Title
Does the finder of alternatives intentionally seek information irrelevant to the trained procedure?

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 45(45)

Authors
Ninomiya, Yuki
Iwata, Tomoyuki
Terai, Hitoshi
et al.

Publication Date
2023

Peer reviewed

Does the finder of alternatives intentionally seek information irrelevant to the familiar solution?

Yuki Ninomiya (ninomiya.yuki.t1@f.mail.nagoya-u.ac.jp)
Institutes of Innovation for Future Society (MIRAI), Nagoya University
Furo-cho, Chikusa-ku, Nagoya, Aichi, 464-8601, Japan

Tomoyuki Iwata (vdc123tykob315dw@gmail.com)
Graduate School of Informatics, Nagoya University
Furo-cho, Chikusa-ku, Nagoya, Aichi, 464-8601, Japan

Hitoshi Terai (terai@fuk.kindai.ac.jp)
Faculty of Humanity-Oriented Science and Engineering, Kindai University
Kaya no Mori 11-6, Iizuka, Fukuoka, 820-8555, Japan

Kazuhisa Miwa (miwa@is.nagoya-u.ac.jp)
Graduate School of Informatics, Nagoya University
Furo-cho, Chikusa-ku, Nagoya, Aichi, 464-8601, Japan

Abstract

Why do humans search for better alternatives when a familiar trained procedure is sufficient to solve the problem? Such a question is important in explaining the flexibility of human thinking. This study investigated whether the finder of an alternative procedure intentionally seeks to access information irrelevant to the trained procedure while solving a problem using the trained procedure. The results show that finders intentionally sought more information irrelevant to the trained procedure, even when solving a problem using the procedure. In addition, differences in intentional search may be caused by resistance to the reinforcement of fixation on the trained procedure. This study provides evidence that the discovery of alternatives involves the tendency to intentionally search for information irrelevant to a familiar solution.

Keywords: Einstellung effect; Discovery; information seeking; mouse tracking; problem solving

Introduction

The fixation on a solution obtained from experience can hinder the discovery of a better alternative (Bilalić et al, 2008a, 2008b; Luchins, 1942). This phenomenon, in which fixation on familiar solutions prevents flexible problem-solving, is called the Einstellung effect (Bilalic et al., 2008a, 2008b; Luchins, 1942).

The water jar task is one typical task used to test the Einstellung effect (Luchins, 1942). In this task, participants are asked to obtain the target amount of water using three water jars A, B, and C with different capacities. The task consists of three phases: Set, Critical, and Inspection. First, in the Set phase, participants are repeatedly presented with task stimuli that can be solved using only a specific procedure (e.g., B – A – 2 C) and learn this procedure. This learned procedure is called the trained procedure. Participants form a mental set in which this problem can be solved using the trained procedure. In the subsequent Critical phase, the problem can be solved either by using the trained procedure or a simpler procedure, such as C – A. This procedure that is more efficient or simpler than the trained procedure is called an alternative procedure. The Inspection phase can be solved using only the alternative procedure. It has been shown that the rate of discovery of alternative procedures in the Critical and Inspection phases is lower, and the time required for discovery is longer than in the case where the trained procedure is not learned (Luchins, 1942).

Studies of the Einstellung effect have shown how difficult it is to discover an alternative procedure. On the other hand, the importance of breaking away from the fixation on knowledge and experience has been pointed out, that is, the discovery of alternative procedures in scientific discovery and creative activity (Bilalić et al., 2008a, 2008b; Jansson & Smith, 1991; Miwa, 2004; Neroni et al, 2017).

Discovery of alternatives under the Einstellung effect

Regarding previous studies on the discovery of alternative procedures, it is necessary to distinguish between the discovery of an alternative procedure in a situation where a problem is successfully solved by a trained procedure, such as in the Critical phase, and where a problem fails to be solved by a trained procedure, such as in the Inspection phase. The former situation is called a success situation, and the latter is called a failure situation. This is because there is a difference between the former and latter discovery processes in terms of whether the need to change the trained procedure is fed back. In a failure situation, participants are forced to depart from the trained procedure because solving the problem using the trained procedure fails. Therefore, the discovery of alternative procedures is more likely to occur in

In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society*.

failure situations than in success situations (Chesney et al., 2013; Sheridan & Reingold, 2013).

The discovery process in failure situations has been discussed in terms of the relationship between the failure of the trained procedure and changes in search, based on the representation change theory in insight research (Chesney et al., 2013; Sheridan & Reingold, 2013; Thomas et al., 2018). In these explanations, the fixation on solutions that inhibit problem solving as a mental set is considered one of the constraints that must be eliminated. Specifically, at first, it fails to solve the problem using a trained procedure recalled as the initial representation. This feedback on failure to solve problems leads to the discovery of an alternative procedure by reducing the bias of attention to the search area related to the fixated trained procedure. We refer to this area as a fixation-relevant area. For example, Sheridan and Reingold (2013) showed that under failure situations, the bias in attention to the fixation-relevant area gradually weakens, leading to discovery by focusing attention on search regions not related to the fixated trained procedure (which we refer to as a fixation-irrelevant area).

Importantly, in failure situations, the focus of attention changes from the time before the discovery of the alternative procedure (Ellis et al., 2011; Knoblich et al., 2001). For example, Ellis et al. (2011) examined the relationship between changes in eye movements and the presence or absence of insight experience (Aha!) in an anagram insight problem. They showed that information seeking related to the solution increased before consciously finding the solution, regardless of the presence or absence of an insight experience.

However, this explanation cannot be directly applied to success situations. This is because, in a success situation, there is no feedback that the trained procedure is not useful in solving the problem. The discovery of alternative procedures in success situations is an important feature of highly skilled experts (Bilalić et al., 2008a, 2008b, 2010). Therefore, the question of how to discover alternative procedures despite successful problem solving is an important issue for examining the flexibility of human thinking.

Discovering the alternative procedure in a success situation

The discovery of an alternative procedure in success situations has been discussed through a comparison between the problem-solving processes of finders and non-finders. For example, Bilalić et al. (2008a, 2008b) showed that super-experts, especially highly skilled experts, can find better alternative procedures in success situations by switching their attention from the fixation-relevant area to the fixation-irrelevant area (Bilalić et al, 2008a, 2008b, 2010; Sheridan & Reingold, 2013). This means that even in a success situation, some participants can discover alternative procedures. To do so, it is necessary to quit using the trained procedure and switch attention to the fixation-irrelevant area.

Previous studies have shown that even in success situations, the distribution of attention changes even before the discovery of the alternative procedure (Ninomiya et al., 2022). Ninomiya et al. (2022) examined when the difference in the distribution of attention between finders and non-finders occurred in the problem-solving process from before obtaining the trained procedure to the discovery of the alternative procedure. Specifically, they distinguished between trials in which the alternative procedure was discovered (hereafter referred to as finding trial) and trials prior to the finding trial in which the finder was solving the problem using a trained procedure (hereafter referred to as pre-finding trial) by conducting a trial-by-trial analysis. They demonstrated that in pre-finding trials, finders' distribution of attention to the fixation-relevant area was already smaller than that of non-finders. This means that finders tended to pay more attention to information unrelated to the trained procedure from the time they were fixated on the trained procedure prior to the discovery of the alternative procedure.

In addition, Ninomiya et al. (2022) demonstrated that this difference was caused by a difference in the amount of change in the distribution of attention from the trial before obtaining the trained procedure to the pre-finding trial. Specifically, they first confirmed that there was no difference in the distribution of attention in the first trial of the Set phase (henceforth referred to as the pre-training trial), in which the fixated trained procedure had not been obtained. Then, they demonstrated that the non-finders' distribution of attention to the fixation-relevant area was enhanced from the pre-training trial to the pre-finding trial, whereas the finders' distribution was not. As already mentioned, in failure situations, the relaxation of constraints by negative feedback occurs before the discovery, leading to the discovery of an alternative procedure (Chesney et al., 2013; Sheridan & Reingold, 2013). In contrast, Ninomiya et al. (2022) suggested that the discovery of an alternative procedure in a success situation involves not only relaxation but also the degree of resistance to the reinforcement of attention to the fixation-relevant area by the positive feedback of successful problem solving using the trained procedure.

Limitations of previous studies

There are two possible explanations for the differences in the distribution of attention between finders and non-finders in previous studies. One is that the finder is more likely than the non-finder to seek information with the intention of obtaining information irrelevant to the trained procedure, even when solving a problem using the procedure. This explanation predicts that the reason for the difference in the distribution of attention is that the finder intentionally seeks information irrelevant to the trained procedure more than the non-finder. In this study, we refer to such a search as intentional seeking.

The other is that finders are simply less focused on the task than non-finders, resulting in a distraction to the fixation-relevant area during the use of the trained procedure. This explanation does not involve the intention to refer to information irrelevant to the trained procedure. In support of this explanation, previous findings have demonstrated that limited attentional resources facilitate the discovery of alternative procedures in the water jar task (Beilock & DeCaro, 2007; Ricks et al., 2007).

Thus, the meaning of the difference in the distribution of attention between finders and non-finders varies depending on whether the participants were intentionally seeking to obtain information irrelevant to the trained procedure. However, in the discussion based on eye movements in Ninomiya et al. (2022), even if a gaze on the fixation-irrelevant area is observed, it would not be clear whether the participants intended to obtain information on that area or not. In the present study, we examine whether the finder intentionally searched for information in the fixation-irrelevant area in the pre-finding trial.

In addition, Ninomiya et al. (2022) demonstrated that the reinforcement of attention bias toward the fixation-relevant area for finders was smaller than that for non-finders. In this study, we also examine how the intentional seeking of the finder and non-finder changes from the pre-training to pre-finding trial.

Approach

In this study, we developed a blind water-jar task to examine participants' intentional seeking in a success situation (Figure 1). This task consisted of three phases—Set, Critical, and Inspection—as in the basic water jar task, and participants were required to answer a formula to draw the required amount of water using five water jars: A, B, C, D, and E. A difference from the original task was that the participant had to point the mouse cursor at the jar that they wanted to check regarding its capacity.

The time and frequency of pointing the mouse cursor at each of the water jars were recorded. Importantly, by blindfolding the jars, participants could not access the information about water jars' capacity simply by directing their attention to them. Therefore, the pointing of their mouse cursor to a water jar irrelevant to the trained procedure can be considered an intentional attempt to access that information.

In addition, in this task, the trained procedure was always C – D – E, and the alternative procedure was always C – A. This allows us to distinguish the water jars in C, D, and E as information relevant to the trained procedure, and the water jars in A and B as information irrelevant to the trained procedure. Therefore, by examining the amount of pointing on the water jars of A and B, we can examine the degree of intentional seeking to access the information in the fixation-irrelevant area.

Objective

In this study, we examine the following research questions (RQs) regarding the intentionality of the finder's information seeking. RQ1 is whether there is a difference in intentional seeking in the pre-finding trial between finders and non-finders of the alternative procedure. To answer RQ1, the amount of mouse pointing on the water jars of A and B in the pre-finding trial was compared between finders and non-finders. RQ2 is how intentional seeking changes from the pre-training trial to the pre-finding trial between finders and non-finders. To answer RQ2, the amount of change in mouse pointing from the pre-training trial to the pre-finding trial was compared between finders and non-finders.

Method

Participants

A total of 117 participants (N_{female} = 73, N_{male} = 44, M_{age} = 38.96, SD_{age} = 8.62) were recruited via crowdsourcing.

Task and procedure

The experimental task was created using jsPsych (de Leeuw, 2015). The task consisted of 41 trials. The first two trials were designed to encourage participants to find the trained procedure. These two trials could be solved only using the trained procedure, and both numbers D and E were divisible by 10 (e.g., C = 71, D = 40, E = 20, and target amount of water = 11). The next seven trials were designed as the Set phase. The Set phase could also be solved only using the trained procedure, which had the same difficulty level as the trained procedure in the subsequent Critical phase. Specifically, all procedures were designed to be two-digit, no carry-up and down, and not divisible by 10. Therefore, participants could not recognize the switch from the Set phase to the Critical phase as long as they solved the problems in the trained procedure. In addition, in the first through fourth trials, participants were given feedback that this trial could be solved using the trained procedure after they responded to encourage them to fixate on it (for a study using feedback, see Van Stockum & DeCaro, 2020). The Critical phase consisted of 30 trials, which could be answered using both the trained and alternative procedures. The order of the trials in the critical phase was randomized. The last two trials (trials 40st and 41st) constituted the Inspection phase, which could be solved only using the alternative procedure. The final trial (41st trial) was designed with C < D, and it was clear that the calculation of C–D was not valid. Therefore, we used the non-use of the trained procedure in the final trial as a screening criterion to confirm that participants responded after performing the calculations.

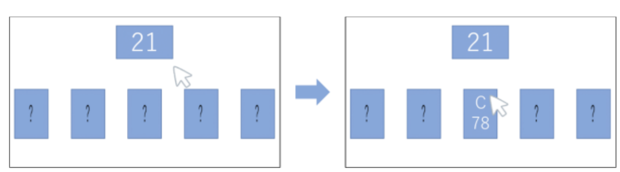

Figure 1: Blind water jar problem

Each trial comprised three elements: a preparation page, a question page, and a response page. First, on the preparation page, when the participant was ready to proceed to the question page, the participant had to press the space key to advance. On the subsequent question page, five water jars from A to E were displayed as blue squares marked with a "?" (Figure 1). The capacity of each water jar could be checked by pointing using the mouse. In addition, a target amount of water was given at the top square, and participants were asked to think of a way to draw the required amount of water using the five water jars. The target was always displayed. When an answer was found, participants pressed the space key to advance to the next page. On the response page, the participants were asked to type their answers in the form of a calculation formula (e.g., C − D + E).

The participants joined the experiment by accessing the task from a URL distributed online. After providing informed consent, participants practiced and completed the task.

Result

Calculation of the amount of mouse pointing pre- and post-finding trial

We eliminated five participants who knew about the water jar task and five who responded to the 41st trial using the trained procedure. In the present study, 17 participants who did not respond using the trained procedure in the two pre-finding trials were excluded from the analysis, in accordance with a previous study (Ninomiya et al., 2022). As a result of the above screening, 90 participants ($N_{female} = 61$, $N_{male} = 29$, $M_{age} = 38.96$, $SD_{age} = 8.23$) were included in the analysis.

The following calculation procedure was carried out: the time spent mouse pointing at water jars A and B (Figure 2, Table 2) and the frequency of mouse pointing at water jars A and B (Figure 3, Table 2) in the finding trial and the two trials before and after the finding trial.

First, the participants were divided into a find group (27 participants) and a not-find group (63 participants) according to whether they had discovered the alternative procedure in the critical phase. In the find group, the amount of mouse pointing in the finding trial and two trials each before and after was used in the analysis. The finding trial is called t, and the two trials before and after it are called t-2, t-1, t+1, and t+2. However, the not-find group does not have the finding trial. Therefore, it is necessary to define the trials of the not-find group, which correspond to t and two trials each before and after t of the find group, in order to compare them with the find group.

In this study, the amount of mouse pointing for the not-find group was calculated using the following procedure to ensure temporal equality of the number of trials between the two groups. First, we estimated a linear approximation between the number of trials and the amount of mouse pointing in the critical phase for each participant in the not-find group. Next, we calculated the average number of trials required to discover for the find group. The average number of attempts required for discovery was 9.22 trials. Then, we estimated the amount of mouse pointing in t and two trials each before and after it in the not-find group by substituting 9.22 into t of the linear approximation formula for each participant.

Amount of mouse pointing pre-finding trial

To examine differences in intentional seeking between the find and not-find groups in pre- and post-finding trials, we performed two-way mixed design ANOVAs with the group as a between-factor (find, not-find) and trial as a within-factor (t-2, t-1, t, t + 1, t + 2) regarding the time and frequency pointed to the water jar in A and B (Figure 2, and Table 1). The results showed that the main effect of group (pointing time: $F(1, 88) = 147.76$, $p < .001$, $\eta_G^2 = 0.47$; pointing frequency: $F(1, 88) = 221.12$, $p < .001$, $\eta_G^2 = 0.61$), the main effect of trial (pointing time: $F(4, 352) = 27.20$, $p < .001$, $\eta_G^2 = 0.13$; pointing frequency: $F(4, 352) = 87.69$, $p < .001$, $\eta_G^2 = 0.27$), and the interaction effects (pointing time: $F(4, 352) = 28.30$, $p < .001$, $\eta_G^2 = 0.13$; pointing frequency: $F(4, 352) = 92.20$, $p < .001$, $\eta_G^2 = 0.28$) were significant for both pointing time and pointing frequency.

Since the purpose of this analysis was to compare mouse pointing between the find and not-find groups, we tested the simple main effect of group for each trial in both time and frequency (Table 1, Figure 2). The results show that both the time and frequency of pointing to the water jars of A and B were greater in the pre-finding trials (t-1 and t-2) for the find group than for the not-find group.

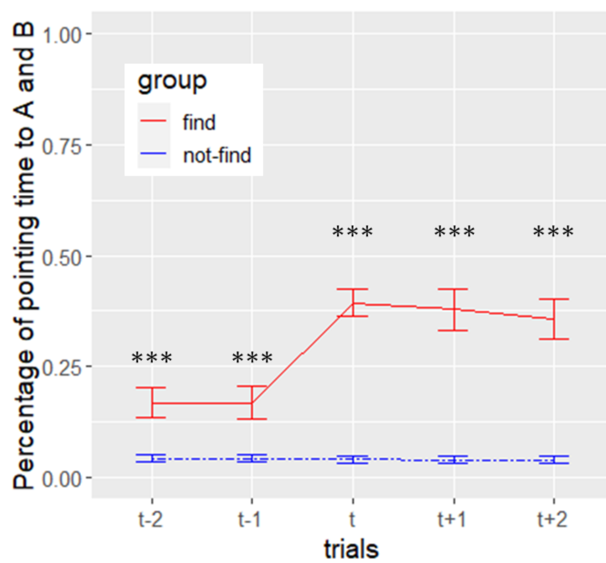

Figure 2: The percentage of pointing time to the water jar of A and B. The pointing frequency is not plotted. Error bars represent standard errors. *** = $p < .001$.

Table 1: Mean (*SD*) and results of multiple comparisons for pointing pre- and post-finding trial

		Mean (*SE*) of Find group	Mean (SE) of Not-Find group	*df*	*F-ratio*	*p*	η_G^2
Percentage of pointing time to A and B	t - 2	.168(.034)	.042(.008)	1, 88	24.34	< .001	0.22
	t - 1	.169(.038)	.042(.008)	1, 88	21.51	< .001	0.20
	t	.394(.031)	.041(.008)	1, 88	222.66	< .001	0.72
	t + 1	.378(.046)	.040(.008)	1, 88	108.27	< .001	0.55
	t + 2	.357(.046)	.039(.008)	1, 88	97.40	< .001	0.53
Percentage of pointing frequency to A and B	t - 2	.178(.032)	.065(.012)	1, 88	16.93	< .001	0.16
	t - 1	.206(.040)	.063(.011)	1, 88	20.62	< .001	0.19
	t	.511(.024)	.062(.011)	1, 88	374.72	< .001	0.81
	t + 1	.508(.035)	.060(.011)	1, 88	249.83	< .001	0.74
	t + 2	.548(.036)	.058(.011)	1, 88	289.88	< .001	0.77

Change in amount of pointing from pre-training trial

Next, we examined whether there was a difference between the two groups in the amount of decrease in pointing to the water jar of A and B from the pre-training trial to the pre-finding trial. First, we confirmed that there were no differences between the two groups at the beginning of the Set phase. Specifically, we conducted a two one-sided test (TOST) using the minimum effect size (*Cohen's d* = 0.58), which rejects the null hypothesis when $1 - \beta = .80$ and $\alpha = .05$ for the sample size (27 in the find group and 63 in the not-find group) as the equivalence boundary. The results showed that both pointing time (upper: $t(46.2) = 2.82$, $p = .004$, lower: $t(46.2) = 2.15$, $p = .018$) and frequency (upper: $t(51.6) = 2.77$, $p = .003$, lower: $t(51.6) = 2.33$, $p = .011$) were significantly equivalent between the two groups. The mean (*SE*) percentage of mouse pointing for the find group was .286 (.170) % for time and .296 (.103) % for frequency, whereas those for the not-find group were .298 (.158) % for time and .301 (.108) % for frequency.

Then, to examine differences in the change in mouse pointing between the two groups, we performed two-way mixed design ANOVAs with the group as a between-factor (find, not find) and trial as a within-factor (preset, t-2, t-1) regarding the time and frequency pointed to the water jars of A and B (Figure 3). The results showed that the main effect of group (pointing time: $F(1, 88) = 14.49$, $p < .001$, $\eta_G^2 = 0.07$; pointing frequency: $F(1, 88) = 16.85$, $p < .001$, $\eta_G^2 = 0.09$), main effect of trial (pointing time: $F(2, 176) = 63.57$, $p < .001$, $\eta_G^2 = 0.28$; pointing frequency: $F(2, 176) = 70.42$, $p < .001$, $\eta_G^2 = 0.27$), and interaction effects (pointing time: $F(2, 176) = 8.80$, $p < .001$, $\eta_G^2 = 0.05$; pointing frequency: $F(2, 176) = 11.10$, $p < .001$, $\eta_G^2 = 0.06$) were significant for both pointing time and pointing frequency.

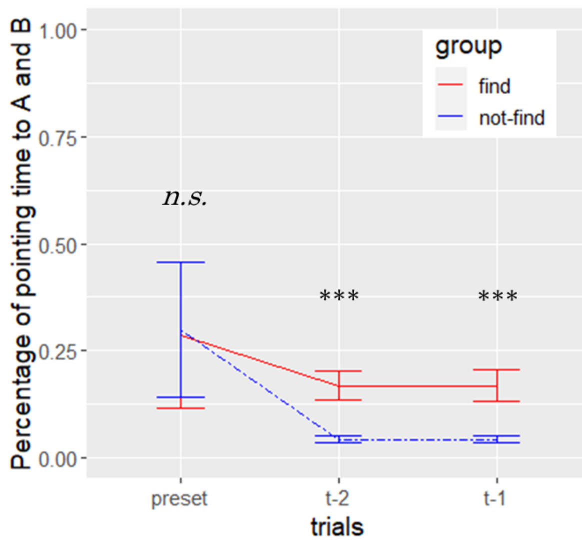

Figure 3: Change in the percentage of pointing time from pre-training to pre-finding trial. The pointing frequency is not plotted. Error bars represent standard errors. *** = $p < .001$. *n.s.* = $p > .10$

The purpose of this analysis was to examine whether the change in the amount of pointing from pre-training trials to pre-finding trials differed between the two groups. Therefore, we tested for a simple main effect of the trial in each group. The results showed that both the find group (pointing time: $F(2, 52) = 5.63$, $p = .006$, $\eta_G^2 = 0.09$; pointing frequency: $F(2, 52) = 4.40$, $p = .017$, $\eta_G^2 = 0.09$) and the not-find group (pointing time: $F(2, 124) = 132.56$, $p < .001$, $\eta_G^2 = 0.57$; pointing times: $F(2, 124) = 233.59$, $p < .001$, $\eta_G^2 = 0.57$) showed a simple main effect in the decreasing direction. The effect size of the simple main effect of the trial, based on Cohen's (1988) criterion (small: $\eta_G^2 = 0.02$, medium: $\eta_G^2 = 0.13$, large: $\eta_G^2 = 0.26$), was small for the find group and large for the not-find group (Bakeman, 2005; Lakens, 2013; Olejnik & Algina, 2003).

Discussion

In this study, we examined intentional seeking before the discovery of an alternative procedure using a blind water jar task to answer the two RQs. The results demonstrated that the find group was more likely to search for information in areas irrelevant to the trained procedure, even in trials in which they were solving the task using the trained procedure. It is important to note that the mouse pointing to the water jar of A and B was caused by intentional seeking to access information irrelevant to the trained procedure, rather than by an attentional distraction, such as accidentally and unintentionally looking at the area. This means that finders intentionally seek to obtain the information irrelevant to the trained procedure more than non-finders, even though the trained procedure can be used to solve the problem.

Differences in intentional seeking were not observed prior to the obtaining of the trained procedure, but only in the pre-finding trials. This suggests that the differences in intentional seeking in pre-finding trials occurred during the process of repeatedly experiencing successful problem solving using the trained procedure.

In addition, it was demonstrated that the decrease in pointing time and frequency to the water jar of A and B from the pre-training trial to the pre-finding trial was greater for non-finders than for finders. This means that finders had a smaller decrease in the amount of intentional access to information irrelevant to the trained procedure than non-finders, despite the success of the trained procedure. The smaller decrease in intentional seeking of information irrelevant to the trained procedure does not mean that the fixation is relaxed and participants are approaching the discovery of the alternative procedure. This is because finders responded using the trained procedure in the pre-finding trial, and their intentional seeking for the fixation-irrelevant area was rather reduced compared to the pre-training trial. Therefore, it is more plausible that the smaller decrease in intentional seeking for finders than for non-finders is interpreted as a resistance to the reinforcement of the fixation on the trained procedure by positive feedback, rather than as a relaxation of the fixation by negative feedback.

This study contributes to the literature by providing evidence that differences in intentional seeking are involved in differences in the distribution of attention prior to the discovery of the alternative procedure. However, it should be noted that the results of this study do not indicate that distraction is not involved in the differences in the distribution of attention prior to discovery. To solve this problem, experiments need to be conducted to control for the degree of distraction by guiding gaze and to discuss the relationship between attentional resources as individual traits, such as vigilance, and the discovery of alternative procedures. Future studies should examine the effect of distraction on the discovery of an alternative procedure in a success situation through these examinations.

Acknowledgments

This research was supported by funding from the JSPS KAKENHI Grant Number JP22H03912.

References

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior research methods*, *37*(3), 379-384.

Beilock, S. L., & DeCaro, M. S. (2007). From poor performance to success under stress: Working memory, strategy selection, and mathematical problem solving under pressure. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(6), 983–998.

Bilalić, M., McLeod, P., & Gobet, F. (2008a). Inflexibility of experts—Reality or myth? Quantifying the Einstellung effect in chess masters. *Cognitive psychology*, *56*(2), 73-102.

Bilalić, M., McLeod, P., & Gobet, F. (2008b). Why good thoughts block better ones: The mechanism of pernicious Einstellung (Set) effect. *Cognition*, *108*, 652-661.

Bilalić, M., McLeod, P., & Gobet, F. (2010). The mechanism of the Einstellung (Set) effect: A pervasive source of cognitive bias. *Current Directions in Psychological Science, 19*(2), 111-115.

Chesney, D. L., McNeil, N. M., Brockmole, J. R., & Kelley, K. (2013). An eye for relations: eye-tracking indicates long-term negative effects of operational thinking on understanding of math equivalence. *Memory & cognition*, *41*(7), 1079-1095.

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd ed.). Routledge.

De Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior research methods*, *47*(1), 1-12.

Ellis, J. J., Glaholt, M. G., & Reingold, E. M. (2011). Eye movements reveal solution knowledge prior to insight. *Consciousness and Cognition*, *20*(3), 768-776.

Jansson, D. G., & Smith, S. M. (1991). Design fixation. *Design studies*, *12*(1), 3-11.

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

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in psychology*, *4*, 863.

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

Miwa, K. (2004). Collaborative discovery in a simple reasoning task. *Cognitive Systems Research*, *5*(1), 41-62.

Neroni, M. A., Vasconcelos, L. A., & Crilly, N. (2017). Computer-based "mental Set" tasks: An alternative approach to studying design fixation. *Journal of Mechanical Design*, *139*(7), 071102.

Ninomiya, Y., Terai, H., & Miwa, K. (2022). Differences in the distribution of attention to trained procedure between finders and non-finders of the alternative better procedure. *Frontiers in Psychology*, *13*.

Olejnik, S., & Algina, J. (2003). Generalized Eta and Omega Squared Statistics: Measures of Effect Size for Some Common Research Designs. *Psychological Methods*, *8*(4), 434–447.

Ricks, T. R., Turley-Ames, K. J., & Wiley, J. (2007). Effects of working memory capacity on mental Set due to domain knowledge. *Memory & Cognition*, *35*(6), 1456-1462.

Sheridan, H., & Reingold, E. M. (2013). The Mechanisms and Boundary Conditions of the Einstellung Effect in Chess: Evidence from Eye Movements. *PLoS one*, *8*(10), e75796.

Thomas, C., Didierjean, A., & Kuhn, G. (2018). It is magic! How impossible solutions prevent the discovery of obvious ones? *Quarterly Journal of Experimental Psychology*, *71*(12), 2481-2487.

Van Stockum Jr, C. A., & DeCaro, M. S. (2020). When working memory mechanisms compete: Predicting cognitive flexibility versus mental set. *Cognition*, *201*, 104313.