ORIGINAL ARTICLE

CrossMark

Closing the gap: connecting sudden representational change to the subjective Aha! experience in insightful problem solving

Amory H. Danek[1] · Joshua Williams[1] · Jennifer Wiley[1]

Abstract

Two hallmarks of insightful problem solving are thought to be suddenness in the emergence of solution due to changes in problem representation, and the subjective Aha! experience. Although a number of studies have explored the Aha! experience, few studies have attempted to measure representational change. Following the lead of Durso et al. (Psychol Sci 5(2):94–97, 1994) and Cushen and Wiley (Conscious Cognit 21(3):1166–1175, 2012), in this study, participants made importance-to-solution ratings throughout their solution attempts as a way to assess representational change. Participants viewed a set of magic trick videos with the task of finding out how each trick worked, and rated six action verbs for each trick (including one that implied the correct solution) multiple times during solution. They were also asked to indicate the extent to which they experienced an Aha! moment. Patterns of ratings that showed a sudden change towards a correct solution led to stronger Aha! experiences than patterns that showed a more incremental change towards a correct solution, or a change towards incorrect solutions. The results show a connection between sudden changes in problem representations (leading to correct solutions) and the subjective appraisal of solutions as an Aha! experience. This offers the first empirical support for a close relationship between two theoretical constructs that have traditionally been assumed to be related to insightful problem solving.

Introduction

The frustration of being unable to solve a vexing problem is a discomfort we have all felt before. Occasionally, the solution unwittingly pops into mind in a flash of insight, accompanied by the so-called "Aha! experience". While this experience may feel special, researchers in the field have long been divided on the distinctiveness of insightful problem solving. There are those who claim it is no different from routine problem solving (Weisberg & Alba, 1981). Conversely, others believe insight is distinct from other types of problem solving, because of an underlying hypothetical process of rapidly restructuring a problem's elements within a mental representation (Duncker, 1926; Metcalfe, 1986; Metcalfe & Wiebe, 1987; Ohlsson, 1984; Wertheimer, 1925, 1959). This cognitive process of restructuring was first postulated by the Gestalt psychologists. Wertheimer used the term "Umkrempelung" which can be roughly translated with "turning inside out" (1925, p. 174), while

Duncker (1926, p. 702) discussed the assumption that "a problem situation contains necessarily a gap" which needs to be closed by a structural change from "bad Gestalt" to "good Gestalt". Moreover, this representational change is thought to occur rapidly, like flipping between interpretations of the Necker Cube, as an "immediate realization" (p. 705). Metcalfe provided first evidence for suddenness in the emergence of solutions by measuring feeling-of-warmth ratings (Metcalfe, 1986; Metcalfe & Wiebe, 1987), but little work has attempted to directly measure changes in problem representations. The goal of the present study was to obtain both measures of restructuring as well as measures of Aha! experiences to clarify their relation, and to test theoretical assumptions that both constructs are connected to insightful problem solving (e.g. Dominowski & Dallob, 1995; Gick & Lockhart, 1995).

The subjective Aha! experience that is often reported upon finding a solution has been taken as a marker for insight (Bowden, 1997; Jung-Beeman et al., 2004; Kaplan & Simon, 1990). Although many studies have attempted to measure Aha!, very few have attempted to measure restructuring (as pointed out by Ash, Cushen, & Wiley, 2009). In those that have, participants have been asked to rate problem elements at multiple time points either with respect to their

✉ Amory H. Danek
  amory.danek@psychologie.uni-heidelberg.de

[1] Department of Psychology, University of Illinois at Chicago, 1007 W. Harrison Street MC 285, Chicago, IL 60607, USA

relatedness or to their importance for solution (Ash & Wiley, 2008; Cushen & Wiley, 2012; Durso, Rea, & Dayton, 1994). In their pioneering study, Durso et al. (1994) presented a verbal puzzle ("A man walks into a bar and asks for a glass of water. The bartender points a shotgun at the man. The man says, "Thank you", and walks out."), and allowed participants to ask yes-or-no questions to gain more information for up to two hours as they attempted a solution. After the end of the solution phase, participants' problem representations were assessed by presenting them with all possible pairings of 14 terms that were either relevant to the story (bartender—man), to the solution that the man had the hiccups (surprise—remedy and relieved—thank you), or were things generally found in a bar (TV—pretzels). Participants then had to rate how related the two words that made up each pair were. A comparison of solvers' and non-solvers' "relatedness ratings" showed quite different problem representations, with solvers making two key connections between words that were not made by non-solvers. In a follow-up experiment, Durso et al. (1994) obtained a more online measure of the changing problem representation by having participants repeat the ratings every 10 min while solving. Only 12 word pairs were rated including two "insight pairs" with the key connections (surprise—remedy and relieved—thank you), plus semantically related and unrelated distractor pairs. The similarity of the related and unrelated word pairs was rated fairly constant across time, while the words from the crucial insight pairs were initially rated as being dissimilar, then as less dissimilar and finally changed to being rated as similar at solution, indicating that a representational change had taken place. Interestingly, the change in representations seen during the solution of the puzzle appeared to be incremental, not sudden. Because two-thirds of the pairwise ratings included words that were related to the solution, it was possible that these ratings might have served as hints, which could have affected solution patterns (as shown by Davidson, 1995).

The Durso et al. study provided an important precedent of how a repeated rating paradigm could be used to track the evolution of problem representations. However, no measure of Aha! was obtained. A subsequent investigation by Cushen and Wiley (2012) tried to measure solvers' subjective Aha! experience by having them rate how surprised they were upon solving and how much it felt like a sudden realization. They also manipulated whether solvers made repeated ratings on all features of a problem, or on only a subset that represented mostly important features. Their results revealed three main findings. First, when solution patterns were explored for individuals, there was evidence that some of the solvers experienced a sudden change in their representations, as shown by a single large jump in their importance ratings for critical problem features before solution. Second, this percentage was higher when

solvers completed ratings for all features of the problem than when they were biased by rating only a subset of the features. When solvers completed ratings on only the subset, they were more likely to show an incremental change in their importance ratings before solution. But, third, the subjective Aha! experience was not seen to differ as a function of whether solvers experienced a sudden change in representation.

The results of both Durso et al. (1994) and Cushen and Wiley (2012) show that repeated ratings of problem features can be a useful approach for measuring representational change. When changes in ratings are explored at an individual level, this approach allows for the detection of both incremental and sudden patterns of restructuring that may precede a solution. Theoretically, one would expect that sudden change patterns should be more likely to be associated with Aha! experiences (e.g. Dominowski & Dallob, 1995; Gick & Lockhart, 1995), even though this result was not obtained in Cushen and Wiley's study. However, the Cushen and Wiley study was done using only a single problem. It also used an Aha! prompt that emphasized "surprise" as part of the Aha! experience. More recent work has suggested that feelings of "surprise" may be misleading, and that suddenness and certainty are more essential facets of the Aha! experience (Danek & Wiley, 2017). For both of these reasons, it was considered worthwhile to conduct a novel test of this hypothesis in the present study using a larger problem set and a different Aha! prompt. To further explore this potential relationship between the cognitive component of sudden representational change and solvers' affective solving experience, the present study obtained measures of both components under the following hypothesis: if the strength of the subjectively reported Aha! varies with the suddenness of representational change, this would provide evidence for the theoretically assumed relationship between affective and cognitive aspects of insightful problem solving. If, however, the Aha! is experienced similarly in both sudden and incremental solution processes, this would cast doubt on the long-standing assumption that the Aha! experience is solely associated with insightful solution processes.

Method

Participants

Participants were 30 undergraduates in Introductory Psychology at a US Midwestern university ($M_{age} = 19.3$ years, range 17–27, 13 males) who received course credit as compensation. Two additional participants were excluded for not following the instructions.

Materials

Magic tricks

Video clips of 18 short magic tricks (see Danek & Wiley, 2017, for descriptions of each trick) were presented to participants as a problem solving task ("Your task is to solve this puzzle and try to see through the magic trick."). Similar to other tasks used to investigate insight, magic tricks typically impose an initial view of the problem that is incorrect and requires a representational change to allow a solution, see Danek et al. for details of the underlying rationale (Danek, Fraps, von Müller, Grothe, & Öllinger, 2014a). This task domain has been shown to trigger strong Aha! experiences (Danek, Fraps, von Müller, Grothe, & Öllinger, 2014b; Danek & Wiley, 2017) and has recently been taken up by others, yielding similar findings (Hedne, Norman, & Metcalfe, 2016). Professional magician Thomas Fraps (Abbott, 2005) performed the tricks (see https://www.youtube.com/watch?v=3B6ZxNROuNw for an example clip).

Verb ratings

A set of verb ratings were developed based on prior work by Ash and Wiley, Cushen and Wiley, and Durso et al. to provide an online assessment of each solver's problem representation at multiple time points during solution. Importantly, solvers were not asked to judge their closeness to solution, nor to reflect on their likelihood of solving each problem, which would represent metacognitive judgments.

Rather, solvers were simply asked to rate each verb for how relevant it seemed for solution, which was meant to capture whether the verb matched a solution that was currently being considered by the solver (see Ash & Wiley, 2008, for more discussion on the distinction between metacognitive and situational judgments).

Paper booklets contained four identical rating sheets for each trick, as shown in the first slide in Fig. 1. On each sheet, participants rated six verbs with regard to how well they described the solution ("How important is this word for the solution?"). During the practice task, they were told to look at each verb individually and make a rating for it by slashing a mark on a 5 in. scale that was anchored by "not important" on the left and "important" on the right. For each trick, there was one target verb that corresponded to the correct solution, one verb that described a false solution, and four distractor verbs. Verb order was fixed for each trick, but varied by trick to avoid, for example, the target verb always being first.

Based on a pilot study in German (Utz, 2013) and previous studies, the set of verbs was developed by analysing each trick and reviewing participants' prior responses. For example, in one trick the magician holds a wine glass in his hand. He places a silk cloth over the glass, lifts it slightly while secretly flipping the glass upside down, and then quickly grabs the stem of the glass with his other hand while taking away the silk to reveal that the glass has vanished. Although the flipping motion was concealed by the silk, participants who deciphered the trick were likely to mention "flipping" in their responses, so it was selected as the target verb. False solution verbs were determined by choosing the most

Fig. 1 Sequence of one trial
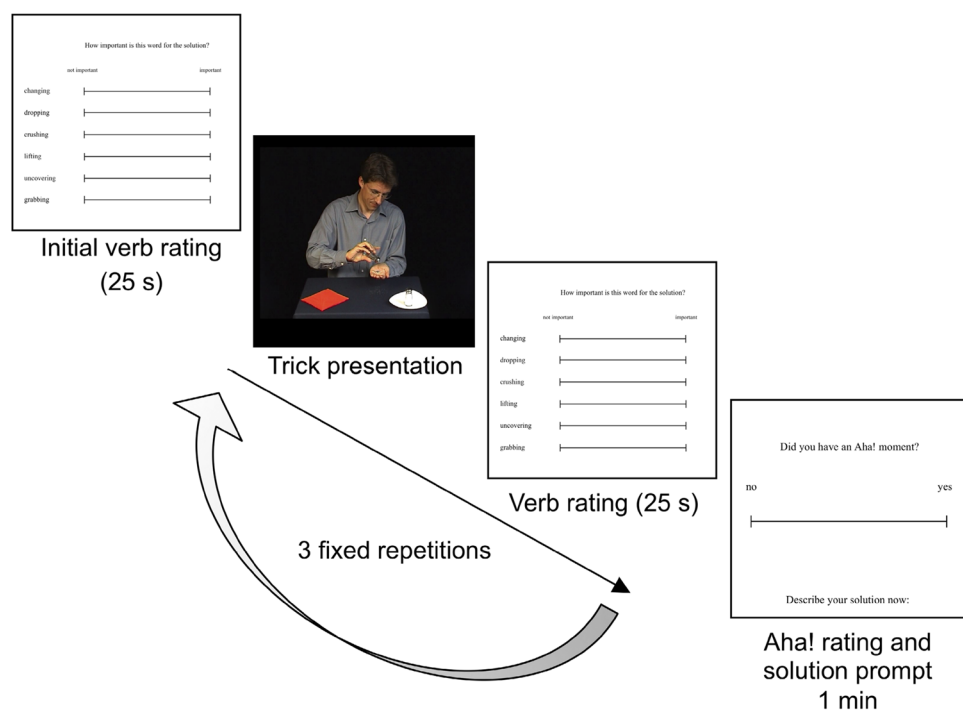
frequently suggested false solution from previous studies. The four distractor verbs were taken from the list of target and false solution verbs for other tricks.

To validate these materials, two independent raters (none of the authors) who were familiar with the solutions to the tricks were presented with the same booklets that would later be used in the study. For each trick, they completed one set of ratings for each verb (as depicted on the first slide in Fig. 1) with regard to how important it was for the solution. The intraclass correlation coefficient, ICC (1, 2) for the two sets of ratings was 0.91. On average, the target verbs were given a rating of $M = 4.72$, SD = 0.95 (on a 5 in. scale), whereas the false solution verbs were rated with $M = 0.80$, SD = 1.44 and the mean across the 4 distractors was $M = 0.16$, SD = 0.48. Further, for 17 of the 18 tricks, both raters gave the target verb the highest rating. However, for one trick, only one of the raters did so (the other gave the false solution verb the highest rating). This trick was removed from analyses.

Aha! ratings and solution prompts

Booklets also contained an Aha! rating for each trick, on the same page that prompted participants to write down their solutions. As shown in Fig. 1, participants were asked to judge on a 5 in. scale from "no" to "yes" the extent to which they had an Aha! moment. During practice, they were provided with this instruction (Danek, Fraps, von Müller, Grothe, & Öllinger, 2013; Jung-Beeman et al., 2004): "Whenever you guess a solution, we want to know whether you experienced an Aha! moment during solving. An Aha! moment is when the solution suddenly dawns on you and everything is clear immediately. In a flash. You are relatively confident that your solution is correct. In contrast, if the solution occurs to you slowly and in steps that would not be an Aha! moment. As an example, imagine a light bulb that is switched on all at once in contrast to slowly turning up the lights. Perhaps you have sometimes experienced an Aha! moment during studying. For each solution, we ask for your subjective rating whether it felt like an Aha! moment or not. There is no right or wrong answer. Just follow your intuition.".

Procedure

Participants were run in a group setting in a classroom. For practice, they were instructed to watch two video clips of magic tricks projected on a screen at the front of the room, and told that they would be asked to write down the solution after the third viewing of each trick in paper booklets. The verb rating task and the Aha! rating task were also explained to them.

After instruction and practice, 18 experimental tricks were shown. Each trick required verb ratings at four time points: one initial rating before the first viewing of the trick, and three more ratings during the solution phase, one after each viewing. Participants were given 25 s to make their ratings each time. After the fourth rating, participants were given 1 min to indicate whether they had experienced an Aha! moment using the Aha! rating scale, and then to describe their solution idea. No feedback about solution correctness was given. This process was repeated for all tricks in fixed order. The procedure took 1.5 h.

Coding

Participants' solutions were coded as correct (method that magician actually used or alternative method) or incorrect (partial solutions, implausible methods or impossible with respect to the conditions seen in the video clip) by two independent raters using a coding manual (compiled with the help of the magician) based on prior work with this problem set. The intraclass correlation coefficient, ICC (1, 2) was 0.95 indicating an excellent level of agreement. Conflicting cases were resolved by a third rater.

Both Aha! and verb ratings were measured using the distance of the mark from the left side of the 5 in. scale. Patterns of changes in importance ratings of the target verbs across the three solution time points (i.e. not including the initial rating) were categorized as being a sudden change in the direction of the correct solution, an incremental change in the direction of the correct solution, flat, or decreasing (i.e. a change away from the correct solution). In a first step, adopting the methodology of Cushen and Wiley (2012), a line graph was created from each participant's target verb ratings across the three solution time points for each individual trick. One rater engaged in a visual analysis of all 457 graphs (all observations with complete data) and made a (sudden increase, incremental increase, flat/decrease, other) judgment for each graph. All patterns that involved an increase in the direction of the correct solution of 2 in. or more between two consecutive viewings were coded as a sudden change, and this matched 100% of sudden patterns detected by the visual analyses. All patterns were coded as an incremental increase if they increased less than 2 in, but more than 0.15 inches. This matched 100% of the incremental patterns detected by the visual analyses. In a second step, the numerical cutoffs derived from these judgments were applied using algorithms to the raw rating data to create the final coding following this decision tree: patterns that increased from initial to final rating were coded as a sudden change toward the correct solution if the rating increased 2 in or more between two consecutive viewings. Patterns that increased from initial to final rating were coded as an incremental change toward the correct solution if each rating

increased less than 2 in, or consecutive ratings increased an equal amount. Patterns that did not increase or decrease more than 0.15 in across all ratings were coded as flat. Those that had a downward trend more than 0.15 in across all ratings were coded as downward. All remaining patterns were coded as "other". This included patterns with a high or low rating at midpoint, i.e. "zigzags", showing no clear increase nor decrease but rather two conflicting dynamics. The patterns assigned by visual analysis and by numerical formulas had an ICC $(1,2) = 0.96$, with all discrepancies occurring between the flat/decrease and "other" categories. The statistical analyses were performed using the patterns as assigned by the numerical algorithms. Because the "other" category comprises a heterogeneous group of patterns (i.e. not a real solution pattern category), it was not included in the main analyses.

Results

In total, 30 participants being presented with 17 tricks yielded 510 observations. Of these, 53 observations were discarded due to missing solutions or ratings. Of the remaining 457 observations, 45.5% (208) were correctly solved, and 54.5% (249) were incorrectly solved. Although prior work has suggested that hints such as those provided by rating tasks can bias solutions (Bowden, 1997; Cushen & Wiley, 2012), these rates appear comparable to those obtained without verb ratings. The solution rate for these 17 tricks in prior studies with this population in which individuals did not complete verb ratings (but also were not forced to view each trick three times before answering) was 51.5% in the dataset published in Danek and Wiley (2017), and 41.0% in another unpublished dataset.

The number of solutions falling into each solution pattern category is shown in Table 1. The overall frequency of correct and incorrect solutions across solution pattern categories was not randomly distributed, $\chi^2(4, 457) = 9.25$, $p = .05$. Follow-up tests showed that incorrect solutions were more likely than correct to be categorized as showing a decrease in

target ratings, $\chi^2(1, 122) = 5.54$, $p < .02$, while correct solutions were more likely than incorrect to be categorized as showing a sudden increase, $\chi^2(1, 46) = 4.26$, $p < .05$.

All remaining analyses were performed using the SPSS MIXED procedure to compute mixed effects models, entering participants as random effect, and fitting random intercepts for participants. These analyses were performed without the observations that fell into the "other" category. Similar proportions of problems were solved after the "other" category was discarded [45.3% (155) correct and 54.7% (187) incorrect out of a total of 342]. Average Aha! ratings for correctly and incorrectly solved problems are shown in Table 2. As in prior research (Danek et al., 2014a; Hedne et al., 2016; Salvi, Bricolo, Kounios, Bowden, & Beeman, 2016; Webb, Little, & Cropper, 2016), correct solutions resulted in higher Aha! ratings than incorrect solutions, $F(1, 340) = 43.42$, $p < .001$ (number of observations = 342).

Ratings for target, false solution, and distractor verbs by correctness

To confirm that the target verbs selected for this study reflected a correct problem representation, solvers' and non-solvers' target verb ratings were compared (number of observations = 342). As shown in Table 2, participants who eventually solved the problems made similar initial ratings on target verbs to those participants who did not eventually solve, $F < 1$. In contrast, the final ratings on target verbs were higher for solvers versus non-solvers, $F(1, 340) = 36.11$, $p < .001$. Consistent with the results from the independent raters, this shows that the target verbs served as a valid index for correct problem representations. Conversely, solvers rated the false solution verbs ($F(1, 340) = 30.83$, $p < .001$) and distractor verbs (using the mean of all four distractors, $F(1, 340) = 4.15$, $p < .05$) lower than non-solvers on the final ratings, although both groups again did not differ in their initial ratings, Fs < 1.

Table 1 Solution pattern frequencies by correctness of solution

	Correct	Incorrect	Total
Decreasing patterns	23.1% (48)	29.7% (74)	25.7% (122)
Other patterns	25.5% (53)	24.9% (62)	25.2% (115)
Flat patterns	26.9% (56)	27.7% (69)	27.4% (125)
Incremental increase patterns	10.1% (21)	11.2% (28)	10.7% (49)
Sudden increase patterns	14.4% (30)	6.4% (16)	10.1% (46)

Actual observed frequencies out of 457 observations appear in parentheses

Table 2 Average Aha! and verb ratings by correctness of solution

	Correct M (SE)	Incorrect M (SE)
Aha! ratings	3.40 (0.19)	2.20 (0.18)
Initial target verb ratings	2.21 (0.17)	2.19 (0.17)
Final target verb ratings	3.68 (0.17)	2.50 (0.15)
Initial false verb ratings	2.13 (0.17)	2.11 (0.17)
Final false verb ratings	2.15 (0.17)	3.29 (0.15)
Initial distractor verb ratings	2.02 (0.16)	2.00 (0.15)
Final distractor verb ratings	1.29 (0.12)	1.49 (0.11)

For distractors, the mean of all four distractors was used. All values are estimated means from the mixed models

Target verb rating patterns

Figure 2 shows the average Aha! ratings for tricks solved correctly (number of observations = 155) as a function of the four solution pattern categories. To test whether sudden change patterns in target verb ratings were more likely to be associated with Aha! experiences, the solution pattern measure (four levels representing the four solution patterns) was entered as a fixed factor in a mixed effects model with Aha! ratings as dependent variable. This factor led to a significant main effect ($F(3, 151) = 5.51$, $p < .01$). To test the main theoretical prediction, a planned comparison showed that solutions following sudden increases in the target verb ratings obtained higher Aha! ratings than incremental patterns ($t(49) = 2.81$, $p < .01$) (number of observations = 51).

Looking at Fig. 2, one notices that flat patterns also receive rather high Aha! ratings which a pairwise comparison, Bonferroni corrected, found was no different from the sudden patterns ($t(151) = 1.19$, $p = .47$). The high Aha! ratings for flat patterns can be better understood by examining the verb ratings that immediately followed the first viewing of the trick in relation to the other three solution patterns as shown in Fig. 3. A mixed effects model using solution pattern as fixed factor (four levels) found significant differences in the target verb rating at the second time point, $F(3, 151) = 14.87$, $p < .001$, with pairwise comparisons using the Bonferroni correction showing ratings that were part of flat patterns were significantly higher immediately after initial viewing than ratings that were part of sudden change patterns ($t(151) = 5.99$, $p < .001$). This suggests that solvers who were categorized as showing a flat pattern may have realized the solution to the trick immediately upon seeing it the first time. The solution may have felt like an Aha! for them, but occurred too immediately for a change in problem representation to be seen across the three post-viewing verb ratings.

Interestingly, the target verb ratings for participants who solved with a sudden pattern (number of observations = 30) tended to decrease from before seeing the trick to after the first exposure to the trick ($t(29) = 2.61$, $p < .05$), leading to lower ratings than for all other patterns. This suggests that the participants who reported the most Aha! might have been getting misled by their initial solution attempts, and then following that, experienced the sudden change toward a correct representation.

Participants did not vary in the number of tricks solved via sudden representational change ($F(28, 126) = 1.11$, $p = .34$), but there was an effect of trick on the likelihood of solving with sudden representational change ($F(16, 138) = 2.28$, $p < .01$). This suggests that the low rate of restructuring observed in this study was more likely due to it occurring on a minority of problems across all participants (rather than on the majority of problems for a minority of individuals).

Discussion

Consistent with other recent work (Danek et al., 2014a; Hedne et al., 2016; Salvi et al., 2016; Webb et al., 2016), this study demonstrated that correct solutions were given higher Aha! ratings than incorrect ones. Further, the effect of solution patterns on Aha! experiences reported here is new: correct solutions following sudden patterns were given higher Aha! ratings than correct solutions following incremental patterns. Note that this is a very tight comparison: in both cases, problem solvers arrived at the same correct solution, but importantly, the underlying solution process (measured via repeated importance-to-solution ratings) differed. The strongest Aha! experiences were reported for solution patterns with a sudden change toward correct

Fig. 2 Mean Aha! ratings for correct solutions as a function of target verb rating pattern. Error bars denote standard error of the mean
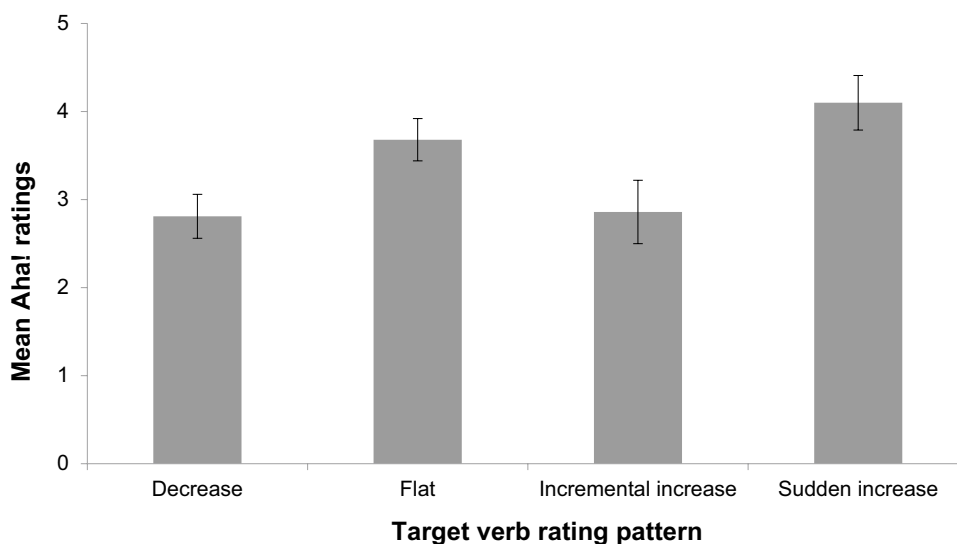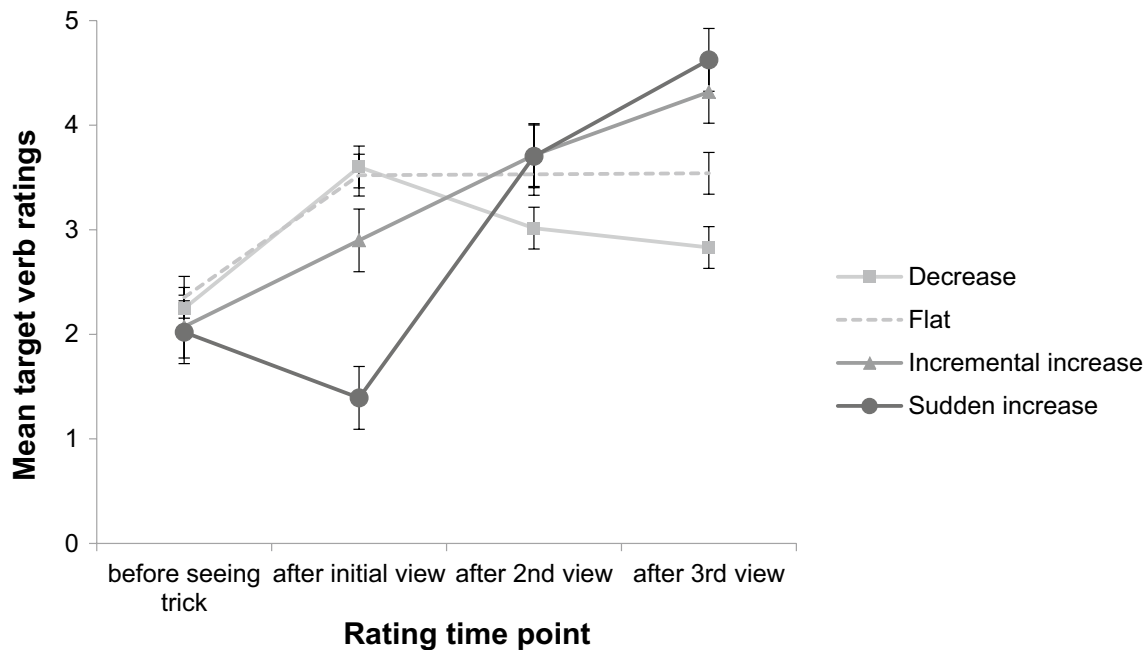
Fig. 3 Changes from initial to final ratings on target verbs for different solution patterns prior to correct solutions. Error bars denote standard error of the mean

problem representations. Further, for these cases, the ratings made after the first viewing of the trick revealed problem representations that were the least accurate, suggestive of fixation or impasse due to an inappropriate initial representation. In line with classic insight theories, the sudden patterns can be interpreted as indicating the sudden restructuring of a problem's elements, whereas the incremental patterns indicate a more routine, step-wise or analytic solution process. The present finding that the strength of the subjectively reported Aha! experience varied with the suddenness of representational change provides evidence for the theoretically assumed relationship between affective and cognitive aspects of insightful problem solving.

A previous study (Cushen & Wiley, 2012) had attempted and failed to find a correspondence between solution patterns and subjective perceptions of the solution experience, making it important to consider the differences between these two studies. First, Cushen and Wiley used only a single object-move problem (Triangle of Circles), while the present study had a substantially more powerful design using 18 problems from the domain of magic tricks. Second, there was an important difference in the way that participants were asked about their solution experiences. In Cushen and Wiley, participants rated whether the solution seemed surprising and sudden, while the present study used an Aha! rating that emphasized suddenness and certainty. Recent work (Danek & Wiley, 2017) has suggested that suddenness and certainty are more essential facets of the Aha! experience, while feelings of surprise can be misleading. Thus, the present study

was in a notably better position to test for this hypothetical relation due to a number of design considerations, and demonstrated a clear connection between solution patterns and subjective experience.

A further key to this study was tracking changes in problem representation for each individual solver on each problem across time by obtaining repeated ratings while using a continuous measure of the Aha! experience. Using this approach, the present study provides a straightforward demonstration that differences in the phenomenology of a solution experience (i.e. strength and likelihood of an Aha! moment) vary with different underlying solution processes. For future studies, this highlights the necessity of measuring both the subjective Aha! experience as well as obtaining information about the dynamics of the underlying solution process for each individual, instead of simply assuming that problems have been solved insightfully by all correct solvers because they are assumed to be "insight problems". A number of studies have shown that these assumptions are problematic (Danek, Wiley, & Öllinger, 2016; Webb et al., 2016). In addition, we are just beginning to understand the factors that influence these dynamics, such as whether solvers are provided with hints (Bowden, 1997; Cushen & Wiley, 2012; Durso et al., 1994), and the degree of constraint relaxation required for solution (Danek et al., 2016). Another possibility for tracking representational change may be eye tracking as used by Ellis, Glaholt, and Reingold (2011) and Knoblich, Ohlsson, and Raney (2001). Such a continuous, unobtrusive

measure allows to follow participants' attention during the solving process, can serve as an indicator of the progress made (e.g. if crucial elements are looked at more often or fixated longer), and provides information about the dynamics of the solving process by revealing either rapid or gradual changes in attention allocation towards crucial problem elements. However, this requires that correct and incorrect representations would result in paying attention to distinctly different areas on the screen as solvers watch the magician perform each trick. Alternatively, obtaining some additional measure that captures the degree of impasse that solvers face could help to clarify whether experiencing impasse early in the solution process is important for either sudden changes toward a correct representation or the Aha! experience.

A final important observation is the low incidence of sudden restructuring in this dataset. Only around 15% of correct solutions showed evidence that they had been reached via sudden representational change. Although some of the flat solutions may also have involved representational change that happened too immediately to be captured by this paradigm, these results suggest that sudden restructuring occurred on a minority of problems. This observation begs the question whether other studies that have collapsed across all correct solutions may have reached misleading conclusions by intermixing sudden, incremental, and other solution patterns.

In designing this study, we were faced with a number of methodological trade-offs: we decided it was better to obtain ratings at all three time points, instead of having participants stop when they thought they had found a solution. Thus, we made participants go through the entire procedure of three viewings and three ratings for each trick. Also, running this study on paper allowed for administration in groups, but did not allow for a reliable measure of the exact solution time point. As a result, the actual time point of solution can thus be only indirectly inferred from the changes in target ratings. This limitation needs to be addressed in future studies.

In general, any problem can be solved either via a sudden or incremental process, and may or may not be accompanied by an Aha! experience. The critical finding of the present study was demonstrating that problems solved with sudden restructuring were more likely to elicit an Aha! experience, which has several important implications. It offers critical evidence for the theoretically assumed relationship between affective and cognitive aspects of insightful problem solving, and suggests that there is a meaningful distinction to be made among different problem solving processes.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Informed consent Informed consent was obtained from all individual participants included in the study.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Data availability The dataset of the present study will be made available at the open repository for psychology data "PsychData" (https://www.psychdata.de/index.php?main=none&sub=none&lang=eng).

References

Abbott, A. (2005). Physics and the public: Science as illusion. *Nature, 434*(7035), 820–820. https://doi.org/10.1038/434820a.

Ash, I. K., Cushen, P. J., & Wiley, J. (2009). Obstacles in investigating the role of restructuring in insightful problem solving. *The Journal of Problem Solving, 2*(2), 6–41. https://doi.org/10.7771/1932-6246.1056.

Ash, I. K., & Wiley, J. (2008). Hindsight bias in insight and mathematical problem solving: Evidence of different reconstruction mechanisms for metacognitive versus situational judgments. *Memory & Cognition, 36*(4), 822–837. https://doi.org/10.3758/MC.36.4.822.

Bowden, E. M. (1997). The effect of reportable and unreportable hints on anagram solution and the Aha! experience. *Consciousness and Cognition, 6*(4), 545–573. https://doi.org/10.1006/ccog.1997.0325.

Cushen, P. J., & Wiley, J. (2012). Cues to solution, restructuring patterns, and reports of insight in creative problem solving. *Consciousness and Cognition, 21*(3), 1166–1175. https://doi.org/10.1016/j.concog.2012.03.013.

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., Fraps, T., von Müller, A., Grothe, B., & Öllinger, M. (2014a). Working wonders? Investigating insight with magic tricks. *Cognition, 130*(2), 174–185. https://doi.org/10.1016/j.cognition.2013.11.003.

Danek, A. H., Fraps, T., von Müller, A., Grothe, B., & Öllinger, M. (2014b). It's a kind of magic—What self-reports can reveal about the phenomenology of insight problem solving. *Frontiers in Psychology, 5*, 1408. https://doi.org/10.3389/fpsyg.2014.01408.

Danek, A. H., & Wiley, J. (2017). What about false insights? Deconstructing the Aha! experience along its multiple dimensions for correct and incorrect solutions separately. *Frontiers in Psychology, 7*, 2077. https://doi.org/10.3389/fpsyg.2016.02077.

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*(1), 47–57. https://doi.org/10.7771/1932-6246.1183.

Davidson, J. E. (1995). The suddenness of insight. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 125–155). Cambridge, MA: MIT Press.

Dominowski, R. L., & Dallob, P. (1995). Insight and problem solving. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 33–62). Cambridge, MA: MIT Press.

Duncker, K. (1926). A qualitative (experimental and theoretical) study of productive thinking (solving of comprehensible problems). *Pedagogical Seminary and Journal of Genetic Psychology, 33*, 642–708.

Durso, F. T., Rea, C. B., & Dayton, T. (1994). Graph-theoretic confirmation of restructuring during insight. *Psychological Science, 5*(2), 94–97.

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

Gick, M. L., & Lockhart, R. S. (1995). Cognitive and affective components of insight. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 197–228). Cambridge, MA: MIT Press.

Hedne, M. R., Norman, E., & Metcalfe, J. (2016). Intuitive feelings of warmth and confidence in insight and noninsight problem solving of magic tricks. *Frontiers in Psychology, 7*, 1314. https://doi.org/10.3389/fpsyg.2016.01314.

Jung-Beeman, M., Bowden, E. M., Haberman, J., Frymiare, J. L., Arambel-Liu, S., Greenblatt, R., et al. (2004). Neural activity when people solve verbal problems with insight. *PLoS Biology, 2*(4), 500–510. https://doi.org/10.1371/journal.pbio.0020097.

Kaplan, C. A., & Simon, H. A. (1990). In search of insight. *Cognitive Psychology, 22*(3), 374–419. https://doi.org/10.1016/0010-0285(90)90008-R.

Knoblich, G., Ohlsson, S., & Raney, G. E. (2001). An eye movement study of insight problem solving. *Memory and Cognition, 29*(7), 1000–1009. https://doi.org/10.3758/BF03195762.

Metcalfe, J. (1986). Premonitions of insight predict impending error. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*(4), 623–634. https://doi.org/10.1037/0278-7393.12.4.623.

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

Ohlsson, S. (1984). Restructuring revisited: II. An information processing theory of restructuring and insight. *Scandinavian Journal of Psychology, 25*, 117–129. https://doi.org/10.1111/j.1467-9450.1984.tb01005.x.

Salvi, C., Bricolo, E., Kounios, J., Bowden, E. M., & Beeman, M. (2016). Insight solutions are correct more often than analytic solutions. *Thinking & Reasoning, 22*(4), 1–18. https://doi.org/10.1080/13546783.2016.1141798.

Utz, F. (2013). *Dimensionen der Einsicht* (Bachelor Thesis). Ludwig-Maximilians-Universität München.

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*, 1424. https://doi.org/10.3389/fpsyg.2016.01424.

Weisberg, R. W., & Alba, J. W. (1981). An examination of the alleged role of "fixation" in the solution of several "insight" problems. *Journal of Experimental Psychology: General, 110*(2), 169–192. https://doi.org/10.1037/0096-3445.110.2.169.

Wertheimer, M. (1925). *Drei Abhandlungen zur Gestalttheorie*. Erlangen: Verlag der Philosophischen Akademie Erlangen.

Wertheimer, M. (1959). *Productive thinking*. New York: Harper.