Cognition 246 (2024) 105746

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Impasse-Driven problem solving: The multidimensional nature of feeling stuck

Wendy Ross [a,*], Selene Arfini [b]

[a] London Metropolitan University, UK
[b] University of Pavia, Italy

ARTICLE INFO

Keywords:
Impasse
Insight
Problem-solving
Feeling of being stuck
Feeling of insight

ABSTRACT

This study reports findings across four preregistered experiments (total $N = 856$) that establish the multidimensional nature of impasse and resolve two paradoxes implicit in the problem-solving literature: how a state of impasse can be at once necessary to solve a problem with insight yet also have appear to have a catastrophic effect on solution rates, and why individuals such as problem-solving and gaming enthusiasts seem to seek out this apparently aversive state. We introduce a new way of measuring impasse based on qualitative reports and subsequently confirmed through quantitative analysis that exploits two aspects of impasse: its dynamic and unstable nature (it can be resolved or unresolved) and its multidimensionality in terms of feelings of cognitive speediness, motivation, and affect. The feeling of being stuck varies between resolved and unresolved impasse in terms of feelings of speediness and positive affect, but not motivation, which remains constant. We demonstrate that the feeling of insight can be reliably elicited by experiencing and resolving impasse but also in the absence of impasse, which suggests that there is more than one path to an insight experience. This adds depths to current proposals of the cognitive mechanisms underlying both insight problem-solving and impasse. Our findings are robust across a range of problem types. The novel conception of impasse in this paper as dynamic and multidimensional has implications for theories of insight problem solving, and also wider implications for understanding how impasse can be resolved across different domains such as education and design.

A problem arises when someone is confronted by a situation that they cannot resolve in the moment; it requires effortful thinking (Gilhooly, 2019; Ormerod, MacGregor, & Chronicle, 2002). Once the problem solver has run out of ideas and is unable to move forward from they are considered to be in a state of impasse. More informally, they are 'stuck'. Without being stuck there is arguably no problem, but rather a smooth progression from the starting state to the goal state. The problem is solved when the solver becomes 'unstuck', either by selecting the correct tools and ways of applying them from an existing cognitive toolbox, or by recruiting tools from outside their own resources.

Problem solving has been considered a core higher cognitive process since the publication of seminal texts such as Newell, Shaw, and Simon (1958), it has been described as one of the first and main topics of interest in psychology (Duncker, 1945), and was also one of the first subjects of investigation for the philosophical research on cognition (Polanyi, 1957). That is why Dominowski and Bourne (1994), in trying to recap the history of "thinking and problem solving," began by admitting that "no matter where you start there are always obvious antecedents" and, they unfortunately needed to "start in the relatively recent past, *just two to three hundred years ago*" to pick an arbitrary beginning (p. 1). Today, problem-solving represents a test bed for different cognitive theories. The so-called toy problems (well-structured and knowledge-lean) are used to shed light on various complex cognitive processes – indeed, they have been called the "fruit-flies" of cognitive research (Ball & Litchfield, 2013, p. 117). In this paper, we draw on this tradition to carefully examine the role of emotions and motivation in cognitive processes through the lens of the state of impasse or 'feeling stuck'.

In both psychological and folk understanding, feeling stuck is commonly ascribed to a knowledge gap - often accompanied by the refrain 'I don't *know* what to do'. In a straight-forward model of problem solving, the problem solver does not have the capacity to successfully solve the problem because they need more knowledge (or, in certain cases, skills). However, in some cases, the state of impasse is "unmerited" (Ohlsson, 1992, p. 4) because the problem solver has all the knowledge and skills that they require to solve the problem but is unable

* Corresponding author at: Psychology Department, London Metropolitan University, 166-220 Holloway Road, N7 8DB, UK.
E-mail address: w.ross@londonmet.ac.uk (W. Ross).

https://doi.org/10.1016/j.cognition.2024.105746
Received 5 July 2023; Received in revised form 13 February 2024; Accepted 16 February 2024
Available online 22 February 2024

to access that knowledge in the moment either because the solver does not know how to deploy the information they have or because they cannot see which information is relevant. This can lead to the common ‘tip of the tongue’ feeling, occurring when the solver feels they know the answer to a problem, but they are also unable to access that knowledge. This also suggests that impasse is more complex than a simple form of ignorance or lack of knowledge.

Unmerited impasse differs from situations in which problem solving is thwarted by limited cognitive resources; rather, it refers to a form of being stuck where the problem solver has all the resources necessary to solve the solution. On being told the answer, problem-solvers can experience a range of emotions, from positive feelings of relief (Danek, Fraps, von Maller, Grothe, & Ollinger, 2014) to frustration at being told something they quickly realise that they already knew (Hill & Kemp, 2018). This has implications for a straightforward account of problem-solving based on traditional rational mechanisms which leads from ignorance to understanding through simple steps. Instead, there seems to be a failure of traditional metacognitive monitoring of existing knowledge (Thompson, Prowse Turner, & Pennycook, 2011) and the participant does not know what they already know. Unmerited impasse should, therefore, be of great interest to those who study creative cognition and the status of knowledge more generally as well as to those who investigate the interaction of cognition, emotion, and motivation. Knowing how this state is resolved is key to understanding how new ideas are generated from existing knowledge. In addition, being in impasse appears to be a cognitive state that orients the problem-solver outwards towards hints and other forms of data-driven restructuring (Fleck & Weisberg, 2013; Ormerod et al., 2002; Ross & Vallée-Tourangeau, 2021a); therefore, understanding this state may support our understanding of moments of inspiration and breakthrough thinking (Perkins, 2001; Ross & Vallée-Tourangeau, 2021c; Weisberg, 2015).

1. A return to impasse-driven problem-solving

Unmerited impasse is commonly investigated using insight problems. These are problems designed (in their traditional form, see Webb, Little, & Cropper, 2016) to lead the problem solver down an unhelpful *cul de sac* leading to a dead end. There are two main competing theories for how these forms of problems are resolved: via a sudden change in problem representation (e.g., Representational Change Theory; RCT, Knoblich, Ohlsson, Haider, & Rhenius, 1999) or via a change in strategy wrought by a failure of maximising moves (e.g., Criterion of Satisfactory Progress Theory; CSPT, MacGregor, Ormerod, & Chronicle, 2001). Both approaches *require* an impasse to precipitate changes that can solve the problem. However, to date, impasse has been understood as a unidimensional “complete cessation of problem-solving activity” (Ohlsson, 1992). This has led to a complex paradox that remains underexplored by current theories: How can a problem-solving model require an element that is defined as the cessation of problem-solving?

The current solution to this paradox is to downplay the importance of impasse in this form of problem- solving. Instead, the focus shifts to the outcome of problem solution rather than its cause. This shift is informed by empirical work that has shed doubt on the *necessary* link between the *feeling* of impasse and the *feeling* of insight at the time of problem solution. For example, when participants were asked about their feelings of impasse through post-task qualitative reports (Danek et al., 2014) or think-aloud protocols during the task itself (Fleck & Weisberg, 2013), there was no consistent evidence of impasse arising prior to a problem solution. Using phenomenological reports of feelings of insight, Webb et al. (2016) showed a negative correlation between reported insight and reported impasse across a range of problem types. These data suggest that insight can be experienced without a conscious experience of feeling stuck and without clear behavioural markers of impasse. This evidence has led to reduced focus on the experience of impasse as a precipitating condition for solving insight problems. For example, in her model of insight, Danek (2018, p. 54) labels impasse as “possible but not necessary”. Weisberg’s integrated model also argues from data collected in Fleck and Weisberg (2004, 2013) that impasse is not necessary, although it appears necessary for what Weisberg defines as “true insight”[1]

Currently, there is an increased interest in the phenomenology of solving insight problems as a clue to the underlying cognitive processes (Bowden & Jung-Beeman, 2003; Bowden, Jung-Beeman, Fleck, & Kounios, 2005; Danek et al., 2014; Webb et al., 2016). This investigation of the so called “aha” moment, the affective dimension of the insight experience, has replaced previous attempts to elicit the cognitive processes of interest through a special class of problems deliberately structured to lead to an impasse. To some extent, this focus has been spurred by the exigencies of multi-trial experimentation aimed at pinning down neural or physiological correlates of the insight experience (Spiridonov, Loginov, & Ardislamov, 2021; Webb et al., 2016). These problems are more likely to elicit “pop out” insight experiences than more traditional problems (Novick & Sherman, 2003, 2008; Spiridonov et al., 2021). In other words, there has been a shift from problems that elicit impasse to those that elicit insight. This shift has been motivated by a deeper exploration of the *feelings* of insight through think-aloud processes or post-task reports.

Indeed, much of the progress in understanding the cognitive mechanisms underlying insight problem-solving has come from a close examination of the affective components through self-report. For example, the experience of insight may act as a metacognitive marker signalling the truthfulness of an idea (Danek & Salvi, 2020; Laukkonen et al., 2022, although see Ross & Vallée-Tourangeau, 2022) or facilitate memory and learning (Danek, Fraps, von Müller, Grothe, & Öllinger, 2013), and may indeed simply feel nice (Danek et al., 2014). This is a clear illustration of the entanglement of cognition and emotion – in this area of research it seems likely that cognitive and affective processes are interdependent and so continued progress requires an assessment of the two.

However, we propose that focusing on the feeling of insight has led us to neglect the stage of impasse in problem-solving. If a participant is not consciously experiencing impasse, then the nature of a search for a correct problem representation will be qualitatively different, and it is unclear that a similar feeling of insight reflects the same processes. Impasse appears to have cognitive, affective, and motivational aspects, and it may be that the current methods of measuring the negative cognitive aspects of impasse by examining negative affective and motivational aspects are not sufficient. We propose that, much as with research on insight problem-solving, the cognitive aspects of unmerited impasse can be elucidated by a clear focus on the affective aspects.

Given that both the main models of insight problem-solving (RCT and CSPT; Weisberg, 2015) require a failure of the problem solving process – either owing to a faulty initial representation or of the maximisation heuristic – its relative neglect is curious. If there is no failure, it is not clear what would precipitate a shift in representation or heuristic. Chu and MacGregor (2011 p. 120) describe the opposite of insight problem-solving – analytical problem-solving – as having “no sense of being blocked and no sudden flash of illumination”. To borrow their words, current models of insight have moved to a “flash of illumination” which follows no “sense of being blocked”.

This significant cognitive change requires further investigation. Investigating impasse-driven problem-solving separately from insightful processes can support our understanding of how this state is resolved. The research in this paper stems from the foundational assumption that the experience of unmerited impasse is valuable as both an affective and a cognitive state, of interest not only to researchers in problem-solving and creative cognition but also to those interested in the acquisition of knowledge and the relationship between cognition and emotion. Therefore, we propose a close examination of impasse-driven problem-

[1] There is a circularity in the argument here – true insight is that achieved after impasse, so of course it is necessary.

solving.

2. Epistemic feelings

Epistemic feelings are studied as elements of the emotional spectrum (Arango-Muñoz, 2014) that have metacognitive functions and affect cognitive abilities, (Sousa, 2009), states (Evans, 2008), and even moral decision making (Terpe, 2016). They can be more or less consciously addressed and they have a general positive or negative valence (Sousa, 2009). Despite the research on how these feelings influence how agents address their own cognitive and metacognitive processes (Carruthers, 2017), they are rarely considered in general theories regarding problem-solving processes. The exception is in studies of insight in which the nature of the insight experience has been extensively profiled (see for example Danek et al., 2014).

Despite its role in a wide range of problem solving phenomena, little research has been conducted on the affective nature of the impasse experience. It is clear that an impasse carries a strong negative phenomenological marker (Beeftink, van Eerde, & Rutte, 2008). Indeed, Shen et al. (2019) assume impasse is the "negative experience component" (p.2) of insight problem solving and their findings support Beeftink et al.'s suggestion that impasse is a metacognitive marker to *give up* attempts at problem solution. This again mirrors Ohlsson's suggestion that impasse is a "complete cessation of problem-solving activity", presumably because the state of being in impasse is so unpleasant that giving up is the preferable option. We return to the original paradox: How can a state defined as complete cessation of problem-solving activity lead to problem solution?

Alongside the paradox outlined above, this definition of impasse raises another contradiction: while impasse is considered to be an aversive state throughout the problem-solving research literature, this literature relies on tasks that are similar in structure to tasks commonly enjoyed by problem solvers as a hobby (e.g., Friedlander & Fine, 2018). This suggests that the experience of impasse may be less aversive than currently theorised. Research in gaming literature confirms that games with challenge are fun and engage intrinsic motivation (Przybylski, Rigby, & Ryan, 2010). The notion that *not knowing* is intrinsically aversive is a hangover of linear theories that prioritise knowledge gain. Rather, impasse is likely to be a relational and subjective phenomenon, with different people enjoying different levels of challenge.

In addition, there has often been a mismatch between self-reports of impasse and behaviours that are assumed to indicate impasse. For example, Ross and Vallée-Tourangeau (2021a) tracked behaviour alongside in-task and post-task reports asking participants the extent to which they had experienced impasse. They found that, even within the same participant on the same trial, clear behavioural markers of giving up (such as pushing the experimental material away, huffing and tapping feet impatiently) did not translate into either in-task or post-task reports. This result supports the observations by Danek et al. (2014) and Fleck and Weisberg (2013) that participants did not consciously report experiencing impasse although they did display behaviours that suggested that they were frustrated or had run out of ideas. Fedor, Szathmáry, and Öllinger (2015) found that there was no link between the subjective feeling of being stuck and the behavioural markers they identified. They concluded that behavioural markers are more reliable as markers of impasse, but this undermines the role of the *feeling* of impasse as an important metacognitive marker suggested elsewhere (Beeftink et al., 2008; Shen et al., 2019). This suggests that there is an underlying multidimensionality to the experience of impasse which warrants further investigation.

An explanation for the contradiction between behavioural and subjective reports is that the feeling of impasse may be more complex than initially considered. Research in learning suggests that there may be two types of impasse: "hopeful" and "hopeless" (D'Mello & Graesser, 2012). The first leads to problem resolution and the second leads to failure. Insight problem-solving research currently fails to consider this dual nature of being "stuck", and this limits efforts to understand how to approach and maybe alleviate these states. The theoretical omission of a positive state of impasse is evident in Shen et al. (2019), where instructions to participants made no distinction between impasse leading to success and impasse leading to giving up. A more granular understanding of impasse may resolve some of these difficulties.

In summary, there are two key incoherencies in the literature regarding feelings of impasse. First, it is defined as a complete cessation of problem-solving activity, and the data suggest that it is a metacognitive marker to give up, yet it is also conceptually necessary for problem solution. Second, behavioural markers of impasse typically do not align with self-report measures, suggesting a more complex phenomenon than initially theorised.

The current explanation for these inconsistencies is to downplay the importance of impasse; however, this leads to the paradox outlined above – that the feeling of insight when problem solving is dependent on experiencing an impasse. An alternative hypothesis is that impasse is currently an umbrella term that encompasses phenomena with different affective, cognitive, and motivational profiles. In the experiments we report below, we provide evidence for this alternative hypothesis. We demonstrate that the association of impasse with problem failure is an artifact of the post-task method of measurement going some way to explain the inconsistencies in measurement outlined above. We also show how the state is multidimensional and varies in motivation, affect and feelings of cognitive speed. We further establish that this is similar across problem-types.

3. The current study

The current study was designed to examine participants' experiences of feeling stuck. We conducted four experiments to explore different aspects of the impasse experience.[2] Qualitative methods are unusual in research into cognitive phenomena which are assumed to operate on a subpersonal level and so may not be amenable to qualitative approaches. However, qualitative methods (either formally in terms of observational methods or informally in terms of post task reports) are growing in popularity as a way to understand complex cognitive processes(see Ball & Ormerod, 2017; Danek et al., 2013; Hill & Kemp, 2018; Ross & Vallée-Tourangeau, 2021b; Steffensen, Vallée-Tourangeau, & Vallée-Tourangeau, 2016).

Initially, we selected a form of riddle known as a "stumper" to elicit the feeling of impasse. Stumpers specifically generate a misleading mental model that causes problem solvers to soon run out of ideas and become "stumped". Ross and Vallée-Tourangeau (2022) reported normative data for 25 stumpers which could guide our selection and they have already been reliably used to elicit impasse (Ross, 2021). Subsequently, in the final study we replicated our findings across different task types.

4. Experiment 1

4.1. Method

4.1.1. Participants

A total of 125[3] participants were recruited from Prolific. Co and were paid £2 to complete the study. One participant did not provide

[2] These initial experiments concentrated on three aspects of impasse experience: the embodied nature of the impasse, the importance of metaphors in understanding cognitive experiences, and the multidimensional nature of the impasse. The focus of this study is the final strand, the multidimensional nature of the impasse experience, and the results reported here are selected to support this focus.

[3] The samples sizes here were preregistered but we did not conduct a power calculation because we were conducting exploratory research.

demographic data. The average age of the women ($N = 62$) was 30.93 ($SD = 10.30$) whereas for men it was 31.35 ($SD = 9.57$).

4.2. Materials and measures

4.2.1. Experimental stimuli

Participants were invited to solve a stumper (see Appendix A), originally detailed by Bar-Hillel (2021) and later by Ross and Vallée-Tourangeau (2022). We selected one that was likely to elicit an impasse and yet also had an appropriate level of success. According to Bar-Hillel, the solution rate of the selected riddle is 48% (time not given), and Ross and Vallée-Tourangeau reported a solution rate of 43% in 90s; therefore, it seems likely that it would yield both impasse followed by success and impasse followed by failure. After being given the answer, participants were asked to assess whether they got the answer correct and were able to answer either "Yes," "No" or "Maybe." These answers were hand checked by the research team and those that had the answer "Maybe" were allocated to either correct or incorrect.

4.2.2. Insight

Participants were briefed on what constituted an "aha" experience using the wording from Danek et al. (2014), adapted from Bowden and Jungbeeman (2007). The following instructions were used for these judgments:

> We would also like to know whether you experienced a feeling of insight when you solve each task: A feeling of insight is a kind of "Aha!" characterised by suddenness and obviousness (and often relief!)—like a revelation. In contrast, you experienced no Aha! If the solution occurs to you slowly and stepwise. As an example, imagine a lightbulb that is switched on all at once in contrast to slowly dimming it up. We ask for your subjective rating whether it felt like an Aha! Experience or not, there is no right or wrong answer. Just follow your intuition.

However, we removed the sentence: "You are relatively confident that your solution is correct without having to check it" to avoid artificially inflating the relationship between certainty and "aha" (Danek & Salvi, 2020). Participants were asked to rate whether or not they had experienced the feeling of insight on a binary (yes/no) and they were also asked to rate the strength of the "aha" element of the insight experience. This allowed us to capture both forms of insight – the binary yes or no measure recommended by Bowden and Jungbeeman (2007); see also Laukkonen & Tangen, 2018) and also the more granular assessment recommended by researchers such as Danek, Fraps, von Maller, Grothe, & Ollinger, 2014; Webb, Little, & Cropper, 2018. Experiencing insight is strongly associated with a correct answer, and, so to disentangle this, all analyses of insight experience (binary and strength of "Aha") take place with just correct answers (Danek & Salvi, 2020; see Ross & Vallée-Tourangeau, 2022).

4.2.3. The feeling of being stuck

There are no commonly accepted methods for assessing how stuck a participant feels. In Experiment 1, we adopted a binary scale of stuck/ not stuck using the following wording: "When trying to solve the problem did you feel stuck?". This instruction was deliberately kept as brief as possible to avoid priming participants. Instead of describing the feeling of being stuck, we asked participants to give us a detailed account of their experience using the following phrasing:

> In your own words, what did the feeling of being stuck feel like? Please try to explain as fully as possible how being stuck made you personally feel.

4.2.4. Procedure

The experiment was conducted using Qualtrics software. After being briefed on the feeling of insight, participants were given 60 s to solve the riddle. They were then rebriefed on the feeling of insight and asked if they had experienced it. They were also asked whether they felt stuck during the problem solving process before being asked to describe this feeling. Finally, they were given the answer and asked to indicate whether they were correct.

4.3. Results

All analyses were conducted using R (R Core Team, 2023). The data and the analysis code for all experiments can be found here: https://osf.io/qbuv8/?view_only=d1c6703968e34f1285427542ffff7c7d

4.3.1. Preregistration

As this experiment was exploratory, we did not preregister any of the quantitative hypotheses. However, we did preregister the following qualitative questions[4]:

> R1: What is it like (phenomenologically) to experience impasse?
>
> R2: Is this experience similar across participants?

We also preregistered the qualitative data analysis process described below. The preregistration can be found at: https://osf.io/avpu7/?view_only=9536335678bc4266bd3d79a3cc6479b1

4.3.2. Performance measures: correct answers, experiencing insight and feeling stuck

In total, 86 of the 125 participants felt stuck, a proportion of 0.69. The proportion of correctly solved problems in those who did not report experiencing impasse was 0.32 while it was much lower in the group who reported experiencing impasse (0.15). The average latency to a correct solution in the group that reported impasse was higher at 57.1 s ($SD = 5.20$, 95% CI [54.0, 60.3) than those who did not report impasse who took an average of 51.9 s ($SD = 11.3$, 95% CI [44.8, 59.1]). We also examined the level of insight. A far higher proportion of those who did not experience impasse reported an insight experience (0.83) than those who experienced impasse (0.23).

We fitted a linear model using lme4 (Bates, Mächler, Bolker, & Walker, 2015) to assess whether being stuck predicted incorrect responses. The explanatory power of the model is weak ($R^2 = 0.04$). The effect of being stuck was statistically significant and negative ($\beta = -0.16$, 95% CI [−0.32, −0.01], $t(122) = -2.13$, $p = .033$; Std. $\beta = -0.41$, 95% CI [−0.79, −0.03]). The model assessing whether being stuck predicted latency to a correct answer explained a statistically non-significant and weak proportion of the variance ($R^2 = 0.09$, $F(1,23) = 2.28$, $p = .145$, adj. $R^2 = 0.05$). The explanatory power of the final linear model to assess whether being stuck predicted experiencing insight was substantial ($R^2 = 0.36$). Within this model, the effect of being stuck was statistically significant and negative ($\beta = -2.81$, 95% CI [−5.08, −0.99], $p < .001$; Std. $\beta = -2.81$, 95% CI [−5.08, −0.99]). These data support other empirical studies in this area and demonstrate that getting stuck has a negative effect on both success rates and reduces the feeling of insight. Interestingly, being stuck did not result in a significantly increased latency to a correct solution supporting suggestions that it is not always matched by behavioural correlates.

4.3.3. Phenomenological measures: what did it feel like to be stuck?

The free-text answers were then read independently by the two authors and the reported emotions were identified independently and refined through discussion. The codes and examples of the responses are listed in Table 1. Initially 10 codes were identified that were reduced

[4] We also preregistered the following research question: R3: (secondary): Is this mediated by expertise? This research question was investigated alongside the other research questions across the four Experiments. The data from this research question are reported in ANONYMISED (in prep).

Table 1
Feelings extracted from the qualitative data with examples and the absolute (percent) number of responses coded.

Code	Example	Number
Feelings		
Fear/panicking/anxiety	Panic, like I should have been able to work it out but the question was quite vague, I needed some visuals of where the painting was and more time. The timer counting down made me panic more	40 (47%)
Feeling Stupid	When I do anagrams like this and more often than not I can't work them out I feel really stupid and thick. As in lacking intelligence. Even when I subsequently read the answer it still never makes sense. I feel like my brain just isn't wired to work this stuff out. I try them because I want to have that 'aha!' moment but ultimately, I just feel stupid.	7 (8%)
Anger/Frustration	I felt frustrated and continue to feel stuck in a loop as I still don't know the answer. I consider myself a clever person, so it annoys me when I can't solve riddles!	26 (31%)
Happiness/hopefulness/ other positive feelings	No major sense of guilt or panic. Excited to solve it and was running ideas through my head. I tried taking my time and thinking of 'the way around' the riddle that is always thrown in there.	35 (41%)
Sadness/hopelessness	It also doesn't feel good, it's a sort of sinking feeling as well as raises feelings of incompetence for myself.	12 (14%)
Other Factors		
Awareness of Time	I kept reading the text, but knowing the timer was going made me feel under pressure more	38 (45%)
Cognitive Speediness	You feel your thoughts speed up	9 (11%)

through discussion to 5 codes. For example, the initial codes of anger and frustration were combined to create one overall code with the answers that displayed self-frustration recoded with the feelings of self-contempt.

In addition to the emotion codes extracted, two further aspects were identified through the iterative, collaborative coding between the first two authors. The first was the importance of the time that was given. Participants showed a clear awareness of the notion of time running out. For example, one participant wrote: "I kept reading the text, but knowing the timer was going made me feel under pressure more". The second was patterns of *cognitive* activity: looping or spiralling through ideas, or being completely at a halt. For example, one participant wrote "it felt like a road block in my mind" while another wrote "I could feel my brain trying to process possible solutions, but none seemed to get. I felt briefly stupid for each solution I came up with that was obviously wrong. I felt frustrated, and continue to feel stuck in a loop as I still don't know the answer".

Independent members of the research lab then read each response and allocated a code each time one of the feeling codes appeared. This allowed us to capture multiple feelings from one response. For example, the response "I felt very frustrated, by the time I was finished reading the question I was almost out of time. Feeling stuck is very frustrating and I felt stupid for not knowing what the question was asking. I also feel annoyed" was coded as feeling stupid, feeling anxious, and feeling angry and frustrated.

4.4. Discussion

Experiment 1 provided evidence that the feeling of being stuck is more reliably related to problem failure than to problem solution (albeit with a small effect size) and that it has a significant impact on feelings of insight even when the answer is correct; experiencing impasse leads to lower solution rates and lower feelings of "aha". In addition, the qualitative data also support the notion of impasse as a catastrophic state, with participants describing strong negative emotions.

This means that current models of insight problem-solving that require impasse do not capture the whole picture: impasse is not a necessary precursor for the feeling of insight. This supports the current empirical findings yet leaves us with a sparse explanation of what causes the change in either representation or strategy that impasse is hypothesised to cause. It also does not allow us to investigate ways of alleviating the state of impasse because it seems that once in impasse, the overall outlook is not good. Additionally, despite this self-report data, the behavioural data showed no significant impact on latency, which would be expected if impasse reflected the cessation of problem-solving activity. The quantitative aspects of Experiment 1 did not progress our understanding but confirmed the already contradictory situation.

Qualitative reports may explain this. These reports indicate that some (albeit not all) people feel positive about this state. In addition, in the affective state, we identified a feeling of cognitive speediness. This latter echoes the model by Danek (2018) who separates "impasse" from "repeated failure" - it may be that alongside feeling hopeful and hopeless (an affective/motivational state), impasse may be experienced differently cognitively with a feeling of running *through* ideas and a feeling of running *out* of them.

5. Experiment 2

Experiment 1 had several limitations. First, a far smaller number of participants answered correctly than those reported in previous research that used this riddle as stimuli. This meant that very few participants experienced impasse but went on to get the riddle correct. This is the group of participants whose process profile would most closely mimic the theoretical process of "true insight' (Spiridonov et al., 2021; Weisberg, 2015). Second, the binary 'stuck or not stuck' question reduced impasse to a unitary one-time experience, whereas the qualitative evidence suggested that it was more complex and could be experienced more as a dynamic process.

Experiment 2 addresses these limitations. We first extended the time available to the 90s used by Ross and Vallée-Tourangeau (2022) to increase success rates. Second, we introduced a third category of being stuck that emphasised feelings across the entire problem- solving experience. We supplemented this with quantitative codes based on emotions and cognitive states identified in the first experiment.

5.1. Method

5.1.1. Participants

A total of 250 participants were recruited from Prolific.co and were paid £2 to participate in the experiment. 3 participants failed to complete the full experiment and their data were excluded from analysis leaving 247 participants. The average age of the women ($N = 120$) was 38.02 ($SD = 14.32$), whereas for men it was 37.50 ($SD = 14.83$). Two participants preferred not to offer their gender, and their ages were 30 and 25 years, respectively. None had seen the problem before.

5.2. Materials and measures

The experimental stimuli and insight briefing were the same as in Experiment 1.

5.2.1. Quantitative measures of feeling stuck

We employed a more granular assessment of the feeling of being stuck using a three-point scale to elicit experience specifically *over the course* of the problem-solving time and to avoid participants reporting their state at the time of asking. We assumed a roughly linear model of impasse where impasse was not yet experienced, experienced and resolved, or still unresolved, to reflect its potentially dynamic nature. The binary scale from Experiment 1 was expanded and consisted of (1) stuck and remaining so [unresolved impasse], (2) stuck but got over it [resolved impasse], and (3) not stuck [no impasse].

In addition, we aimed to confirm the findings of qualitative coding. Five dimensions were presented to participants as a choice to describe their feelings: (i) fear/panic/anxiety (ii) anger (iii) self-contempt/shame (iv) happiness/hopefulness (v) sadness/hopelessness.

5.2.2. Procedure

The procedure was the same as in Experiment 1, except for the following: we extended the time to 90s to increase the solution rate. After the problem time elapsed, for those participants who reported feeling stuck, we replaced the qualitative measures with quantitative questions that asked participants to select the one feeling from the five dimensions that they felt *the most* during the feeling of being stuck, whereas for those participants who did not feel stuck, we asked them to describe how they felt when trying to solve the problem.

5.3. Results

5.3.1. Preregistration

We did not preregister any quantitative hypotheses, but we did preregister the experimental design. https://osf.io/avpu7/?view_only=9536335678bc4266bd3d79a3cc6479b1

5.3.2. Performance measures: correct answers, experiencing insight and feeling stuck

In total, 72 of the 247 participants answered correctly, with a proportion of 0.29. Although this is lower than expected from previously published research, it is higher than in Experiment 1, although not significantly so, F (1, 370) = 3.62, p = .058, $\eta^2 < 0.01$, 95% CI [0.00, 1.00]. However, despite the lack of significance, we were confident that we achieved a greater proportion of correct answers in our analysis, which was the primary aim of extending the time. Supporting this finding, the average latency to a correct solution was higher than the total time (60s) allowed in Experiment 1 (M = 78.03 s, SD = 16.99, 95% CI [75.91, 80.14]).

Initially, we grouped those who felt stuck and continued to do so and those who resolved the feeling of being stuck. A total of 183 of the 247 participants experienced being stuck, with a proportion of 0.74. This was greater than in Experiment 1, although again not significantly, F (1, 370) = 3.62, p = .282, $\eta^2 < 0.01$, 95% CI [0.00, 1.00].

5.3.3. The relationship between feeling stuck and a correct outcome

Again, unresolved impasse had a detrimental effect on solution rates, with only 15 of the 138 participants who felt like this (11%) getting the correct answer. This compared with 33 of the 64 participants who did not experience impasse (52%). However, 24 (53%) of the 45 participants who reported resolved impasse (i.e. they felt stuck but got over it) went on to declare a correct answer. An ANOVA demonstrated that there was an overall significant difference, F(2, 244) = 31.46, p < .001, $\eta^2 = 0.21$, 95% CI [0.13, 1.00], and pairwise comparisons demonstrated that while unresolved impasse was significantly different from resolved (p_{Bonf} < .001) and no impasse (p_{Bonf} < .001*),* there was no significant difference between resolved and no-impasse conditions (p_{Bonf} > .999). The data show that resolving an impasse is no more detrimental to success than not experiencing an impasse at all.

In addition, experiencing an impasse did influence the latency to a correct solution. Those who did not experience an impasse solved the problem in an average of 69.28 s (SD = 19.42, 95% CI [62.40, 76.17]), while those who resolved their impasse solved the problem in an average of 84.08 s (SD = 11.91, 95% CI [79.04, 89.11]), and those who did not resolve their impasse took even longer (M = 87.58 s, SD = 11.91, 95% CI [84.65, 90.51]). A one-way ANOVA indicated that this was significant, F(2, 244) = 41.89, p < .001, $\eta^2 = 0.26$, 95% CI [0.18, 1.00], and post-hoc tests showed that this difference lies between both resolved and non-resolved impasses and no impasses (both p_{Bonf} < .001), but not between the two stages of impasse (p_{Bonf} > .999).

5.3.4. Feeling stuck and experiencing insight

Sixteen of the 24 participants (67%) who answered correctly and resolved an impasse experienced insight, which dropped to 45% for those who did not experience an impasse. Not resolving impasse had a catastrophic effect on rates of insight, with only one of the 15 participants with a correct answer in this group experiencing insight (7%). An ANOVA demonstrated that there was an overall significant difference, F (2, 244) = 7.95, p < .001, $\eta^2 = 0.20$, 95% CI [0.13, 1.00], and pairwise comparisons demonstrated that while unresolved impasse was significantly different from resolved (p_{Bonf} < .001) and no impasse (p_{Bonf} = .024*),* there was no significant difference between resolved and no-impasse conditions (p_{Bonf} = .265). Again, the data showed that the resolved impasse functioned in the same manner as the unresolved impasse.

5.3.5. Problem-solving emotions

5.3.5.1. The feeling of being stuck. Table 2 first shows each emotion was experienced by at least 10% of each category of those experiencing impasse validating the use of the emotion codes as a way of capturing experience. The two stages of impasse differed significantly across emotional type, χ^2 (4) = 14.05, p = .007.

Those who resolved their impasse were more likely to be in a highly aroused emotional state, that is, a higher proportion reported feeling angry or fearful, panicky, anxious, or even happy. Unresolved impasse was associated with lower arousal states such as sadness/hopelessness and self-contempt/shame. It is worth noting that a higher proportion of those who experienced impasse and did not resolve it selected happiness/hopefulness and positive feelings than sadness and hopelessness, again suggesting that being in a state of unresolved impasse is not necessarily an aversive state.

5.3.5.2. Phenomenological measures: what did it feel like to not be stuck. The free-text answers were read independently by the two authors, and three additional emotion codes were extracted and refined. The codes and examples of the responses are listed in Table 3. As before, these were coded by naïve members of the research team to ensure that they appeared at least once in the data set.

5.4. Discussion

Asking participants if they had resolved their impasse appeared to reliably elicit reports of different stages of the impasse experience. In terms of process, this suggests that participants can experience an impasse and successfully resolve it, which aligns with canonical theories of insight and increases the chances of finding successful interventions to resolve the state of impasse. Moreover, resolving impasse had the same effect on solution rates and levels of insight as not experiencing impasse at all. However, it had a significant effect on the latency to the correct solution; behaviourally, it led to the same time lag as not resolving impasse. In other words, the additional category seemed to induce the behavioural correlate missing from Experiment 1 with the

Table 2
The Main emotion experienced by participants as a function of stage of impasse.

	Unresolved impasse	Resolved Impasse
Anger	5 (4%)	5 (11%)
Fear/panic/anxiety	49 (36%)	20 (44%)
Happiness/hopefulness/other positive feelings	29 (21%)	14 (31%)
Sadness/hopelessness	28 (20%)	1 (2%)
Self-contempt/shame	27 (20%)	5 (11%)

Table 3
Problem -solving feelings extracted from the qualitative data.

How did you feel when trying to solve the problem?		Number of Responses
Excitement	I felt quite excited as if I was achieving something worthwhile. It made me take my mind off other things and I was able to concentrate on this problem.	6 (10%)
Determination	I relished the challenge but was a little worried that I might not be able to solve the riddle. I hate not being able to do something intellectual and at 83 I worry a lot about my abilities, both physical and intellectual, especially knowing that I am lazy physically, although not so much mentally.	14 (23%)
Satisfaction	But when something popped into my mind I felt good about it, then other ideas popped up too. Very satisfying (whether it was right or not!)	13 (21%)

binary scale, suggesting that this category led participants to focus more on the process than on the outcome. Additionally, it is not implausible that the lack of difference in latencies in Experiment 1 was because those situations where impasse was resolved were collapsed into those where it was not experienced at all.

This then adds nuance to the current empirical work: experiencing impasse is not detrimental, but failing to resolve it is. In hindsight, this is almost trivial but has not been captured by the insight and problem-solving literature research to date. The lack of difference across all measures save latency between not experiencing impasse and resolving it supports the view that impasse is not *essential* to insight. In other words, these data hint at two different paths to the same outcome (impasse-driven and non-impasse-driven). It is also important to note that not experiencing impasse or resolving impasse was not a guarantee of the correct solution – still around half of the participants failed to get the correct answer. While we cannot rule out the possibility that the self-report reflects how participants felt at the time, the high number of participants who still failed to answer correctly even when reporting that they did not experience impasse or they resolved it suggests that the reports of the process captured here are not entirely outcome dependent.

There is a complex mix of emotions associated with being stuck. It appears that negative emotions such as anger, fear, and panic were more strongly associated with being stuck and resolving it, suggesting that these active feelings could act as a metacognitive spur for problem solving. Other negative emotions such as hopelessness were lower in those who went on to resolve them. Unexpectedly, it is notable that the feeling of hopelessess, even among those who did not resolve the feeling of being stuck, was selected by a lower percentage of people (albeit almost the same) than happiness. This suggests that impasse can elicit positive affect (not only after the resolution of the impasse state, but also when it is not resolved yet).

6. Experiment 3

Experiment 3 aimed to investigate the feeling of being stuck at a more granular level. First, we systematically manipulated the time experience of the participants to investigate the effect of time pressure on the feeling of impasse. This allowed us to assess whether the feeling remained stable across different time conditions. This was motivated by the time difference between Experiments 1 and 2 and the importance of time for the qualitative reports. We also aimed to replicate the findings of Experiment 2 that a more dynamic sense of impasse can be captured by the introduction of a third category, which is linked to both problem success and the feeling of insight. Finally, we drew on the findings of Experiments 1 and 2 to develop an impasse dimension scale.

We were also consistently getting lower rates of solution for the stumper we selected than reported by Ross and Vallée-Tourangeau, even with an increased time allowance. Thus, the number of participants who fell into the key groups of interest – resolved impasses and a correct answer – was too small to draw meaningful or stable conclusions. Therefore, an easier set of stumpers was selected. To account for differences in the difficulty of the effects we aimed to replicate, we repeated the experiment on a set of stumpers matched to the difficulty of the original stumpers. This led to two simultaneous experiments, one with easy stimuli and one with hard stimuli. The preregistered hypotheses all related to the easier stimulus set.

Finally, we combined the qualitative reports to generate an impasse scale which would combine positive and negative feelings to assess in a more detailed way the feelings encountered during problem-solving.

6.1. Method

6.1.1. Participants

Participants were recruited from Prolific.co. and were paid £2 for their participation. We originally preregistered and recruited 120 participants for each experiment based on a power calculation conducted using GLIMMPSE (Kreidler et al., 2013) but data loss led to119 participants in each condition. 58 women with an average age of 41.38 ($SD =$ 12.51) and 61men with an average age of 40.85 ($SD = 13.37$) took part in the easy experiment. 62 women with an average age of 38.39 ($SD =$ 12.04) and 57 men with an average age of 45.04 ($SD = 13.82$) took part in the hard experiment.

6.1.2. Experimental design

The experimental design had one within-subject variable, time (30, 60, 90, and 120 s). Two simultaneous experiments were conducted: One on an easier set of stimuli to induce a higher proportion of participants to experience resolved impasse and one of a set of stimuli matched for difficulty with the riddle used in Experiments 1 and 2.

6.2. Materials and measures

6.2.1. Experimental stimuli

We selected four stumpers from Ross and Vallée-Tourangeau (2022) with an average reported success rate of 61.5% for the easy trials and with an average success rate of 44% for the hard trials. Stumpers and normative data are reported in the Appendix.

6.2.2. Impasse scale and additional measures

We refined the multidimensional nature of the impasse experience based on the results of Experiments 1 and 2. We removed the reference to anger because few participants selected it in the second experiment and incorporated the findings from participants who were not stuck. All participants were asked to rate their position across the following seven dimensions: fearful/anxious-calm/relaxed, self-contempt-self-contentedness, sad-happy, hopeless-hopeful, bored-excited, indifferent-determined, and disappointed-satisfied. Half of the pairs were presented as negative to positive, and half as positive to negative. The scores were then converted such that a low number indicated a negative experience for all pairings (i.e., fearful, self-contempt, sad, hopeless, bored, indifferent, and disappointed). These dimensions collapsed to yield an overall average score for the valence of the problem-solving experience.

We also aimed to distinguish between the findings in Experiment 1 of impasse as slow and impasse as fast by asking people whether they characterised the feeling of being stuck as "running through ideas but none of them seems to fit" and "running out of ideas and coming to a halt".

6.2.3. Procedure

The experiment was administered using PsychoPy via Pavlovia to control the time category more carefully. However, the procedure remained the same as in Experiments 1 and 2. Participants were invited to solve all four riddles, time conditions were counterbalanced across participants, and the four stimuli were counterbalanced across these time conditions. A second experiment was run directly after with different participants drawing on harder riddles. The results for both experiments are reported below.

6.3. Results

All results reported here were analysed at the trial level with the participant and riddles as random intercepts (formula = (1|participant) + (1+ time|riddlecode), where the participant only had one trial in each time condition but the riddles were counterbalanced across these). For H1a,b and c, we had no specific hypotheses across time conditions, so we included time condition as a random factors (formula = (1|participant) + (1|time) + (1+ time|riddlecode). For H2a and b, the time condition may have a significant effect on the feeling of cognitive speediness, so it was included as a fixed factor. These models were then compared to a null model using likelihood ratio testing (LRT) from the afex package (Singmann, Bolker, Westfall, Aust, & Ben-Shachar, 2023), as recommended by Brown (2021).

6.3.1. Correct answers

Participants were invited to solve problems across four different time periods (30s, 60s, 90s, and 120 s). With the main set of stimuli (the easier set), participants were correct 56% of the time in the 30s condition, rising to 64% when they were given 60s, a proportion which dropped to 60% with the 90s group and with 120 s to solve the problem 72% of the participants answered correctly. A model with the time condition as a fixed variable was not a significantly better fit for the data (χ^2 (3) = 6.84, p = .077). Thus, while more people got the correct answer with more time, this was not a significant difference. This replicates the findings of the differences between Experiments 1 and 2. Time is beneficial, but not a panacea. With the harder problems, time was more useful: 36% of people got the correct answer with 30s, 43% got the correct answer with 60s, 66% of people got the correct answer with 90s and slighter fewer with 120 s, 55%. This time a model with time as fixed variable was a better fit for the data, χ^2 (3) = 28.20, $p < .001$) and post hoc tests with a Bonferroni correction show us that the differences lie between 30s and 90s ($p < .001$), 30s and 120 s ($p = .011$) and between 60s and 90s (p = . 002). All other comparisons were not significant (lowest $p = .385$).

6.3.2. Feeling stuck

The proportion of people experiencing each stage of impasse as a function of time condition is presented in Table 4. The number of those failing to resolve their impasse decreased as time increased, whereas the number that did not experience an impasse increased. Those who resolved their impasse fell across time conditions.

The χ^2 test demonstrated no overall significant difference from what would be expected in the easy puzzles: χ^2 (6) = 2.99, p = .810. This pattern of experiencing an impasse did not replicate in the hard puzzles: a χ^2 test demonstrated an overall significant difference from what would be expected, χ^2 (6) = 23.27, $p < .001$. For these harder puzzles, post hoc tests with Bonferroni correction demonstrated that the significant differences lie in 90 s conditions for no impasse ($p_{Bonf} = .011$, next lowest $p_{Bonf} = .271$), while resolved impasse had no significant change to expected distribution (lowest $p_{Bonf} = .405$), and the only significant difference was in the 30 s condition for the unresolved impasse group ($p_{Bonf} = .048$, next lowest $p_{Bonf} = .076$). Therefore, the relationship between time and impasse should be treated with caution but it seems that in the main the distribution of these three impasse states remained roughly consistent across time conditions.

Table 4
The proportion of people experiencing each of stage of impasse as a function of time condition (absolute number and percentge).

	30 Seconds	60 Seconds	90 Seconds	120 Seconds
Easy Problems				
No Impasse	49 (45%)	55 (48%)	57 (51%)	62 (55%)
Resolved Impasse	33 (30%)	31 (27%)	27 (24%)	26 (23%)
Unresolved Impasse	27 (25%)	28 (25%)	28 (25%)	24 (21%)
Hard Problems				
No Impasse	35 (30%)	40 (35%)	60 (53%)	46 (40%)
Resolved Impasse	28 (24%)	27 (23%)	27 (24%)	39 (34%)
Unresolved Impasse	52 (45%)	48 (42%)	27 (24%)	30 (26%)

6.3.3. Latency to solution of correct answers

The average latency to a correct solution is presented in Table 5. Latency gradually increased across all time conditions.

The LRT suggests that a model with the main effect of impasse stage is statistically significant and medium, χ^2 (2) = 18.65, $p < .001$. Unsurprisingly, the main effect of time was also statistically significant, χ^2 (3) = 33.66, $p < .001$. The interaction was not significant, χ^2 (6) = 9.17, $p = .164$. There were differences between resolved and no impasse ($p_{Bonf} < .001$) but not between no impasse and unresolved impasse (p_{Bonf} = .161) or resolved and unresolved impasse ($p_{Bonf} > .999$). In terms of time, all pairwise comparisons were significant (highest $p_{Bonf} = .050$) In terms of time, all pairwise comparisons were significant (highest p_{Bonf} = .050) except between 60 and 90 s ($p_{Bonf} > .999$), and 90 and 100 s ($p_{Bonf} = .375$). For the harder problems, the LRT suggests that a model with the main effect of impasse stage is statistically significant and medium, χ^2 (2) = 20.78, $p < .001$. Again, the main effect of time was also statistically significant, χ^2 (3) = 47.33, $p < .001$. The interaction was not significant, χ^2 (6) = 9.93, $p = .128$. There were differences between resolved and no impasse ($p_{Bonf} < .001$) and between no impasse and unresolved impasse ($p_{Bonf} = .020$) but not between resolved and unresolved impasse ($p_{Bonf} >$ .999). We replicated the findings from Experiments 1 and 2 that an impasse is associated with a longer latency to solution.

6.3.4. Preregistered hypotheses

The preregistered hypotheses can be found here: https://osf.io/qbuv8/?view_only=d1c6703968e34f1285427542ffff7c7d

6.3.4.1. H1a: resolved impasse will lead to higher solution rates than non-resolved impasse.

In the easy problem set, the overall success rate was highest in those who did not report impasse (84%), slightly lower in those who resolved their impasse (63%), and, in line with findings from Experiment 2, the lowest in those who had not resolved their impasse (21%). LRT showed that in a model with stage of impasse as a fixed factor was a significant fit, χ^2 (3) = 124.80, $p < .001$. Pairwise comparisons showed that all three stages of impasse were significantly different from one another (p_{Bonf} across all comparisons <.001). In the

Table 5
Average (SD) latency to a correct solution as function of stage of impasse and time condition across hard and easy problem sets.

	30 Seconds	60 Seconds	90 Seconds	120 Seconds
Easy Problems				
No Impasse	22.69 (7.10)	30.19 (13.91)	33.37 (18.48)	35.71 (23.22)
Resolved Impasse	22.21 (7.59)	36.85 (13.93)	42.65 (26.84)	48.48 (23.83)
Unresolved Impasse	21.84 (8.78)	37.13 (7.74)	38.64 (9.66)	40.71 (21.06)
Hard Problems				
No Impasse	26.98 (4.05)	41.57 (12.95)	35.50 (17.82)	49.22 (16.56)
Resolved Impasse	26.73 (4.34)	50.77 (6.73)	46.04 (18.71)	67.63 (23.75)
Unresolved Impasse	24.95 (4.38)	45.94(9.31)	76.61 (n/a)[a]	59.32 (31.83)

[a] Only one participant solved correctly while failing to resolve their impasse in this time condition.

hard problem set, success rates were highest in those who did not experience impasse, 77%, followed by 60% in those that resolved impasse and finally 11% of those who were still in impasse. LRT showed that including the main effect of impasse stage was a significantly improved fit, χ^2 (2) = 152.92, p < .001. Again, pairwise comparisons showed that all three stages of impasse were significantly different from one another (p_{Bonf} across all comparisons <.009). H1a is supported: a resolved impasse leads to higher solutions than non-resolved impasse, but not experiencing an impasse is still better for solution rates.

6.3.4.2. H1b resolved impasse will lead to higher levels of affective "aha" in correct solutions than non-impasse trials. We then sought to replicate the analysis from Experiment 2 that the trials in which people resolved their feelings of impasse were associated with higher levels of insight than those in which there were no feelings of impasse. Again, we examined only those in which they gave the correct answer. The average proportion and strength of the insights are presented in Table 6.

LRT testing showed that in a model for predicting experiencing insight with stage of impasse as fixed factor, the stage of impasse was a significant improvement, χ^2 (2) = 26.54, p <.001. As in Experiment 2, there was a significant difference between those who were still experiencing impasse and those who did not experience it or resolved it (both p_{Bonf} < .001) but not between those who experienced impasse and resolved it and those who did not (p_{Bonf} = .446). This pattern was repeated with the strength of the insight experience: the stage of impasse provided a significantly better fit, χ^2 (2) = 47.15, p <.001, but the difference lay only between unresolved impasse and the other two conditions (p_{Bonf} < .001), and the difference between resolving impasse and not experiencing it was not significant (p_{Bonf} > .999). In the hard problems, this replicated. LRT testing showed that in a model for predicting experiencing insight with stage of impasse as fixed factor and time condition as a random factor, the stage of impasse was a significant improvement, χ^2 (2) = 18.82, p < .001. Again, there was a significant difference between those who were still experiencing impasse and those who did not experience it (p_{Bonf} = .006) or resolved it (p_{Bonf} = .003) but not between those who experienced impasse and resolved it and those who did not (p_{Bonf} < .999). In terms of insight strength, again, stage of impasse was a significantly better fit, χ^2 (2) = 38.93 p < .001 and there was a significant difference between those who were still experiencing impasse and those who did not experience it or resolved it (all p_{Bonf} < .001) but not between those who experienced impasse and resolved it and those who did not (p_{Bonf} < .999).

The data do not support H1b – the insight experience was stronger in resolved impasse conditions but not significantly so; rather, they suggest that the occurrence of insight and the strength of that insight is similar where impasse is experienced or not.

6.3.4.3. H1c: resolved impasse will be experienced as pleasant by problem solvers. Next, we investigated whether pleasantness was predicted by the impasse stage. In hindsight, H1c was poorly phrased; therefore, we used the no-impasse group as a baseline measure. 51% of participants who did not experience an impasse found the experience pleasant, compared with 48% of those who resolved their impasse and 16% of those who remained in impasse. Again, the model with the stage of impasse was a significant improvement, χ^2 (2) = 35.11, p ≤.001. Post hoc tests showed that there were no significant differences between no impasse and resolved impasse (P_{Bonf} = .956), but there were significant differences between non-impasse and unresolved impasse (P_{Bonf} < .001), and resolved and unresolved impasse (P_{Bonf} < .001). We established therefore that both resolving impasse and not experiencing it at all were experienced as similarly pleasant across our participants.

With the hard puzzles, 43% of participants who did not experience an impasse found the experience pleasant, compared with 25% of those who resolved their impasse and 10% of those who remained in impasse. Again, the model with the stage of impasse was a significant improvement, χ^2 (2) = 50.31, p < .001. Unlike with easy stumpers, there were significant differences between all three groups (highest P_{Bonf} < .03 for the comparison between resolved and unresolved impasse).

6.3.4.4. H2: there will be two distinct forms of impasse – "fast" and "slow" impasse. Fast impasse will be associated with (a) higher solution rates and (b) higher levels of insight. Drawing on the findings from Experiment 2, we first established that alongside the emotional dimensions, impasse could be defined cognitively as "fast" or "slow". We aimed to quantify this by asking those who were stuck whether they experienced the feeling as being fast or slow. For the easy problems 53% of participants in the 30 s condition reported the feeling as fast, rising to 56% in the 60 s condition, 55% in the 90 s condition and 66% in the 120 s condition. In addition, we examined whether these forms were associated with different stages of impasse. 79% of those who resolved their impasse described it as fast, compared to 34% of those who did not resolve their impasse. There was a significant difference in distribution, χ^2 (1) = 44.37, p < .001). In the hard problem set, 41% of those in the 30 s condition described it as fast, 39% in the 60s condition, 46% in the 90s condition and 58% in the 120 s condition. In terms of stage of impasse, 73% of those who resolved their impasse described it as fast compared with 25% of those who did not which was again a significant difference, χ^2 (1) = 61.24, p < .001. Therefore, the first part of H2 is supported: There are two easily distinguishable cognitive forms of impasse.

6.3.4.5. H2a Fast impasse will be associated with higher solution rates. We then assessed whether there were differences in the success rates between the two types of fast slow impasses. Fast impasse was associated with higher levels of success across all four conditions in easy and hard problems (see Table 7). The LRT with time and type of impasse (fast/slow) as fixed effects showed that including time did not significantly improve the null hypothesis (easy = χ^2 (3) = 2.39, p = .496, hard = χ^2 (3) = 4.90, p = .179) but the type of impasse did (easy = χ^2 (1) = 16.92, p < .001, hard = χ^2 (1) = 28.40, p < .001. Hypothesis 2a was supported.

6.3.4.6. H2b Fast impasse will be associated with higher levels of insight. The relationship between fast or slow impasse, insight and time conditions is displayed in Table 7. Again, we only looked at the partcipants who gained a correct answer to account for confounds between experiencing insight and a correct answer. The LRT suggests that the main effect of impasse speed was significant, χ^2(2) = 10.70, p < .001, but the main effect of time was not, χ^2 (3) = 0.97, p = .810. We then looked at the strength of insight which followed the same pattern. Cognitive speed was significant, χ^2(2) = 16.55, p < .001 but the effect of time condition was not, χ^2(3) = 2.88, p = .411. The data from the easy trials support H2b. However, from the hard trials, the LRT suggests that the main effect of impasse speed was not significant, χ^2(2) = 1.14, p = .285, but the main effect of time was, χ^2 (3) = 0.97, p = .810. We then looked at the strength of insight where both feelings of Cognitive speed was significant, χ^2(2) = 16.55, p < .001 and the effect of time condition was not, χ^2(3) = 2.88, p = .411. We note however that the number of trials where participants both got the answer correct and encountered impasse was small in the hard stimuli set (only 90 trials in total across all time conditions) supporting our choice to use easier stimuli.

Table 6
Percent of people experiencing insight in each impasse experience and Average (SD) strength of the insight experience.

	No Impasse	Unresolved Impasse	Resolved Impasse
Easy Problems			
Insight	74%	32%	84%
Strength of Insight	6.56 (2.32)	3.88 (1.58)	6.66 (2.05)
Hard Problems			
Insight	81%	28%	84%
Strength of Insight	7.01 (2.32)	3.03 (1.58)	6.65 (2.28)

Table 7
Proportion of people getting the answer correct and experiencing insight along with strength of insight as a function of time condition and feelings of speediness.

	Fast			Slow		
	Percent Correct	Percent Experienced Insight	Insight Strength	Percent Correct	Percent Experienced Insight	Insight Strength
Easy problems						
30s	44%	57%	5.81 (2.21)	29%	75%	5.66 (2.48)
60s	67%	86%	6.50 (2.29)	23%	33%	4.01 (1.36)
90s	47%	100%	7.56 (1.10)	28%	29%	4.57 (2.65)
120 s	61%	80%	6.24 (2.31)	29%	40%	4.40 (1.84)
Hard Problems						
30s	39%	69%	4.98 (2.54)	11%	60%	4.54 (3.00)
60s	41%	64%	5.43 (2.38)	28%	54%	4.20 (2.37)
90s	68%	100%	8.29 (1.44)	7%	100%	6.44 (0.66)
120 s	50%	80%	6.71 (2.25)	28%	50%	4.77 (2.82)

6.3.4.7. H3: those trials where impasse is not reported will have a larger number of "pop-out" solutions, that is solutions which are answered in the first 10 s. Across all the trials, only 18 in the easy condition and 3 in the hard condition were solved in less than 10s, which meant that this hypothesis could not be clearly tested.

6.3.5. Impasse scale

Experiments 1 and 2 used qualitative analysis to assess the different dimensions of impasse. These qualitative explorations yielded seven potential emotional dimensions of the problem solving experience: i) self-contempt – self-contentedness, ii) sadness–happiness, iii) hopelessness–hopefulness, iv) boredom–excitement, v) determination–indifference, vi) disappointment–satisfaction, and vii) fearfulness – calmness.

In Experiment 3, we quantitatively tested these across each trial. Participants were invited to rate their experiences on a scale of 1–7. These were presented as counterbalanced but converted so that 1 represented a negatively valenced emotion and 7 represented a positively valenced emotion. Trials with missing data were excluded from the analysis. The overall scale showed good reliability ($\alpha = 0.79$).

To assess whether the feeling of being stuck was uni-or multidimensional, we ran an exploratory factor analysis on the datasets that collapsed across time conditions and hard and easy trials. Bartlett's test of sphericity, which tests the overall significance of all the correlations within the correlation matrix, was significant χ (21) = 922.70, $p < .001$, indicating that it was appropriate to use the factor analytic model on this set of data. The Kaiser-Meyer-Olkin measure of sampling adequacy indicated that the strength of the relationships among variables was high (KMO = 0.76); thus, it was acceptable to proceed with the analysis. Two factors with eigenvalues greater than one were extruded. Factor One had an eigenvalue of 3.11 and accounted for 27% of the variance. Factor Two had an eigenvalue of 1.31 and it accounted for 23% of the variance. Factor One ($\alpha = 0.73$) was identified as an affective factor and Factor Two ($\alpha = 0.75$) as a motivational factor. The relationships are illustrated in Table 8.

Table 8
Factor loadings for each of the two factors.

Item	Factor Loading	
	Factor1	Factor2
Factor 1: Affective		
Self-Contempt ➔ Content	**0.339**	0.248
Sad ➔ Happy	**0.727**	0.030
Disappointed ➔ Satisfied	**0.725**	0.083
Fearful ➔ Calm	**0.805**	−0.256
Factor 2: Motivation		
Hopeless ➔ Hopeful	0.272	**0.516**
Bored ➔ Excited	−0.025	**0.716**
Indifferent ➔ Determined	−0.163	**0.833**

6.4. Discussion

We report three key findings. First, we established that impasse itself, whether resolved or unresolved, is multidimensional and differs across feelings of processing speed, affect, and motivation. This goes some way to resolve the second paradox of why a seemingly aversive state is sought by habitual problem-solvers. Impasse which is characterised by a fast cognitive state, is higher in general affect and motivation and is more likely to lead to a correct answer and a feeling of insight. This reflects the notion of "hopeful" impasse.

Second, the addition of an intermediate stage of impasse (resolved impasse) elicited the same affective and performance-based behaviours as not experiencing impasse. This across all time conditions and difficulty of riddle. This finding resolves the first paradox: the empirical evidence for impasse has been poor because it has not tapped into this resolved state. The state of resolved impasse aligns theoretical and empirical observations and also has the behavioural correlate of longer latencies. However, because it only differs from the no-impasse state in terms of latency, this suggests that there are multiple routes to similar affective responses. This means that the use of insight experience to assume a single cognitive process may be less reliable than previously assumed.

The time manipulation did not significantly alter solution rates or feelings of insight. Additionally, there is inconsistent evidence regarding the effect of time on the feeling of being stuck. There were no significant differences in the time conditions on stage of impasse in the easy puzzles or hard puzzles. The data suggest that the stage of impasse is more important than the time allocated to solving the problem.

7. Experiment 4

Experiment 4 built on the findings of Experiment 3 to investigate the feeling of being stuck across multiple task types. We followed the same procedure as in Experiment 3, but we varied the problem stimuli to encompass four forms of problem: (a) a traditional verbal insight puzzle, (b) a traditional verbal analytical puzzle, (c) a word problem from the UK GCSE mathematics curriculum, and (d) the original stumper. We selected these problems to illustrate a wide range of different forms of problems that might be encountered in experimental studies.

7.1. Method

7.1.1. Participants

To support the large samples needed for factor analysis, we preregistered and recruited a total of 250 participants from Prolific.co. Each was paid £2.75 for participating. Data were saved for 246 participants with an average age of 39.91 (*SD* = 13.08). A total of 124 participants were men, 120 were women, and two were classified as non-binary or other.

7.2. Material and measures

We selected four different types of problems, as presented in the Appendix. Verbal insight and analytical problems were taken from Webb et al. (2016), whereas the mathematical problem was selected from the AQA guide to mathematical problems, and the stumper was the same as that used in Experiments 1 and 2.

7.2.1. Procedure

The procedure was the same as in Experiment 3, except that we followed Webb et al., and gave the participants 4 min to solve each of the problems.

7.3. Results

7.3.1. Preregistered hypotheses

We preregistered the following hypotheses: https://osf.io/qbuv8/?view_only=d1c6703968e34f1285427542ffff7c7d. First, trials in which the impasse is resolved will (a) have the same solution rate as no-impasse trials, (b) have the same levels of insight as no-impasse trials, (c) have the same level of motivation as no-impasse trials, (d) have lower affect than no-impasse trials, and (e) have longer latencies than no-impasse trials. Second, trials in which an impasse is described as fast will (a) have higher solution rates than trials in which it is described as slow, (b) have higher levels of insight than slow trials, (c) have higher levels of motivation than slow trials, (d) have higher affect than slow trials, and (e) have the same latency to a correct solution as slow trials.

7.3.2. Factor analysis

We performed another EFA to assess whether the two factors were still present in the data. Bartlett's test of sphericity was significant $\chi^2(21) = 1024.22$, $p < .001$ and the Kaiser-Meyer-Olkin measure of sampling adequacy indicated that the strength of the relationships among variables was high (KMO = 0.72) Again, two factors with eigenvalues greater than1 were extruded which mapped onto the same factors seen in Experiment 3.

7.3.3. The process of impasse

First, we assessed whether the proposed three-stage impasse process was sustained across all problem types. As shown in Table 9, all problem types experienced all three stages of the impasse-related processes.

7.3.4. Preregistered hypotheses

All hypotheses were tested with linear mixed models with participant and problem type as a random factor and the main effect specified in the hypothesis (either impasse stage or speediness of impasse) as fixed factors using the formula (1|participant +1|problem_type). These models were then compared to a null model using likelihood ratio testing (LRT) from the afex package (Singmann et al., 2023), as recommended by Brown (2021).

Trials in which the impasse is resolved will:

a. Have the same solution rate as no impasse trials

The proportion of correct answers is displayed in Table 10. Solution rates across all four problem types were generally highest in the resolved impasse condition and were lowest in the unresolved impasse condition. We fitted a linear mixed model with participants and problem type as random factor to test the main effect of impasse stage: A likelihood-ratio test indicated that this was significant, $\chi^2(2) = 179.95$, $p < .001$. Post hoc tests with a Bonferroni correction suggest that there were significant differences in solution across all three types of impasse, lowest $p_{Bonf} = .009$ for no impasse - resolved impasse. Hypothesis 1a was not supported however there were still higher solution rates for resolved impasse compared to unresolved impasse as we have seen in Experiment 2 and 3.

Table 9
Proportion of each problem types experiencing each stage of impasse.

	Analytical	Insight	Maths	Stumper
No Impasse	41 (18%)	89 (36%)	52 (23%)	73 (30%)
Resolved Impasse	72 (32%)	88 (36%)	76 (33%)	69 (29%)
Unresolved Impasse	110 (49%)	67 (27%)	99 (44%)	99 (41%)

Table 10
Descriptive statistics for each of the preregistered hypotheses relating to stage of impasse.

	Analytical	Insight	Mathematical	Stumper
Proportion Correct				
No Impasse	0.34	0.65	0.63	0.79
Resolved Impasse	0.33	0.56	0.51	0.64
Unresolved Impasse	0.17	0.06	0.05	0.27
Proportion Experienced Insight				
No Impasse	0.50	0.81	0.61	0.66
Resolved Impasse	0.50	0.94	0.74	0.68
Unresolved Impasse	0.26	0.50	0.20	0.30
Strength of insight				
No Impasse	5.27(3.02)	6.46(3.05)	5.12(3.35)	5.75(2.95)
Resolved Impasse	4.55(2.6)	7.06(2.17)	6.02(2.34)	5.66(2.53)
Unresolved Impasse	3.25(2.57)	5.37(3.13)	4.01(2.48)	3.09(2.41)
Strength of Motivation				
No Impasse	4.28(1.36)	4.25(1.6)	4.75(1.48)	4.29(1.34)
Resolved Impasse	4.09(1.24)	4.28(1.14)	4.26(1.06)	4.34(1.2)
Unresolved Impasse	4.28(0.9)	4.19(0.77)	4.08(0.93)	4.19(0.96)
Direction of Affect				
No Impasse	4.75(0.94)	5.11(0.94)	5.18(0.93)	4.89(0.94)
Resolved Impasse	4.35(0.82)	4.55(0.93)	4.31(0.83)	4.5(0.76)
Unresolved Impasse	3.8(0.77)	3.96(0.75)	3.53(0.9)	3.85(0.91)
Latency				
No Impasse	35.47 (26.38)	36.58(33.3)	61.65(62.2)	40.07 (21.14)
Resolved Impasse	65.11 (43.62)	55.61 (47.18)	87.99(63.86)	64.66 (37.22)
Unresolved Impasse	83.96 (75.27)	119.48 (85.36)	85.31(58.51)	59.84 (54.92)

b. Have the same levels of insight as no impasse trials

The proportion of correct trials in which participants reported an insight experience is displayed in Table 10. There were slightly higher levels of insight trials associated with resolved impasse than those not experiencing impasse across the problem types. As before, not resolving impasse was catastrophic in terms of the proportion of trials experiencing insight. The same pattern was seen for the strength of insight (also Table 7): the insight experience was strong when impasse was either not experience or was resolved.

A likelihood-ratio test indicated that the main effect of impasse stage was significant on the binary measure, $\chi^2(2) = 30.35$, $p < .001$. Post hoc tests with a Bonferroni correction demonstrate a significance difference between unresolved impasse and no impasse ($p_{Bonf} = .001$) and between unresolved impasse and resolved impasse ($p_{Bonf} < .001$) but not between resolved and no impasse ($p_{Bonf} = .458$). In terms of strength of insight, a likelihood-ratio test indicated that the main effect of impasse stage was significant, $\chi^2(2) = 33.44$, $p < .001$. Post hoc tests with a Bonferroni correction demonstrate a significance difference between unresolved impasse and no impasse ($p_{Bonf} < .001$) and between unresolved impasse and resolved impasse ($p_{Bonf} < .001$) but not between resolved and no impasse ($p_{Bonf} > .999$). Hypothesis 1b was supported.

c. Have the same level of motivation as the no-impasse trials.

The average motivation is displayed in Table 10. Motivation was broadly similar across all problem types and impasse stage. The LRT indicated that the main effect of impasse stage was not significant, $\chi^2(2) = 2.69$, $p = .261$.Hypothesis 1c was supported.

d. Have lower affect than no impasse trials.

The average affect is displayed in Table 10. Affect was highest in no-impasse trials and then resolved impasse trials, before being lowest in the unresolved impasse trials across all problem types. A likelihood ratio test indicated that the main effect of impasse stage was significant, $\chi^2(2) = 275.36$, $p < .001$. Post hoc tests with a Bonferroni correction demonstrate a significant difference between all three stages of impasse (all $p_{Bonfs} < .001$) Hypothesis 1d was supported.

e. Have longer latencies than no impasse trial.

The average latency to a correct solution is shown in Table 10. Latency increased sharply on experiencing impasse across all problem types and was highest when that impasse was not resolved. A likelihood ratio test indicated that the main effect of impasse stage was significant, $\chi^2(2) = 28.30$, $p < .001$. Post hoc tests with a Bonferroni correction demonstrate a significance difference between no impasse and unresolved impasse ($p_{Bonf} < .001$) and no impasse and resolved impasse ($p_{Bonf} < .001$) but not between resolved and unresolved impasse ($p_{Bonf} = .331$). The behavioural marker of impasse remains consistent. Hypothesis 1e was supported.

Trials where impasse is described as fast will have:

a. Higher solution rates than those where it is described as slow

The proportion of correct answers is shown in Table 11. A likelihood-ratio test indicated that the main effect of impasse speed was significant, $\chi^2(1) = 30.19$, $p < .001$. Hypothesis 1a was therefore upheld.

b. Higher levels of insight than when it is it is described as slow

The proportion of correct trials in which participants reported an insight experience is displayed in Table 11. The LRT suggests that for the experience of insight, the feeling of speediness, $\chi^2(1) = 15.23$, $p = .001$, was a significant main effect. The strength of insight (Table 11) was also significantly different as a function of feelings of speed, $\chi^2(1) = 32.46$, $p < .001$.

c. Higher levels of motivation than when it is described as slow

Table 11
Descriptive statistics for each of the preregistered hypotheses relating to feeling of speed of impasse.

	Analytical	Insight	Mathematical	Stumper
Proportion Correct				
Fast	0.33	0.38	0.32	0.47
Slow	0.11	0.23	0.06	0.33
Proportion Experienced Insight				
Fast	0.49	0.94	0.68	0.68
Slow	0.00	0.87	0.50	0.27
Strength of insight				
Fast	4.68(2.64)	7.21(1.83)	6.34(2.23)	5.93(2.66)
Slow	1.08(0.05)	6.51(2.8)	4.23(3.11)	2.7(2.11)
Strength of Motivation				
Fast	4.31(1.1)	4.05(1.11)	4.18(1.03)	4.16(0.9)
Slow	4.19(0.89)	4.16(0.78)	4.06(0.9)	4.23(1.04)
Direction of Affect				
Fast	4.17(0.86)	4.4(0.89)	3.98(0.93)	4.3(0.8)
Slow	3.72(0.78)	4.02(0.74)	3.59(0.87)	3.85(0.96)
Latency				
Fast	79.27(62.23)	52.39(53.07)	83.28(71.14)	55.54(40.18)
Slow	56.22(43.92)	64.22(61.86)	100.35(29.06)	62.88(50.39)

The average motivation is displayed in Table 11. Examining the data, we can see that motivation goes up with a slow impasse for insight problems and stumper problems and down for analytical and mathematical problems. A likelihood ratio test indicated that the main effect of cognitive speed was not significant, $\chi^2(1) = 0.01$, $p = .941$. Hypothesis 1c was not supported.

d. Lower affect than when it is described as slow

The average affect is displayed in Table 11. Affect was highest in the fast trials and lowest in the slow trials across all problem types. A likelihood ratio test indicated that the main effect of cognitive speed was significant, $\chi^2(1) = 32.30$, $p < .001$, Hypothesis 1d was upheld. There is a relationship between experiencing impasse as a cognitively speedy and positive affect.

e. The same latencies than when it is described as slow

The average latency to a correct solution is shown in Table 11. Latency was broadly similar across fast and slow cognitive speeds, and an LRT suggests that the model with cognitive speed as a main effect, $\chi^2(1) < 0.01$, $p = .964$, is not a better fit than the null. Hypothesis 2e was supported – cognitive speediness is a subjective feeling rather than a reflection of how objectively fast the problems are actually solved.

7.4. Discussion

The majority trends established in the first three experiments were replicated across the problem types. All hypotheses relating to the impasse stage were upheld or partially upheld. Resolving the impasse stage did not have a catastrophic affect on solution rates (although it was still better to not encounter impasse at all) and solutions were obtained and with the same level of insight. The latencies were longer when experiencing impasse whether resolved or not confirming a behavioural element to the impasse experience. Experiencing impasse did not seem to influence motivation, although it was associated overall with lower affect. This latter finding suggests that motivation can still be present in the face of failure and low affect.

A feeling of cognitive speediness was associated with higher solution rates. The occurrence of insight and the strength of insight followed the same pattern as in Experiment 2. Unexpectedly, cognitive speediness and motivation were not associated but speediness was associated with higher affect. Again, the feeling of cognitive speediness and actual latency to solution was not related. Thus, the feeling of cognitive speediness although not actually leading to faster solution times did lead to a higher solution rate and a more pleasant problem-solving experience with higher levels of insight.

8. General discussion

The current conceptualisation of impasse in problem-solving as a simple aversive state leading to problem failure leads to two paradoxes. First, impasse is both required for problem solving with phenomenon of insight and yet also likely to lead to failure to solve. Second, this state is regularly sought out by people who enjoy recreational problem-solving or gaming. A series of experimental studies presents evidence that goes some way to resolving these paradoxes; first by showing that the experience of impasse is multi-dimensional and can be predictive of both failure and success, and second, by revealing a separation between motivational, cognitive and affective dimensions of the impasse experience. In Experiment 1, when impasse was measured post-task as a binary construct as it is traditionally, it was related to problem failure. Even if participants were successful, impasse was not clearly associated with a feeling of insight, suggesting that it is not a necessary precursor for insight in insight problem-solving theories. However, when (in Experiments 2 to 4) we introduced the small change into the post-task question that acknowledged that the impasse state can change across the task, the link between insight and impasse was reliably restored, and impasse was no longer catastrophic, suggesting that it still holds importance. Rather, there are two forms of impasse – one which leads to problem failure and one which leads to success. Behaviourally, both these stages lead to significantly longer latencies to solution than not experiencing impasse even as they differ in cognitive and affective outcomes.

Along with underlining the importance of considering the dynamics of impasse in problem solving, we also introduced a series of questions in Experiment 3 that explored the findings with regards to the epistemic feelings extracted from the qualitative reports in Experiments 1 and 2. Responses to these questions indicate there are two separate factors for impasse experience; affective and motivational. These factors vary differently across stages of impasse: affect changes depending on the resolution or otherwise of impasse, while motivation remains stable whether impasse is resolved or not. In addition, we established that the feeling of cognitive speediness also varied in a way that was not related to actual latencies to a correct answer. The variability in the experience of impasse appears to be related to impasse resolution and problem success. Unpacking the relations between the multi-factorial experiences of impasse and the outcomes of problem-solving performance may offer promising avenues for understanding how to generate a *productive* state of impasse.

The findings were broadly replicated across different time conditions and problem types, suggesting that there is a robust underlying pattern to this cognitive state. The fast and slow distinction we drew upon echoes Danek's (2018) model of insight problem solving which separates "repeated failure" and "impasse", but we suspect that the process is more dynamic than this model suggests or that our findings can so far illustrate. What is clear is that describing the state of an impasse as a necessarily aversive state is unlikely to capture its full complexity. In addition, negatively valenced affect does not necessarily have a catastrophic impact on motivation. The cognitive implications are clear: traditional methods of understanding cognitive states such as response times or solution rates do not capture differing affective–cognitive dynamics leading to the same behavioural outcomes. Complex cognitive processes require an understanding of the accompanying feelings.

8.1. Implications for understanding insight problem solving

Although there is a focus in problem-solving on the moment of insight and its underlying cognitive correlates, it is worth noting that the *feeling* of insight is not conceptually necessary although in the past, it has been deemed methodologically useful for discriminating different cognitive paths. This does not mean that the phenomenon of insight is not interesting: Its metacognitive role appears to be an important indicator of confidence and may have other motivational and epistemic effects. However, a focus on this feeling means that we have moved away from research into how people resolve the state of impasse to how they experience insight. This is a subtle but very important shift which has left the field of impasse-driven problem solving in a state of neglect.

However, models of "true" insight problem solving (Weisberg, 2015) *do* require impasse even while they do not require the feeling of insight. The findings we report here suggest that there is more than one route to the same feeling of insight: impasse-driven or non-impasse-driven. It seems likely that understanding how to resolve or avoid an impasse may be more profitable for interventions to improve creative thinking than investigating the dynamics that lead us to experience insight, not least because one appears more instantly tractable than the other; we can more reliably elicit impasse than insight.

At its heart, the domain of insight problem-solving is tasked with understanding how people solve hard problems, in other words how they resolve the state of impasse. This is what the two competing models (CSPT and RCT) are aiming to explain. For both, the switch – either from one maximising strategy to another or from a faulty representation to helpful one – occurs in the stage of impasse. Therefore, while our data do not allow us to adjudicate more clearly between these two models, returning the focus to the state of impasse is likely to support a return to these theories. Interestingly, only CPST does not require an impasse to be accompanied by the feeling of being stuck, describing it as the simple failure of a maximising heuristic rather than the cessation of problem-solving activity. That these forms of problems can be solved *without* experiencing impasse may support those to newer and updated models that can embrace a more incremental solution process. However, it may also be that only more integrated theories (e.g. Weisberg, 2015) can reflect the idiosyncratic pathways to solution our data suggest.

8.2. Implications for other domains

The complex nature of unmerited impasse has broader implications than may at first appear. Impasse has been downplayed in the problem-solving literature because it is associated with failure through omission, in other words, giving up. Giving up carries a low cost in experimental tasks, especially multi-trial tasks within a short time period. However, for many other tasks, giving up in the face of uncomfortable feelings that accompany a state of impasse carries a higher cost and is not always an option. Indeed, the key role of unmerited impasse is not confined to traditional problem-solving research. Since the 1980s, there has been an increase in the incorporation of problem-solving strategies across educational systems, especially in the domain of mathematics (Liljedahl, Santos-Trigo, Malaspina, & Bruder, 2016; Verschaffel, Schukajlow, Star, & Van Dooren, 2020). Here, the focus is on the application of learned knowledge; in theory, problem solvers have the knowledge required to solve the problem but are still stuck. Thus, these experiments, which aim to increase our understanding of impasse as an affective, motivational, and cognitive state, may also lead to a clearer understanding of the ways to better scaffold success also in these areas.

Alongside this, the current findings indicate that there is a complex interplay between resolving impasse and affective and motivational states. It is notable that not resolving impasse did not seem to affect the feelings of motivation, although it had an impact on the overall affect. This suggests that affect and motivation may not be as tightly linked as previously thought. Understanding what motivates people to persevere even when they encounter negative affect is important not only for creativity and innovation but also for fields such as game design and education. A focus on how an impasse is resolved and the role of motivation in this may reorient research efforts to how people solve difficult problems.

8.3. Limitations

Post-task reports are convenient. This is especially true when administering increasing amounts of online research. We have already noted how the exigencies of neuroscientific investigations into insight have led to a focus on the feeling of insight rather than experiencing impasse. With something as dynamic as the feeling of impasse, post task reports are a necessarily limited tool. They compress the problem-solving experience into a smaller space and rely on the metacognitive skills of the participant. As Danek et al. (2014) caution in their work on the dynamics of insight, we cannot know the basis upon which participants have made their judgments. The current work needs to be substantiated with behavioural measures and paradigms that can test the dynamics of problem-solving experience. However, the main findings here have replicated across different task types and experimental conditions and this gives us confidence in their robustness. While we acknowledge the contradiction, we provide a solid foundation for further exploration.

In addition, the relationship between the feeling of impasse and the feeling of insight is likely to be further complicated when the full spectrum of the insight experience is taken into account. In this paper, while we measured the strength of insight, we did not take into account the different facets that make up the feeling of insight. Danek et al. (2014) suggest that there are five aspects – impasse, suddenness, certainty, surprise and happiness. It may be that a closer look at these aspects of insight in relation to the dimensions of impasse elucidated in this paper can provide a highly granular yet generalisable pathway to problem solution.

8.4. Directions for future research

Understanding the dynamics of impasse experiences opens up the opportunity to develop a manipulation (affective, motivational, embodied, etc.) that can affect the development of a state of impasse. If, for example, the feeling of cognitive speediness is associated with higher affect and increased solution rates, future research should focus on eliciting this cognitive feeling perhaps through regular motivational interventions in the process. This would be particularly important when it comes to longer and more complex problems in which the option to "give up" is less easy. Indeed, this motivation could also come from the level of intrinsic motivation to solve the problem. We therefore suggest moving from the tightly constrained problems we used in this series of studies to more complex problems or those where the problem solvers are more highly motivated. Alongside this, propensity to engage in this form of problem-solving can also be investigated (Novick & Sherman, 2003).

In addition, the feeling of impasse is related to faulty metacognitive monitoring – the affective markers suggest that the participant does not have the necessary knowledge when they do - but it is not clear what the direction of monitoring is. Does the affective state reflect or drive the cognitive? Our data hint at a feedback loop in which affective and cognitive states feed into each other but this would need to be supported by further research.

9. Conclusion

The relationship between theory and empirical evidence is complex. In the case of the role of impasse in insight problem solving, the tension between theory and empirical data led to paradoxical claims. Relying only on the "feeling of being stuck" to describe the impasse state misrepresents the complexity of the feelings that the subjects actually experience during this phase. This study represents a first attempt to provide clarity regarding the emotional, motivational, and cognitive complexity of impasse, which is necessary to structure further investigations on how insight and creative reasoning and solutions emerge.

CRediT authorship contribution statement

Wendy Ross: Writing – review & editing, Writing – original draft, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Selene Arfini:** Writing – review & editing, Writing – original draft, Project administration, Methodology, Data curation, Conceptualization.

Data availability

All datasets and analysis codes are available on the OSF: https://osf.io/qbuv8/?view_only=d1c6703968e34f1285427542ffff7c7d

Acknowledgements

We thank the audience of the AISC 2022 midterm conference for thoughtful discussions on an early draft of the data presented here.

We also thank Noemi Hrvatin, Elena Tobol, Raul Scarabusci, Nafia Sherazi and Diana Tudoroiu for support coding the qualitative studies.

The raw data and analysis scripts alongside the preregistrations can be found here: https://osf.io/qbuv8/?view_only=d1c6703968e34f1285427542ffff7c7d

These experiments investigated several broad research questions; aspects of the data collected in these experiments that are not relevant to this manuscript will be reported elsewhere. In addition to the empirical work reported here and elsewhere, two complementary theoretical manuscripts (Arfini & Ross, 2023, in prep) have been prepared.

The research reported in this manuscript was made possible by internal funding from London Metropolitan University (awarded to WR) and the University of Pavia (awarded to SA).

Appendix A. Stumper used in experiment 1 and 2

Dame Dora owns an Old Masters painting in a heavy gilded frame. The cord for hanging the painting, as old as the painting itself, is made of thick 3-ply hemp, and is somewhat frayed. Dame Dora was thinking of replacing it. But before she could, a couple of hungry little mice invaded her mansion. Sneaking behind the painting, they chewed right through the cord. For a while nobody noticed because the painting didn't budge. Explain the painting's stability briefly (from (Bar-Hillel, 2021; Ross & Vallée-Tourangeau, 2022)

Appendix B. Normative Data for Stumpers Used in Experiment 3 from Ross and Vallée-Tourangeau (2022) Split in Terms of Correct and Incorrect Trials (Standard Deviations Are Reported in Brackets)

		Correct						Incorrect					
	Proportion Correct	Latency		Confidence		Aha		Latency		Confidence		Aha	
Hard Problems													
A hungry horse is tied by its neck to a 10-m-long chain. A bale of hay is 13.8 m away from it. Explain briefly how the horse reaches the hay with the chain intact.	0.43	48.51	(17.21)	57.06	(29.16)	43.08	(31.19)	60.06	(18.63)	29.33	(25.12)	23.71	(22.03)
Dame Dora owns an Old Masters painting in a heavy gilded frame. The cord for hanging the painting, as old as the painting itself, is made of thick 3-ply hemp, and is somewhat frayed. Dame Dora was	0.43	56.59	(16.84)	44.92	(31.56)	33.05	(25.61)	64.31	(16.91)	25.02	(26.15)	21.54	(25.31)

(continued on next page)

(*continued*)

	Proportion Correct	Correct						Incorrect					
		Latency		Confidence		Aha		Latency		Confidence		Aha	
thinking of replacing it. But before she could, a couple of hungry little mice invaded her mansion. Sneaking behind the painting, they chewed right through the cord. For a while nobody noticed because the painting didn't budge. Explain the painting's stability briefly.													
A man in town married 20 women in the town. He and the women are still alive, and he has had no divorces. He is not a bigamist and is not a Mormon and yet he broke no law. How is that possible?	0.44	40.65	(20.33)	75.83	(26.12)	64.56	(27.33)	58.65	(18.14)	25.78	(30.31)	21.67	(25.72)
Denise is a pretty good tennis player. She made a bet that she could hit a regular tennis ball, send it flying off in the air, and after a bit, it would turn around 180 degrees and fly right back to her – without making contact with any other object on its way. She won the bet. Explain how in a few sensible words.	0.46	54.92	(18.18)	69.26	(30.83)	53.21	(31.53)	60.74	(17.89)	29.97	(27.08)	26.62	(27.16)
Easy Problems													
Bob's driver's license was recently revoked, following a string of severe traffic violations. Just a few days later, a cop spotted the unlicensed Bob yet again, entering a one-way street against the direction of the traffic. This was the same cop who had cited Bob before. Explain briefly how come the cop did not stop him, and just gave him a smile.	0.60	39.14	(16.32)	71.45	(25.79)	49.81	(29.90)	57.85	(16.90)	35.21	(31.27)	38.52	(31.19)
A clerk at a butcher shop stands five feet ten inches tall and wears size 13 sneakers. What does he weigh?	0.62	32.46	(16.75)	67.67	(29.75)	62.20	(30.97)	51.63	(21.01)	20.16	(22.97)	20.04	(26.48)
Farmer Joe eats two fresh eggs from his own farm for breakfast every day. Yet there are no chickens on his farm. Where does Farmer Joe get his eggs?	0.62	43.83	(19.90)	44.36	(30.89)	29.28	(26.71)	42.14	(20.99)	47.17	(33.06)	35.38	(27.1)
Laura took a multiple-choice test. She barely speaks, reads, or understands English, and had nobody who could translate for her. Explain briefly how Laura scored nearly 100% on the test, completely legitimately.	0.62	38.64	(17.80)	65.25	(27.67)	45.04	(28.36)	44.30	(19.01)	37.96	(28.45)	24.08	(26.31)

Appendix C. Problems used in experiment four

Riddle	Type	Source
A dealer of antique coins received an offer to buy a beautiful bronze coin by an unknown man. The coin had an emperor's head on one side and the date 544 BCE stamped on the other side. The dealer examined the coin, but instead of buying it, he called the police to arrest the man. What made him realise that the coin was fake?	Insight	Webb et al. (2016)
The police were convinced that either A, B, C, or D had committed a crime. Each of the suspects, in turn, made a statement, but only one of the four statements was true. A said, "I didn't do it." B said, "A is lying." C said, "B is lying."	Analytical	Webb et al. (2016)

(*continued on next page*)

(*continued*)

Riddle	Type	Source
D said, "B did it." Who is telling the truth? and who committed the crime?		
Marcus thinks of a number between 25 and 35. He divides the number by 2 and then subtracts 0.5. He takes his answer, divides it by 2 and then subtracts 0.5. He repeats this process a number of times and gets zero. What number did he start with?	Maths	AQA GCSE Mathematics (2015)
Dame Dora owns an Old Masters painting in a heavy gilded frame. The cord for hanging the painting, as old as the painting itself, is made of thick 3- ply hemp, and is somewhat frayed. Dame Dora was thinking of replacing it. But before she could, a couple of hungry little mice invaded her mansion. Sneaking behind the painting, they chewed right through the cord. For a while nobody noticed because the painting didn't budge. Explain the painting's stability briefly.	Stumper	Ross and Vallée-Tourangeau (2022)

References

Arango-Muñoz, S. (2014). The nature of epistemic feelings. *Philosophical Psychology, 27*(2), 193–211. https://doi.org/10.1080/09515089.2012.732002

Arfini, S., & Ross, W. (2023). Feeling stuck. Epistemic and emotional dimensions of impasse and doubt. *Sistemi intelligenti, 2023*(3), 587–610.

Ball, L. J., & Litchfield, D. (2013). Interactivity and embodied cues in problem solving, learning and insight: Further contributions to a "theory of hints". In S. J. Cowley, & F. Vallée-Tourangeau (Eds.), *Cognition beyond the brain* (pp. 232–239). Springer.

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.

Bar-Hillel, M. (2021). Stumpers: An annotated compendium*. *Thinking & Reasoning, 27*(4), 536–566. https://doi.org/10.1080/13546783.2020.1870247

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1). doi: 10/gcrnkw.

Beeftink, F., van Eerde, W., & Rutte, C. G. (2008). The effect of interruptions and breaks on insight and impasses: Do you need a break right now? *Creativity Research Journal, 20*(4), 358–364. https://doi.org/10.1080/10400410802391314

Bowden, E., & Jungbeeman, M. (2007). Methods for investigating the neural components of insight. *Methods, 42*(1), 87–99. https://doi.org/10.1016/j.ymeth.2006.11.007

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

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

Brown, V. A. (2021). An introduction to linear mixed-effects modeling in R. *Advances in Methods and Practices in Psychological Science, 4*(1). https://doi.org/10.1177/2515245920960351, 251524592096035.

Carruthers, P. (2017). Are epistemic emotions metacognitive? *Philosophical Psychology, 30*(1–2), 58–78. https://doi.org/10.1080/09515089.2016.1262536

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

Danek, A. H. (2018). Magic tricks, sudden restructuring and the aha! Experience. In F. Vallée-Tourangeau (Ed.), *Insight: On the origin of new ideas* (pp. 51–79). Routledge.

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

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

Danek, A. H., & Salvi, C. (2020). Moment of truth: Why Aha! Experiences are correct. *The Journal of Creative Behavior, 54*(2), 484–486. https://doi.org/10.1002/jocb.380

D'Mello, S., & Graesser, A. (2012). Dynamics of affective states during complex learning. *Learning and Instruction, 22*(2), 145–157. https://doi.org/10.1016/j.learninstruc.2011.10.001

Dominowski, R. L., & Bourne, L. E. (1994). CHAPTER 1—History of research on thinking and problem solving. In R. J. Sternberg (Ed.)*, Vol. 2. Thinking and problem solving* (pp. 1–35). Academic Press. https://doi.org/10.1016/B978-0-08-057299-4.50007-4.

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

Evans, J. S. B. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology, 59*, 255–278.

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. (2004). The use of verbal protocols as data: An analysis of insight in the candle problem. *Memory & Cognition, 32*(6), 990–1006. https://doi.org/10.3758/BF03196876

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

Gilhooly, K. J. (2019). *Incubation in problem solving and creativity: Unconscious processes*. Routledge.

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

Knoblich, G., Ohlsson, S., Haider, H., & Rhenius, D. (1999). Constraint relaxation and chunk decomposition in insight problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(6), 1534–1555. https://doi.org/10.1037/0278-7393.25.6.1534

Kreidler, S. M., Muller, K. E., Grunwald, G. K., Ringham, B. M., Coker-Dukowitz, Z., Sakhadeo, U. R., … Glueck, D. H. (2013). GLIMMPSE: Online power computation for linear models with and without a baseline covariate. *Journal of Statistical Software, 54*(10). https://doi.org/10.18637/jss.v054.i10

Laukkonen, R. E., & Tangen, J. M. (2018). How to detect insight moments in problem solving experiments. *Frontiers in Psychology, 9*, 282. https://doi.org/10.3389/fpsyg.2018.00282

Laukkonen, R. E., Webb, M. E., Salvi, C., Tangen, J. M., Slagter, H. A., & Schooler, J. (2022). Eureka heuristic: How feelings of insight signal the precision of a new idea [preprint]. *PsyArXiv.*. https://doi.org/10.31234/osf.io/ez3tn

Liljedahl, P., Santos-Trigo, M., Malaspina, U., & Bruder, R. (2016). Problem solving in mathematics education. In P. Liljedahl, M. Santos-Trigo, U. Malaspina, & R. Bruder (Eds.), *Problem solving in mathematics education (pp. 1–39)*. Springer International Publishing. https://doi.org/10.1007/978-3-319-40730-2_1.

MacGregor, J. N., Ormerod, T. C., & Chronicle, E. P. (2001). Information processing and insight: A process model of performance on the nine-dot and related problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(1), 176–201. https://doi.org/10.1037//0278-7393.27.1.176

Newell, A., Shaw, J. C., & Simon, H. A. (1958). Elements of a theory of human problem solving. *Psychological Review, 65*, 151–166. https://doi.org/10.1037/h0048495

Novick, L. R., & Sherman, S. J. (2003). On the nature of insight solutions: Evidence from skill differences in anagram solution. *The Quarterly Journal of Experimental Psychology, 56*(2), 351–382. https://doi.org/10.1080/02724980244000288

Novick, L. R., & Sherman, S. J. (2008). The effects of superficial and structural information on online problem solving for good versus poor anagram solvers. *Quarterly Journal of Experimental Psychology, 61*(7), 1098–1120. https://doi.org/10.1080/17470210701449936

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

Ormerod, T. C., MacGregor, J. N., & Chronicle, E. P. (2002). Dynamics and constraints in insight problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(4), 791–799. https://doi.org/10.1037//0278-7393.28.4.791

Perkins, D. N. (2001). The eureka effect: The art and logic of breakthrough thinking. *Norton*.

Polanyi, M. (1957). Problem Solving. *The British Journal for the Philosophy of Science, 8*(30), 89–103.

Przybylski, A. K., Rigby, C. S., & Ryan, R. M. (2010). A motivational model of video game engagement. *Review of General Psychology, 14*(2), 154–166. https://doi.org/10.1037/a0019440

R Core Team. (2023). *R: A language and environment for statistical computing. [Computer software]*. R Foundation for Statistical Computing. https://www.R-project.org/.

Ross, W. (2021). Feeling stumped: Investigating dimensions of impasse [preprint]. *PsyArXiv.*. https://doi.org/10.31234/osf.io/4zhse; https://web.archive.org/web/20211209194922/https://psyarxiv.com/4zhse/

Ross, W., & Vallée-Tourangeau, F. (2021a). *Accident and agency: A mixed methods study contrasting luck and interactivity*. Thinking & Reasoning.

Ross, W., & Vallée-Tourangeau, F. (2021b). Kinenoetic analysis: Unveiling the material traces of insight. *Methods in Psychology.*. https://doi.org/10.36850/e4

Ross, W., & Vallée-Tourangeau, F. (2021c). Microserendipity in the creative process. *Journal of Creative Behavior, 55*(3), 661–672.

Ross, W., & Vallée-Tourangeau, F. (2022). Insight with stumpers: Normative solution data for 25 stumpers and a fresh perspective on the accuracy effect. *Thinking Skills and Creativity, 101114*. https://doi.org/10.1016/j.tsc.2022.101114

Shen, W., Yuan, Y., Lu, F., Liu, C., Luo, J., & Zhou, Z. (2019). Unpacking impasse-related experience during insight. *The Spanish Journal of Psychology, 22*. https://doi.org/10.1017/sjp.2019.40

Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. (2023). *afex: analysis of factorial experiments* (R package version 1.3-0) [computer software]. https://CRAN.R-project.org/package=afex.

Sousa, R. (2009). Epistemic feelings. *Mind and Matter, 7*(2), 139–161.

Spiridonov, V., Loginov, N., & Ardislamov, V. (2021). Dissociation between the subjective experience of insight and performance in the CRA paradigm. *Journal of Cognitive Psychology, 1–15*. https://doi.org/10.1080/20445911.2021.1900198

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

Terpe, S. (2016). Sentiments epistemològics en les experiències morals i dinàmiques morals en la vida quotidiana. *Digithum, 18*. https://doi.org/10.7238/d.v0i18.2874. Article 18.

Thompson, V. A., Prowse Turner, J. A., & Pennycook, G. (2011). Intuition, reason, and metacognition. *Cognitive Psychology, 63*(3), 107–140. https://doi.org/10.1016/j.cogpsych.2011.06.001

Verschaffel, L., Schukajlow, S., Star, J., & Van Dooren, W. (2020). Word problems in mathematics education: A survey. *ZDM, 52*(1), 1–16. https://doi.org/10.1007/s11858-020-01130-4

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, S. 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

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