COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 43 (2019) e12771

ISSN: 1551-6709 online
DOI: 10.1111/cogs.12771

When the Solution Is on the Doorstep: Better Solving Performance, but Diminished Aha! Experience for Chess Experts on the Mutilated Checkerboard Problem

Merim Bilalić,[a] Mario Graf,[b] Nemanja Vaci,[c] Amory H. Danek[d]

[a]*Department of Psychology, University of Northumbria at Newcastle*
[b]*Institute of Psychology, University of Klagenfurt*
[c]*Department of Psychiatry, University of Oxford*
[d]*Department of Psychology, University of Heidelberg*

Received 18 June 2018; received in revised form 10 June 2019; accepted 14 June 2019

Abstract

Insight problems are difficult because the initially activated knowledge hinders successful solving. The crucial information needed for a solution is often so far removed that gaining access to it through restructuring leads to the subjective experience of “Aha!”. Although this assumption is shared by most insight theories, there is little empirical evidence for the connection between the necessity of restructuring an incorrect problem representation and the Aha! experience. Here, we demonstrate a rare case where previous knowledge facilitates the solving of insight problems but reduces the accompanying Aha! experience. Chess players were more successful than non-chess players at solving the mutilated checkerboard insight problem, which requires retrieval of chess-related information about the color of the squares. Their success came at a price, since they reported a diminished Aha! experience compared to controls. Chess players’ problem-solving ability was confined to that particular problem, since they struggled to a similar degree to non-chess players to solve another insight problem (the eight-coin problem), which does not require chess-related information for a solution. Here, chess players and non-chess players experienced the same degree of insight.

Keywords: Insight; Aha! Experience; Restructuring; Problem representation; Expertise; Chess

1. Introduction

Insight problems are notoriously difficult. Typically, the first attempts to solve such problems are not successful, because they require a new and non-obvious solution which does not readily emerge from the way the problem is presented (Dominowski & Dallob,

Correspondence should be sent to Merim Bilalić, Department of Psychology, University of Northumbria at Newcastle, Ellison Building, Newcastle upon Tyne NE1 8ST, UK. E-mail: merim.bilalic@northumbria.ac.uk

1995). It was shown that insight problems mislead the problem solver into setting up an inappropriate initial problem representation, in contrast to non-insight problems such as standard arithmetic problems which have no misleading structure and can be solved by directly applying familiar routines (Ash, Jee, & Wiley, 2012). Take, for example, the eight-coin insight problem presented in Fig. 1B (Ormerod, MacGregor, & Chronicle, 2002). The task is to move only two coins so that each coin in the new constellation touches exactly three other coins. Stacking coins on top of each other solves the problem. The presentation of the problem as a two-dimensional picture, however, leads to the implicit assumption that one can only move the coins around each other, in the same plane. This inappropriate representation makes it difficult to think of the stacking solution, which requires the third dimension.

Classical theories of insight postulate that insight problems automatically activate prior knowledge and assumptions (e.g., solving only in the plane), which prevent finding the right solution (Duncker, 1945; Ohlsson, 1984, 1992a, 1992b). Solvers need to disengage from their initial, inappropriate view of the problem, in order to notice the crucial, but remote information and relate it adequately to the given problem. This process is called restructuring and refers to change in the mental representation of the problem: A fundamental alteration in how the problem is perceived and encoded (Ohlsson, 1984, p. 119). According to the Representational Change Theory (Knoblich, Ohlsson, Haider, & Rhenius, 1999; Ohlsson, 1992b, 2011), possible mechanisms to accomplish restructuring are elaboration (adding new information), re-encoding (categorizing or perceiving the problem elements differently), and constraint relaxation (overcoming self-imposed assumptions). For example, in the eight-coin problem, the realization that a 3D representation of the problem is possible would represent constraint relaxation (Öllinger, Jones, Faber, & Knoblich, 2013; Ormerod et al., 2002). It has been shown that the difficulty of a problem is directly determined by the degree of necessary constraint relaxation (Öllinger, Jones, & Knoblich, 2006).

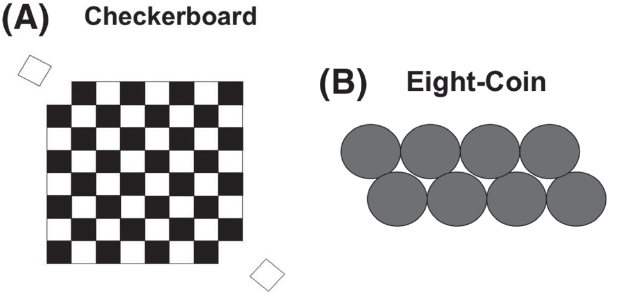

Fig. 1. Insight problems used in the study. (A) The mutilated checkerboard problem, where the task is to check whether one can cover the remaining 62 squares (after removing two) with 31 dominoes, if one domino always covers two squares. It is impossible to cover all the squares, since one domino would always cover one black and one white square. The removed squares are both white, which leaves 30 white squares and 32 black squares. (B) The eight-coin problem, where the goal is to find a way of moving two coins so that each coin in the new constellation touches precisely three other coins. The solution is to move the two middle coins onto the top of the remaining two groups of three coins.

Most researchers agree that, besides the cognitive component (restructuring), the insight phenomenon is also characterized by an affective component, a strong positive response (Bowden, Jung-Beeman, Fleck, & Kounios, 2005; Danek, 2018; Duncker, 1945; Gick & Lockhart, 1995; Kaplan & Simon, 1990). When solutions to insight problems are eventually reached, they tend to be rather sudden and surprising, often leaving people with a strong feeling of Aha! (Gick & Lockhart, 1995; Gruber, 1995). Karl Duncker, one of the key figures in the Gestalt movement, first noticed the connection between the cognitive and affective component: “The decisive points in thought-processes, the moment of sudden comprehension, of the ‘Aha!,’ of the new, are always at the same time moments in which such a *sudden restructuring* [emphasis added] of the thought-material takes place” (Duncker, 1945, p. 29). Many insight theories therefore propose that the process of restructuring is one of the main factors driving the Aha! experience (Dominowski & Dallob, 1995; Gick & Lockhart, 1995; Kaplan & Simon, 1990). For example, a recent model of insight (Danek, 2018) postulates that if the initial problem representation is correct, no Aha! will be triggered because a search within the initial representation allows to retrieve the crucial problem element, without any restructuring. Fleck and Weisberg’s four-stage model of problem solving also makes the same prediction, namely that if any given problem can be matched with information that is easily accessible in memory, the solution transfers and there should be no Aha! experience (Fleck & Weisberg, 2013). Knoblich and colleagues (Knoblich, Ohlsson, & Raney, 2001) propose that the sudden emergence of crucial information in working memory constitutes an Aha! experience. In their eye-tracking study (Knoblich et al., 2001), they demonstrated that, shortly before the solution was found, successful solvers allocated their attention to those elements of the problem that had been ignored previously (for a similar finding, see also Grant & Spivey, 2003).

So far, there is hardly any empirical evidence for the theoretical link between the cognitive and affective component of insight. An exception is a recent study which demonstrated the difference in the subjective Aha! ratings of solutions arising from objectively sudden or incremental changes in solvers’ problem representations (Danek, Williams, & Wiley, 2018). Sudden changes produced stronger subjective Aha! rating than incremental changes. One reason for this lack of research is the long-held assumption that solving insight problems inevitably leads to an Aha! experience, without ever measuring the affective component (for a critique of this approach, see Weisberg, 1995). For example, the eye tracking study (Knoblich et al., 2001) discussed above never asked about the actual experience of the solvers. Only recently have researchers started to measure the affective or phenomenological component of insight, which is necessary to determine whether insight was indeed subjectively "experienced" by the solver (e.g., Bowden et al., 2005; Danek, Fraps, Von Mueller, Grothe, & Öllinger, 2014). This line of research commonly employs self-reports and ratings on a number of different aspects, which are believed to constitute the Aha! experience (Danek et al., 2014; Danek & Wiley, 2017; Bowden & Grunewald, 2018; Topolinski & Reber, 2010; Webb, Little, & Cropper, 2016). Although a consensus about what constitutes an Aha! experience has not yet been reached, there are theoretical and empirical reasons to focus on the dimensions of *Suddenness*, *Surprise*, *Pleasure*, and *Certainty*.

Gick and Lockhart (1995) argue that the Surprise aspect stems from the fact that the new, changed problem representation is fundamentally different from the one used in initial solution attempts. A feeling of Suddenness arises because this new representation (if correct) yields the solution more or less immediately. Insight problem solving has often been related to feelings of pleasure or happiness (e.g., Gruber, 1995). A recent study (Danek et al., 2014) found that the feeling of pleasure featured prominently both in solvers' open-ended self-reports about how an Aha! moment felt, and in quantitative ratings, where it was more strongly endorsed than any other dimension. Finally, a feeling of confidence that an insightful solution is correct (before any conscious verification) is often reported not only in the context of scientific discoveries where experts are supposed to have an "intuitive sense of success" (Gick & Lockhart, 1995, p. 215), but also in experimental settings involving non-expert problem solvers (e.g., Danek et al., 2014; Hedne, Norman, & Metcalfe, 2016; Webb et al., 2016).

These four dimensions (i.e., Suddenness, Surprise, Pleasure, Certainty) were chosen for this study not only for theoretical reasons, but also on the basis of recent empirical evidence. In a study using magic tricks as problem-solving tasks, Danek and Wiley (2017) asked participants after each solution attempt to rate their solution experience on six different dimensions, together with an overall Aha! rating. They found that Pleasure, Suddenness, and Certainty all contributed unique variance to the overall Aha! rating and concluded that these are core aspects of the Aha! phenomenology. The role of Surprise is currently unclear, as the same study (Danek & Wiley, 2017) found that Surprise had the lowest correlation with Aha! compared to the other dimensions. However, in another recent study, Surprise was highly correlated with Aha! ratings (Webb et al., 2016).

Here, we investigate the hypothesis of a close relation between restructuring and the Aha! experience by making the crucial piece of knowledge more accessible. We employ the checkerboard problem, in which participants need to prove whether or not it is possible to cover the remaining 62 squares with 31 dominoes after the two diagonal corner squares have been removed (Fig. 1A). The checkerboard problem is difficult, as evidenced by the unfortunate student from the study of Kaplan and Simon (1990) who spent almost a full day trying to devise a mathematical proof for the problem. This anecdote illustrates how prior knowledge and its initial activation may lead the solver onto the wrong path, because the problem triggers an initial representation that hinders a solution. The initial representation usually includes the concept of "covering" and, at first glance, it seems possible to cover 62 squares with 31 dominoes. This leads problem solvers to an exhaustive testing of possible coverings, which in turn is rather taxing on working memory. Solving the checkerboard problem requires a switch from the incorrect "covering representation" to a "parity representation." Parity means taking into account the alternation of the squares (i.e., black–white–black–white) and realizing that the two removed squares are of the same color which renders the task impossible (see Fig. 1A).

We argue that the critical feature of parity will be more readily available to chess players because they engage with this property on a daily basis. The color of the squares also plays an important role in the game of chess. For example, some chess pieces such as bishops can access the squares of only one color. The color feature is engrained in the chess mind to such an extent that proficient players can retrieve the content of a particular square, including its color, without having the chessboard in view (Oliver & Ericsson, 1984, reported in Ericsson & Staszewski, 1989). Therefore, chess players will be more likely to set up a correct initial representation, which allows them to quickly solve the problem. In contrast, the problem was shown to be scarcely solvable without hints (Kaplan & Simon, 1990) for non-chess players. We therefore hypothesize that we should find higher solution rates for chess players compared to non-chess players. The neutral eight-coin problem (Fig. 1B) should, however, be similarly difficult for both groups of participants.

The second hypothesis tested here is that the availability of the crucial information will have an impact on the solving process and thus on the likelihood and strength of the Aha! experience. We hypothesize that, on the checkerboard problem, chess players who solve the problem will experience fewer Aha! than non-chess players, as operationalized by self-reports of Suddenness, Pleasure, and Surprise. In contrast, both groups should experience the Aha! phenomenon to a similar extent on the neutral eight-coin problem. We also assume that chess players should be more certain of the checkerboard solution due to their expertise. No such differences were expected for the eight-coin problem.

Finally, we gave hints during the problem-solving period to those participants who could not find the solution within the first 5 min. This intervention had a twofold purpose. On the one hand, it shortened solution times, which may otherwise be too long for laboratory testing. On the other hand, hints presented an additional check as to whether the parity feature really constituted the crucial solution element. Should the information provided by the hints be the crucial one, we should see a considerable rise in solution rates after hints.

2. Method

Please note that the stimuli, raw data, and bootstrapping code including its results can be found at the Open Science Framework site: https://osf.io/g7fvq/?view_only=001931f1366644f3b6cbad80377c8100

2.1. Participants

There were 85 students from Klagenfurt University and Ludwig-Maximilians-Universität München[1] (82% female, $M_{age} = 23.7 \pm SD_{age} = 6$ years) who participated voluntarily or for course credit. One participant was excluded from the final sample because it was not possible to track his responses, while two others were not shown the checkerboard task due to technical problems. Therefore, 82 were tested on the checkerboard problem, while 84 were tested on the eight-coin problem (see Stimuli section below). There were 34 chess players (94% male, $M_{age} = 29.6 \pm 8.4$ years) who were recruited from local chess clubs in Klagenfurt, Tübingen, Tuzla, Sarajevo, and Newcastle. The differing number of the participants in the two groups reflects their availability. While it is not difficult to get access to undergraduate students, the most common participant group in psychology, skilled chess players are much rarer. Expertise studies often involve only a dozen participants because experts are by definition rare (Bilalić, 2017). Our chess players, for example, were particularly skilled, as they were better than 90% of all chess players. Chess skill is measured by an interval scale where the theoretical average is 1,500 Elo points and the theoretical standard deviation is 200 Elo points (Elo, 1978; Vaci & Bilalić, 2017; Vaci, Gula, & Bilalić, 2014, 2015). Players had an average rating of 2,083 ± 192 Elo (range: 1,710–2,416), which makes them three standard deviations above average chess players. The small number of experts in expertise studies is also reflected in the typically large effect sizes for the differences between experts and novices (Bilalić, 2017; Campitelli & Speelman, 2013). We also provide an additional analysis to demonstrate that the larger sample size of non-chess players is not responsible for the results (see the Results section, as well as the OSF link above). All participants signed a written consent and the local ethics committee in Klagenfurt approved the study.

2.2. Stimuli, task, and design

Two insight problems were used. For the mutilated checkerboard problem (see Fig. 1A), participants were given 5 min to solve the problem and were instructed to indicate that they had found a solution and then to describe it. If they did not solve the problem within 5 min, the first hint was presented: "*It is impossible to cover all the squares. Can you see why*?". If a further 2 min elapsed without a solution, the second hint was given (at the 7-min mark): "*Consider the color of the removed squares*." Finally, the third hint was given after an additional 2 min (at the 9-min mark): "*Compare the number of black and white squares*." Participants were allowed to continue attempting to find a solution for another 2 min. Altogether, including the three hints, the maximum time allowed was 11 min (660 s). All solutions that clearly explained the impossibility of covering the remaining squares were considered correct. Here all the correct solutions involved the color property of the removed squares.

The second problem was the eight-coin problem (see Fig. 1B). Again, participants were given 5 min and, if no solution was found, they were provided with the following hint: "*Please consider all three spatial dimensions*." We could not come up with useful additional hints without repeating the information already given. Therefore, the eight-coin problem was considered not solved if two more minutes had passed after the hint was given (maximum time 7 min, or 420 s).

After each problem, participants who had solved the problem filled in an Aha! phenomenology questionnaire about their subjective perception of the solution process, addressing the following four dimensions of the Aha! experience (adopted from Danek & Wiley, 2017): Suddenness ("*The solution came to me…stepwise/suddenly*"), Surprise ("*The solution came to me…surprisingly/expectedly*"), Certainty ("*I felt about the solution…uncertain/certain*"), and Pleasure ("*At the moment of solution, my feelings were…pleasant/ neutral*"). Answers were given on a 5-point Likert scale (no numbers shown) with only the extremes being labeled (see Fig. 3).

The students were tested at the Department of Psychology in Klagenfurt and the Department of Psychology at LMU. Eight of the chess players were tested in the laboratory while the other chess players were tested at the chess club (a space away from the main playing hall was used).[2] All participants were tested individually. Participants were given instructions and a couple of warm-up unrelated problems. The order of presentation of the two problems was randomly assigned to each participant, but overall half of the participants in each of the two groups were given the eight-coin problem first, and the other half the checkerboard problem. The order of presentation was not a statistically significant factor in any of the reported analyses. For student non-chess players, the two insight problems were given as part of a larger set of problem-solving tasks (the two problems were always presented one after the other and at the beginning of the session). For chess players, other chess-specific problems were presented instead of the problems solving tasks that the non-chess players were solving (as with control participants, the problems presented here followed each other and were presented at the beginning of the session). After testing, participants were debriefed on the nature of the study and asked about any specific strategies used for solving the problems.

3. Results

Fig. 2 presents the cumulative solving rates across time for chess players and controls on both problems. Chess players were more successful than controls on the checkerboard problem in the first 5 min (e.g., 20 or 59% chess players vs. 11 or 13% controls). This difference persisted throughout the entire problem-solving period (33 or 97% of chess players solved the problem in the end, compared to 68 or 83% of controls). In contrast, there was little difference in the eight-coin problem (e.g., 6 or 18% of chess players vs. 10 or 13% of controls in the first 5 min, and 19 or 56% vs. 51 or 61% after 7 min).

Chess knowledge had an impact on performance in one problem (checkerboard), but not in the other (eight-coin). To see whether the interaction between chess knowledge and problem type was indeed significant, we ran a logistic regression with the two variables, group (chess and non-chess players) and problem (checkerboard and eight-coin), on the binary outcome of the success in the first 5 min (i.e., before hints). This analysis showed a significant main effect of the group factor ($b = 2.02$, $SE = 0.46$, $z = 4.39$, $p < .001$; odds-ratio = 7.58). The results indicated that chess players were more than seven times more likely to find the solution in the first 5 min compared to the control group. The main effect of the group factor was driven by the chess players' success on the checkerboard problem because the problem type factor did not reach significance ($b = -.22$, $SE = 0.44$, $z = .51$, $p = .61$; odds-ratio = 0.8) while the interaction between the group and problem type was significant ($b = -1.67$, $SE = 0.72$, $z = 2.32$, $p = 0.02$; odds-ratio = 0.19, Nagelkerke $R^2 = 0.18$). Not surprisingly, when we ran two separate logistic regressions on the two different problems, the chess players were significantly better at solving the checkerboard problem ($b = 2.03$, $SE = 0.46$, $z = 4.31$, $p < 0.001$;

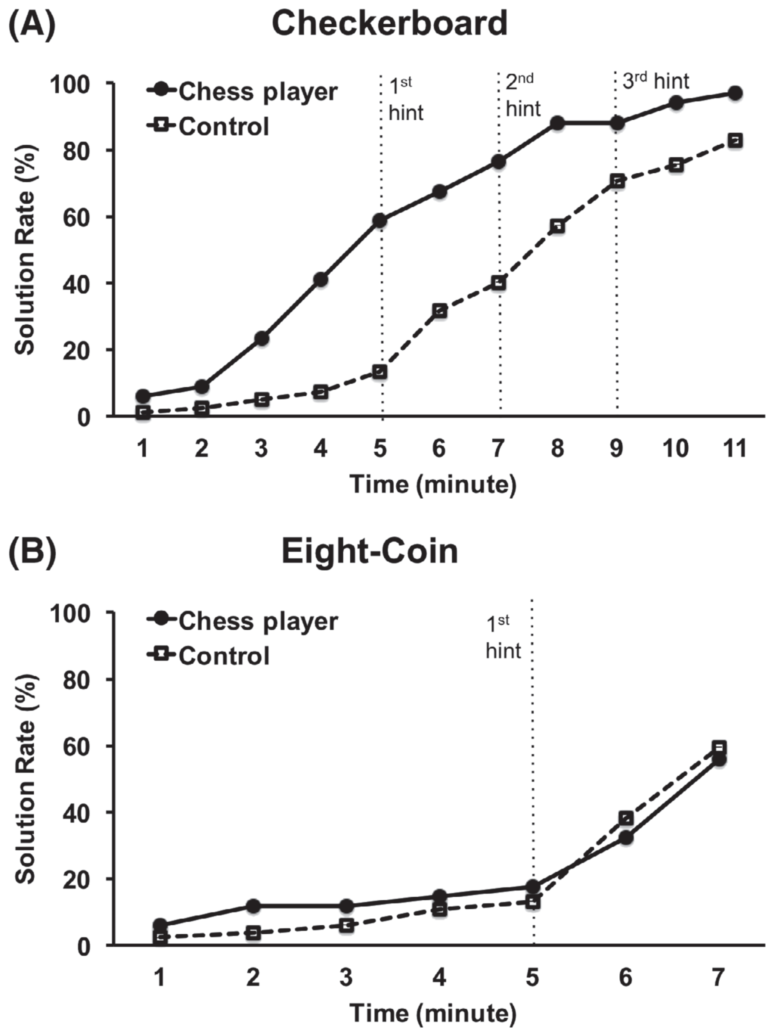

Fig. 2. Cumulative solution rates in percentage over time. (A) The mutilated checkerboard problem, where the first hint was given after 5 min (It is impossible to cover all the squares. Can you see why?), the second hint after 7 min (Consider the color of the removed squares), and the third and final hint after 9 min (Compare the number of black and white squares). Chess players (solid circles, n = 34) were much more successful than controls (open squares, n = 82 for checkerboard and n = 84 for the eight-coin problem). (B) The eight-coin problem, with only one hint, which was given after 5 min (Please consider all three spatial dimensions). There was no difference in success between the two groups of participants.

odds-ratio = 7.58, Nagelkerke R^2 = 0.24), whereas there was no significant difference on the eight-coin problem (b = 0.35, SE = 0.55, z = 0.64, p = 0.52; odds-ratio = 1.42, Nagelkerke R^2 = 0.01). Overall, we can say that chess experience in combination with problem type changed the probability of solving correctly in the first 5 min (see the Appendix for the analysis on the time and hint(s) needed to solve the problems).

The Aha! phenomenology questionnaire revealed differences between the two groups of participants with regard to their subjective solving experience. Fig. 3 presents the average answers for the participants who solved the checkerboard (left) and the eight-coin problem (right) at any time. Fig. 3A shows that chess players (dark gray bars) who solved the checkerboard problem felt that the solution came to them less suddenly than controls (dotted white bars) who solved the same problem (independent samples t test, $t(99) = 2.5$, $p = .01$; $d = .54$). In contrast, for the eight-coin problem, the subjective feeling of Suddenness did not differ between the two groups ($t(68) = 0.6$, $p = .52$; $d = .18$).

It is important to note that the Suddenness ratings on the two insight problems had opposite patterns in chess players and controls (Fig. 3A). Chess players reported significantly less sudden solutions in the checkerboard problem than in the eight-coin problem (dependent samples t test, $t(18) = 3.8$, $p < .001$; $d = .86$). In contrast, there were no differences among controls with regard to Suddenness in the checkerboard and the eight-coin problem ($t(44) = .3$, $p = .77$; $d = .04$). Solving the two insight problems felt similar for controls, but different for the chess players.

The dimension Surprise produced similar results (see Fig. 3B). Solutions in the checkerboard problem were significantly less surprising for chess players than for controls

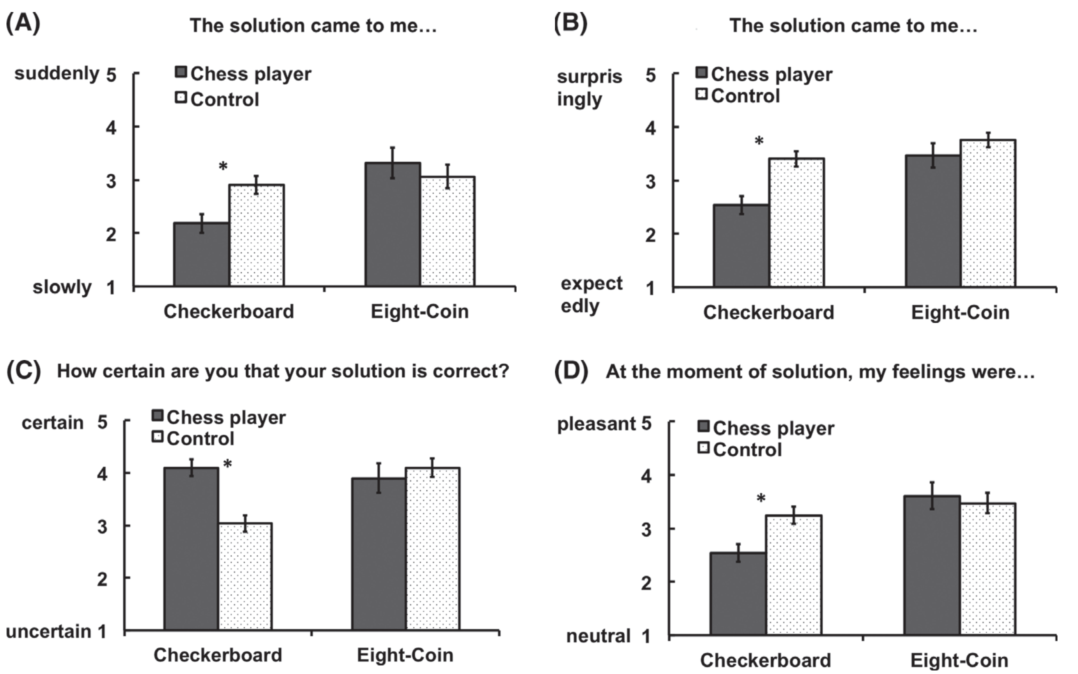

Fig. 3. Aha! Phenomenology Questionnaire. Average ratings (independent from the actual solution time point) as a function of problem type (checkerboard and eight-coin) and test group (chess players and non-chess players) on four dimensions of the Aha! experience: (A) Suddenness, (B) Surprise, (C) Certainty, and (D) Pleasure. *$p < .05$; Error bars represent standard error of the mean (*SEM*).

(independent samples t test, $t(99) = 3.8$, $p < .001$; $d = .80$), while there were hardly any differences between the two groups on the eight-coin problem ($t(68) = 1.1$, $p = .30$; $d = .28$). Further, chess players reported significantly less Surprise upon solving the checkerboard problem than for the eight-coin problem (dependent samples t test, $t(18) = 2.4$, $p = .026$; $d = .55$). Similar differences were found among the controls, but these did not quite reach the significance level ($t(44) = 1.9$, $p = .06$; $d = .28$).

Participants also differed in how certain they were about their solutions (Fig. 3C). Chess players were significantly more certain than controls regarding the checkerboard problem (independent samples t test, $t(99) = 4.2$, $p < .001$; $d = .89$), but there were no differences for the eight-coin problem ($t(68) = .4$, $p = .68$; $d = .11$). Controls were less certain about the checkerboard problem than for the eight-coin problem ($t(44) = 3.3$, $p = .002$; $d = .50$). Chess players displayed a reversed pattern where they were more certain about the solution on the checkerboard problem than for the eight-coin problem, but this difference did not reach significance (dependent samples t test, $t(18) = 1.6$, $p = .13$; $d = .36$).

Finally, controls experienced a stronger feeling of Pleasure (Fig. 3D) after solving the checkerboard problem than did chess players (independent samples t test, $t(99) = 2.7$, $p = .009$; $d = .56$). Again, the differences between the two groups were negligible on the eight-coin problem ($t(68) = .4$, $p = .69$; $d = .11$). There was no difference between the ratings of Pleasure among controls between the checkerboard and eight-coin problems (dependent samples t-test, $t(44) = 1$, $p = .31$; $d = .15$), but chess players were significantly more pleased when they solved the neutral eight-coin problem than the chess-related checkerboard problem ($t(18) = 2.4$, $p = .026$; $d = .56$).

In summary, chess players clearly differed from controls in their solution experience on the checkerboard problem, but not on the neutral eight-coin problem. As hypothesized, they experienced less Suddenness, less Surprise, and less Pleasure, but more Certainty in their solutions.

It is known that hints can change the insight process and subjective ratings (Bowden, 1997; Cushen & Wiley, 2012). To rule out the possibility that the insight questionnaire ratings are influenced by hints, we analyzed the ratings depending on when the solvers found the solution (and received the hints). There were no significant correlations for Suddenness (correlation between time and Suddenness for checkerboard: $r(101) = -.13$, $p = .19$ and for eight-coin: $r(70) = -.06$, $p = .63$) and for Surprise (checkerboard: $r(101) = .09$, $p = .35$; eight-coin problem: $r(70) = .17$, $p = .16$). Participants who needed less time for the solution felt more certain about it in the checkerboard problem (checkerboard: $r(101) = -.36$, $p < .001$) but not in the eight-coin problem ($r(70) = -.17$, $p = .15$). There was a marginally significant effect where the feeling of Pleasure was lower with more time needed for the solution on the checkerboard (checkerboard: $r(101) = -.20$, $p = .048$) but again, this was not observed in the eight-coin problem ($r(70) = -.03$, $p = .81$).

Although there was a trend that ratings for certainty and pleasure changed after the hints (and consequently the time needed to solve the problems), most of the rating differences between the two participant groups persisted over time. We can therefore be certain that the differences found between chess players and controls in their subjective perception of the solution process (Aha! phenomenology questionnaire) were not a direct product of the differing time needed to solve the problems.

As an additional test, we used bootstrapping to investigate whether the differences found between the two groups were influenced by the differing sample sizes. We equalized the number of participants per group (chess players or non-chess players) by randomly sampling non-chess players and calculating all individual statistical tests. This procedure was repeated 1,000 times, resampling participants from the non-chess player group in each run (see https://osf.io/g7fvq/?view_only=001931f1366644f3b6cbad80377c8100). The results confirm the differences between chess and non-chess players on the checkerboard problem, where the significant differences were constantly reproduced during bootstrapping. In contrast, the analyses on the eight-coin problem reached the significance level only a handful of times. Overall, the bootstrapping procedure fully confirmed the results reported throughout the study.

4. Discussion

In contrast to the traditional conception that prior knowledge is a hindrance for insight problem solving (Duncker, 1945), this study demonstrates that prior knowledge can also be helpful. We manipulated whether restructuring was required or not for successful problem solving by comparing two test groups with differing prior knowledge. Chess players, who were familiar with the chessboard's critical feature of parity (black–white characteristics of the squares), had higher solution rates than non-chess players for the checkerboard problem, for which the color of the squares is crucial to a correct solution. In the neutral eight-coin problem, both groups struggled to the same extent. This supports the assumption that chess players had set up a correct initial representation that allowed them to quickly solve the problem without restructuring. This result aligns with increased solution rates also for control participants once they had received hints which pointed to the crucial information in the checkerboard problem. This indicates that attending to the crucial element was necessary for restructuring and solving the problem. Hints also helped both groups to more readily solve the eight-coin problem (see Fig. 2).

The second main finding of our study is that there is a connection between the necessity of restructuring an initially incorrect problem representation and the subjective solving experience. When the crucial information was readily available, participants experienced less insight: Chess players had higher solution rates, but lower insight ratings than controls in the checkerboard problem. Apparently, chess players did not need to restructure their problem representation, because their prior knowledge allowed them to start out with a correct view of the problem, taking into account the critical feature of parity, which was overlooked by controls. This supports one of the key tenets of insight theories, namely that restructuring is a driving factor for Aha! experiences. Arriving at the solution presumably without any restructuring elicited weaker feelings of Suddenness, Surprise, and Pleasure for the chess player group than for controls. In contrast, controls who needed to restructure their problem representation reported higher Suddenness, Surprise, and Pleasure ratings after solving. We argue that this indicates that the availability of the crucial element altered the subjective phenomenology of the solution process (see also, Danek, 2018; Fleck & Weisberg, 2013). Participants who did not require restructuring because they could transfer their prior knowledge reported less Aha! (i.e., less Suddenness, Surprise, and Pleasure) than those who had to restructure. This finding is in accordance with a recent study by Danek et al. (2018) that offers the first empirical support for a close relationship between patterns of restructuring and the Aha! experience.

The specific pattern revealed in the questionnaire analysis highlights the importance of assessing the individual dimensions of the multifaceted Aha! experience separately, as proposed by Danek and Wiley (2017), instead of obtaining only one global "Aha!" measure. Key dimensions of Aha! such as Suddenness and Pleasure were lower for the chess group in comparison to controls, but their Certainty was higher. These opposing trends would have been obscured in a global Aha! rating. That the pattern for Surprise fitted the overall trend of lower ratings for chess players is in accordance with the findings of Webb et al. (2016), although this dimension was previously reported not to be a good predictor of Aha! experiences (Danek & Wiley, 2017). Certainty, which in prior studies was always rated in alignment with the other dimensions (Danek & Wiley, 2017), showed the opposite pattern here. At first glance, this may suggest that Certainty may not be an important dimension of the Aha! experience after all. However, we argue that this is a special case for experts, who are likely to be very confident in solutions to which they could apply their expertise (as in the case of the checkerboard problem), as opposed to non-experts. For the neutral problem, experts and non-experts showed the same level of Certainty.

There is a caveat that prevents us from claiming with certainty that the differing accessibility of the crucial problem feature is the main reason for the differences with regard to the Aha! experience. After all, we did not obtain any online information on solution strategies. Think-aloud protocols are a potentially useful method of gaining access to participants' thoughts, but we decided against using them because they tend to be disruptive on insight problems (Schooler, Ohlsson, & Brooks, 1993; although there are opposing findings by Fox, Ericsson, & Best, 2011). However, when we asked solvers during the debriefing procedure how they came up with the solution, they all inevitably mentioned the color of the removed squares. This may look obvious, but it excludes the possibility that an exhaustive strategy of covering the checkerboard with the dominoes was employed. Given that chess players as good as those in the present study (almost three standard deviations above average chess players) can automatically access the color of the square without actually looking at a chessboard (Ericsson & Staszewski, 1989), it is plausible to assume that they would notice more easily the color parity of the removed squares. Similarly, the large differences in success rate and solution time indicate that the crucial feature was easier to access for chess players than for non-chess players. It should be noted, however, that chess players also needed 2 min to solve the problem (and 40% of them needed at least one additional hint). This speaks against the possibility of that for chess players, the checkerboard problem does not really constitute a problem, because they might solve it by simple retrieval of the crucial solution element and its application in chess players.

Alternative explanations for the present differences between the two test groups need to be considered. One could argue that playing chess is a problem-solving exercise not unlike insight problem solving (Saariluoma, 1995), so that in general, a better performance for chess players on insight problems could be expected. But this idea can be refuted because not only is there no clear connection between insight problems and other more analytic problems (Fleck & Weisberg, 2013; Gilhooly & Fioratou, 2009; Gobet, Chassy, & Bilalic, 2011; Wegbreit, Suzuki, Grabowecky, Kounios, & Beeman, 2012), but chess players performed on the same level as controls on the neutral eight-coin problem. This speaks against the assumption that chess players' generally higher cognitive ability compared to non-chess players (Burgoyne et al., 2016) might have helped them. Similarly, the comparable performance on the neutral problem is evidence against other factors, such as age and gender, which were differently distributed in the chess players and controls.

One could still argue that the high difficulty of the eight-coin problem, and its low solution rates before hints, may have masked a possible overall advantage of chess players (floor effect). Other, less difficult insight problems may have revealed the difference. We do not believe this is the case as the hints, which improved the performance, resulted in near-equal performance of both test groups (Fig. 2B). Similarly, the fact that controls had similar solution rates on the eight-coin problem as on the checkerboard problem further indicates that the problems are of similar difficulty. Although using one single comparison problem may not completely remove doubts about possible pre-existing group differences, it goes a long way toward accounting for them. The advantage of chess players over non-chess players on the checkerboard problem is very large, not only in this study, but also compared to the typically extremely low solution rates for that problem (e.g., 6.5% as reported by Gick & McGarry, 1992). Should chess players indeed be superior solvers of insight problems, one would expect that some of this advantage would be retained on neutral insight problems. Instead, the differences are negligible. Given what we know about chess players and non-chess players (see the preceding paragraph), it would be rather surprising if the differences on the checkerboard problem were not the product of chess players' expert knowledge in the domain of chess.

So far, the manipulation of prior knowledge has mainly been achieved by hints (Kaplan & Simon, 1990; Öllinger et al., 2013), but the present approach of choosing test groups with differing prior knowledge presents another way of tackling the issue of restructuring. Future research could expand this approach to other traditional insight problems, such as matchstick arithmetic problems (Knoblich et al., 2001; Öllinger et al., 2006) where mathematical knowledge may significantly alter the solution process and experience. Mathematical knowledge may also influence the way people solve and experience Luchins's classical water jar paradigm (Bilalić & McLeod, 2014; Luchins, 1942), which is essentially an arithmetic problem.

The present finding of a diminished Aha! experience concurrent with higher solution rates demonstrates the importance of measuring the subjective feeling of Aha! on several dimensions to obtain a more differentiated picture of the individual problem-solving experience, as suggested by Danek and Wiley (2017). Ultimately, this study shows that only the investigation of both insight components, cognitive as well as affective, and also the interaction between them, is likely to lead to a full understanding of the complex insight phenomenon.

Acknowledgments

This work was supported by OeNB Jubiläumsfondsprojekt 16449 to MB and the DFG (German Research Foundation) grant DA 1683/1-1 to AD. We thank Iris van Rooij, Stellan Ohlsson, Michelle Arnold, Jim MacGregor, and Trina Kershaw for their comments on an earlier draft of the paper. The help and cooperation from the participants is greatly appreciated, as is Matthew Bladen's contribution to preparing the article for publication.

Notes

1. There were no significant differences between the two groups of students on any relevant variables (e.g., problem solving rates, solution times, or Aha! phenomenology questionnaire).
2. Testing in a chess club may have provided additional unwanted hints about the square color. We cannot exclude this possibility, but we note that most of the chess players who were tested in the laboratory solved the checkerboard problem within five minutes and needed a similar amount of time to those who were tested in the chess club.

References

Ash, I. K., Jee, B. D., & Wiley, J. (2012). Investigating insight as sudden learning. *The Journal of Problem Solving*, *4*(2), 1–27.

Bilalić, M. (2017). *The neuroscience of expertise*. Cambridge, UK: Cambridge University Press.

Bilalić, M., & McLeod, P. (2014). Why good thoughts block better ones. *Scientific American*, *310*(3), 74–79.

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.

Bowden, E. M., & Grunewald, K. (2018). Whose insight is it anyway?. In F. Vallee-Tourangeau (Ed.), *Insight* (pp. 28–50). London: Routledge.

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

Burgoyne, A. P., Sala, G., Gobet, F., Macnamara, B. N., Campitelli, G., & Hambrick, D. Z. (2016). The relationship between cognitive ability and chess skill: A comprehensive meta-analysis. *Intelligence*, *59*, 72–83.

Campitelli, G., & Speelman, C. (2013). Expertise paradigms for investigating the neural substrates of stable memories. *Frontiers in Human Neuroscience*, *7*, 740. https://doi.org/10.3389/fnhum.2013.00740.

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. (2018). Magic tricks, sudden restructuring and the Aha! experience: A new model of non-monotonic problem solving. In F. Vallée-Tourangeau (Ed.), *Insight: On the origins of new ideas* (pp. 51–78). London: Routledge.

Danek, A. H., Fraps, T., Von Mueller, A., Grothe, B., & Öllinger, M. (2014). Working wonders? Investigating insight with magic tricks. *Cognition*, *130*(2), 174–185.

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., Williams, J., & Wiley, J. (2018). Closing the gap: Connecting sudden representational change to the subjective Aha! Experience in insightful problem solving. *Psychological Research Psychologische Forschung*, 1–9.

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

Elo, A. E. (1978). *The rating of chessplayers, past and present* (Vol. 3). London: Batsford.

Ericsson, K. A., & Staszewski, J. J. (1989). Skilled memory and expertise: Mechanisms of exceptional performance. *Complex Information Processing: The Impact of Herbert A. Simon*, *2*, 235–267.

Fleck, J. I., & Weisberg, R. W. (2013). Insight versus analysis: Evidence for diverse methods in problem solving. *Journal of Cognitive Psychology*, *25*(4), 436–463.

Fox, M. C., Ericsson, K. A., & Best, R. (2011). Do procedures for verbal reporting of thinking have to be reactive? A meta-analysis and recommendations for best reporting methods. *Psychological Bulletin*, *137* (2), 316.

Gick, M. L., & McGarry, S. J. (1992). Learning from mistakes: Inducing analogous solution failures to a source problem produces later successes in analogical transfer. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(3), 623.

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.

Gilhooly, K. J., & Fioratou, E. (2009). Executive functions in insight versus non-insight problem solving: An individual differences approach. *Thinking & Reasoning*, *15*(4), 355–376. https://doi.org/10.1080/13546780903178615.

Gobet, F., Chassy, P., & Bilalic, M. (2011). *Foundations of cognitive psychology*. New York: McGraw-Hill.

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

Gruber, H. E. (1995). Insight and affect in the history of science. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 397–431). 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.

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., 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 & Cognition*, *29*(7), 1000–1009.

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

Ohlsson, S. (1984). Restructuring revisited: II. An information processing theory of restructuring and insight. *Scandinavian Journal of Psychology*, *25*(2), 117–129.

Ohlsson, S. (1992a). 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.

Ohlsson, S. (1992b). Information-processing explanations of insight and related phenomena. *Advances in the Psychology of Thinking*, *1*, 1–44.

Ohlsson, S. (2011). *Deep learning: How the mind overrides experience*. New York: Cambridge University Press.

Öllinger, M., Jones, G., Faber, A. H., & Knoblich, G. (2013). Cognitive mechanisms of insight: The role of heuristics and representational change in solving the eight-coin problem. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(3), 931–939. https://doi.org/10.1037/a0029194.

Öllinger, M, Jones, G., & Knoblich, G. (2006). Heuristics and representational change in two-move matchstick tasks. *Advances in Cognitive Psychology*, *2*(4), 239–253.

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

Saariluoma, P. (1995). *Chess players' thinking: A cognitive psychological approach*. London: Routledge.

Schooler, J. W., Ohlsson, S., & Brooks, K. (1993). Thoughts beyond words: When language overshadows insight. *Journal of Experimental Psychology: General*, *122*(2), 166.

Topolinski, S., & Reber, R. (2010). Gaining Insight Into the "Aha" Experience. *Current Directions in Psychological Science*, *19*(6), 402–405. https://doi.org/10.1177/0963721410388803.

Vaci, N., & Bilalić, M. (2017). Chess databases as a research vehicle in psychology: Modeling large data. *Behavior Research Methods*, *49*(4), 1227–1240.

Vaci, N., Gula, B., & Bilalić, M. (2014). Restricting range restricts conclusions. *Frontiers in Cognition*, *5*, 569. https://doi.org/10.3389/fpsyg.2014.00569.

Vaci, N., Gula, B., & Bilalić, M. (2015). Is age really cruel to experts? Compensatory effects of activity. *Psychology and Aging*, *30*(4), 740–754. https://doi.org/10.1037/pag0000056.

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.

Wegbreit, E., Suzuki, S., Grabowecky, M., Kounios, J., & Beeman, M. (2012). Visual attention modulates insight versus analytic solving of verbal problems. *The Journal of Problem Solving*, *4*(2), 94–115. https://doi.org/10.7771/1932-6246.1127.

Weisberg, R. W. (1995). Prolegomena to theories of insight in problem solving: A taxonomy of problems. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 157–196). Cambridge, MA: MIT Press.

Appendix

Time and Hint Analyses

Fig. 4 presents the time needed to solve the problem. If participants did not solve the problem within the first 5 min, we attributed to them the maximum time (300 ms) so as to keep them in the analysis. A two-way ANOVA with participants as a group factor (chess players vs. control) and problem type as a within factor (checkerboard vs. eight-coin) confirmed the differing patterns between chess players and controls (interaction group × problem: $F_{1,114} = 5.53$, $MSE = 14496$, $p = .024$, $\eta_p^2 = .05$). Participants needed less time for the checkerboard problem than for the eight-coin problem, but the effect was not significant (main effect problem: $F_{1,114} = 3.1$, $MSE = 8563$, $p = .08$, $\eta_p^2 = .03$). The chess players were generally faster on both problems (main effect group: $F_{1,114} = 11.5$, $MSE = 53014$, $p < .001$, $\eta_p^2 = .09$). A t-test for independent groups confirmed that chess players were significantly faster in solving the checkerboard problem ($t(114) = 4.3$, $p < .001$; $d = .88$) but that there was no significant difference between the two groups at the eight-coin problem ($t(114) = 1.3$, $p = .21$; $d = .25$).

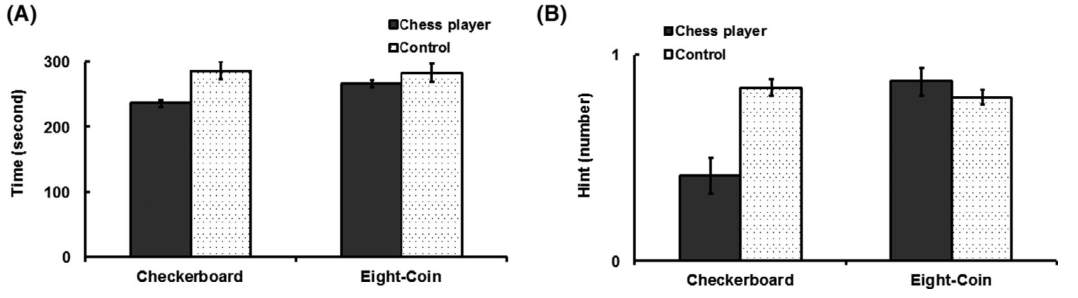

Fig. A1. Time and hints needed. (A) Chess players (dark bars) were much faster than controls (white bars) on the mutilated checkerboard problem (left) while there were no differences on the eight-coin problem (right) in the first 5 min of the problem-solving period. (B) The same pattern was observed in the number of participants that needed a hint during the first 7 min. Fewer chess players needed a hint than controls when solving the mutilated checkerboard problem (left) while there were no differences on the eight-coin problem (right). Error bars represent standard error of the mean (*SEM*).

Unsurprisingly, the same pattern of results was found when we checked the number of hints the participants received (Fig. 4). Considering the first 7 min (one hint maximal), fewer chess players needed a hint in general (main effect group: $F_{1,114} = 17.9$, $MSE = 3.02$, $p < .001$, $\eta_p^2 = .14$). However, this effect was driven by their better performance on the checkerboard (interaction group × problem: $F_{1,114} = 11.7$, $MSE = 1.54$, $p < .001$, $\eta_p^2 = .10$). Fewer participants in general needed a hint on the eight-coin problem (main effect problem: $F_{1,114} = 15.2$, $MSE = 1.99$, $p < .001$, $\eta_p^2 = .12$). The *t* test confirmed that fewer chess players than controls needed a hint when solving the checkerboard problem ($t(114) = 5.1$, $p < .001$; $d = 1.04$), but the difference between the two groups was not significant for the eight-coin problem ($t(114) = 1.02$, $p = .31$; $d = .20$).