Spontaneous use of external resources in verbal problem solving is rare but beneficial

Wendy Ross (w.ross@londonmet.ac.uk)
Department of Psychology, London Metropolitan University, London, N7 8DB.

Selene Arfini (selene.arfini@unipv.it)
Department of Humanities – Philosophy Section, University of Pavia, Piazza Botta 6 - 27100, Pavia, Italy.

Abstract

There are two foundational assumptions that underlie research in interactivity. First, that resources external to the human agent should support problem-solving and other cognitive activities and second, that human agents naturally engage in this form of offloading when they are allowed to. We aimed to test whether participants would naturally engage with external resources, without prompting, in four types of simple verbal problems and whether the level of engagement was affected by expertise or the experience of impasse. We found that very few people naturally engaged external resources apart from with mathematical problems where it had a benefit. There was no difference in expertise in problem-solving between those who did and those who did not use external props and nor was there a significant difference in the proportion of people using external resources as a function of experiencing impasse. These results suggest that researchers in interactivity need to focus on how and when interactivity is both engaged and provides a benefit.

Keywords: interactivity; problem solving; extended mind; impasse; external resources.

In the past 20 years, theories of extended cognition have posed a challenge to traditional internalist views of thinking. The Extended Mind Thesis (EMT; Clark & Chalmers, 1998) argues that objects in the environment act as cognitive extensions. In the classic paper introducing this theory, for example, the external rotation of Tetris shapes replaced internal rotations of mental representations (drawing on evidence from Kirsh & Maglio, 1994). Other forms of extensions have been proposed from scaffolding memory (Sutton, 2005) to mind wandering (Bruineberg & Fabry, 2022).

Under these frameworks, cognition has been seen as something necessarily dependent on or constituted by bodily processes or aspects of the external world. The argument from proponents of this view is that a position of methodological solipsism, that is assuming that thinking can be researched in isolation from the surrounding environment (Fodor, 1980), would not accurately reflect cognitive processes. Instead, cognition should be seen as a systemic activity dependent on more than simply skull-bound activity (F. Vallée-Tourangeau & G. Vallée-Tourangeau, 2020).

The theories of EMT have most commonly found their way into research on problem-solving through interactivity – that is the recruitment of external resources to support complex cognitive processes. For example, in research on mental arithmetic, addition has been scaffolded by movable numbered tokens representing the numbers (Ross, F. Vallée-Tourangeau & Van Herwegen, 2020). Other times, the external resources are more complex such as when participants are presented with model animals and pipe cleaners to solve a problem involving animals and pens (Steffensen, F. Vallée-Tourangeau & G. Vallée-Tourangeau, 2016).

Experimental research in the tradition of interactivity tends to contrast two conditions: A high interactivity condition which presents participants with movable artifacts which can be recruited to support their thinking and a low interactivity condition which restricts interaction in some way. The levels of restriction vary from participants sitting on their hands to being restrained from moving objects but otherwise able to move freely. The assumption is that the behaviour of participants varies in line with the conditions although this has only been verified recently (G. Vallée-Tourangeau, Abadie & F. Vallée-Tourangeau 2015) and, as might be expected, verification has shown the experimental manipulation has been inconsistent across participants (Ross & F. Vallée-Tourangeau, 2021a). Indeed, Maglio, Matlock, Raphaely, Chernicky & Kirsh (1999) report that, in an interactive word generation task, roughly one third of the participants did not, in fact, use their hands or used their hands very briefly despite being at complete liberty to do and despite this being the key experimental manipulation.

The aim of this research programme has been to show that moving external representations of the problem components confers a benefit to the problem-solver. These findings have been robustly established in the domain of mental arithmetic in both adults (F. Vallée-Tourangeau, Sirota & G. Vallée-Tourangeau 2016) and children (Ross et al, 2020) and also in planning (Guthrie, F. Vallée-Tourangeau, G. Vallée-Tourangeau & Howard 2015) and word finding tasks (F. Vallée-Tourangeau & Wrightman, 2010). The well-defined nature of these tasks means that external props are likely to accrue a benefit as placeholders or cognitive offloads. Participants in these tasks often know what to do but lack the computational capacity to do so efficiently.

However, the evidence has been less conclusive in the domain of insight problem-solving. Insight problem solving refers to a class of problems structured in such a way to lead the participant to an impasse which can only be broken by trying something different. While there have been some reports of interactivity conferring a benefit in the form of problem-solving (Henok, F. Vallée-Tourangeau & G. Vallée-

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society*.

Tourangeau, 2020; F. Vallée-Tourangeau, Ross, Ruffatto Rech and G. Vallée-Tourangeau 2020; F. Vallée-Tourangeau, 2014) other studies have shown no benefit (Chuderski, Jastrzębski & Kucwaj, 2020) and indeed, in some case interactivity hinders problem solution (Ross & F. Vallée-Tourangeau, 2021b).

A likely explanation for this is that insight problem-solving concerns problem-solving activity on a class of problems specifically designed to lead the problem-solver down an unhelpful solution route. These problems differ from more structured or analytical problems because the problem-solver is often at an impasse caused by not knowing what to do. In contrast to problems where the participant is stymied by a lack of cognitive resources, insight problems are "easy once you know the answer". If a problem-solver knew what to recruit to solve the problem, then the problems would be trivial. Therefore, the use of external resources is unlikely to support problem-solving through the same mechanisms as in well-defined problems.

However, to counter this explanation, insight problems tend to elicit an unmerited impasse – broadly the feeling of being stuck. There is some speculation that the state of impasse leads participants to naturally extend their problem-solving into the external environment. Interaction with concrete things has been seen to break the fixedness that can result from an impasse (F. Vallée-Tourangeau, Euden & Hearn, 2011) and also generate felicitous unplanned and unthought of moves (Ross & F. Vallée-Tourangeau, 2022). Therefore, while human agents may direct attention to internal operations when they are making progress, it is plausible that at impasse, they will recruit more resources.

To further complicate matters, there appears to be an interaction between expertise and the use of interactivity in experimental situations. For example, Ross et al (2020) found that the benefits of interactivity in a mathematical task accrued to those who were higher in numerical skill and, in a naturally occurring situation, Ormerod and Gross (2023) found that experts were more likely to engage with external resources and use them to structure their thinking. So, it seems likely that expertise in a domain affects both the tendency to engage with external resources and the benefits of so doing.

A final complication is that the problem types in this research have been selected to maximize the benefits of interactivity and can be easily translated into abstracted movable representation types. For example, research on insight problem-solving has tended towards problems that can be represented with movable objects such as those that require participants to rearrange coins or links in a necklace with varying constraints. However, while sometimes problems encountered outside the laboratory have this structure (for example, Watson claims to have discovered the structure of the double helix through this kind of interactive manipulation of cardboard representations; Watson & Stent, 1998), other problems are more complex and are not easily reduced to external representations in this way. In order to test the claim that cognition naturally extends, research in this area needs to establish if this claim holds when the problem and environment are less facilitatory.

In the current study, we had a very simple aim but one we argue is essential to progress research in this field. We aimed to test whether participants would naturally engage with external resources without prompting in four types of simple verbal problems. We also aimed to assess whether the use of external resources was influenced by experiencing impasse or by the expertise of the problem-solver. This allowed us to test the following hypotheses:

H1: The natural use of external resources would yield a performance benefit.
H2: Participants would be more likely to rely on external resources when they encountered impasse.
H3: There will be a relationship between expertise and use of external resources. we did not hypothesize a direction because it is plausible that experts will not be sufficiently challenged by the problems to rely on scaffolds or equally may recruit resources more than non-experts.

Method

Participants

We 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.

Materials and Measures

We selected four different types of problems: A verbal insight problem, an analytical problem, a mathematical and a form of riddle known as a stumper. The full wording of each problem can be seen in Table 1.

Participants were also profiled on their expertise, borrowing from a modified version of the expertise scale in Novick & Sherman (2003; see also Ross & Vallée-Tourangeau, 2021). They were asked to rate how often they solved problems (every day, a few times a week, once a week, a few times a month, once a month, a few times a year, never) and how they would rate their anagram expertise in relation to other students (on a scale of 1 to 10 with 10 being average). Following Novick and Sherman, the two items were collapsed to create a composite experience score.

Procedure

After being asked for demographic information and to complete the modified expertise scale, participants were given a maximum of 4 minutes to solve each of the problems. The order of problems was counterbalanced across participants.

After each problem participants were asked if they felt stuck at any point during the problem-solving process and were offered three options (a) no, (b) yes but I got over it and (c) yes and I still am. These three options represent no impasse, resolved impasse and unresolved impasse (Ross & Arfini, 2024). For the purpose of our research question, we

Table 1: Experimental stimuli used and the source

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 B.C. 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." D said, "B did it." Who is telling the truth? and who committed the crime?	Analytical	Webb et al. (2016)
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 & Vallée-Tourangeau (2022)

collapsed these into experienced impasse (resolved or unresolved) or did not experience impasse.

At the end of the presentation of all four problems, participants were asked if they used external props with the following wording:

> When you were trying to solve the problems, did you use anything else such as a pen and paper? There are no right or wrong answers and it doesn't affect your payment

They were then given the option to select "yes" or "no" for each of the four problem types. They were then thanked and debriefed.

Results

Proportion of correct answers

Over half of the participants got the correct answer for the stumper (54%) followed by 45% who got the correct answer for the insight problem. Solution rates were lower for the mathematical (34%) and analytical problems (26%).

The use of external scaffolds during problem solving

Very few participants reporting using external support during the problem-solving process although this varied by problem types. Only one participant reported using external support for the stumper (< 0%), 2% used support for the insight problem, 8% for the analytical problem but 30% of participants reported using scaffolds for the mathematical problem. We constructed a linear mixed model with type of problem as a fixed effect and participant as random slope and intercept. A Likelihood Ratio Test suggested that this was a better fit for the data, $\chi^2(3) = 169.87$, p <.001 than a null model. Post hoc tests with a Bonferroni correction show that there was a significant difference between the mathematical problem and all other problems (all p_{Bonf} <.001), a significant difference between the analytical problem and the stumper (p_{Bonf} = .011) but not between the analytical problem and the insight problem (p_{Bonf} = .087). There was no significant difference in use of external props between the insight and stumper task (p_{Bonf} > .999).

The effect of external scaffolds on solution rates

Figure One shows the effect of external scaffolds on solution rates across problem types. Solution rates were generally higher when props were used although they were lower for the insight task.

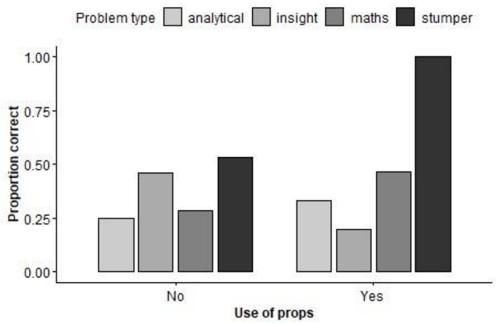

Figure 1: The proportion of correct answers as function of prop use and type of problem.

The numbers of people opting to use external scaffolds in all conditions save the mathematical problem question were too low for statistical analysis. However, for the mathematical problem, opting to use external props generated a significant benefit, $F(1, 225) = 7.01$m, $p = .009$, $\eta^2 = .03$.

Impasse and external scaffolds

The proportion of participants using props for each of the problem types as a function of whether they experienced impasse is displayed in Figure Two.

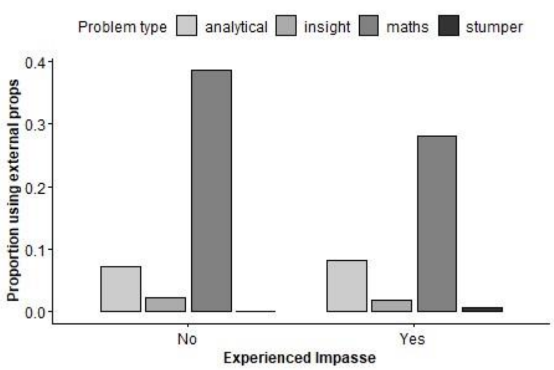

Figure 2: The proportion of participants using props as function of problem type and experiencing impasse.

We constructed a linear model with experiencing impasse as a fixed effect and problem type and participant as random slopes and intercept. A LRT suggested that this was not a better fit for the data, $\chi^2(1) = 0.86$, $p = .353$. We repeated the analysis for solely the mathematical problem and there was also no difference in use of external props whether the participant experienced impasse or not, $F(1, 225) = 2.07$, $p = .151$, η^2 <..001. Hypothesis Two was not supported.

Expertise and external scaffolds

The average level of expertise as a function of whether props were used or not is displayed in Figure Three.

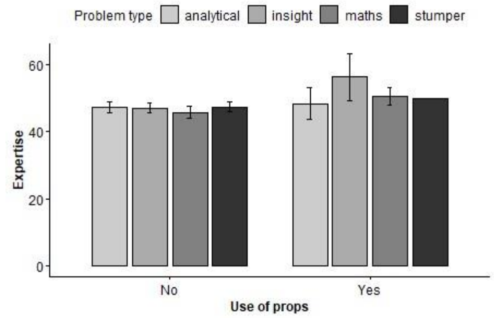

Figure 3: Average expertise across those who used props and those who didn't as a function of problem type (error bars represent SEM).

We constructed a linear model with expertise as a fixed effect and problem type and participant as random slopes and intercept. The effect of expertise was not a significant contributor to the model, $\beta <.001$, $p = .241$. We repeated the analysis for solely the mathematical problem and expertise was again not a significant predictor, $\beta = .010$, $p = .220$. Hypothesis Three was not supported.

Discussion

The current set of results showed that when given free choice very few of our participants opted to use external scaffolds. However, there was not a floor effect: One third of participants opted to use external support for the mathematical problem. This suggests that while the experimental situation did not encourage interactivity, it did not prevent it. When they used external props, performance tended to be better. The use of props was not predicted by experiencing the state of impasse or by the level of problem-solving skill of the participants. Research into interactivity in problem-solving to date has assumed that when allowed, participants would naturally gravitate towards using external props. The research reported here demonstrates that this assumption deserves closer examination, and that any observed tendency may relate to the affordances of the experimental situation.

The results here may appear trivial. The effort of switching from a computer-based task to pen and paper is clearly greater than remaining within the pen and paper environment and it is not unreasonable that people avoid that. However, it is increasingly the case that people interact with a digital environment without easy access to external resources. Consider, for example, the computer-focused nature of many current jobs. Although the experimental situation demonstrated a lack of clear control, these results do give us an indication of how human agents interact with external resources in their everyday lives.

Heuristics and interactivity

In short, these results support the idea that interactivity is worth pursuing when the costs do not outweigh the benefits (Kirsh, 2009; Rowlands, 2010). In those cases, it is likely subjects used what in the ecological rationality literature is known as a fast and frugal heuristic (Gigerenzer & Goldstein, 1996), which is a cognitive strategy used to "exploit the representation and structure of information in the environment to make reasonable judgments and decisions" (Gigerenzer, 2000, p. 57).

In this case, the participants were not told to have external resources at the ready and their use of these resources would have constituted a cognitive and motivational challenge requiring modality switching from screen to (presumably) pen and paper and so it is strong support for the importance of considering interactivity in human performance that one third of participants opted to do this for mathematical problems.

Ecological rationality principles share with EMT the idea that rational patterns of reasoning emerge when the agents can balance off their limited resources (as time, computational capacity, or information) by using, scaffolding, or exploiting environmental cues to make decisions and find adaptive solutions to the problems at hand (Clark 2001). Discussing what interactivity brings to the rationality table in this context is particularly interesting because the focus of both EMT and ecological rationality programmes have been on how the interaction with the subjects' surrounding environment can compensate the management and use of their limited resources in a task-related problem – as time, computational capacity and information. In this case it is interesting to note that contrary to our hypothesis, interactivity did not seem to consist of compensatory actions in the environment that could balance off the problematic affective traits of the state of impasse.

Tacit Knowledge

One key limitation of the current study is that we did not measure the *types* of interactivity that took place nor when in the process they occurred. Studies on the emergence of tacit knowledge support the idea that subjects will be alert and look for chances to interact with the environment in a focused and directed way the more competence and expertise they have in a certain domain (Chen and Wall 2022). Less expert subjects not only adopt a more exploratory than a focus approach when they can scaffold external resources to aid cognitive processes, but they may also fail to recognize some of their actions as examples of interaction with the environment (as their gaze movements – Ball and Litchfield 2013), or as helpful interactive strategies that helped their cognitive processes. While then, the quantity of interaction with the environment might be similar across levels of expertise, the quality of the interactivity involved may be the reason for the relevance of interactivity on solution rates.

While determining how the use of props can aid the overcoming of an impasse state, the research focus should not only be held on *if* those interactive actions happened, but also on whether subjects were aware of them as affordable procedures, if they perceived actual possibilities in the environment to compensate cognitive issues, or even if they realized their moving gaze could provide them clues to manipulate the environment to their advantage. Even knowing if subjects were alone or with other people in the room when they tried to solve the problems may have affected their capacity to perceive useful ways to interact with the environment, since even exchanging glances with other people can scaffold environmental cues to engage in problem-solving activities (Ball and Litchfield 2013; Ross & F. Vallée-Tourangeau, 2021b). In this respect, the current experiment allows for a limited discussion over the quality of the subjects' interaction with the environment, but this seems an important line to pursue.

To conclude, we submit that these findings open up the research field in interactivity to pursue several new lines of enquiry. Rather than establishing a scaffolding effect of the use of external representations, this field can now turn to the conditions under which people sample their environment and the interaction between that and expertise. That so many people do not sample the environment even though it acts as a useful resource suggests that such a research programme is necessary to scaffold more skilled environmental interactions.

References

Ball, L.J. and Litchfield, D. (2013). Interactivity and Embodied Cues in Problem Solving, Learning and Insight: Further Contributions to a "Theory of Hints. In S.J. Cowley and F. Vallée-Tourangeau (Eds.) *Cognition beyond the brain: Computation, interactivity and human artifice.* Springer.

Bruineberg, J., & Fabry, R. (2022). Extended mind-wandering. *Philosophy and the Mind Sciences*, *3*. https://doi.org/10.33735/phimisci.2022.9190

Chen, M. & Wall, E. (2022). Perception of skill in visual problem solving: An analysis of interactive behaviors, personality traits, and the Dunning-Kruger effect. *Workshop on TRust and EXpertise in Visual Analytics (TREX, at VIS'22).*

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

Clark, A. (2001). Reasons, robots and the extended mind' *Mind & Language, 16*(2), 121–145. https://doi.org/10.1111/1468-0017.00162.

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

Fodor, J. A. (1980). Methodological solipsism considered as a research strategy in cognitive psychology. *Behavioral and Brain Sciences*, *3*(1), 63–73. https://doi.org/10.1017/S0140525X00001771

Gigerenzer, G. and Goldstein, D.G. (1996). Reasoning the fast and frugal way: Models of bounded rationality', *Psychological Review*, *103*(4), 650–669.

Gigerenzer, G. (2000). *Adaptive thinking: Rationality in the real world*. Oxford University Press.

Guthrie, L. G., Vallée-Tourangeau, F., Vallée-Tourangeau, G., & Howard, C. (2015). Learning and interactivity in solving a transformation problem. *Memory & Cognition*, *43*(5), 723–735. https://doi.org/10.3758/s13421-015-0504-8

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

Kirsh, D. (2009). Problem solving and situated cognition. In P. Robbins & M. Ayede (Eds.), *The Cambridge handbook of situated cognition*. Cambridge University Press.

Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science*, *18*(4), 513–549. https://doi.org/10.1207/s15516709cog1804_1

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

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

Ormerod, T. C., & Gross, H. (2023). Expertise overcomes impasse to yield far transfer and insight in problem-solving. *Thinking & Reasoning*, 1–25. https://doi.org/10.1080/13546783.2023.2177730

Ross, W., & Arfini, S. (2024). Impasse-driven problem solving: The multidimensional nature of feeling stuck. *Cognition*.

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

Ross, W., & Vallée-Tourangeau, F. (2021b). Rewilding cognition: Complex dynamics in open experimental systems. *Journal of Trial and Error*, *2*.

Ross, W., & Vallée-Tourangeau, F. (2022). Accident and agency: A mixed methods study contrasting luck and interactivity in problem solving. *Thinking & Reasoning*, *28*(4), 487–528. https://doi.org/10.1080/13546783.2021.1965025

Ross, W., Vallée-Tourangeau, F., & Van Herwegen, J. (2020). Mental arithmetic and interactivity: The effect of manipulating external number representations on older children's mental arithmetic success. *International Journal of Science and Mathematics Education*, 985–1000. https://doi.org/10.1007/s10763-019-09978-z

Rowlands, M. (2010). *The new science of the mind: From extended mind to embodied phenomenology*. MIT Press.

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

Sutton, J. (2005). Memory and the extended mind: Embodiment, cognition, and culture. *Cognitive Processing*, *6*(4), 223–226. https://doi.org/10.1007/s10339-005-0022-x

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

Vallée-Tourangeau, F., Euden, G., & Hearn, V. (2011). Einstellung defused: Interactivity and mental set. *Quarterly Journal of Experimental Psychology*, *64*(10), 1889–1895. https://doi.org/10.1080/17470218.2011.605151

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

Vallée-Tourangeau, F., Sirota, M., & Vallée-Tourangeau, G. (2016). Interactivity mitigates the impact of working memory depletion on mental arithmetic performance. *Cognitive Research: Principles and Implications*, 1–26. https://doi.org/10.1186/s41235-016-0027-2

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

Vallée-Tourangeau, F., & Wrightman, M. (2010). Interactive skills and individual differences in a word production task. *AI & SOCIETY*, *25*, 433–439. https://doi.org/10.1007/s00146-010-0270-x

Vallée-Tourangeau, G., Abadie, M., & Vallée-Tourangeau, F. (2015). Interactivity fosters Bayesian reasoning without instruction. *Journal of Experimental Psychology: General*, *144*(3), 581–603. https://doi.org/10.1037/a0039161

Watson, J. D., & Stent, G. S. (1998). *The double helix: A personal account of the discovery of the structure of DNA*. Scribner.

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