Original Papers

Polish Psychological Bulletin
2021, vol. 52(4) 311–321
DOI: 10.24425/ppb.2021.139164

Angelika Olszewska[*]
Agata Sobkow[**]

Can observing a Necker cube (really) make you more insightful? The evidence from objective and subjective indicators of insight

Abstract: Changing a problem's representation is a crucial process when solving insight problems. Recently, Laukkonen and Tangen (2017) found that observing ambiguous figures such as a Necker Cube before solving problems can increase insight frequency. In our research, we extended their procedure by including measures of feelings of insight (e.g., confidence and pleasure). This approach allowed us to test the replicability of relationships between perceptual switching and insight frequency in terms of both accuracy of problem solutions and insight phenomenology. The research took the form of two studies using two different samples (N_A = 68 and N_B = 198) using online platforms. Our results consistently showed no effect of prior Necker cube perception on accuracy. However, we found a significant difference in self-reported insight (1 - non-aha! experience to 5 – a very strong aha! experience) in our Sample B study. The results suggest the possibility that viewing ambiguous figures may not have a triggering effect on insight problem-solving performance but that it may trigger stronger insight experiences when solving insight problems.

Keywords: *problem-solving, insight*

INTRODUCTION

Have you ever been in a situation where you were struggling with a problem, and then suddenly and spontaneously the answer just popped into your mind? If so, you probably felt an accompanying wave of pleasure and huge confidence in the correctness of your solution. Such experiences are termed moments of *insight* or *aha! moments*, and are defined as "occurring when sudden comprehension, realization or solution of a problem arises from a process whereby the reorganization of elements of one's mental representation of a stimulus, situation, or event yields a non-obvious or non-dominant interpretation" (Kounios & Beeman, 2014, p.71). Classic Gestalt psychologists such as Duncker (Newell, 1985) and Köhler (1947) believed mental problem representations to be just one possible way of interpreting problems, and that restructuring of problem representations can occur without cognitive control.

"Can observing a Necker cube make you more insightful?" The studies of Laukkonen and Tangen (2017)

Laukkonen and Tangen (2017) conducted two interesting studies examining the relationship between insight problem-solving and perception of the reversal of bistable figures. The first study examined correlations between perception of the reversal of bistable figures and the solving of both insight problems and analytical problems, and significant correlations were found between all of these ($0.38 < r < 0.48$). The second study was an experimental study in which observing a bistable figure was shown to have a positive effect on insight problem-solving performance.

The authors gave three possible reasons as to why bistable image reversal seems to be linked to insight. First, they suggested that the "aha!" experiences of perceptual change occurring in response to a bistable image and during insight problem-solving have similar phenomenological characteristics. Second, they proposed that bistable image reversals and some insightful experiences are preceded by changes in the representation or interpretation of problem elements or assumptions. Thus, Laukkonen and Tangen suggested that both problem-solving insights and bistable image reversals result from the same underlying processes described by Gestalt psychologists. Third, the authors speculated that the ability to change perspectives when viewing ambiguous images seems to be associated with the ability to solve creative problems (Laukkonen and Tangen, 2017).

The first explanation was examined in Laukkonen and Tangen's (2017) first study. Here, the authors used

[*] SWPS University of Social Sciences and Humanities, Wroclaw, Poland; ORCID iD - 0000-0002-7953-3904
[**] SWPS University of Social Sciences and Humanities, Wroclaw, Poland; ORCID iD - 0000-0002-5357-744X
Corresponding author: Angelika Olszewska, aolszewska3@swps.edu.pl

Metcalfe's (1986) warmth scale to measure suddenness when making a comparison between phenomenological experiences of suddenness when perceiving the reversal of ambiguous figures and experiencing insight while solving insight problems. Their results showed that both of these phenomena were experienced more suddenly than the appearance of solutions during analytical problem-solving. For both insight problem-solving and ambiguous figures, representational changes of problem structures / figures occurred in the absence of feelings that solutions / switches in perception were becoming increasingly imminent.

Laukkonen and Tangen's second explanation involves representational change theory (Ohlsson, 1984). This theory suggests that insight occurs immediately after a problem's representation is changed. The theory posits that a problem solver constructs an initial problem representation and that their search for a solution is bound by this representation. The situation does not determine the problem's representation: a problem can be represented in many different ways, depending on how the solver encodes the initial situation. Ohlsson further distinguished between external and internal searching for a solution: internal searches operate in the context of an initial state based on prior knowledge and mental imagery; external searches are based on trial and error methods. In both cases, changes in representation are crucial.

The third explanation seems vague. The authors speculate that the relationship between insight and bistable image reversals may involve more than a simple analogy, and they suggest that the ability to change perspective may be associated with the ability to solve creative problems. While Laukkonen and Tangen stressed that this explanation is in want of an explanatory mechanism and further evidence, they wrote the following: "Our more specific, but also more speculative hypothesis, is that experiencing conflict with the Necker cube would elicit conflict monitoring and cognitive control mechanisms, which would lead to better performance in the subsequent insight problem." (Laukkonen & Tangen, 2017, p. 203). With respect to conflict monitoring, they proposed that the reinterpretation processes used in both insight problem solution and ambiguous figure perception are partly accounted for by the Conflict Monitoring System described by Botvinick et al. (2001).

Laukkonen and Tangen based their third explanation on a previous study demonstrating that the role of conflict monitoring and cognitive control during insight problem-solving is to detect competing options (Kounios & Beeman, 2014). It is assumed that bistable image reversal reflects a switching between two competing good gestalts or two competing interpretations (Kornmeier & Bach, 2005; Long & Toppino, 2004). Moreover, activation of the anterior cortex prior to problem-solving is associated more with solving insight problems than analytic problems (Kounios et al., 2008).

To evaluate their three explanations, Laukkonen and Tangen examined the causal relationship between experiencing conflict during the viewing of ambiguous figures and performance when solving insight problems. To do this, they conducted an experimental study (Laukkonen & Tangen, 2017; Study 2) in which they used the Necker cube as an example of an ambiguous figure. The Necker cube is a two-dimensional drawing of a cube, which is interpreted as being three-dimensional.

The observer can experience two different views of the cube: it can be perceived as having either the lower-left square or the upper-right square as its front side. An ordinary Necker cube was presented in an experimental (conflict) condition, while in a control (non-conflict) condition, the authors used two possible interpretations of the cube and switched them externally. The non-conflict condition involved two images, one with an unambiguous cube pointing left and down, and one with an unambiguous cube pointing right and up. These images were presented one after the other, so that a switch occurred while the images were being viewed. In the conflict condition a single image where both interpretations were possible was presented. After observing the cube (conflict or non-conflict, depending on the condition), participants were asked to solve an insight problem. Results showed that insight problems were more likely to be solved when preceded by the original Necker cube.

It is worth noting that it is not entirely clear why Laukkonen and Tangen (2017) decided to examine the above causal relationship. The authors did not explain why engagement of the Conflict Monitoring System during Necker cube perception should have an impact on insight problem-solving. However, there are studies showing that more creative people report more frequent reversals between possible percepts when viewing ambiguous figures, although further studies are necessary.

Similar studies

The classic approaches to insightful problem-solving, on which the assumptions of Laukkonen and Tangen's (2017) studies were based, stress the role of spontaneous restructuring of problem representations. Duncker (1926) suggested that a problem situation consists of a gap which needs to be closed by structurally changing a "bad Gestalt" to a "good Gestalt". This change occurs rapidly, and he compared the process by which an immediate realization of a second perspective is reached to the flipping between interpretations of a Necker cube.

The Necker cube is not the only reversal figure used in previous research. For example, Wiseman et al. (2011) and Olteteanu et al. (2019) have examined relationships between insight and different ambiguous figures such as Jastrow's rabbit-duck and the girl-saxophonist illusion.

Wiseman et al. (2011) found a positive relationship between self-reported creativity and self-reported ambiguous figure reversal. These authors conducted two studies using different methodologies. In the first study, participants were asked if they would describe themselves as artistically creative and as creative problem solvers using 5-point Likert scales, whereas the second study used the Alternative Uses Task (Guilforfd, 1967) to measure creativity. A relationship between creativity and self-reported ease of ambiguous figure reversal was found in both studies.

In their study, Olteteanu et al. (2019) found significant correlations between ambiguous figure reversal and performance on both the Pattern Meanings Test (Wallach & Kogan, 1965) and the Alternative Uses Test (Guilford, 1967), but not between ambiguous figure reversal and insight problems. These results are interesting because it is possible that participants occasionally experience switches between categories in the Alternative Uses Test (Reiter-Palmon, Forthmann, & Barbot, 2019) and we might expect a similar phenomenon in insight problem-solving. However, Olteteanu et al. (2019) hypothesized three levels of reinterpretation (the first level being re-representation of features as different sets of objects or images, the second level being re-representation of objects and their properties as different objects, and the third level being re-representation of objects and scenes under different problem templates), thus the lack of a relationship between ambiguous figure reversal and insight problem performance. They used their ambiguous figures task as a measure of the first level, the Pattern Meanings Test and the Alternative Uses Test as measures of the second level, and insight problems as a measure of the third level – each of these tasks required the reinterpretation of an initial thought, idea, or problem structure.

The present study. A replication and extension of Laukkonen and Tangen's (2017) second study

Our study revisited the findings of Laukkonen and Tangen's (2017) Study 2, and extended their work by including measurement of insight phenomenology. We chose Laukkonen and Tangen's work for several reasons. First, we appreciated the idea of using two versions of the Necker cube as an experimental manipulation, thus providing the opportunity to manipulate bistable images while retaining a difference between conditions. Second, the inconsistency of results for relationships between ease of reversal of bistable images and insight problem-solving performance suggested the need for further studies. Third, the small effect size relating to the impact of Necker cube manipulation on insightfulness required replication. And finally, as Popper (2005) claimed, a true effect should be repeatable.

We aimed to clarify the relationship between perceptual reinterpretations of ambiguous figures and reinterpretations during problem-solving. We expected to find differences in accuracy and insight experiences between experimental conditions. Specifically, we expected to observe more puzzle solutions after participants' perceptions switched between the two possible Necker cube percepts when a cube was presented prior to puzzles. Additionally, we expected that participants in a conflict condition would report stronger insight experiences than those in a non-conflict condition. Our criterion for a successful replication was the observation of a significant difference in accuracy between the conflict and non-conflict conditions.

Our extension used measures of insight experience which were more comprehensive than those used in Laukkonen and Tangen study. These authors used the warmth scale created by Metcalfe (1986) and a self-report question about the experience of insight, where participants answered "yes," "no," or "other." But, we considered four separate dimensions from previous studies (Danek, Fraps, von Müller, Grothe, & Öllinger, 2014; Danek & Wiley, 2017; Webb, Little, & Cropper, 2016). These were suddenness, pleasure, confidence, and the aha! experience. Suddenness was included because insightful solutions are experienced very suddenly: as shown by Metcalfe (1986), prior to a solution arising, participants do not have a feeling that they are approaching a solution. Note that it is not only insight that is experienced suddenly, a recent study by Danek et al. (2018) also showing the sudden character of representational changes during problem-solving. A pleasure scale was used because positive affect accompanies feelings of insight (Danek et al., 2014; Gick & Lockhart, 1995). Also, subsequent to insight, a solution seems obvious, and problem solvers report having strong confidence in the correctness of solutions, so, similarly to Danek et al. (2014), we used a confidence scale. Finally, we also obtained participants' self-reports of insight. Although Danek et al. (2014) observed some instances of false insight in their study, false insight (feelings of insight when a solution is wrong) is rare, and this rarity suggests that the phenomenon of insight may be more connected with changing a problem's representation and finding a new path to a solution than with actually obtaining a correct answer, and we thought that self-reports would cast further light on this issue.

Summarizing, we conducted a conceptual replication of Laukkonen and Tangen's (2017) second study, measuring insight experiences as an extension. We hypothesized that we would observe more correct solutions of verbal puzzles in a conflict condition (where participants switched between two possible Necker cube percepts) than in a non-conflict condition (where participants attended to only one of the two possible Necker cube percepts). Additionally, we expected that participants in the conflict condition would report more robust insight experiences than those in the non-conflict condition.

METHOD

Two studies were conducted using two independent samples (A and B) on two different online platforms. However, the methods and procedures were largely the same in both studies. The only difference was that for Sample B an attention check was used as described in the Procedure and Materials section below.

Participants

Sample A

The Sample A study consisted of 68 undergraduate psychology students participating in exchange for course credits (M_{age} = 26.55 years, SD_{age} = 7.34 years;

54 females). Data for an additional 18 participants were excluded from analysis because these participants did not complete the whole experimental procedure; these data were excluded before coding of correct answers to the verbal puzzles was performed. Participants were randomly assigned to one of two between-subjects conditions, these conditions being named the *conflict condition* (n = 36) and the *non-conflict condition* (n = 32) in accordance with Laukkonen and Tangen (2017).

Sample B

Participants for the Sample B study were drawn from the Polish population (N = 198; 52 females; M_{age} = 24.10 years, SD_{age} = 6.63 years) and were recruited for this online study via the Prolific commercial panel. Participants took part in the study in exchange for a financial reward (GBP = 4.00). Of the participants, 34% had graduated, 56% had finished their upper secondary education, 6% had completed the first stage of education (the equivalent of secondary school), and 3.5 % had a vocational education. Participants were randomly assigned to one of two between-subjects conditions named as for Sample A: a *conflict condition* (n = 98) and a *non-conflict condition* (n = 100). The above description does not include three participants who were excluded before the coding of answers because they had duplicate records and had probably taken part in the study twice, six participants who failed an attention check, and 14 participants who did not finish the whole experimental procedure.

Power analysis

Based on Laukkonen and Tangen's (2017) paper, we calculated a Cohen's d effect size estimate of 0.26. This small effect size required either a very large sample size or the application of another solution if we were to reliably replicate Laukkonen and Tangen's results. We opted for the latter: based on Simonsohn's (2015) suggestions for designing replication studies, we used the N x 2.5 rule, i.e., we sought to obtain a sample size 2.5 times larger than the original study would have required to attain 80% power in detecting an effect 33% as large as that actually obtained in the study. This led to a sample size estimate of N = 200. However, a priori power analysis was only conducted for the study involving Sample B. Post-hoc power analysis was conducted for the Sample A study using G*Power 3.1.9.4 (Faul, Erdfelder, Lang, & Buchner, 2007), which, for α = .05 – one-tailed, an effect size of d = .26, and subsample sizes of n_1 = 36 and n_2 =33, showed that the study was underpowered in that power was only 28%.

Procedure and materials

Except where specified, the procedure and materials were the same for each study. Participants in both conditions received training using Jastrow's classic duck–rabbit reversal figure before performing the main task. The training comprised twenty trials in which participants were instructed to press the space bar on a keyboard each time their perception switched from duck to rabbit or from rabbit to duck. This training was provided so that participants fully understood the instruction to press the space bar after ambiguous figure reversal. The rabbit–duck image was chosen because its reversal is easy to explain.

Participants were randomly assigned to one of the two between-subjects experimental conditions. The classic Necker cube was presented in the conflict condition, while in the non-conflict condition only one of the possible percepts of the Necker cube was used. In each of 10 trials (see Figure 2), participants were instructed to observe the cube (in the presence of conflict or non-conflict) for 90 seconds on a computer monitor and to press a space bar on a computer keyboard whenever they experienced a reversal of the cube's position. Next, they were given 60 seconds in which to solve a verbal puzzle. The time limits were identical to the times used by Laukkonen (2017) in study 2. The same set of ten insight puzzles was used in random order in both conditions. At the end of each trial, participants were asked about their experience when solving the puzzle. Finally, participants responded to demographic questions about their gender and age.

MEASURES

The Necker cube

The nature of the Necker cube viewed was the independent variable in our studies. In the conflict condition, we used the classic Necker cube as used in the Laukkonen and Tangen (2017) study. This is a well-known ambiguous figure where there are two perceptual possibilities. A picture of an unambiguous cube pointing left and down was attached to the left arrow key of a computer keyboard, and a picture of an unambiguous cube pointing right and up was attached to the right arrow key. In the non-conflict condition, only a picture of the cube pointing left and down was used (the cubes used are presented in Figure 1).

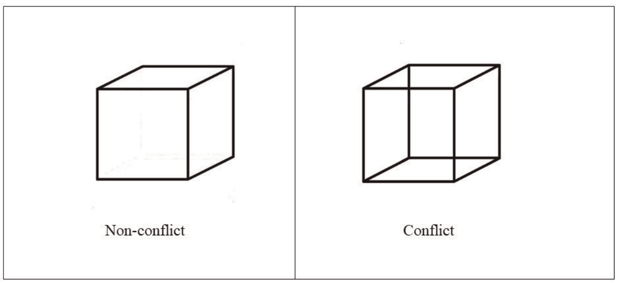

Figure 1. The Necker cube figures used in the non-conflict and conflict conditions

Verbal puzzles

Participants had one minute to solve each puzzle. Because of the possible influence of culture and language on solving the puzzles, insight problems were selected in two pilot studies. The verbal puzzles from Laukkonen and Tangen's (2017) study were translated into Polish and merged with puzzles previously used by Chuderski and Jastrzębski (2018a), and Karwowski (2014). The final set of verbal puzzles contained seven items from Laukkonen

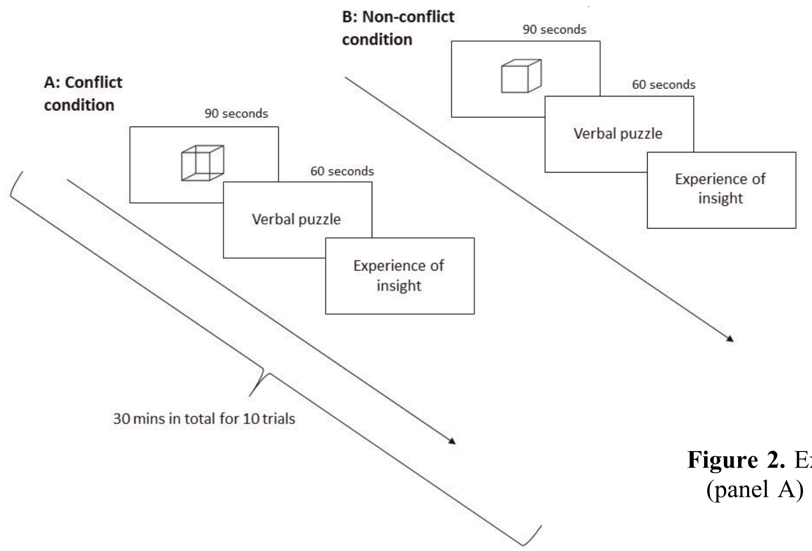

Figure 2. Example trials for the conflict condition (panel A) and non-conflict condition (panel B)

and Tangen's study, two from Karwowski's study, and one from Chuderski and Jastrzębski's study.

Each participant solved ten verbal puzzles. These puzzles took the form of open-ended questions, participants having to type out answers. If a response was not given within 60 seconds the next trial was presented automatically. Subsequent to data collection, participants' answers were coded by three independent competent judges (0 – a wrong answer/no answer, 1 – a correct answer). The judges were blind to each participant's experimental condition: they received a datasheet including only a participant's ID, and a participant's answers. The criterion for coding was that at least two judges had to agree.

The sum of a participant's correct answers served as an index of accuracy and was used as the dependent variable in our replication. The set of insight puzzles used had acceptable reliability for both samples (McDonald's $\Omega_A = .63$ and McDonald's $\Omega_B = .61$)[1], and can be found in the Supplementary materials.

Insight phenomenology

Participants' insight experiences concerning their answers were measured by four questions inspired by the multidimensional approach to qualitative insight phenomena (Danek et al., 2014; Webb, Little, & Cropper, 2017). After responding to each puzzle, participants answered these questions on a five-point Likert scale. The four questions covered confidence (unsure to very sure), pleasure (very unpleasant to very pleasant), an aha! experience (no-aha! experience to a very strong aha! experience), and suddenness (a solution came gradually to a solution came all of a sudden). Each dimension was analyzed separately, each variable being considered as a different dependent variable.

Attention check

An attention check was only conducted in the Sample B study, this being done after the main tasks (shown in Figure 2) but before the measurement of other metrics. A Necker cube consistent with a participant's condition (conflict or non-conflict) was presented for 10 seconds and participants were subsequently asked to type in the word "cube" in an open-ended response window. We only analyzed data for participants who were paying attention.

RESULTS

Analyses for the Sample A and Sample B studies were conducted separately but were the same for each study. Similarly to Laukkonen and Tangen (2017), we aggregated observations across the 10 trials for each participant by calculating a mean value. Table 1 presents a comparison of descriptive statistics from our studies and those of Laukkonen and Tangen. For both Sample A and Sample B, no difference was observed between the results of our studies and those of Laukkonen and Tangen in terms of the median number of switches in the conflict condition. However, in the non-conflict condition, there was far less switching in our studies than in Laukkonen and Tangen's study. This is probably because there was only one possible percept of the Necker cube in our control group. Using (non-parametric) Mann-Whitney tests because of some negative skewness in switching distributions, we observed significantly more frequent switching in the conflict condition than in the non-conflict condition in both samples ($U_A = 860$, $p < .001$, and $U_B = 7508.5$, $p < .001$), and can conclude that our manipulation was effective in both studies.

Descriptive statistics for the insight experience measures are presented in Tables 2A and 2B. Average values for all the puzzles were compared between the two conditions within both samples. No differences were found between the two conditions with respect to accuracy, $t_A(66) = 0.89$, $p = .38$, $d = 0.214$, and $t_B(196) = 0.11$, $p = .91$, $d = 0.017$. However, we found a significant difference in self-reported insight (1 – no-aha! experience to 5 – a very strong aha! experience) in Sample B, $t_B(196) = 2.57$, $p < .001$, $d = 0.365$. Figure 3 presents a comparison of effects. We observed no significant differences for the other insight phenomenology dimensions.

[1] Cronbach's $\alpha_A = 0.61$ and Cronbach's $\alpha_B = 0.61$

Table 1 Comparison of the present results with those of Laukkonen and Tangen (2017)

	Sample A		Sample B		Laukkonen & Tangen	
	Conflict	**Non-conflict**	**Conflict**	**Non-conflict**	**Conflict**	**Non-conflict**
Switching	*Mdn* = 24.90 *SD* = 22.03	*Mdn* = 4.50 *SD* = 14.85	*Mdn* = 20.25 *SD* = 21.87	*Mdn* = 3.75 *SD* = 23.11	*Mdn* = 8.20 *SD* = 26.40	*Mdn* = 26.10 *SD* = 4
Accuracy	*M* = 4.33 *SD* = 2.61	*M* = 3.82 *SD* = 2.21	*M* = 4.89 *SD* = 2.10	*M* = 4.86 *SD* = 2.40	*M* = 4.24 *SD* = 1.87	*M* = 3.76 *SD* = 1.82

Table 2A Descriptive statistics and t-test results for Sample A measures of insight phenomenology

	Conflict		Non-conflict		*t*(66)	*p*	*d*
	M	***SD***	***M***	***SD***			
Confidence	2.79	0.97	2.66	0.89	0.59	.56	0.142
Pleasure	3.04	0.60	3.03	0.79	0.07	.95	0.016
Suddenness	3.02	0.83	3.10	0.74	-0.43	.67	-0.104
Aha! experience	2.59	0.86	2.43	0.94	0.72	.47	0.176

Table 2B Descriptive statistics and t-test results for Sample B measures of insight phenomenology

	Conflict		Non-conflict		*t*(196)	*p*	*d*
	M	***SD***	***M***	***SD***			
Confidence	3.10	0.63	3.10	0.67	0.01	.99	0.002
Pleasure	3.48	0.58	3.37	0.68	1.28	.26	0.160
Suddenness	3.33	0.69	3.23	0.59	1.09	.28	0.155
Aha! experience	2.79	0.71	2.54	0.66	2.57	.01	0.365

Figure 3. A comparison of differences between conditions showing means and standard deviations

We also examined relationships between all the measures for both samples (see Table 3 and Table 4). There were no significant Pearson's correlations between perceptual switching and either accuracy or insight experience. However, we observed positive relationships among dimensions of insight phenomenology. Interestingly we observed significantly weaker correlations among aha! experience and confidence, pleasure and suddenness in sample B compared to sample A (see supplementary materials, Table S4).

DISCUSSION

The present study aimed to replicate and extend the findings of Laukkonen and Tangen's (2017) research. In their study, Laukkonen and Tangen found that looking at a Necker cube had an impact on accuracy in solving insight problems. But both of our studies found no support for this effect. We found no difference between the conflict and non-conflict conditions in the number of correct answers participants provided to verbal puzzles. We also found no relationship between number of switches and average number of correct answers in the conflict condition. However, in an extension of Laukkonen and Tangen's study, we found a significant difference in self-reported insight for Sample B but not for Sample A. For Sample B, the aha! experience was stronger in the conflict condition than in the non-conflict condition. Laukkonen and Tangen (2017) found similar results in their study. After each puzzle, Laukkonen and Tangen asked participants if they experienced insight using 3-point scale (1 – no, 2 – other, 3 – yes), and they found that participants in the conflict condition experienced insight more frequently than those in the non-conflict condition. This effect is consistent with the result we observed for our extended self-reported insight scale (we used a 5-point Likert scale from 1 – no-aha! experience to 5 – a very strong aha! experience).

It was surprising that we found no differences between the conflict and non-conflict conditions with respect to number of correct answers and the suddenness, pleasure, and confidence dimensions of insight phenomenology. There are at least two possible explanations for these results. First, it is possible that there was a significant difference for the aha! experience but not accuracy because the experiencing of insight is related to the type of problem-solving involved rather than solving problems per se, and classical insight problems can be solved both in a manner involving insight and analytically (Danek, Wiley, & Öllinger, 2016; Webb et al., 2016). Second, previous studies have observed differences in self-reported insight between different types of problems in the absence of differences in other phenomenological dimensions. Thus, Webb et al. (2016) observed differences across insight and non-insight tasks using an aha! experience scale but did not observe differences on other insight dimensions. Similarly, Chuderski et. al. (2020), compared insight problems with Raven's matrices problems and observed differences with respect to self-reported intuition but not pleasure, suddenness, and certainty. Finally, Danek et al. (2020) found that a stronger aha! experience is related to sudden changes in problem representation. Moreover, we would like to highlight the limitations of

Table 3 Pearson's r correlations for Sample A measures

	1	2	3	4	5
1. Confidence					
2. Pleasure	.69**				
3. Suddenness	.60**	.56**			
4. Aha! experience	.76**	.65**	.66**		
5. Accuracy	.59**	.27*	.30*	.48**	
6. Switching	.13	-.05	.09	.09	.23[a]

Note: ** $p < .001$, * $p < .05$, [a]p <.1

Table 4 Pearson's r correlations for Sample B measures

	1	2	3	4	5
1. Confidence					
2. Pleasure	.53**				
3. Suddenness	.58**	.38**			
4. Aha! experience	.28**	.37**	.31**		
5. Accuracy	.51**	.36**	.23**	.02	
6. Switching	-.08	-.05	-.12	.08	-.02

Note: ** $p < .001$, * $p < .05$

aha! experience scale. As we observed in sample B this scale has slightly different properties than in sample A. We observed it in correlation comparison between samples (see supplementary materials, Table S4), the aha! experience scale is the only scale with different correlation coefficients in sample B. We can find a few possible explanations for this effect. First, the different motivations (credit points in sample A vs monetary reward in sample B) may affect the awareness of the Likert scale. The participants in sample A were less reflective about the phenomenology and more consistent with the other. Second, they were psychology students who may have a different understanding what aha! experience is. Finally, when we conducted the sensitivity calculation for this effect using pwr-package (Champely, 2020), assuming 80% power and effect size $d = 0.36$, 122 (per group) where required for observing this effect. Thus, the sample A may be underpowered.

It is worth noting that, despite a significant difference in number of correct answers between conditions, the mean numbers of correct answers for the two conditions observed for Sample A were more similar to those in Laukkonen and Tangen's study than were those for Sample B. For Sample A, an average of 38% correct answers occurred in the non-conflict condition compared to 37.6% in Laukkonen and Tangen's study, and in the conflict condition an average of 43% of correct answers occurred compared to the 42.4% observed by Laukkonen and Tangen. However, for Sample B, we observed more correct answers for both conditions (49% for the conflict condition, and 48.6% for the non-conflict condition). This inconsistency may be explained by the nature of participants in Sample B (non-students sample rather than psychology undergraduate students) or by a difference in participants' motivations: Sample B participants obtained a financial reward. It might be suspected that one reason for the higher proportion of correct answers in our study is that three puzzles were added to the seven (translated) original puzzles used by Laukkonen and Tangen, and that these three extra puzzles were easier. However, a test of whether this was the case showed that the average proportion of correct answers for the new puzzles did not differ from the average proportion of correct answers for the translated Laukkonen and Tangen puzzles in each of our samples.

Our replication has limitations in that we did not replicate Laukkonen and Tangen's study (2017) completely. First, they used a within-subjects design whereas we used a between-subjects design. Second, we used only seven of Laukkonen and Tangen's puzzles, and these were translated into a different language. Finally, we used a slightly different Necker cube manipulation in that we used only one of the two possible percepts in the non-conflict condition.

Although the above differences may have affected our results, it seems reasonable to argue that the manipulation used in our two studies was effective. We observed a significant difference for both samples. The number of switches in our conflict condition was comparable with that reported by Laukkonen and Tangen (2017). However, the number of switches in our non-conflict condition (Mdn_A = 4.5; Mdn_B = 3.75) was far lower than in Laukkonen and Tangen's study (Mdn = 26.1). Surprisingly, participants effectively experienced the switching of a non-ambiguous figure roughly once every 30 seconds, suggesting the possibility that participants may have experienced some perceptual changes. It would be useful to explore the phenomenological and qualitative experience of such perceptual changes. In ending, this brief discussion of switching, it is worth highlight the fact that the relationship between the number of switches participants experienced and accuracy was non-significant in both the original study and our replication.

CONCLUSIONS

To conclude, the present study did not replicate Laukkonen and Tangen's (2017) observation that prior Necker cube perception has an effect on accuracy during insight problem-solving. Furthermore, we observed no relationships between different dimensions of insight experience and switching when observing a Necker cube. However, we would argue that the manipulation we used was effective (as shown by significant differences in the number of switches across our conflict and non-conflict conditions), and that our set of puzzles showed good reliability and validity (in the form of significant correlations with insight experience measures). Our results do not allow us to conclude that observing a Necker cube triggered problem-representation changes. A comparison of our findings with the inconsistent results of previous studies (Olteteanu et al., 2019; Wiseman et al., 2011) leads us to suggest the possibility that viewing ambiguous figures may not have a triggering effect on insight problem-solving performance but that it may trigger stronger insight experiences when solving insight problems.

ACKNOWLEDGMENTS

This research was carried out with the financial support of SWPS University. We would like to thank the lead author of the original study, Ruben Laukkonen, for providing all the information required to conduct our replication, e.g., sharing correct answers for the puzzles he used, and providing effect size and confidence interval information relating to the main effect observed in his work.

REFERENCES

Botvinick, M. M., S, B. T., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict Monitoring and Cognitive Control. *Psychological Review*, *103*(3), 624–652.

Champely, S., (2020). pwr: Basic Functions for Power Analysis. R package version 1.3-0. https://CRAN.R-project.org/package=pwr

Chuderski, A., Jastrzębski, J., Kroczek, B., Kucwaj, H., & Ociepka, M. (2020). Metacognitive experience on Raven's matrices versus insight problems. *Metacognition and Learning*, 31–44. https://doi.org/10.1007/s11409-020-09236-6

Danek, A. H. (2018). Magic tricks, sudden restructuring, and the Aha! experience: A new model of nonmonotonic problem solving. In F. Valee-Tourangeau (Ed.), Insight (pp. 51-78). Routledge.

Danek, A. H., Fraps, T., von Müller, A., Grothe, B., & Öllinger, M. (2014). It's a kind of magic-what self-reports can reveal about the phenomenology of insight problem solving. *Frontiers in Psychology*, *5*, 1–11. https://doi.org/10.3389/fpsyg.2014.01408

Danek, A. H., & Wiley, J. (2017). What about false insights? Deconstructing the aha! experience along its multiple dimensions for correct and incorrect solutions separately. *Frontiers in Psychology*, *7*, 1–14. https://doi.org/10.3389/fpsyg.2016.02077

Danek, A. H., Wiley, J., & Öllinger, M. (2016). Solving classical insight problems without aha! experience: 9 Dot, 8 Coin, and matchstick arithmetic problems. *Journal of Problem Solving*, *9*(1), 47–57. https://doi.org/10.7771/1932-6246.1183

Danek, A. H., Williams, J., & Wiley, J. (2020). Closing the gap: connecting sudden representational change to the subjective Aha! experience in insightful problem solving. *Psychological Research*, *0*(0), 1–9. https://doi.org/10.1007/s00426-018-0977-8

Duncker, K. (1926). A qualitative (experimental and theoretical) study of productive thinking (solving of comprehensible problems). *Pedagogical Seminary and Journal of Genetic Psychology*, *33*, 642–708. https://doi.org/10.1080/08856559.1926.10533052

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

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

Guilforfd, J. P. (1967). The nature of human inteligence. McGraw-Hill.

Köhler, W. (1947). *Gestalt Psychology. The definitive statement of the Gestalt Theory*. New York: Liveright.

Kornmeier, J., & Bach, M. (2005). The Necker cube - An ambiguous figure disambiguated in early visual processing. *Vision Research*, *45*(8), 955–960. https://doi.org/10.1016/j.visres.2004.10.006

Kounios, J, & Beeman, M. (2014). The cognitive neuroscience of insight. *Annual Review of Psychology*, *65*. https://doi.org/10.1146/annurev-psych-010213-115134

Kounios, John, Fleck, J. I., Green, D. L., Payne, L., Stevenson, J. L., Bowden, E. M., & Jung-Beeman, M. (2008). The origins of insight in resting-state brain activity. *Neuropsychologia*, *46*(1), 281–291. https://doi.org/10.1016/j.neuropsychologia.2007.07.013

Laukkonen, R. E., & Tangen, J. M. (2017). Can observing a Necker cube make you more insightful? *Consciousness and Cognition*, *48*, 198–211. https://doi.org/10.1016/j.concog.2016.11.011

Long, G. M., & Toppino, T. C. (2004). Enduring interest in perceptual ambiguity: Alternating views of reversible figures. *Psychological Bulletin*, *130*(5), 748–768. https://doi.org/10.1037/0033-2909.130.5.748

Metcalfe, J. (1986). Feeling of knowing in memory and problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*(2), 288–294. https://doi.org/10.1037/0278-7393.12.2.288

Newell, A. (1985). Duncker on thinking: An inquiry into progress in cognition. *A Century of Psychology as Science.*, 392–419. https://doi.org/10.1037/10117-032

Ohlsson, S. (1984). Restructuring revisited: II. An information processing theory of restructuring and insight. *Scandinavian Journal of Psychology*, *25*(2), 117–129. https://doi.org/10.1111/j.1467-9450.1984.tb01005.x

Olteteanu, A. M., Schöttner, M., & Bahety, A. (2019). Towards a multi-level exploration of human and computational re-representation in unified cognitive frameworks. *Frontiers in Psychology*, *10*(APR), 0–34. https://doi.org/10.3389/fpsyg.2019.00940

Popper, K. (2005). The Logic of Scientific Discovery. In *Hume's Problem*. https://doi.org/10.1093/0198250371.003.0009

Reiter-Palmon, R., Forthmann, B., & Barbot, B. (2019). Scoring divergent thinking tests: A review and systematic framework. *Psychology of Aesthetics, Creativity, and the Arts*, *13*(2), 144–152. https://doi.org/https://doi.org/10.1037/aca0000227

Wallach, M. A., & Kogan, N. (1965). A new look at the creativity-intelligence distinction 1. *Journal of Personality*, *33*(3), 348–369. https://doi.org/10.1111/j.1467-6494.1965.tb01391.x

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*, 1–13. https://doi.org/10.3389/fpsyg.2016.01424

Webb, M. E., Little, D. R., & Cropper, S. J. (2017). Once more with feeling: Normative data for the aha experience in insight and noninsight problems. *Behavior Research Methods*, *50*(5), 2035–2056. https://doi.org/10.3758/s13428-017-0972-9

Wiseman, R., Watt, C., Gilhooly, K., & Georgiou, G. (2011). Creativity and ease of ambiguous figural reversal. *British Journal of Psychology*, *102*(3), 615–622. https://doi.org/10.1111/j.2044-8295.2011.02031.x

VERBAL PUZZLES USED IN STUDY[2]

1. Bezrobotna kobieta, która nie miała przy sobie prawa jazdy zignorowała nakaz stopu na skrzyżowaniu, następnie nie zważając na znak informujący o ulicy jednokierunkowej ruszyła pod prąd. To wszystko obserwował policjant na służbie, który nic z tym nie zrobił. Dlaczego?
2. Morderca został skazany na karę śmierci. Musi wybrać jedno z trzech pomieszczeń. Pierwszy pokój jest wypełniony płomieniami ognia, drugi jest pełen zabójców z naładowaną bronią, z kolei w trzecim są lwy, które nie jadły od 3 lat. Który pokój jest dla niego najbezpieczniejszy?
3. Więzień planował ucieczkę z więzienia. W swojej celi znalazł linę, która wystarczała do połowy wysokości budynku. Podzielił linę na pół, połączył dwie części i uciekł. Jak to możliwe?
4. Jan mył okna wielopiętrowego biurowca, kiedy nagle się poślizgnął i spadł z 18-metrowej drabiny na betonowy chodnik. Na szczęście nic mu się nie stało. Jak to możliwe?
5. Dwóch mężczyzn rozegrało pięć pełnych partii w warcaby. Każdy z nich wygrał równą liczbę partii, przy czym żadna z gier nie zakończyła się remisem ani rezygnacją któregoś z graczy. Jak to możliwe?
6. Jeśli w szufladzie masz wymieszane czarne i brązowe skarpetki w proporcji 4 do 5 to ile musisz wyjąć skarpetek, aby mieć pewność, że masz parę w tym samym kolorze?
7. Emerytowany Profesor prowadził swój stary samochód, kiedy nagle auto samo zmieniło bieg. Profesor nie zwrócił na to uwagi i kontynuował jazdę, jakby nic się nie stało. Dlaczego się nie zaniepokoił?
8. Brat, jego siostra oraz mąż ze swoją żoną znaleźli 4 monety. Każde z nich wzięło po jednej i jedna im została jedna moneta. Jak to możliwe?"
9. Co to jest: antyczny wynalazek wciąż używany w większości krajów świata, umożliwiający patrzenie przez ściany?
10. Kierowca Fiata ma brata, lecz brat kierowcy Fiata nie ma brata. Kim kierowca Fiata jest dla brata?

ENGLISH VERSION OF VERBAL PUZZLES[3]

1. An unemployed woman did not have her driver's license with her. She failed to stop at a railroad crossing, then ignored a one-way traffic sign and traveled three blocks in the wrong direction down the one-way street. All this was observed by a policeman, who was on duty, yet he made no effort to arrest the woman. Why?
2. A murderer is condemned to death. He has to choose among three rooms. The first is full of raging fires, the second is full of assassins with loaded guns, and the third is full of lions that haven't eaten in 3 years. Which room is safest for him?
3. A prisoner was attempting to escape from a tower. He found in his cell a rope that was half long enough to permit him to reach the ground safely. He divided the rope in half, tied the two parts together, and escaped. How could he have done this?
4. Mr. Hardy was washing windows on a high-rise office building when he slipped and fell off a sixty foot ladder onto the concrete sidewalk below. Incredibly, he did not injure himself in any way. How is this possible?
5. Two men played five full games of checkers and each won an even number of games, with no ties, draws, or forfeits. How is that possible?
6. If you have black socks and brown socks in a drawer, mixed in a ratio of 4 to 5, how many socks will you have to take out to make sure that you have a pair of the same colour?
7. Professor Bumble, who is getting on in years, was driving along in his old car when suddenly it shifted gears by itself. He paid no attention and kept on driving. Why wasn't he concerned?
8. A brother with his sister and a husband with his wife found 4 coins. Each of them took 1 coin, but there was still 1 coin left. How is that possible?
9. What is it: an antique invention still used in most countries of the world which makes it possible to look through walls?
10. A car driver has a brother, but the brother of the driver does not have a brother. Who is the driver for the brother?

[2] Puzzles with numbers 1-7 were used in oryginal Laukkonen and Tangen's (2017) study. Puzzles 8-9 were used by Karwowski (2014) and puzzle 10 was used by Chuderski and Jastrzębski (2018a).

[3] Puzzles with numbers 1-7 were used in oryginal Laukkonen and Tangen's (2017) study. Puzzles 8-9 were used by Karwowski (2014) and puzzle 10 was used by Chuderski and Jastrzębski (2018a).

Table S1 Insight phenomenology measures used in study

dimension	question	answers (5-point Likert scale)
nagłość (Suddenness)	W jaki sposób odpowiedź pojawiła się w Twoim umyśle?	1 - stopniowo byłem/-am coraz bliżej rozwiązania 5 - rozwiązanie pojawiło się nagle
przyjemność (Pleasure)	W jakim stopniu rozwiązywanie tej zagadki było dla Ciebie przyjemne?	1 - bardzo nieprzyjemne 5 - bardzo przyjemne
pewność (Confidence)	W jakim stopniu jesteś pewny/-a swojej odpowiedzi?	1 - nie jestem pewny/-a, 5 - jestem bardzo pewny/-a
olśnienie (Aha! experience)	W jakim stopniu doznałeś/aś olśnienia? Olśnieniem nazywa się moment, w którym nagle wpadasz na rozwiązanie jakiegoś problemu i spontanicznie mówisz „Aha!”.	1 - nie doznałem-am olśnienia; 5 - całkowicie mnie olśniło

Table S2 Insight phenomenology measures in English

dimension	question	answers (5-point Likert scale)
suddenness	How did the answer come to your mind?	1 - solution came gradually 5 - all of a sudden
pleasure	How pleasant was solving this puzzle?	1 - very unpleasant 5 - very pleasant
confidence	How strong are you sure that your answer is correct?	1 - unsure 5 - very sure
aha! experience	How strongly did you feel an insight? Insight is called the moment when you realized the solution suddenly and often spontaneously say “aha!”.	1 - non-aha experience 5 - very strong aha experience

Table S3 Factor loadings, model fit indices, and internal consistency measures for the subjective insight measure for Sample A and Sample B

	Sample A	**Sample B**
	Factor loading	Factor loading
1. Confidence	0.868	0.795
2. Pleasure	0.774	0.655
3. Suddenness	0.725	0.673
4. Aha! experience	0.875	0.447
Model fit (EFA, one factor solution)	X^2 (2) = 1.31; p = .519	X^2 (2) = 9.79; p = .007
	RMSEA = 0.00 [0.00; 0.21]	RMSEA = 0.14 [0.06; 0.23]
	TLI = – 1.01	TLI = 0.86

Table S4 Comparison[4] of significant correlations observed in two samples

Correlation between	Sample A	Sample B	Z	p
Confidence - pleasure	.69	.52	*1.86*	*.063*
Confidence - suddenness	.60	.58	*0.30*	*.761*
Confidence – aha! experience	**.76**	**.28**	***4.97***	***<.001***
Confidence - accuracy	.59	.51	*0.86*	*.391*
Pleasure - suddenness	.56	.38	*1.65*	*.099*
Pleasure – aha! experience	**.65**	**.37**	***2.76***	***.005***
Pleasure - accuracy	.27	.36	*-0.69*	*.489*
Suddenness – aha! experience	**.66**	**.31**	***3.24***	***.001***
Suddenness - accuracy	.29	.23	*0.47*	*.637*

Sample A – *r* Pearson coefficient for the Sample A,
Sample B – *r* Pearson coefficient for the Sample B

[4] Diedenhofen, B. & Musch, J. (2015). cocor: A Comprehensive Solution for the Statistical Comparison of Correlations. PLoS ONE, 10(4): e0121945. doi: 10.1371/journal.pone.0121945 Available: http://dx.doi.org/10.1371/journal.pone.01219