measure of impasse. The triangulation of data here taken from mid task self-report, post-task self-report and systematic observation of video data indicates two things: First, that the linear model of impasse followed by restructuring also hinges on motivation which supports the data from Fedor et al. (2015) who found a significant effect of impasse on success - those who declared impasse were much more likely to fail. A declaration of impasse perhaps was not a feeling of being stuck but rather tantamount to giving up. Second, that impasse may not be consciously experienced by an experimental participant while in the ‘flow’ of problem-solving supporting Fleck and Weisberg’s (2013) observation that restructuring can occur without conscious impasse. It is also worth noting in this respect that while there was a significant difference between the conditions in post task feelings of “being stuck” this did not replicate in the self-reporting of impasse during the task suggesting that each are measuring different phenomenon. Externalising cognitive processes appears to support the affective dimension of problem solving (Bilda et al., 2006; Guthrie & Vallée-Tourangeau, 2015). Perhaps the slowness to declare impasse in a materially rich situation is because the very material situation does not invite it - the menu of problem-solving options is far richer and more inviting in an interactive or even a shuffle condition than non-interactive environment - or perhaps it was also because the very definition of impasse may be conceptually problematic.

There is very little research on the quality or otherwise of the epistemic feeling of impasse in contrast to the feeling of insight, but it seems likely that impasse is experienced differently—either as motivator to problem search or leading someone to giving up entirely—with differing implications for problem-solving success. We suggest sustained attention to behavioural and self-report data may help to clarify some of the inconsistencies in this area.

**Kineoetic Analysis**

If we approach research from the position of interactivity (i.e., problem solving as a process that unfolds in the coordination of people and things) then the particular model of mixed method research we have followed here provides empirical evidence for the nature of that coupling and the manner in which the reciprocal, looping process of thought unfolds. Furthermore, such a method supports more detailed understanding of insight problem solving alongside the use of eye tracking measures (Bilalić et al., 2019) and think aloud protocols (Fleck & Weisberg, 2013). As Christensen and Friis-Olivarius (2020) argue, the use of movable objects allows us to track thoughts through action without the need for expensive technology.

The method of research that we present here involves the combination of three phases: an experimental phase assessing aggregated effects of experimental manipulations, supported by inductive behavioural coding of a full data set and finally close analysis of a selected set of case studies. This form of analysis is only made possible when the participants are invited to interact with instrumentalised movable pieces and thought is moved outside of the head. This allows researchers to adopt a qualitative analysis which focuses on changes in objects, rather than the conjectured changes in the participants mental representation of these objects, what we would call a kinenoetic analysis (see Ross & Vallée-Tourangeau, in press).

Such an analysis is predicated on the underlying position that thoughts and objects are mutually co-constituted and that action on a material object is the antecedent of knowledge or meaning, not as its consequent and that traditional notions of intentionality and agency focus to much on the internal untraceable cognitive trajectories. This detailed attention to environmental chance and complexity we have demonstrated here which moves the focus away from the psychometric properties of the person, has important implications for the design of environments to maximise problem solving success. This type of deeper and more finely grained analysis becomes particularly important when we have little control over aspects of the experimental set up. This is particularly salient each time we embed participants in a complex, materially rich world. There is a paucity of qualitative research in cognitive psychology (Ball & Ormerod, 2017; Steffensen, 2016) and we argue that the object tracing method we employ here is essential to help to understand complex cognitive processes.

**References**

Bakeman, R., & Gottman, J. (2009). *Observing interaction: An introduction to sequential analysis*. Cambridge University Press.

Bakeman, R., & Quera, V. (2011). *Sequential analysis and observational methods for the behavioral sciences*. Cambridge University Press.

Ball, L. J., & Ormerod, T. C. (2017). Cognitive psychology. In C. Willig & W. Stainton Rogers (Eds.), *The Sage handbook of qualitative research in psychology..* (pp. 572–588). SAGE.

Bardone, E., & Magnani, L. (2013). Turning down a chance: An argument from simplicity. *Advances in Chance Discovery*, 19–31. https://doi.org/10.1007/978-3-642-30114-8

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bowden, E., 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

Carter, J. A., & Peterson, M. (2017). The modal account of luck revisited. *Synthese*, *194*(6), 2175–2184. https://doi.org/10.1007/s11229-016-1047-7

Charmaz, K. (2014). *Constructing grounded theory* (2nd edition). Sage.

Christensen, B. T., & Friis-Olivarius, M. (2020). How do initial ideas evolve into final ones? Exploring the cognitive size, structure and life of ideas using sticky notes. In B. T. Christensen, K. Halskov, & C. N. Klokmose (Eds.), *Sticky Creativity: Post-it note cognition, computers and design.* (pp. 53–75). Elsevier. https://doi.org/10.1016/B978-0-12-816566-9.00003-3

Chuderski, A., Jastrzębski, J., & Kucwaj, H. (2020). How physical interaction with insight problems affects solution rates, hint use, and cognitive load. *British Journal of Psychology*. https://doi.org/10.1111/bjop.12442

Clark, A. (2010). Material surrogacy and the supernatural: Reflections on the role of artefacts in ‘off-line’ cognition. In L. Malafouris, & C. Renfrew (Eds.), *The cognitive life to things* (pp. 23–28). Cambridge: McDonald Institute for Archaeological Research.

Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, *58*(1), 7–19. https://doi.org/doi:10.1093/analys/58.1.7

Coffman, E. J. (2009). Does luck exclude control? *Australasian Journal of Philosophy*, *87*(3), 499–504. https://doi.org/10.1080/00048400802674677

Copeland, S. M. (2019). On serendipity in science: Discovery at the intersection of chance and wisdom. *Synthese*, *196*, 2385–2406. https://doi.org/10.1007/s11229-017-1544-3

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (Third Edition). SAGE.

Csikszentmihalyi, M. (1996). *Creativity: The psychology of disovery and invention*. Harper Collins.

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

Ellis, J. J., Glaholt, M. G., & Reingold, E. M. (2011). Eye movements reveal solution knowledge prior to insight. *Consciousness and Cognition*, *20*(3), 768–776. https://doi.org/10.1016/j.concog.2010.12.007

Fedor, A., Szathmáry, E., & Öllinger, M. (2015). Problem solving stages in the five square problem. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.01050

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

Friedlander, K. J., & Fine, P. A. (2018). “The penny drops”: Investigating insight through the medium of cryptic crosswords. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.00904

Gallucci, M. (2019). *GAMLj: General analyses for linear models [jamovi module]*. https://gamlj.github.io

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

Gilhooly, K. J., & Webb, M. E. (2018). Working memory in insight problem solving. In F. Vallée-Tourangeau (Ed.), *Insight: On the origins of new ideas.* (pp. 105–120). Routledge.

Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory: Strategies for Qualitative Research*. Aldine Publishing Company.

Griffith, M. (2010). Why agent caused actions are not lucky. *North American Philosophical Publications*, *47*(1), 43–56.

Heersmink, R. (2013). A taxonomy of cognitive artifacts: function, information, and categories. *Review of philosophy and psychology*, *4*(3), 465-481. https://doi.org/10/1007/s13164-013-0148-1

Henok, N., Vallée-Tourangeau, F., & Vallée-Tourangeau, G. (2020). Incubation and interactivity in insight problem solving. *Psychological Research*, *84*(1), 128–139. https://doi.org/10.1007/s00426-018-0992-9

Knoblauch, H. (2013). Videography: Focused ethnography and video analysis. In H. Knoblauch, H.-G. Soeffner, J. Raab, & B. Schnettler (Eds.), *Video analysis: Methodology and methods. Qualitative audiovisual data analysis in sociology*. Peter Lang GmbH, Internationaler Verlag der Wissenschaften.

Kounios, J., Fleck, J. I., Green, D. L., Payne, L., Stevenson, J. L., Bowden, E. M., & Jung-Beeman, M. (2009). The origins of insight in resting-state brain activity. *Neuropsychologia*, *46*(1), 281–291. https://doi.org/10.1016/j.neuropsychologia.2007.07.013

Maglio, P. P., Matlock, T., Raphaely, D., Chernicky, B., & Kirsh, D. (1999). Interactive skill in Scrabble. In M. Hahn & S. C. Stoness (Eds.), *Proceedings of the 21st Annual Conference of the Cognitive Science Society* (pp. 326–330). Lawrence Erlbaum Associates, Publishers.

Makri, S., & Blandford, A. (2012). Coming across information serendipitously – Part 1: A process model. *Journal of Documentation*, *68*, 684–705. https://doi.org/10.1108/00220411211256030

Ohlsson, S. (2018). The dialectic between routine and creative cognition. In F. Vallée-Tourangeau (Ed.), *Insight: On the origins of new ideas* (pp. 8–28).

Pritchard, D. (2014). The modal account of luck. *Metaphilosophy*, *45*(4–5), 594–619. https://doi.org/10.1111/meta.12103

Rescher, N. (1995). *Luck: The brilliant randomness of everyday life*. Farrar, Straus and Giroux.

Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences*, *20*(9), 676–688. https://doi.org/10.1016/j.tics.2016.07.002

Ross, W. (forthcoming a). Heterocalar serendipity and the importance of accidents. In W. Ross & S. Copeland (Eds.), *The Art of Serendipity*. Palgrave MacMillan.

Ross, W. (forthcoming b). Serendipity and creative cognition: Towards a systematic consideration of the serendipitous genesis of a new idea. In S. M. Copeland, M. Sand, & W. Ross (Eds.), *Serendipity Science*. Springer.

Ross, W. (2020). Serendipity. In V. P. Glǎveanu (Ed.), *The Palgrave encylopedia of the possible*. Palgrave MacMillan.

Ross, W., & Vallée-Tourangeau, F. (under review a). Kinenoetic analysis: Unveiling the material traces of insight. *Methods in Psychology*.

Ross, W., & Vallée-Tourangeau, F. (forthcoming). Rewilding cognition: Complex dynamics in open experimental systems. *Journal of Trial and Error*.

Ross, W., & Vallée-Tourangeau, F. (2021). Catch that word: Interactivity, serendipity and verbal fluency in a word production task. *Psychological Research*, *85*(2), 842-856. https://doi.org/10.1007/s00426-019-01279-y

Ross, W., & Vallée-Tourangeau, F. (2020). Microserendipity in the creative process. *Journal of Creative Behavior*.

Rubin, V. L., Burkell, J., & Quan-Haase, A. (2011). Facets of serendipity in everyday chance encounters: A grounded theory approach to blog analysis. *Information Research*, *16*. https://doi.org/[Available at http://InformationR.net/ir/16-3/paper488.html]

Saldaña, J. (2016). *The coding manual for qualitative researchers*. SAGE.

Sand, M. (2020). Did Alexander Fleming deserve the Nobel prize? *Science and Engineering Ethics*, *26*(2), 899–919. https://doi.org/10.1007/s11948-019-00149-5

Schubert, C. (2013). Video analysis of practice and the practice of video analysis. In H. Knoblauch, H.-G. Soeffner, J. Raab, & B. Schnettler (Eds.), *Video analysis: Methodology and methods qualitative audiovisual data analysis in sociology*. Peter Lang GmbH, Internationaler Verlag der Wissenschaften. <http://nbn-resolving.de/urn:nbn:de:101:1-201311054294>

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception*, *28*, 1059–1074.

Steffensen, S. V. (2016). Cognitive probatonics: Towards an ecological psychology of cognitive particulars. *New Ideas in Psychology*, *42*, 29–38. https://doi.org/10.1016/j.newideapsych.2015.07.003

Steffensen, S. V. (2017). Human interactivity: Problem-solving, solution-probing, and verbal patterns in the wild. In S. J. Cowley & F. Vallée-Tourangeau (Eds.), *Cognition beyond the brain* (pp. 85–113). Springer. https://doi.org/10.1007/978-3-319-49115-8

Steffensen, S. V., & Vallée-Tourangeau, F. (2018). An ecological perspective on insight problem solving. In F. Vallée-Tourangeau (Ed.), *Insight: On the origins of new ideas.* (pp. 169–190). Routledge.

Steffensen, S. V., Vallée-Tourangeau, F., & Vallée-Tourangeau, G. (2016). Cognitive events in a problem-solving task: A qualitative method for investigating interactivity in the 17 Animals problem. *Journal of Cognitive Psychology*, *28*(1), 79–105. https://doi.org/10.1080/20445911.2015.1095193

Terai, H., & Miwa, K. (2013). A chance favors a prepared mind: Chance discovery from cognitive psychology. *Advances in Chance Discovery*, 33–48. https://doi.org/10.1007/978-3-642-30114-8\_3

The jamovi project. (2020). *Jamovi* (1.2) [Computer software]. https://www.jamovi.org

Urquhart, C., Lehmann, H., & Myers, M. D. (2009). Putting the ‘theory’ back into grounded theory: Guidelines for grounded theory studies in information systems: Guidelines for grounded theory studies in information systems. *Information Systems Journal*, *20*(4), 357–381. https://doi.org/10.1111/j.1365-2575.2009.00328.x

Vallée‐Tourangeau, F. (2014). Insight, interactivity and materiality. *Pragmatics & Cognition*, *22*(1), 27–44. https://doi.org/10.1075/pc.22.1.02val

Vallée-Tourangeau, F. (2018). Introduction. In *Insight: On the origins of new ideas* (pp. 1–7). Routledge.

Vallée-Tourangeau, F., Ross, W., Ruffatto Rech, R., & Vallée-Tourangeau, G. (2020). Insight as discovery. *Journal of Cognitive Psychology*, 1–20. https://doi.org/10.1080/20445911.2020.1822367

Vallée-Tourangeau, F., & Vallée-Tourangeau, G. (2020). Mapping systemic resources in problem solving. *New Ideas in Psychology*. https://doi.org/10.1016/j.newideapsych.2020.100812

Vallée-Tourangeau, G., & Vallée-Tourangeau, F. (2017). Cognition beyond the classical information processing model: Cognitive interactivity and the systemic thinking model (SysTM). In S. J. Cowley & F. Vallée-Tourangeau (Eds.), *Cognition beyond the brain* (pp. 133–155). Springer. https://doi.org/10.1007/978-3-319-49115-8

Valueva, E., Lapteva, E., & Ushakov, D. (2016). Aha-cueing in problem solving. *Learning and Individual Differences*, *52*, 204–208. https://doi.org/10.1016/j.lindif.2016.02.003

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*. https://doi.org/10.3389/fpsyg.2016.01424

Webb, M. E., Little, D. R., & Cropper, Simon. J. (2018). Once more with feeling: Normative data for the aha experience in insight and noninsight problems. *Behavior Research Methods*, *50*(5), 2035–2056. https://doi.org/10.3758/s13428-017-0972-9

Willig, C. (2013). *Introducing qualitative research in psychology* (3. ed). Open Univ. Press.

1. The cheap necklace problem requires participants to make a complete closed loop (necklace) out of 12 links of chain, with the starting point being 4 smaller, 3-link chains. A cost constraint (2 cents to break a link, and 3 cents to join a link) is imposed. The correct solution involves breaking all three links of one of the 3-link chains, and using the individual links to connect the three remaining 3-link chains together [↑](#footnote-ref-1)
2. How to distribute 17 animals in four enclosures such that there are an odd number of animals in each enclosure. The solution requires the participants to overlap sets, permitting the double counting of some of the animals. [↑](#footnote-ref-2)
3. ‘A dealer in antique coins got an offer to buy a beautiful bronze coin. The coin had an emperor’s head on one side and the date 544BC stamped on the other. The dealer examined the coin, but instead of buying it, he called the police. Why?’ Perkins (1981), cited in Fleck and Weisberg (2013) [↑](#footnote-ref-3)
4. using the RANDINT function in Excel [↑](#footnote-ref-4)
5. As suggested by a reviewer the idea of “closeness” here requires a lexical realism which may not be sustained on closer analysis. A word either is a word when recognised by a participant or is not and therefore the status of these letter strings is relational to the person observing them. This is also problematic when we are assessing the moments of “missed serendipity” because they are only missed in relation to a problem solver who already knows the answer. We raised similar concerns in Ross and Vallée-Tourangeau (2021). This reveals the complexities of assessing process when the researchers know the answer but the participants do not. [↑](#footnote-ref-5)