

Interacting physically with insight problems does not affect problem solving process

Jan Jastrzębski

Hanna Kucwaj

Institute of Psychology, Jagiellonian University
Ingardena 6, 30-060 Krakow, Poland

Adam Chuderski

Institute of Philosophy, Jagiellonian University
Grodzka 52, 31-044 Krakow, Poland

Abstract

So-called insight problems are believed to tap into sudden, creative thinking that is crucial for real problems. In contrast, recent findings suggest that solving insight problems depends on the same cognitive mechanisms that underpin systematic, analytical thinking. However, existing studies may have low ecological validity, because insight problems were usually presented in static formats (on paper, computer screen) which allowed no physical interaction with the problem elements. This study administered 8 established insight problems either in the static or interactive variants. It also probed two markers of analytical thinking: working memory capacity and reasoning ability. Virtually no difference in performance was observed between the static and interactive variants of insight problems with regard to (1) solution rate, (2) subjective experience of suddenness, pleasure, and relief accompanying the solutions, as well as (3) correlations with the working memory capacity and analytical reasoning tests. These results suggest that externalized/embodied/situated factors play no substantial role in insight problem solving and the crucial parts of this process seem to occur in the mind of a solver.

Keywords: insight problem solving; analytical thinking; working memory; interactivity.

Introduction

An important category of problems investigated in the problem solving literature is so-called insight problems. Such problems are defined in the vague and misleading way that suggests a typical but wrong problem representation, so following this representation often results in an impasse. The correct solution can be found only when the problem is viewed from a novel perspective and can be appropriately restructured. Especially difficult are problems that require rejecting one strongly believed and subjectively obvious assumption that, however, is not implicated by the problem description (Knöblich, Ohlsson, Haider, & Renius, 1999). For example, when instructed to transform an incorrect equation including Roman numerals made of matchsticks: “VI = VI + VI” into a correct equation by moving just one matchstick (without adding or removing any matchsticks), people must realize that equations do not necessarily include only one equation sign and that two such signs can also be allowed, here resulting in the tautology “VI = VI = VI”.

Insight problems have been studied intensively in cognitive science and psychology because many authors believe that they tap into mental processes that also play a role in “full-blown” creative cognition, leading to great masterpieces, discoveries, and inventions (Ohlsson, 2011).

The crucial controversy is whether the processing underpinning insight problem solving is distinct from solving so-called analytical problems, such as complex but typical arithmetic equations, which are defined in a more precise way, and require more systematic construction of the problem representation, while including no tricky obstacles. Some evidence suggested that insight problem solving involves idiosyncratic processes: constraint relaxation, defocusing attention, and uncontrolled spread of activation in memory (Knöblich et al., 1999; Kounios & Beeman, 2014), and so relies minimally on cognitive resources such as executive control and working memory capacity that typically determine success on analytical problems (see Wiley & Jarosz, 2012). Other evidence highlighted a large overlap of attentional, control, memory, reasoning, and imagery processes for insight and non-insight problems (MacGregor, Ormerod, & Chronicle, 2001; Weisberg, 2015). Specifically, two recent meta-analyses suggested that individual success on insight problems is strongly correlated with performance on analytical problems as well as with executive control and working memory tasks (Chuderski & Jastrzębski, 2018a; Gilhooly & Webb, 2018).

However, such a similarity of insight and analytical thinking might result from the fact that most of the experiments to date presented insight problems in a static format, usually printed on a paper sheet or shown on a computer screen, and participants were not allowed to interact with the problem by manipulating its elements. For instance, in a typically administered matchstick arithmetic problem, there are no actual matchsticks to be manipulated; all transformations of the equation must proceed in the mind, and the potential solution has to be written down. This lack of interaction with the problem may to some extent impede more spontaneous, “fuzzy” cognition that might be crucial for creative solutions. Participants, forced to represent and explore the problem space solely in the mind, might be prone to using more systematic, gradual problem solving strategies typical for analytical problems, while in the contexts that are

more externalized/embodied/situated they switch to less systematic strategies, such as trial-and-error, remote associations, etc. Obviously, the former strategies are more strictly constrained by available attentional resources and working memory capacity, while cognitive load might be largely reduced when artefacts can be used. Also, as many real-life problems seem to be situated to a large extent (see Clark and Chalmers, 1998; Cowley & Vallée-Tourangeau, 2010), investigating insight problem solving using non-interactive paradigms may yield low ecological validity.

Interactive insight problem solving

Indeed, a few studies by F. Vallée-Tourangeau, who applied insight problems in such a way that problems elements could be manipulated, as compared to static variants, have shown that solutions occur more frequently when the problem can be interacted with. Substantial effects, reaching the doubled solution rates, have been reported for the well-known insight problems such as the cheap necklace (Henok, Vallée-Tourangeau, & Vallée-Tourangeau, 2018; see also Fioratou & Cowley, 2009), the triangle of coins (Vallée-Tourangeau, 2017), the anagrams (Vallée-Tourangeau & Wrightman, 2011), the animals in pens (Vallée-Tourangeau, Steffensen, Vallée-Tourangeau, & Sirota, 2016), Luchins' water jars (Vallée-Tourangeau, Euden, & Hearn, 2011), and matchstick arithmetic (Weller, Villejoubert, & Vallée-Tourangeau, 2011). Also, some studies reported no difference in working memory capacity between solvers and non-solvers in the situated context. All this suggests that cognitive processing may change substantially in the embodied/situated contexts.

Besides the fact that virtually all these data (except for Fioratou & Cowley, 2009) come from one and the same lab, and thus require independent replication, existing evidence needs to be extended for at least three reasons. First, each study examined a single insight problem, applied either in the computerized/paper format or in the interactive format. As different samples of participants were used in consecutive studies, it is not possible to compare across the problems the size of presumed benefit from interactivity. (Do all problems benefit equally?) Second, recent studies (Danek, Wiley, & Öllinger, 2016; Fleck & Weisberg, 2013) probed experience during solution (asked how sudden and surprising it was), and suggested that many insight problems, originally designed to require sudden restructuring, by some participants could be solved in a fully systematic, gradual way. Thus, probably no insight problem always elicits "pure" insight. Unfortunately, so far subjective measures of insight have not been combined with examination of interactivity. Examining if interactivity can affect the subjective experience of insight might reveal mechanisms facilitating solutions. Finally, because to date, single problems were studied, the resulting binary dependent variables prohibited a proper analysis of correlations between performance on insight problems, analytical problems, and working memory tests. (Do interactive variants correlate with cognitive aptitude more weakly than the static variants?) All

these research goals have important ramifications for our understanding of insight problem solving.

To tackle these three goals, the present study applied 8 popular insight problems. They were organized in 4 pairs of comparable problems. In each pair, one problem was shown in a typical, paper-and-pencil format, while the other problem was applied in a way that allowed manipulating the artifacts comprising this problem. Which problem from each pair was applied in the static format, and which was applied in the interactive way, was randomized across the sample. This fact allowed the within-subjects manipulation of the presentation format that gave control over group differences in general performance. Moreover, the size of the expected interactivity effect could be compared across the problems, in order to see if the problems differ in how strongly they benefit from externalizing. Additionally, after each solution given to an insight problem, the four-dimensional scale that probed the subjective experience of suddenness, pleasure, relief, and certainty accompanying the solution, was applied in order to see if the surplus solutions, which were expected to occur in the interactive problem format, would consist primarily of solutions assessed subjectively as the Aha! experience. Finally, an established working memory test and a hallmark analytic reasoning test were applied in order to compare whether the correlations of these two measures with the interactive variants could really be weaker than the respective correlations with the static variants, the latter presumed to load more substantially on cognitive resources.

The study

Participants

The total sample included 64 people (34 females; aged 19 to 39, $M = 25.8$ y, $SD = 5.3$ y). All participants were recruited from the general population via internet adverts and paid an equivalent of 12 USD in local currency. They signed a written consent to participate, were screened for normal or corrected-to-normal vision and no history of neurological problems, and were informed that they could stop the experiment and leave the lab at will. Data were anonymized. All other procedural aspects of the study conformed to the WMA's Declaration of Helsinki.

Insight problems

Matchstick arithmetic. Two matchstick arithmetic problems consisted of incorrect arithmetic equations written using Roman numerals. One problem was the above described "VI + VI = VI" equation. The second problem (I = II - II) required introducing a negative number (not a typical Roman numeral) by changing one of the sticks into the minus sign. The instructions were: "This equation consists of Roman numerals made of sticks. Unfortunately, the equation is wrong! Move exactly one stick so that the equation becomes correct. The allowed operations are „-“, „+“ and „=.“ You can't remove any stick. Upright sticks and tilted sticks are not

interchangeable („|” is not „|”).” In order to familiarize the participants with the Roman numerals, the instruction contained also a table linking each Arabic number with its Roman equivalent, up to number ten. In the interactive format, the equations were constructed out of plastic sticks.

Triangle of coins/Eight coins. In the first problem, the participants were presented with a triangle facing upwards composed of 10 coins, and their task was to “Move exactly 3 coins to make the triangle point downward.” In the eight coins problem, the participants were presented with a figure composed of 8 coins, and they had to “Move exactly 2 coins so that each of the 8 coins touches exactly 3 coins,” which requires realizing that the coins have to form 3D piles. Both configurations require breaking constraints (of the X-axis rotation and 2D solution, respectively). In the interactive formats of the two tasks, the initial configurations were composed of real coins that could be manipulated freely. In the triangle problem, the response included presenting to the research assistant all the steps that had led to the solution.

Sheep in pens/Nine dots. In the first problem, the task was to “Close the 11 sheep in 4 pens so that in each pen there is an odd number of sheep.” In the interactive format, the participants were given 11 small cloth figures of sheep and 4 pieces of string. As it is impossible to divide the number 11 into any combination of 4 odd numbers, the solution required embedding at least one of the pens inside another pen. In the nine dots problem, a 3×3 array of dots was presented and the task was to “Connect all the 9 dots with a broken line composed of 4 straight lines so that each following straight line begins at the end of the preceding line.” In the correct solution, the lines should extend beyond the square shape of the array, but most people constrain themselves to explore only lines that fit within the array. In the interactive format, the participants were given tacks, 4 pieces of string, and a piece of paper with 9 dots printed.

Card split/Figure split. In the first task, participants were instructed to “Cut a hole in the card so that you can put your head through.” In the figure split, problem participants were presented with an L-shape figure and the task was to “Divide the figure into four identical parts.” Both problems require non-standard topological solutions. In the interactive formats of the tasks, participants were given scissors and several card/L-shape figures made of thick paper to experiment with.

Problems administration. In all the 8 problems, there was an identical instruction for each variant, and the variants differed only in the presentation method and response format. In the static variant, problems were given on a sheet of paper, with a blank space for making notes and drawings, and for providing a solution using a pen. In the interactive variants, the participants were given a cardboard box with respective objects placed on it, but were not provided with anything to write with, so they had to physically manipulate the objects provided. Participants were tested in individual cabins. The time limit for each problem was 5 minutes.



Fig. 1: The eight problems in initial configurations presented to the participants in the interactive format (left column) together with the sample correct solutions (right column). In the static format (not depicted) each problem elements were printed on a paper sheet, and a solution had to be drawn.

Subjective experience scale

The scale was modelled after Danek and Wiley (2016), who tested which dimensions of subjective experience best predict correct solutions to insight problems (suddenness, pleasure, relief, certainty). Here, the instruction was “Please describe your subjective experience at the moment when you found the solution to this problem”, and the four questions were: “The solution came to me...” (Gradually – Suddenly) “At the moment of finding the solution my feelings were...” (Unpleasant – Pleasant), “When I realised the solution I felt...” (Tension – Relief) “My feeling that the solution was correct was...” (Uncertain – Certain)

Ratings were recorded on a 19-point graphical scale (line of cells) for which the contradicting words (e.g. Unpleasant – Pleasant) occupied the extremes. Point “10” was marked with the text “Don’t know”, and served for inconclusive cases.

Working memory task

The letter complex span required memorizing 4, 6, or 8 letters, which were drawn from 9 possible stimuli and were presented using a computer for 1.2 s apiece. After each letter presentation, participants indicated with a mouse button if a simple arithmetic equation (e.g., $2 \times 3 - 1 = 5?$) was correct. Then, they were to recall the letters in proper order. Five trials for each set size (in increasing order) were presented. The response procedure employed as many 3×3 matrices as was a particular set size. Each matrix contained all possible letters. Those letters that had been presented in a sequence should be selected in the matrices in the correct order. There was no time limit for responding. The dependent variable was the proportion of correctly selected letters.

Reasoning test

Raven Advanced Progressive Matrices (RAPM; Raven, Court, & Raven, 1983) consists of items that include a 3×3 matrix of figural patterns which is missing the bottom-right pattern, and 8 response options presenting the potentially matching patterns. The goal was to discover the rules that govern the distribution of patterns and to choose the response option including the correct pattern that completed the matrix according to these rules. The 18 odd-numbered items were given with the 20-min. time limit.

Procedure

Participants were tested in groups of 5 to 9 people. They first undertook RAPM and the letter complex span (as well as several other cognitive tests unrelated to the present study). Then, they attempted the 8 insight problems in the fixed order. A random half of the sample attempted the odd-numbered problems in the interactive variant and the even-numbered problems in the paper-and-pencil variant. The other half used the paper and pencil for the odd-numbered problems and the interactive formats for the even-numbered problems. The entire session lasted about 2 hours.

Results

No one was able to solve correctly the Card split and Letter split problems, so the analysis pertained to the 6 remaining problems. Participants admitted familiarity with 11 out of 384 problems applied, and these 11 problems were excluded from further analysis. Fig. 2 presents the number of correct solutions for each problem, for the static versus interactive format, separately. The Triangle of coins problem was the easiest one, solved by 37.5% of participants. In contrast, the 8 coins and the 9 dots problems were most difficult, solved only by 7.8% of the sample. These solution rates matched some existing data for the same problems (e.g., Chuderski & Jastrzębski, 2018b, 2018c). Importantly, for no problem the difference between the static and the interactive format was statistically significant. The largest numerical difference was observed for the Triangle of coins problem, which was solved by 15 people (out of 32) in the static format, and by 9 (out of 32) in the interactive format, but even this difference was far from reaching statistical significance, $\chi^2(1) = 1.50, p = .220$. Overall, 41 problems were solved with the paper and pencil, while 37 problems were solved in the interactive way, which is a totally non-significant difference.

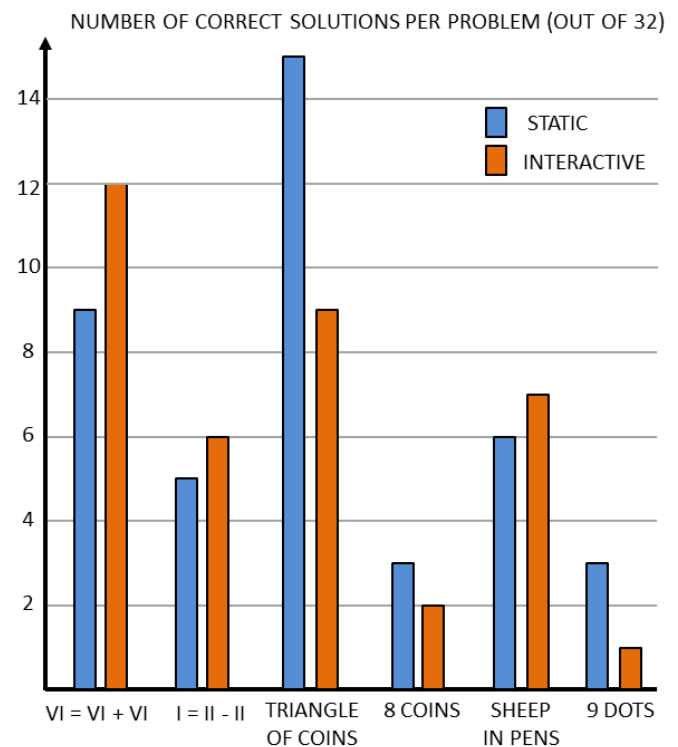


Fig.2: Number of insight problems (only those reported by the participants to be unfamiliar to them) correctly solved by the group who attempted a given problem in the static format (blue bars) vs. the group who undertook it in the interactive format (red bars). For no problem the difference between the conditions was significant at the $p < .05$ level.

Next, as it was possible that even though overall problem solving accuracy was not affected by the problem format, but it changed the way of processing the problems (at least the way subjectively experienced, and later reported, by the participants). Fig. 3 presents mean ratings for 4 indicators of insight: suddenness, pleasure, relief, and certainty, for 41 problems solved in the static format versus 37 interactive problems. Mean ratings ranging from 11 to 16 suggest that solutions yielded experience more typical for insight than for gradual, analytical processing. These ratings were submitted to MANOVA, with the problem variant (static vs. interactive) as a factor. Wilks' $\lambda = 0.917$ suggested no significant multivariate difference in experience between problem variants, $p = .173$. Second, single ratings were compared, with the Tukey correction for multiple comparisons. The only significant difference between the problem variants was noted for certainty, $F(1, 76) = 4.45$, $p = .038$, $\eta^2 = .06$, with interactive variants yielding 20% higher certainty of the correctness of the solution, as compared to the static variants. For the single problems, only the "VI = VI + VI" and the Triangle of coins problems yielded enough solutions (>20) so the accompanying reports could be compared meaningfully. For the former problem, significantly higher ratings in the interactive variant were observed for pleasure, $F(1, 20) = 7.58$, $p = .012$, $\eta^2 = .27$. No significant difference in ratings between variants was observed for the latter problem.

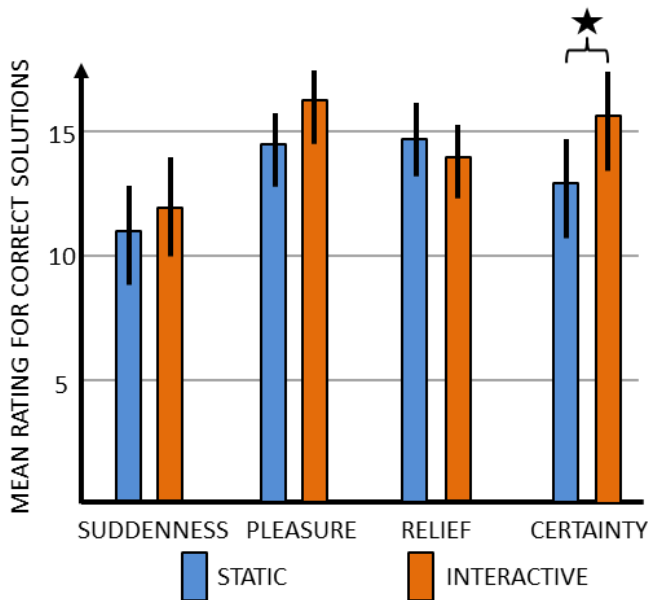


Fig. 3: Mean ratings for the reported subjective experience of suddenness, pleasure, relief, and certainty during correct solutions of insight problems, separately for the 41 problems in static variants versus the 37 problems in interactive variants. The star indicates a weak difference significant at the $p < .05$ level (for the experience of certainty). The three other differences were non-significant.

Finally, for each participant her or his score on all the 6 problems, the 3 problems applied in the static format, and the 3 problems applied in the interactive formats were calculated. The Spearman rank correlation was computed to assess the relationship between the number of problems solved and the letter complex span and RAPM scores. The resulting correlations are presented in the Table.

Table: Matrix of Spearman correlations between variables

Variable	1.	2.	3.	4.
1.All 6 problems	1			
2.Static variants	.756	1		
3.Interactive variants	.765	.195	1	
4.Complex span	.404	.316	.282	1
5.RAPM	.640	.520	.465	.353

Note. $N = 64$. All correlations significant at $p < .05$ except for the correlation between static and interactive variants.

Overall, correlations between the insight problem scores and the complex span ($\rho \approx .3$) and RAPM ($\rho \approx .5$) were substantial. However, the difference in correlation with the complex span between the scores on static versus interactive variants equaled only $\Delta\rho = .034$ that was far not significant. The analogical difference for RAPM equaled $\Delta\rho = .055$, which was not significant, either.

Discussion

The present study aimed to examine the role of interactivity in the process of insight problem solving. More specifically, it aimed to test (1) whether insight problems could be more frequently solved when presented in the interactive format allowing physical manipulation of the problem elements, as compared to the static format; (2) whether solutions in the former format could yield different subjective experience of suddenness, pleasure, relief, and certainty than yielded by solutions in the static format; and (3) how much performance on the interactive variants depended on cognitive resources, in comparison to the static variants.

A variety of established insight problems were used, which ranged in difficulty from a complete floor up to over one-third of correct solutions. Given existing evidence, the present results were quite surprising. There was virtually no difference in problem solving accuracy, regardless of the format used. Subjective experience reported, especially the suddenness of solution, linked closely to actual insight, was comparable for both problem formats. One exception was slightly increased certainty in interactive problem variants, which might have resulted from the fact that interactively delivered solutions were more concrete, so they could be more directly evaluated than the solutions written on paper. Importantly, both the static and the interactive problem variants substantially depended on working memory capacity and analytic reasoning and did it in a fully comparable way. Consequently, no evidence was found for any substantial

effects of interactivity on the process of insight problem solving on the sample of diverse and established insight problems. As both static and interactive conditions substantially relied on working memory/reasoning ability, no evidence was found for the decreased role of analytic thinking in the interactive format. Thus, it seems that physically manipulating problem elements did not substantially decrease cognitive load or affect the use of strategies in the process of problem solving.

On the other hand, it may be arguable to what extent the paper-and-pencil format, at least in case of some particular problems, is fully static, i.e. it does not provide any external support that may help in the process of solving. For example, making drawings and sketches may help to test hypotheses, keep track of the progress and perform simple trial-and-error strategies compared to problem solving without any external support provided. Thus, comparing interactive, static paper-and-pencil and the “pure” static format at the same time should be considered in future studies.

We cannot fully exclude the possibility that the null effects observed in the present study resulted from the selection of the particular problems or the specific way the interactive format was implemented in the given problems. If interactivity substantially affects only specific insight problems under some specific conditions it would be valuable to identify the characteristics of such problems that moderate this effect. Also, one limitation of the present study is that although no substantial effects of interactivity were observed, the relatively low statistical power prevented the detection of any potential small effects. However, if the effects of interactivity are negligible or observed only in very specific problems or circumstances, it would be questionable whether the role of interactivity in the process of insight problem solving is an important research topic at all.

Summing up, no evidence was found that manipulating physically problem elements when solving problems presumably involving insight helps to reach the correct solution. Neither the influence on the course of the problem solving process (the extent of its suddenness) nor on its affective consequences (pleasure, relief) were observed. Interactivity did not decrease a substantial reliance of the problem solving process on working memory capacity and reasoning ability, either. The only observed effect was a small increase in certainty about the solution (a meta-cognitive consequence). Thus, at least for the problems applied in the present study, externalized/embodied/situated factors played no substantial role in finding solutions, and the results are in line with the key role of analytic reasoning in solving insight problems. Still, more research is needed to comprehensively examine the potential role of interactivity in the process of insight problem solving.

References

- Chuderski, A. & Jastrzębski, J. (2018a). The relationship of insight problem solving to analytical thinking: Evidence from psychometric studies. In F. Vallée-Tourangeau (Ed.), *Insight: On the origin of ideas* (pp. 120-142). Abingdon, UK: Routledge.
- Chuderski, A. & Jastrzębski, J. (2018b). Much ado about Aha! Insight problem solving is strongly related to working memory capacity and reasoning ability. *Journal of Experimental Psychology: General*, 147, 257-281.
- Chuderski, A., & Jastrzębski, J. (2018c). No role of initial problem representation in insight problem solving. *Creativity Research Journal*, 30:4, 428-438.
- Clark, A., & Chalmers, D. J. (1998). The extended mind. *Analysis*, 58, 7–19.
- Cowley, S. J., & Vallée-Tourangeau, F. (2010). Thinking in action. *AI & Society*, 25, 469–475.
- Danek, A. H., Wiley, J., & Öllinger, M. (2016). Solving classical insight problems without aha! experience: 9 dot, 8 coin, and matchstick arithmetic problems. *The Journal of Problem Solving*, 9, 4.
- Fioratou, E., Cowley, S. J. (2009). Insightful thinking: Cognitive dynamics and material artifacts. *Pragmatics & Cognition*, 17, 549–572.
- Fleck, J. I., & Weisberg, R. W. (2013). Insight versus analysis: Evidence for diverse methods in problem solving. *Journal of Cognitive Psychology*, 25, 436–463.
- Gilhooly, K. & Webb, M. E. (2018). Working memory in insight problem solving. In F. Vallée-Tourangeau (Ed.), *Insight: On the origin of ideas* (pp. 105-119). Abingdon, UK: Routledge.
- Henok, N., Vallée-Tourangeau, F. & Vallée-Tourangeau, G. (2018). Incubation and interactivity in insight problem solving. *Psychological Research*.
- Knöblich, 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, 1534–1555.
- Kounios, J., & Beeman, M. (2014). The cognitive neuroscience of insight. *Annual Review of Psychology*, 65, 71–93.
- 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, 176–201.
- Ohlsson, S. (2011). *Deep learning: How the mind overrides experience*. Cambridge, UK: Cambridge University Press.
- 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, 791–799.
- Raven, J. C., Court, J. H., & Raven, J. (1983). *Manual for Raven's Progressive Matrices and Vocabulary scales* (Section 4: Advanced progressive matrices). London, UK: H. K. Lewis.
- Vallée-Tourangeau, F.(2017). Interactivity and ego depletion in insight problem solving. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. J. Davelaar (Eds.), *Proceedings of the 39th Annual Conference of the Cognitive Science Society* (pp. 1248-1253). Austin, TX: Cognitive Science Society.
- Vallée-Tourangeau, F., Euden, G., & Hearn, V. (2011). Einstellung defused: Interactivity and mental set. *Quarterly Journal of Experimental Psychology*, 64, 1889–1895.
- Vallée-Tourangeau, F., Steffensen S. V., Vallée-Tourangeau, G., Sirota, M. (2016). Insight with hands and things. *Acta Psychologica*, 170, 195-205.
- Vallée-Tourangeau, F., & Wrightman, M. (2010). Interactive skills and individual differences in a word production task. *AI & Society*, 25, 433– 439.
- Weisberg, R. W. (2015). Toward an integrated theory of insight in problem solving. *Thinking & Reasoning*, 21, 5-39.
- Weller, A., Villejoubert, G., & Vallée-Tourangeau, F. (2011). Interactive insight problem solving. *Thinking & Reasoning*, 17, 424-439
- Wiley, J., & Jarosz, A. F. (2012). How working memory capacity affects problem solving. *Psychology of Learning and Motivation*, 56, 185–227.