Psychological Research (2013) 77:659–669
DOI 10.1007/s00426-012-0454-8

ORIGINAL ARTICLE

Aha! experiences leave a mark: facilitated recall of insight solutions

Amory H. Danek · Thomas Fraps · Albrecht von Müller · Benedikt Grothe · Michael Öllinger

Received: 8 May 2012 / Accepted: 16 July 2012 / Published online: 25 September 2012

Abstract The present study investigates a possible memory advantage for solutions that were reached through insightful problem solving. We hypothesized that insight solutions (with Aha! experience) would be remembered better than noninsight solutions (without Aha! experience). 34 video clips of magic tricks were presented to 50 participants as a novel problem-solving task, asking them to find out how the trick was achieved. Upon discovering the solution, participants had to indicate whether they had experienced insight during the solving process. After a delay of 14 days, a recall of solutions was conducted. Overall, 55 % of previously solved tricks were recalled correctly. Comparing insight and noninsight solutions, 64.4 % of all insight solutions were recalled correctly, whereas only 52.4 % of all noninsight solutions were recalled correctly. We interpret this finding as a facilitating effect of previous insight experiences on the recall of solutions.

A. H. Danek (✉)
Graduate School of Systemic Neurosciences, Ludwig-Maximilians-Universität Munich, Großhaderner Str. 2, 82152 Planegg-Martinsried, Germany
e-mail: amory.danek@biologie.uni-muenchen.de

T. Fraps
Trick 17 Magic Concepts, Neureutherstr. 17, 80799 Munich, Germany

A. von Müller · M. Öllinger
Parmenides Foundation, Kirchplatz 1, 82049 Munich, Germany

B. Grothe
Department Biology II, Ludwig-Maximilians-Universität Munich, Großhaderner Str. 2, 82152 Planegg-Martinsried, Germany

M. Öllinger
Department of Psychology, Ludwig-Maximilians-Universität Munich, Großhaderner Str. 2, 82152 Planegg-Martinsried, Germany

Introduction

In contrast to analytical problem solving, insight problems are characterized by a sudden, unexpected solution that is often accompanied by a so-called "Aha! experience" (Metcalfe 1986; Sternberg & Davidson, 1995). Implicitly, it is often assumed that insightful experiences lead to strong memory effects. For example, insight tasks like the nine-dot problem (Scheerer, 1963) are never presented without controlling for prior exposure to the problem. Already at the dawn of insight research, Köhler (1921) reported that his apes were more efficient (shorter solution times compared to the first attempt) in re-solving problems to which they had previously found an insightful solution.

Still, there is a scarcity of studies explicitly addressing this question (Dominowski & Dallob, 1995). An exception is a recent study from Dominowski and Buyer (2000) revealing near-perfect performance in several insight problems which had been successfully solved 1 week before. This "re-solution effect" was not present if participants had failed to solve a problem and then had been shown the solution. This finding was explained by differences between solvers and nonsolvers with regard to their mental representation of a problem, with the solvers building a better integrated and more complete representation.

The idea that Aha! experiences lead to a facilitation of later recall was first posited by Auble, Franks, and Soraci (1979). They presented participants with initially incomprehensible sentences, followed by a cue that revealed the meaning of the sentence. There was a facilitating effect on immediate recall of sentences with Aha! in contrast to sentences without Aha! (Aha! was defined as initial non-comprehension of a sentence followed by comprehension—please note that this differs substantially from the present conceptualization of an Aha! experience, as outlined below). Since participants were not asked about their understanding of the sentence until after the cue was presented 5 s later, possible attempts to solve the problem without cue were not assessed. Therefore, Auble's findings are limited to cued solutions. We agree with Luo and Knoblich (2007, p. 79) who state that "without doubt, the phenomenon of interest is internally generated insight." Consequently, we decided to use a paradigm in which no cues were provided, so that all solutions might be found by participants themselves (self-generated).

Wills, Soraci, Chechile, and Taylor (2000) investigated self-generated insight with pictorial stimuli in the context of the "generation effect", a memory advantage for self-generated over provided items. They found a facilitating effect on immediate recall for connect-the-dot pictures that were drawn by participants (connecting dots until the full picture appeared) in contrast to a presentation of the already complete picture. It was argued that this result was mediated by the Aha! that participants could only experience in the first condition, when the figure became suddenly identifiable during drawing.

Indirect evidence for the generation effect comes from a series of class-room experiments on college students by Weisberg and Alba (1981) who reported low recall rates for provided solutions. The students could not solve the presented task (the nine-dot problem) on their own, even after a cue, therefore the solution was demonstrated to all of them. Approximately 5 months after this initial encounter with the problem, Weisberg and Alba asked for an unexpected recall of the solution. 5 of the 12 participants could not recall the solution, the others needed several minutes to successfully recall it. Independent of their recall performance, all participants could, however, remember that the virtual boundaries of the nine-dot square must be transgressed. Despite this correctly retained information, participants were not able to successfully recall the entire solution. Later on, Kershaw and Ohlsson (2004) were able to explain this puzzling finding by suggesting multiple causes of difficulty for the nine-dot problem whereas at the time of Weisberg and Alba's study (1981), it was still assumed to have only one source of difficulty (i.e. transgressing the virtual boundaries). In a second experiment with a 1-week delay, Weisberg and Alba found 8 of 23 participants unable to recall the solution. This is in accordance with Dominowski's finding of no re-solution effect for provided solutions (2000), also after 1 week. We infer from the reported studies that a possible memory effect of insightful solutions can only be expected if participants solve the problems by themselves.

To conclude this overview, there is first evidence that insightful experiences facilitate recall of initially uncomprehended stimuli in the verbal (sentences) and visual (pictures) domain. An impressive recall performance was also demonstrated for solved classical insight problems, but only if participants generated the solutions by themselves (in contrast to being shown the solution). However, from our point of view, there is one limitation common to all reported studies: a failure to distinguish between insight and noninsight solutions. We agree with Bowden and colleagues (Bowden, Jung-Beeman, Fleck, & Kounios, 2005; Bowden & Jung-Beeman, 2007) in their argument that any problem can be solved with insight, but also without insight. Following this rationale, a direct assessment of the occurrence of insightful experiences is required which poses methodological challenges, but might lead to more valid results. We regard the subjective Aha! experience as the clearest defining characteristic of insight problem solving (Gick & Lockhart, 1995) that accompanies only insightful solutions, and follow Bowden's recommendation (2005) of using participants' subjective reports of Aha! experiences to sort solutions into insight (Aha! reported) and noninsight solutions (Aha! not reported). The occurrence of an insight experience is thus directly determined by the problem solver. If its occurrence is simply assumed without direct feedback from the problem solver, as it was done in previous work (Auble et al., 1979; Dominowski & Buyer, 2000; Wills et al., 2000), it is difficult to claim that the reported findings actually stem from insightful experiences. To our knowledge, the present study is the first one that, with the purpose of investigating memory effects, differentiates between insight and noninsight solutions, using direct Aha! judgments by participants.

In the present work, we will investigate a possible memory effect in the problem-solving domain. We address the question whether gaining sudden insight into the solution of a difficult problem yields strong and long-lasting memory effects, facilitating subsequent recall of a solution.

Besides empirical findings, our hypothesis is theoretically motivated by Knoblich's account (Knoblich, Ohlsson, Haider, & Rhenius, 1999) who claimed that the representational change (Ohlsson, 1992) underlying insightful experiences leads to persisting changes in the representation of a problem (transfer hypothesis). Confirming this claim, at least for short time intervals, they could show positive transfer effects from one solved problem to others with the same source of difficulty, using matchstick arithmetic tasks. The basic idea is that the successful solution of an insight problem is preceded by a change of the problem and/or goal representation and that those representational changes persist over time (Knoblich et al., 1999). For example, in Katona's Triangle Problem (1940), participants are asked to build four equilateral triangles with only six matchsticks. Solvers will realize that this problem is unsolvable in a two-dimensional problem representation, but can easily be solved in three dimensions by building a tetrahedron. Once this constraint (2D) is relaxed, it remains relaxed. Therefore, initially successful solvers will be able to remember later on that the problem must be solved in a 3D presentation (i.e. transfer takes place).

Besides positive transfer effects, also negative transfer can be found if problems with differential solution requirements (different types of constraints) are used: for example, presenting a large student sample with two groups of matchstick arithmetic problems, each problem group requiring a different type of constraint to be relaxed, Kershaw, Braasch, and Flynn (2010) found negative transfer effects, i.e. prior successful solving of one problem type led to longer solution times for the other problem type. In another study placed within the framework of mental set (Luchins, 1942), both positive and negative transfer could be demonstrated (Öllinger, Jones, & Knoblich, 2008).

In the present study, we will try to extend these findings not by investigating short-time transfer from one problem block to another within one experimental session, but by introducing a rather long delay (14 days) and, like Dominowski and Buyer (2000), asking for a recall of solutions to previously solved problems. The present time span of 14 days was chosen because with a delay of 1 week, Dominowski and Buyer (2000) found near-perfect performance (98 %). We wanted to avoid performance at ceiling by introducing a longer delay.

To reach this aim, we will use our newly developed problem-solving task of magic tricks. Participants watch magic tricks and are asked to find out how the magician achieves the magic effect. We argue that magic tricks are especially well suited to investigate representational change, because in order to gain insight into the magicians' secret method, observers must overcome implicit constraints. This was already shown in a previous paper (Danek, Fraps, von Müller, Grothe, & Öllinger, submitted manuscript). Specifically, a number of processes that have been identified as playing a crucial role in insight problem solving can ideally be addressed within the task domain of magic tricks: for example, Grant and Spivey (2003) showed that visually guiding participants' attention towards the critical feature of Duncker's radiation problem (Duncker, 1945) significantly increased solution rates (compare also Knoblich, Ohlsson, & Raney, 2001). This corresponds to the magicians' often-used method of misdirecting observers' attention (e.g. Fraps, 2006; Kuhn, Tatler, & Cole, 2009) to irrelevant locations, e.g. guiding attention away from the secret action that is performed with the left hand by loudly clicking the fingers of the right hand. Similar to Grant and Spivey's study (2003), participants trying to solve a magic trick that is based on misdirection must change their problem representation by re-directing their attention to the relevant locations. Another well-known concept that Duncker developed to explain the difficulty of another one of his classical insight problems (the Candle Problem), is that of functional fixedness (Duncker, 1935). Functional fixedness is exploited by magicians when they present everyday objects like a glass of wine and therefore automatically activate observers' implicit knowledge about these objects (e.g. that glasses break when dropping to the floor). In the case of a magic trick, this knowledge usually turns out to be wrong (the glass remains intact because it is a gimmick, and not a real glass) and that constitutes the surprising effect. Participants in our study must overcome these implicit constraints in order to be able to solve the trick. There is a third process discussed in the insight literature that can easily be manipulated through magic trick stimuli: mental set (Luchins, 1942). Magicians force their audience into a certain mental set by wrapping a story around their magic effect. For example, an observer is asked to put his freely selected card back into a deck of cards and shuffle them. Great emphasis is put on the fact that the card should be put deeply in the middle of the stack and that the cards must be very carefully shuffled, for a very long time etc. The mental set induced in this case would be that the cards are now in completely random order (of course, later on the magician will effortlessly find the card in question right on top of the stack or even behind the observer's ear). Only if the observer achieves to overcome this mental set can the solution of the trick be found out. We conclude from these examples that in the framework of representational change (Ohlsson, 1992; Knoblich et al., 1999), magic tricks can be regarded as insight problems.

Furthermore, we have previously shown that solving a magic trick can lead to subjective Aha! experiences (Danek et al., submitted manuscript). The Aha! experience is known to be associated with strong positive emotional responses (Gick & Lockhart, 1995; Jung-Beeman et al., 2004) which might further strengthen the memory trace (for a review, see LeDoux, 1996, 2000).

Specifically, we hypothesize that insight solutions (accompanied by an Aha! experience) will be remembered better than noninsight solutions lacking the Aha! experience.

We base our hypothesis on two arguments: first, in noninsight solutions, no representational change occurred. Second, the Aha! experience together with its typical positive emotional response is lacking. If representational change and the experience of Aha! are indeed factors that lead to strong and long-lasting memory effects, we may expect better recall for insight than noninsight solutions.

Method

Participants

Fifty students (mean age 24.4; 16 male) were paid for participation according to established levels of payment and were tested individually. Participants were only included if they agreed to come to a second appointment exactly 14 days later and if they gave informed consent to participate in the study. Two participants were excluded because they did not solve any of the presented tasks, resulting in a final sample size of 48.

Testing material

We investigated possible memory effects of insight in the new domain of magic tricks that was previously shown to trigger Aha! experiences (Danek et al., submitted).

The testing material consisted of 37 video clips of magic tricks that had been performed by a professional magician (TF) and recorded in a standardized setting. Clips ranged from 6 to 80 s. The magic stimuli covered a wide range of different magic effects (e.g. transposition, restoration, vanish) and techniques (e.g. misdirection, gimmicks, optical illusions). Stimulus development, a complete list of the magic tricks and the experimental rationale are described in detail elsewhere (Danek et al., submitted).

The magic tricks were presented to participants as a problem-solving task (“Please try to find out how the trick works!”).

Procedure

There were two separate testing sessions with exactly 14 days delay. Session 1 consisted of solving magic tricks, and session 2 of the solution recall. Both sessions lasted about 2 h. Note that in each session, an additional quantitative and qualitative assessment of participants’ individual Aha! experiences was conducted after the end of the experiment. This data is reported in detail in Danek et al. (submitted), but not relevant for the present analysis and therefore not considered further.

Session 1: solving magic tricks

Participants were seated in a distance of 80 cm in front of a computer screen. After filling in an informed consent, participants were orally instructed by the experimenter. Their task was to watch video clips of magic tricks and to find the secret method used by the magician to achieve the magic effect, i.e. the solution to this problem. Following Bowden and Jung-Beeman’s approach (2007), participants were asked to categorize their solution experiences into insight and noninsight solutions. The instruction for these judgements read as follows (adapted from Jung-Beeman et al., 2004): “We would like to know whether you experienced a feeling of insight when you solved a magic trick. A feeling of insight is a kind of “Aha!” characterized by suddenness and obviousness. Like an enlightenment. You are relatively confident that your solution is correct without having to check it. In contrast, you experienced no Aha! if the solution occurs to you slowly and stepwise, and if you need to check it by watching the clip once more. As an example, imagine a light bulb that is switched on all at once in contrast to slowly dimming it up. We ask for your subjective rating whether it felt like an Aha! experience or not, there is no right or wrong answer. Just follow your intuition.”

After three practice trials, a randomized sequence of 34 magic tricks was presented. If a trick was solved, participants had to indicate on a trial-by-trial basis whether they had experienced an Aha! during the solution. If participants failed to solve the trick, the video clip was repeated up to two more times while solving attempts continued.

As soon as they had found a potential solution, participants were required to press a button. The button press stopped the video clip and terminated the trial. A dialog with the following question appeared: did you experience an Aha! moment? Participants indicated Yes or No with a mouse click. Subsequently, they were prompted to type in their solution on the keyboard and gave a certainty rating of how confident they felt about the correctness of their solution on a scale from 0 to 100 %. Figure 1 illustrates the procedure.

Please note that participants never received any feedback about the accuracy of their solutions. To control for familiarity of tricks, at the end of the experiment participants received a questionnaire with screenshots from all 34 tricks (i.e. two stills, one from the beginning and one from the end of the clip) and were asked to indicate whether the solution to a trick was previously known to them. These tricks were excluded on an individual level and handled as missing data.

After this, individually for each participant, solutions to their unsolved tricks were revealed by presenting the same screenshots as in the familiarity questionnaire with a line of text that explained the solution. Although we were only interested in the recall of self-generated solutions (compare introduction), this procedure was necessary to ensure that the memory load during recall was the same for each participant. In this way, all participants had to recall the solutions to 34 magic tricks during the second testing session. The recall data from unsolved tricks was not taken into account for the analysis.

Session 2: solution recall

After 14 days, another testing session was conducted. Participants had not been informed about the nature of the second testing session, the solution recall was therefore completely unexpected. They were instructed in the following way: “Two weeks ago, you observed magic tricks and were asked to find out how the magician achieves the magic effect. Either you found it out by yourself or you were shown the solution. Now we would like to know

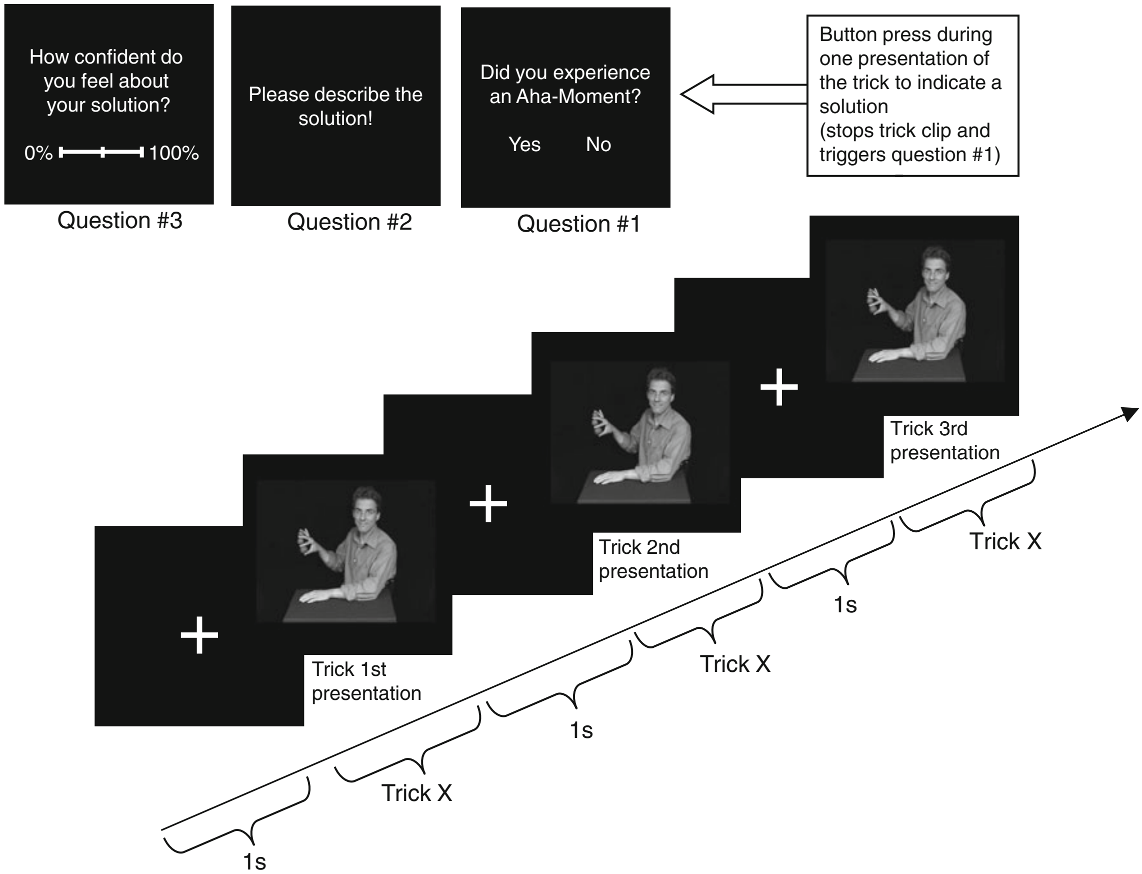

Fig. 1 Procedure of one trial. Different phases and timing are displayed. Note that individual tricks vary in length

which of the solutions you can still remember. Please look at the pictures carefully and try to remember the solution, then type it in. If you can't remember the solution, simply write 'Forgot solution'. It is also possible that you have forgotten not only the solution, but the entire trick—please indicate these cases by writing 'Forgot trick'." Furthermore, it was stressed that they should not generate any new solutions, but only rely on memory. Again, the pictures used to control for familiarity (described above) were used as a reminder of the trick (a second viewing of the entire trick clip was avoided in order to prevent participants from creating new solutions while watching the clip). Participants viewed the trick pictures, typed their answers directly below the picture on the screen and proceeded to the next trick by pressing the enter button.

Results

Response coding and data analysis

Participants solved magic tricks and categorized their solutions into insight (with Aha!) and noninsight solutions (without Aha!), therefore the condition (Solution Type, either insight or noninsight) was determined by participants' responses (binary data). Solution Rate (number of solved tricks), Solution Accuracy (true or false) and Recall (matched or failed) were the dependent variables. A small percentage of all trials (5.2 %) had to be discarded because the tricks were familiar.

Participants' solutions were coded off-line as true or false by two independent raters (Cronbach's alpha as a measure of inter-rater reliability was 0.99). *True solutions* were identical with the procedure that the magician had actually used. *False solutions* consisted of methods that were impossible with respect to the conditions seen in the video clip. If no solution at all had been suggested, the tricks were coded as unsolved. Recall performance was coded as matched and failed by the same two raters (Cronbach's alpha of 0.99). *Matched recall* means that a participant recalled the same true or false solution as in session 1. The *failed recall* category comprised three cases: forgot trick, forgot solution or false memory (if solutions from session 1 and 2 did not match).

To correct for the varying solution rates across participants (ranging from 7 to 27 solved tricks), the mean number of solved tricks for each factor level was weighted, participant-wise, with the participant's individual solution rate. This means, for each participant, the absolute frequency of solved tricks in each category (e.g. true or false, insight or noninsight solution) was divided by the respective participant's individual solution rate. The resulting value indicates which percentage of the individual total number of solved tricks falls into each category. For example, if participant A solved a total of 16 tricks, and 8 of them with insight, this would yield a percentage of 50 % insight solutions, and the remaining 50 % would consist of noninsight solutions. In this way, it could be assured that each participant contributed equally to the statistical analyses. The same rationale applies to the recall rate. Since every participant had solved a different number of tricks, the number of correctly recalled solutions (=matched recall) was dependent on the individual solution rate. For example, a participant who had only solved three tricks could not reach a higher number of correctly recalled tricks than three. To correct for this, the recall rate was also weighted with the individual solution rate, i.e. divided by it.

We were interested in the influence of insight experiences (Solution Type) on subsequent recall of these solutions. Therefore, we asked if the ratio of failed versus matched recall would be different in each of the two Solution Type categories (insight and noninsight). Assuming a memory advantage for insight solutions, we expected that there should be more matched than failed recall events in the insight category, but not in the noninsight category. In addition, the influence of Solution Accuracy had to be taken into account. Therefore, a 2 × 2 × 2 repeated measures ANOVA of the weighted number of solved tricks was conducted, followed by paired *t* tests. All *p* values are Greenhouse–Geisser corrected.

Data from session 1: problem-solving performance

Table 1 provides an overview on the data obtained in session 1. In 45.8 % of all trials, participants watched the magic trick three times without providing any solution. 5.2 % of trials were discarded because participants had indicated that they were already familiar with the solution of the trick. In the remaining 49 % of trials, participants suggested a solution. Solution rates across participants ranged from 7 to 27 solved tricks, with $M = 16.7$ and $SD = 5.03$. Further analyses of solution rates, solution accuracy, certainty and influence of demographic variables on the solution rate are not relevant for the aims of this study, but are presented in Danek et al. (submitted). The present study investigates self-generated insight and, specifically, the influence of Aha! experiences on subsequent recall. Therefore, trials with no solution were excluded from the analysis because no Aha! experiences could occur (compare method). The present analysis is based on the 49 % solved trials (see Table 1).

Table 1 Solution rates collapsed into different categories

Outcome	Frequency (Σ = 1,632)	Percentage of all trials (*n* = 1,632)	
Not solved	747	45.8	
Discarded trials	85	5.2	
True insight solution (with Aha!)	254	15.6	49 % solved
False insight solution (with Aha!)	75	4.6	
True non-insight solution (without Aha!)	263	16.1	
False non-insight solution (without Aha!)	208	12.7	

34 tricks × 48 participants yielded a total of 1,632 trials. 51 % of them were either not solved or discarded due to familiarity of the trick (see first two rows) and 49 % of all trials were solved (see four last rows). False solutions refer to implausible or even physically impossible solution suggestions. Note that absolute frequencies are depicted here, together with percentages of all 1,632 trials

Data from session 2: recall performance

In the second session, participants had to recall their solutions. Figure 2 depicts the weighted mean number of solved tricks for each factor level (note that the percentages always refer to the total number of solved trials) and illustrates that overall, of all solved trials, more were recalled correctly (matched recall, 55 %, grey background) than incorrectly (failed recall, 45 %, white background). Furthermore, there were more noninsight solutions (61 %, plain colour) than insight solutions (39 %, striped) and more true (65 %, right half of the circle) than false solutions (35 %, left half).

For a direct comparison between insight and noninsight solutions with regard to the number of matched recall events, it is necessary to correct for the overall lower occurrence of insight trials (only 39 % vs. 61 % noninsight trials). In doing this, we found that 64.4 % of all insight solutions were recalled correctly, whereas only 52.4 % of all noninsight solutions were recalled correctly.

However, for statistical analyses, we had to take into account the variable Solution Accuracy and then only corrected for the individual varying solution rates (as described in the data analysis section), but not for the lower occurrence of insight. A 2 × 2 × 2 ANOVA for repeated measures with the factors Solution Type (insight vs. noninsight), Accuracy (true vs. false) and Recall (matched vs. failed) was conducted, with the weighted number of solved tricks as dependent variable. It revealed significant main effects for two factors: Solution Type with $F(1, 47) = 10.78$, $p = .002$, $\eta^2_{\text{partial}} = .19$ (more noninsight solutions with 61 % than insight solutions with only 39 %) and

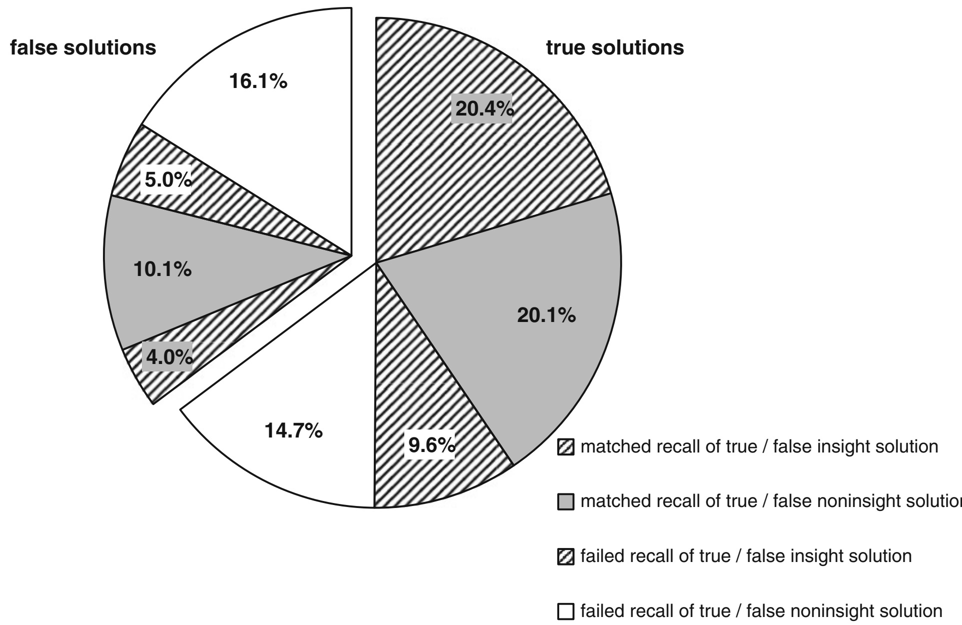

Fig. 2 Overview on the data from session 2. Weighted mean number of solved tricks (in %) and their proportion of matched and failed recall in the insight and noninsight categories, depicted separately for false and true solutions

Accuracy with $F(1, 47) = 45.99$, $p < .001$, $\eta^2_{\text{partial}} = .50$ (more true, 65 %, than false solutions, 35 %). There was no significant main effect for Recall (matched recall, 55 % vs. failed recall, 45 %). There were significant two-way interactions between Solution Type and Accuracy: $F(1, 47) = 6.96$, $p = .011$, $\eta^2_{\text{partial}} = .13$, Solution Type and Recall: $F(1, 47) = 4.64$, $p = .036$, $\eta^2_{\text{partial}} = .09$ and finally, between Accuracy and Recall: $F(1, 47) = 36.03$, $p < .001$, $\eta^2_{\text{partial}} = .43$.

Note that taking into account the mean viewing time for each trick as a covariate for this ANOVA does not significantly change the results.

We wanted to investigate the influence of insight experiences (factor Solution Type) on subsequent recall of these solutions. But due to the significant interaction between Accuracy and Recall, the two factor levels of Accuracy (i.e. true and false solutions) must be considered separately, therefore Fig. 3 refers only to false solutions and Fig. 4 only to true solutions.

Note that the number of solved tricks in each category (compare Fig. 2) was weighted by the individual solution rate of each participant, as described in the data analysis section. Grey bars indicate failed recall, black bars matched recall.

A comparison of the two figures shows that, in general, false (i.e. impossible) solutions (Fig. 3) are more likely to be forgotten (grey bars > black bars) whereas true solutions (Fig. 4) are more likely to be recalled correctly (black bars > grey bars).

Due to the significantly higher number of noninsight (61 %) relative to insight (39 %) solutions, we could not directly compare the number of matched/failed recall events between insight and noninsight solutions. For example, in

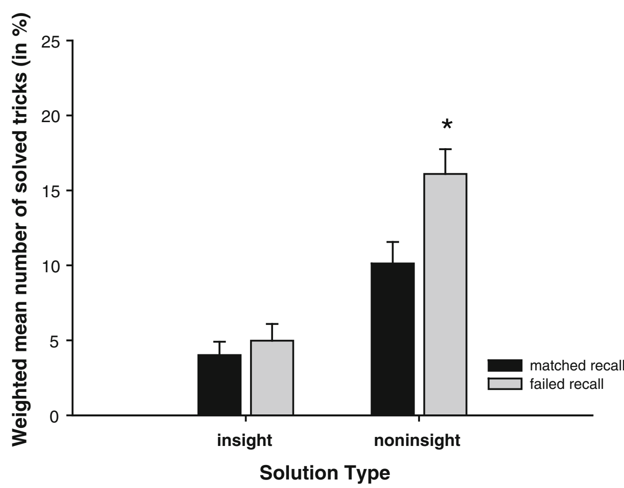

Fig. 3 Recall of false solutions. The weighted mean number of solved magic tricks (in % of all 800 solved trials) is depicted as a function of Solution Type and Recall. Note that only false solutions are presented (see Fig. 4 for true solutions). *Error bars* denote standard errors of the mean. Significant differences between matched (*black bars*) and failed (*grey bars*) recall are marked with an *asterisk*

Fig. 4, the percentage of matched recall events is about the same for insight and noninsight, but the two categories are not directly comparable because the insight category is based on a much lower number of trials (it would be necessary to correct for the lower occurrence of insight trials, as presented above, see p.16). Therefore, we asked if the ratio of failed/matched recall would be different in each of the two categories (insight and noninsight). Assuming a memory advantage for insight solutions, there should be more matched than failed recall events in the insight category, but not in the noninsight category. This was analysed with post hoc *t* tests.

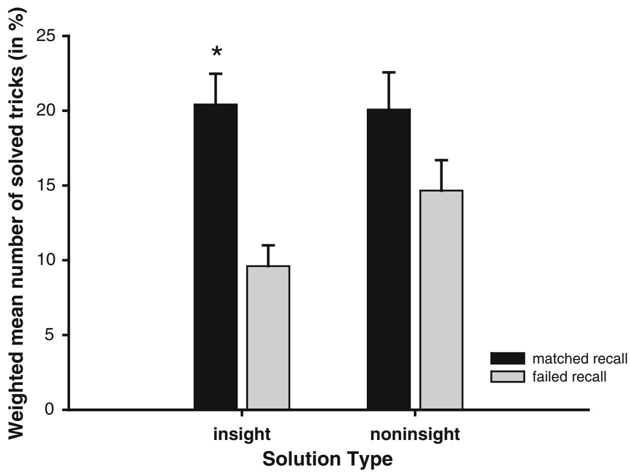

Fig. 4 Recall of true solutions. The weighted mean number of solved magic tricks (in % of all 800 solved trials) is depicted as a function of Solution Type and Recall. Note that only true solutions are presented. *Error bars* denote standard errors of the mean. Significant differences between matched (*black bars*) and failed (*grey bars*) recall are marked with an *asterisk*

First, only false solutions are considered (Fig. 3): follow up paired t tests yielded a significant difference between matched and failed recall solely in the noninsight category [$t(47) = 2.63$, $p = .011$, Cohen's dz = .39]—in contrast to the insight category with no difference. Only 10.1 % of noninsight, false solutions were recalled correctly, and for 16.1 % the recall failed.

Second, only true solutions are taken into account (Fig. 4): here, follow up paired t tests showed that insight solutions included a significantly higher percentage of matched recall (20.42 %) than failed recall (9.61 %) with [$t(47) = 4.6$, $p < .001$, Cohen's dz = .61]. For noninsight solutions, the percentages did not differ significantly.

These results confirmed our hypothesis. The present analysis revealed differential ratios of failed/matched recall in the two Solution Type categories with more matched than failed recall events for insight solutions (in the case of true solutions) and more failed than matched recall events for noninsight solutions (in the case of false solutions). We take this as evidence for a facilitating effect of previous insight experiences on the recall of solutions.

Discussion

The present work addressed the question whether gaining sudden insight into the solution of a difficult problem would facilitate the recall of these solutions relative to noninsight solutions. On average, participants were able to solve 49 % of the presented tasks, and 39 % of these solutions were classified as insightful (Aha! experience reported).

The rather low solution rate of 49 % was expected, because we used real magic tricks that, of course, are difficult to solve. Previously, we had decided against eliminating extremely difficult ones, hoping that these might trigger especially strong Aha! experiences if solved. The present recall rate of 55 % is much lower than the 98 % found by Dominowski and Buyer (2000). We explain this by the higher number of problems implemented. Compared to the six insight problems used by Dominowski and Buyer, our participants had to keep in mind 34 different magic tricks including solutions. In contrast to their design, we did not present the entire problem again during recall, but only showed a reminder of the trick (a still from the video). This increased the number of failed recall events, because participants had forgotten a substantial amount of tricks. The low recall rate could also be attributed to the longer time delay (14 days instead of seven days).

The data obtained confirmed our hypothesis. We predicted that insight solutions (39 % of all solved trials) would be remembered better than noninsight solutions (61 %). Regarding only true solutions, it was found that insight solutions included a significantly higher percentage of matched recall (20.42 %) than failed recall (9.61 %). According to Cohen (1988), the corresponding effect size of dz = .61 can be regarded as a strong effect. For noninsight solutions, no significant difference existed. Therefore, the previously experienced Aha! seems to lead to a memory advantage.

This finding extends the earlier stated transfer hypothesis (Knoblich et al., 1999) to a much longer time delay (14 days) and to the recall of solutions to previously solved problems. It provides support for the proposal that the representational change (Ohlsson, 1992) underlying insightful solving experiences leads to long-lasting changes in the representation of a problem that lead to full retention of the problem's solution.

One possible explanation for this finding could be the high emotional involvement during an Aha! experience that might facilitate the retention of insight solutions in memory (as suggested by Seifert, Meyer, Davidson, Patalano, & Yaniv, 1995), in contrast to noninsight solutions for which the emotional response is lacking. Following the somatic marker hypothesis with its claim that a somatic state can become linked to a memory content (Damasio, 1996), a possible mechanism could work like this: the emotional state experienced during the insight moment (recognized and classified by participants as "Aha! experience") becomes linked to a cognitive state, namely the information about the solution of the problem. It is now generally acknowledged that emotional events are remembered with greater accuracy and vividness than neutral events (Reisberg & Hertel, 2004). The most prominent examples are flashbulb memories (Brown & Kulik, 1977). However, the functional mechanisms of this memory enhancement are widely debated and still not clarified (Levine & Pizarro, 2004; Van Giezen, Arensman, Spinhoven, & Wolters, 2005). For example, it remains unclear which type of information is better recalled: specific visual details about the objects causing the emotional arousal or rather context information (Kensinger, Garoff-Eaton, & Schacter, 2006). At present, the impact of positive emotions (thought to be involved in Aha! experiences, Gick & Lockhart, 1995) is less well documented than that of negative emotions, for which there is plenty of evidence for a memory enhancement effect. Cases of posttraumatic stress disorder illustrate how strong negative emotions, experienced during the traumatic event, can hinder forgetting (e.g. Van der Kolk, 1994). A recent example is provided by Pezdek (2003) who questioned US college students about their memory of the events on September 11th (World Trade Center attacks) and found more accurate event memory in the New York sample (shown to be more distressed by this event) than control samples from Hawaii and California. At the same time, the autobiographical memory of the New York group was least accurate. Similarly, Smith, Bibi, and Sheard (2003) reported a positive correlation of event memory with the level of emotional arousal. However, other studies have shown detrimental effects of strong emotions on recall performance (e.g. Adolphs, Denburg, & Tranel, 2001), at least on the memory for details. In summary, the effects of emotion on recall performance are not completely understood yet.

Nevertheless, the possible link between strong emotional reactions and memory for insightful solutions would also be meaningful from an anatomical perspective. It has been suggested (Öllinger, 2005) that during insight problem solving, there is an intense interplay between the hippocampus and the amygdala. The amygdala is known as a crucial region for processing of emotionally relevant stimuli (e.g. Davidson & Irwin, 1999) whereas the hippocampus subserves memory consolidation (e.g. Tulving & Markowitsch, 1998). In an fMRI study, Luo and Niki (2003) detected hippocampus activity during insight problem solving of Japanese riddles. In the framework of the representational change theory (Ohlsson, 1992; Knoblich et al., 1999), this proposal can be further elaborated: it was claimed recently that the dorsolateral prefrontal cortex (DLPFC) might play an important role in determining the goal representation of a given problem (Frith, 2000). The basic idea is that the DLPFC biases the response space by activating a set of potential solution strategies which initially seem to be appropriate to solve the problem at hand (Frith, 2000). However, because these solution strategies are based on wrong assumptions (i.e. constraints), the representation is incorrect and no solution can be found (Ohlsson, 1992) unless the constraints are relaxed (compare introduction). There is evidence for this proposal from a study on brain-lesioned patients (Reverberi, Toraldo, D'Agostini, & Skrap, 2005) where it was impressively demonstrated that patients with lesions to the lateral frontal lobe were more successful than healthy controls in solving very difficult insight problems (matchstick arithmetic tasks) that required many constraints to be relaxed. This was explained by the representational change theory (Ohlsson, 1992; Knoblich et al., 1999) as well as by Frith's (2000) account. Therefore, the DLPFC might be the brain site where constraints are activated by prior knowledge. Just like other insight problems, magic tricks require the observer to overcome constraints that are induced by the magician, as detailed in the introduction. We assume that there is a mismatch between the initially activated, biased problem representation and the observed magic effect, e.g. a flying table. In this example, one constraint would consist of initially representing the table as an ordinary, normal object—and flying is clearly not compatible with our prior knowledge about tables (including e.g. their weight). The crucial brain site for detection of such cognitive conflict might be the anterior cingulate cortex (ACC), as shown by Mai, Luo, Wu, and Luo (2004) in an ERP study on Chinese riddles and also by Luo, Niki, and Phillips (2004) in an imaging study. There is further evidence for an implication of the ACC in insight processes (Starchenko, Bekhtereva, Pakhomov, & Medvedev, 2003; Kounios et al., 2006; Aziz-Zadeh, Kaplan, & Iacoboni, 2009). It is conceivable that the detected mismatch triggers the emotional arousal that is mediated by the amygdalae and that facilitates the encoding of such a newly gained insight into the memory system. However, since the neural basis of insightful experiences still is not completely clarified (see Dietrich & Kanso, 2010, for a very thorough review), further research is warranted to confirm these speculative explanations.

A pertaining methodological challenge in insight research is the question of how the occurrence of insight can be accurately assessed (Haider & Rose, 2007; Luo & Knoblich, 2007; Ash, Cushen, & Wiley, 2009). The present study demonstrates that the method of obtaining direct insight judgements from participants can reveal interesting differences between insight and noninsight events. Of course, a consequence of this procedure is that only some trials are classified as "insightful", compared to the traditional approach of simply treating all solved trials as insightful events. But in our opinion, this is the only way to make sure that the subjective insight experience is actually measured and this might help us to get a firmer grip on this elusive phenomenon.

A drawback of the present study is the rather high number of false solutions (17.3 % of all trials) that made it impossible to simply dismiss them. We therefore analysed true and false solutions separately and found differential ratios of failed/matched recall for insight and noninsight solutions in both cases. Interestingly, both patterns speak for the same facilitating effect of insight solutions on recall performance, but one pattern is reversed to the other. In the case of true solutions, there are more matched than failed recall events for insight solutions. In the case of false solutions, there are more failed than matched recall events for noninsight solutions.

In summary, using a new problem-solving paradigm, the present work demonstrates that insight solutions are remembered better than noninsight solutions. This finding is in accordance with the theoretical assumption of representational change and the resulting transfer of knowledge about a solution. A replication of these results, also with stimuli from classical insight task domains, must be awaited and future studies addressing the neural basis of this effect are needed to clarify the role of the insight experience in memory.

Acknowledgments We thank Matus Simkovic for help with programming the experiment.

References

Adolphs, R., Denburg, N. L., & Tranel, D. (2001). The amygdala's role in long-term declarative memory for gist and detail. *Behavioral Neuroscience, 115*(5), 983–992.

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.

Auble, P. M., Franks, J. J., & Soraci, S. A. (1979). Effort toward comprehension: Elaboration or "aha!"? *Memory and Cognition, 7*(6), 426–434.

Aziz-Zadeh, L., Kaplan, J. T., & Iacoboni, M. (2009). "Aha!": The neural correlates of verbal insight solutions. *Human Brain Mapping, 30*(3), 908–916.

Bowden, E. M., & Jung-Beeman, M. (2007). Methods for investigating the neural components of insight. *Methods, 42*(1), 87–99.

Bowden, E. M., Jung-Beeman, M., Fleck, J. I., & Kounios, J. (2005). New approaches to demystifying insight. *Trends in Cognitive Sciences, 9*(7), 322–328.

Brown, R., & Kulik, J. (1977). Flashbulb memories. *Cognition, 5*(1), 73–99.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Hillsdale: Lawrence Erlbaum.

Damasio, A. R. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophical Transactions: Biological Sciences, 351*(1346), 1413–1420.

Davidson, R. J., & Irwin, W. (1999). The functional neuroanatomy of emotion and affective style. *Trends in Cognitive Sciences, 3*(1), 11–21.

Dietrich, A., & Kanso, R. (2010). A review of EEG, ERP, and neuroimaging studies of creativity and insight. *Psychological Bulletin, 136*(5), 822–848.

Dominowski, R. L., & Buyer, L. S. (2000). Retention of problem solutions: The re-solution effect. *The American Journal of Psychology, 113*(2), 249–274.

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: MIT Press.

Duncker, K. (1935). *Zur Psychologie des produktiven Denkens*. Berlin: Springer.

Duncker, K. (1945). On problem-solving. *Psychological Monographs*, *58*, Whole No. 270.

Danek, A. H., Fraps, T., von Müller, A., Grothe, B., & Öllinger, M. (2012). Working wonders? Investigating insight with magic tricks (submitted manuscript).

Fraps, T. (2006). The blind spot. *Magic, 15*(6), 1–6.

Frith, C. D. (2000). The role of dorsolateral prefrontal cortex in the selection of action as revealed by functional imaging. In S. Monsell & J. Driver (Eds.), *Control of cognitive processes: Attention and performance XVIII* (pp. 549–565). Cambridge: MIT Press.

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: MIT Press.

Grant, E. R., & Spivey, M. J. (2003). Eye movements and problem solving. *Psychological Science, 14*(5), 462–466.

Haider, H., & Rose, M. (2007). How to investigate insight: A proposal. *Methods, 42*(1), 49–57.

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.

Katona, G. (1940). *Organizing and memorizing: Studies in the psychology of learning and teaching*. New York: Columbia University Press.

Kensinger, E. A., Garoff-Eaton, R. J., & Schacter, D. L. (2006). Memory for specific visual details can be enhanced by negative arousing content. *Journal of Memory and Language, 54*(1), 99–112.

Kershaw, T. C., Braasch, J. L. G., & Flynn, C. K. (2010). Negative transfer in matchstick arithmetic insight problems. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the 32nd Annual Meeting of the Cognitive Science Society* (pp. 1505–1510). Austin: Cognitive Science Society.

Kershaw, T. C., & Ohlsson, S. (2004). Multiple causes of difficulty in insight: The case of the nine-dot problem. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 30*(1), 3–13.

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

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

Köhler, W. (1921). *Intelligenzprüfungen am Menschenaffen*. Berlin: Springer.

Kounios, J., Frymiare, J. L., Bowden, E. M., Fleck, J. I., Subramaniam, K., Parrish, T. B., et al. (2006). The prepared mind. *Psychological Science, 17*(10), 882–890.

Kuhn, G., Tatler, B. W., & Cole, G. G. (2009). You look where I look! Effect of gaze cues on overt and covert attention in misdirection. *Visual Cognition, 17*(6), 925–944.

LeDoux, J. E. (1996). *The emotional brain*. New York: Simon & Schuster.

LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience, 23*(1), 155–184.

Levine, L. J., & Pizarro, D. A. (2004). Emotion and memory research: A grumpy overview. *Social Cognition, 22*(5), 530–554.

Luchins, A. S. (1942). Mechanization in problem solving: The effect of Einstellung. *Psychological Monographs, 54*(6), 1–95.

Luo, J., & Knoblich, G. (2007). Studying insight problem solving with neuroscientific methods. *Methods, 42*(1), 77–86.

Luo, J., & Niki, K. (2003). Function of hippocampus in "insight" of problem solving. *Hippocampus, 13*(3), 316–323.

Luo, J., Niki, K., & Phillips, S. (2004). Neural correlates of the Aha! reaction. *NeuroReport, 15*(13), 2013–2017.

Mai, X.-Q., Luo, J., Wu, J.-H., & Luo, Y.-J. (2004). "Aha!" effects in a guessing riddle task: An event-related potential study. *Human Brain Mapping, 22*(4), 261–270.

Metcalfe, J. (1986). Feeling of knowing in memory and problem solving. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 12*(2), 288–294.

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

Öllinger, M. (2005). *Untersuchungen über den Einfluss von Einstellungseffekten und Heuristiken beim Lösen von Einsichtsproblemen*. Berlin: Wissenschaftlicher Verlag Berlin.

Öllinger, M., Jones, G., & Knoblich, G. (2008). Investigating the effect of mental set on insight problem solving. *Experimental Psychology, 55*(4), 270–282.

Pezdek, K. (2003). Event memory and autobiographical memory for the events of September 11, 2001. *Applied Cognitive Psychology, 17*(9), 1033–1045.

Reisberg, D., & Hertel, P. (2004). *Memory and Emotion*. New York: Oxford University Press.

Reverberi, C., Toraldo, A., D'Agostini, S., & Skrap, M. (2005). Better without (lateral) frontal cortex? Insight problems solved by frontal patients. *Brain, 128*(12), 2882–2890.

Scheerer, M. (1963). Problem-solving. *Scientific American, 208*(4), 118–128.

Seifert, C. M., Meyer, D. E., Davidson, N., Patalano, A. L., & Yaniv, I. (1995). Demystification of cognitive insight: Opportunistic assimilation and the prepared-mind perspective. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 33–62). Cambridge: MIT Press.

Smith, M. C., Bibi, U., & Sheard, D. E. (2003). Evidence for the differential impact of time and emotion on personal and event memories for September 11, 2001. *Applied Cognitive Psychology, 17*(9), 1047–1055.

Starchenko, M. G., Bekhtereva, N. P., Pakhomov, S. V., & Medvedev, S. V. (2003). Study of the brain organization of creative thinking. *Human Physiology, 29*(5), 652–653.

Sternberg, R. J., & Davidson, J. E. (1995). *The nature of insight*. Cambridge: MIT Press.

Tulving, E., & Markowitsch, H. J. (1998). Episodic and declarative memory: Role of the hippocampus. *Hippocampus, 8*(3), 198–204.

Van der Kolk, B. A. (1994). The body keeps the score: Memory and the evolving psychobiology of posttraumatic stress. *Harvard Review of Psychiatry, 1*(5), 253–265.

Van Giezen, A. E., Arensman, E., Spinhoven, P., & Wolters, G. (2005). Consistency of memory for emotionally arousing events: A review of prospective and experimental studies. *Clinical Psychology Review, 25*(7), 935–953.

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.

Wills, T. W., Soraci, S. A., Chechile, R. A., & Taylor, A. H. (2000). "Aha" effects in the generation of pictures. *Memory and Cognition, 28*(6), 939–948.