

Avoiding the Conflict: Metacognitive Awareness Drives the Selection of Low-Demand Contexts

Kobe Desender
Vrije Universiteit Brussel

Cristian Buc Calderon
Université Libre de Bruxelles

Filip Van Opstal
Université Libre de Bruxelles and University of Amsterdam

Eva Van den Bussche
Vrije Universiteit Brussel

Previous research attempted to explain how humans strategically adapt behavior in order to achieve successful task performance. Recently, it has been suggested that 1 potential strategy is to avoid tasks that are too demanding. Here, we report 3 experiments that investigate the empirically neglected role of metacognitive awareness in this process. In these experiments, participants could freely choose between performing a task in either a high-demand or a low-demand context. Using subliminal priming, we ensured that participants were not aware of the visual stimuli creating these different demand contexts. Our results showed that participants who noticed a difference in task difficulty (i.e., metacognitive aware participants) developed a clear preference for the low-demand context. In contrast, participants who experienced no difference in task difficulty (i.e., metacognitive unaware participants) based their choices on variables unrelated to cognitive demand (e.g., the color or location associated with a context), and did not develop a preference for the low-demand context. Crucially, this pattern was found despite identical task performance in both metacognitive awareness groups. A multiple regression approach confirmed that metacognitive awareness was the main factor driving the preference for low-demand contexts. These results argue for an important role of metacognitive awareness in the strategic avoidance of demanding tasks.

Public Significance Statement

In this study, participants were free to perform a task in either a high-demand context or in a low-demand context. In three experiments, it was observed that only participants who were aware that the task was more demanding to perform in one of both contexts developed a preference for the low-demand context. Participants who were not aware of this difference in cognitive demand did not, on average, develop any preference. Importantly, both metacognitive awareness groups displayed similar task performance, ruling out that differences in performance can account for this effect. We conclude that metacognitive awareness is crucial for the avoidance of cognitive demand.

Keywords: metacognition, cognitive control, demand avoidance, conflict, subjective experience

Supplemental materials: <http://dx.doi.org/10.1037/xhp0000391.supp>

Because our cognitive resources are limited (e.g., Kahneman, 1973), humans try to match the cognitive effort they put in a task with that required by the task. This ability to control behavior (i.e.,

cognitive control) can be broadly defined as the set of mechanisms required to efficiently perform a task. Cognitive control is especially needed when automatic or habitual behavior needs to be

This article was published Online First April 3, 2017.

Kobe Desender, Department of Psychology, Vrije Universiteit Brussel; Cristian Buc Calderon, Center for Research on Cognition and Neurosciences and Neurosciences Institute, Université Libre de Bruxelles; Filip Van Opstal, Center for Research on Cognition and Neurosciences and Neurosciences Institute, Université Libre de Bruxelles and Department of Psychology, University of Amsterdam; Eva Van den Bussche, Department of Psychology, Vrije Universiteit Brussel.

This work was supported by a grant of the Research Foundation Flanders, Belgium (FWO-Vlaanderen) awarded to Kobe Desender (Grant 11H3415N). Cristian Buc Calderon is supported by a fellowship

from the National Fund for Scientific Research (FRS—FNRS Belgium, Grant I.A.188.13F). Eva Van den Bussche is supported by a grant of the Research Foundation Flanders, Belgium (FWO-Vlaanderen, Grant G023213N). We thank Sarah Beurms, Kilian Amelinckx, Wim van Gorp, Charlotte Muscarella and Elke Van Lierde for their help in coding the data, and Gethin Hughes and two anonymous reviewers for useful comments on a draft of the manuscript.

Correspondence concerning this article should be addressed to Kobe Desender, Faculty of Psychology and Educational Sciences, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium. E-mail: kobe.desender@vub.ac.be

overruled in favor of more appropriate responses (Botvinick, Braver, Barch, Carter, & Cohen, 2001; Verguts & Notebaert, 2009). When a task is highly automated and relies purely on habitual behavior, task performance is optimal without investing much control. However, an increased investment in cognitive control is necessary when an automatic response needs to be overruled in favor of another, more appropriate, response.

Recently, it was suggested that cognitive control comes at a cost (Botvinick, 2007). Building on the notion that physical effort is avoided whenever possible (Hull, 1943), it was shown that cognitive effort is also avoided when given the choice. This was demonstrated using the demand selection task (Kool, McGuire, Rosen, & Botvinick, 2010), in which participants can freely decide on each trial in which of two possible contexts they want to perform a task. In this paradigm, the task requires high control investment when participants choose one context (e.g., with a high proportion of conflict trials; high-demand context), whereas the task can be performed more automatically when the other context (e.g., with a low proportion of conflict trials; low-demand context) is chosen. As expected, participants tend to avoid the high-demand context and prefer the low-demand context (Kool et al., 2010). Interestingly, Kool and colleagues (2010) observed that even participants who were unaware of the experimental manipulation, measured by a postexperimental questionnaire, developed a preference for the low-demand context. This observation is in line with earlier theoretical and empirical work arguing that awareness does not play an important role in decision making. Indeed, it has been suggested that deciding between multiple strategies to perform a task does not depend on awareness, but rather is driven by implicit processes (Reder, 1988; Reder & Schunn, 1996). For example, participants tend to continue a successful strategy without being aware of doing so (Reder & Schunn, 1996). More recently, converging evidence was provided by several studies showing that task choices can be reliably influenced outside of awareness by means of unconscious cues (Lau & Passingham, 2007; Mattler, 2003; Reuss, Kiesel, Kunde, & Hommel, 2011) or by means of sequences in the switches between tasks that participants learned to use to predict the upcoming task, without being aware of doing so (Gotler, Meiran, & Tzelgov, 2003; Koch, 2001). These findings have led to the suggestion that awareness is not a prerequisite for high-level cognition, such as choice behavior (e.g., Hassin, 2013).

However, there are reasons to doubt the conclusion of Kool et al. (2010) that awareness is not involved in the avoidance of cognitive demand. In their study, it was assessed whether participants were aware of the experimental manipulation itself, rather than of the difference in difficulty between both contexts (for similar approaches, see Blais, Harris, Guerrero, & Bunge, 2012; Schmidt, Crump, Cheesman, & Besner, 2007). It is, however, widely acknowledged that people are not always aware of the underlying reason *why* they feel or do things (Hurlburt & Heavey, 2001; Nisbett & Wilson, 1977). The questionnaire used to assess awareness in the study by Kool and colleagues (2010) might thus target the wrong content of awareness. While participants might not be aware of the experimental manipulation, it is plausible that they were aware of the fact that the task was more difficult to perform in one context than in the other. To understand the relation between awareness and choice behavior, it is therefore crucial to investigate if the avoidance of cognitive demand depends on the *metacognitive* appreciation of the processing costs. Metacognition,

often defined as cognition about cognition, is used here to denote the conscious experience of task performance. For instance, in the fluency literature, a number of studies have revealed how the readability of font type has a large influence on the feeling of fluency (Song & Schwarz, 2008, 2009). When reading a book, one is typically unaware of the exact font that is being used. Nevertheless, when the font is difficult to read it will feel much harder to stay focused on the reading activity. As a consequence, it is more likely that one will put the book aside to avoid cognitive effort. Although this avoidance behavior might be misattributed to something else than the font type: for example, to the content of the book, the underlying trigger for this avoidance is the metacognitive experience that the book is difficult to read. Clearly, it would be hard-pressed to claim that the decision to put the book aside (cf., demand avoidance) was made without awareness, based on the observation that participants are unaware of the influence of the font type (cf., the experimental manipulation).

The above raises the possibility that the absence of a relation between awareness and demand avoidance in previous work was caused by using a questionnaire that targeted the wrong contents of awareness. In fact, there are good reasons to assume that demand avoidance does depend on metacognitive awareness. It is known that tasks with high processing costs lead to experiences of high subjective effort (Desender, Van Opstal, & Van den Bussche, 2014; Morsella, Gray, Krieger, & Bargh, 2009; Westbrook, Kester, & Braver, 2013). Also, there is evidence from other domains, such as memory and reasoning, that subjective experiences guide strategic behavior (Metcalf & Finn, 2008; Thompson, Evans, & Campbell, 2013; Thompson, Turner, et al., 2013). Moreover, recent theoretical developments also emphasize the crucial involvement of subjective experiences of cognitive events in choice behavior, and cognitive control more generally (Kurzban, Duckworth, Kable, & Myers, 2013; see also Hockey, 2011; Inzlicht & Schmeichel, 2012).

To investigate whether the avoidance of cognitive demand depends on the metacognitive appreciation of the demand, we ran three experiments in which participants could freely choose between performing a task in either a high-demand or a low-demand context. Although the task was the same in both contexts, the proportion of response conflict differed between contexts and could either be high (creating a high-demand context) or low (creating a low-demand context). We used subliminal priming to create response conflict, which ensured that task performance and perceived difficulty are influenced without participants being aware of the visual stimuli driving these changes (Chambon & Haggard, 2012; Desender et al., 2014; Wenke, Fleming, & Haggard, 2010). By carefully questioning for metacognitive awareness of the difference in difficulty between both boxes after completion of the experiment, we aimed to contrast the avoidance of high-demand contexts between metacognitive aware participants and metacognitive unaware participants. This allows us to examine whether subjective experience is a crucial factor in the avoidance of high-demand contexts.

Experiments 1 and 2

Experiments 1 and 2 were designed to show that metacognitive awareness plays an important role in the selection of low-demand contexts. The difference between both experiments lies in the instructions given to the participants. In Experiment 1, participants

were explicitly instructed that the experimental task would be easier in one of the two contexts, and they were encouraged to select the low-demand context more often. Because this raises the possibility that a preference for the low-demand context was induced by the instructions, in Experiment 2, participants were given instructions that remained neutral with respect to the difference in difficulty between both contexts.

Method

Participants. Thirty-two students (13 males, mean age: 19.6 years, *SD* = 1.8, range 18–23) participated in Experiment 1, and 32 students (7 males, mean age: 19.6 years, *SD* = 3.3, range 18–33) in Experiment 2. All participated for course requirements. Participants provided written informed consent and all reported normal or corrected to normal vision and were naive with respect to the hypothesis.

Stimuli and apparatus. All stimuli were presented on a black background on a 15-in. 85 Hz CRT monitor, synchronized with the vertical refresh rate. During the experiment, two boxes were presented, one at the top and one at the bottom of the screen. Both boxes had their center at 4.6° above or below the middle of the screen. Boxes were 16.9° wide and 7.5° high, with a black center (7.6° wide and 3.7° high). In Experiment 1, one box was always orange (RGB: 247, 150, 70, CIE-lab values: 70.65, 16.84, 33.05) and the other blue (RGB: 79, 129, 189, CIE-lab values: 65.92, -9.13, -26.93). The color of the boxes was counterbalanced between participants. The color of both boxes was changed to gray (RGB: 166, 166, 166, CIE-lab values: 74.38, 4.71, -3.10) in Experiment 2, because the use of different colors in Experiment 1 sometimes obscured a clear interpretation of the origin of the participants’ choice behavior (see below). Primes were arrows

(1.5° wide and 0.7° high) pointing to the left or to the right. These primes fit perfectly within the contours of the targets (3.3° wide and 1.4° high) which also pointed to the left or right (i.e., metac-contrast masking; Klotz & Neumann, 1999). Responses were collected using a Cedrus response box (RB-840).

Procedure. Participants were instructed that their main task was to respond quickly and correctly to the left or right direction of target arrows, by respectively pressing a left or right button on a response box. They were told that the experiment was divided into several blocks, each block comprising 20 trials. Before the start of each block, participants could freely choose in which of two boxes they wanted to perform the task by pressing the button with the corresponding color (Experiment 1) or the corresponding label (i.e., “below” or “above”; Experiment 2) on their response box. In each block, the target arrows to which participants had to respond only appeared in the box that was selected. Meanwhile the other box remained empty. In Experiment 1, participants were instructed that, although it would not be clear why, they might notice that the task was easier to perform in one of the two boxes. It was explained that if they would notice this, they were encouraged to select the “easy” box more often. In Experiment 2, participants were simply told that they were free to choose whichever box they wanted and the difference in difficulty between both boxes was not mentioned in the instructions.

Each trial started with a centrally presented fixation cross for 1,000 ms. Next, a prime arrow was presented for 23 ms, followed by a blank screen for 23 ms, and finally a target arrow for 118 ms. Responses were recorded up to 3,000 ms after target onset (Figure 1). The intertrial interval lasted 800 ms. Feedback about the accuracy was not provided. Prime and target could trigger the same response (i.e., a congruent trial) or different responses (i.e., an

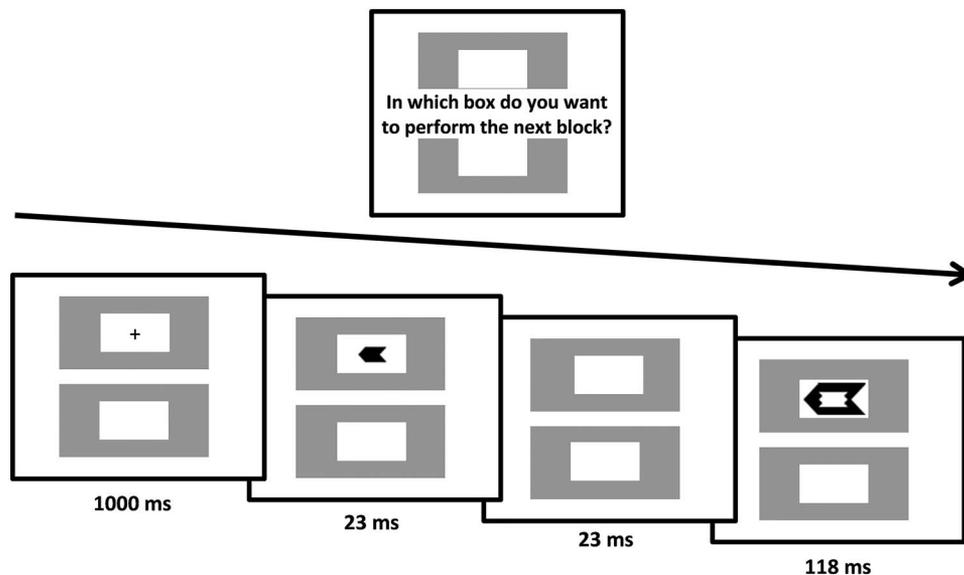


Figure 1. Example of a congruent trial in the main part of Experiments 1, 2 and 3. Boxes were colored orange and blue in Experiment 1. In Experiment 2 and 3, boxes were gray. After each block of 20 trials, participants freely chose in which box they wanted to perform the task for the next block. The task required low control investments in one box (low-demand context; 20% incongruent trials) and high control investments in the other box (high-demand context; 80% incongruent trials).

incongruent trial). Typically, congruent trials result in faster reaction times (RTs) compared with incongruent trials, a difference known as the congruency effect. Importantly, the proportion of congruency between prime and target was determined by the position of the boxes. In the upper box, prime and target triggered the same response on 16 out of 20 trials and a different response on the remaining 4 trials (i.e., 80% congruent trials; low-demand context). In the lower box, prime and target triggered the same response on four out of 20 trials and different responses on the remaining 16 trials (i.e., 20% congruent trials; high-demand context). The relation between position and demand context was counterbalanced across participants. In Experiment 1, participants performed 30 blocks of 20 trials, which took them approximately 25 min. Because the effect appeared rather quickly and remained constant across time (see the Results section), Experiment 2 was shortened to 16 blocks of 20 trials, taking participants approximately 12 min.

After performing all blocks, participants completed a questionnaire assessing their experiences during the task (Table 1). They were asked to answer the questions as accurately and as detailed as possible. The questionnaire was similar to that used by Kool and colleagues (2010), but was changed to probe the experience of difficulty, and participants were instructed not to restrict themselves to very short or to yes/no answers, but to provide a concise and deliberate answer to each question. In the supplemental materials, we added four samples of filled-out questionnaires: two for which the metacognitive status of the participants was very clear, and two for which the voting of the raters was two to one.

Finally, objective prime visibility was assessed by a detection task that was identical to the task during the main experiment, except that participants were instructed to respond to the direction of the prime arrows instead of the target. During the detection task, targets were neutral (i.e., not pointing to the left or right) to ensure that participants were not accidentally responding to the target. It has been shown that these neutral targets provide a more sensitive test of prime visibility, compared with targets that are congruent or incongruent with the primes (Vermeiren & Cleeremans, 2012). Each participant first performed 50 detection trials in the upper box, followed by 50 detection trials in the lower box (counterbalanced across participants).

Statistical approach. Besides the *default* frequentist approach to analyze the data, the Bayes factor (BF) associated with an effect is reported whenever a null finding was obtained that was theoretically expected. Compared with classical p values, a BF has the

advantage that it can dissociate between data in favor of the null hypothesis ($BF < 1/3$), data that is not informative ($BF \approx 1$), and data in favor of the alternative hypothesis ($BF > 3$; based on Jeffreys' convention; Jeffreys, 1961). Thus, this allows determining whether nonsignificant effects simply reflect a lack of statistical power, or whether the data actually provide evidence for the null hypothesis. BFs were computed using the default priors in the BayesFactor package (Morey & Rouder, 2014).

Results

Metacognitive awareness. To assess whether participants were aware of the difference in difficulty between both boxes, the answers to the questionnaire were coded by three independent raters. All raters were blind to the results of the main experiment. The raters were informed that the task was a masked priming experiment, that there was a proportion congruency manipulation, and that participants were free to choose in which box they wanted to perform the task. Their task was simply to assess for each participant whether s/he subjectively experienced a difference in difficulty between performing the task in the upper versus the lower box. Among the raters, it was agreed to integrate responses to all questions in the assessment of metacognitive awareness. When participants were judged to genuinely have experienced a difference in difficulty between both boxes, they were coded as being *metacognitive aware*. When participants were judged not to show any indication of experiencing a difference in difficulty they were coded as *metacognitive unaware*.

In Experiment 1, 17 participants were classified as metacognitive aware and 15 as metacognitive unaware. The three coders unanimously agreed on the classification of 22 participants. Of the remaining 10, seven were classified as metacognitive aware. The interrater reliability (Fleiss' kappa for three raters) was .58. Interestingly, all but one metacognitive aware participants indicated the low-demand box to be the easiest (16 out of 17, $\chi^2(1, N = 17) = 13.23, p < .001$), thus displaying *accurate* metacognitive awareness. Several metacognitive unaware participants did not report any differences between the boxes (7 out of 15), and for those who did indicate that one of the boxes was easier than the other (e.g., based on color or location), half of them indicated the *high-demand* context to be easier (4 out of 8, $\chi^2(1, N = 8) = 0, p = 1$). After classification, we additionally explored whether certain reasons for choosing one box over the other emerged from the questionnaire. Participants who were classified as metacognitive aware responded that in one of both boxes the task simply felt easier ($n = 10$), they made less mistakes ($n = 4$), it was easier to respond ($n = 2$), or easier to keep focus ($n = 1$). Participants who were classified as metacognitive unaware responded that they liked one of both colors ($n = 7$), that it was easier to view the screen at one of both locations ($n = 1$), or that they didn't notice any difference at all ($n = 7$).

In Experiment 2, 14 participants were classified as metacognitive aware and 18 as metacognitive unaware. The three coders unanimously agreed on the classification of 24 participants. Of the remaining eight, three were classified as metacognitive aware. The interrater reliability (Fleiss' kappa for three raters) was .66, which shows that there was substantial agreement (Landis & Koch, 1977). Similar to Experiment 1, all metacognitive aware participants displayed accurate metacognitive awareness, identifying the

Table 1
Questionnaire Used to Assess Participants' Experience During the Main Experiment

1. What was it like to perform the task?
2. On which information did you base your choice?
3. Did you develop a preference for one of the two boxes?
4. Was there a difference between the two boxes?
5. Did you feel that the task was easier in one of the two boxes? If so, in which one?
6. If you answered positively to the previous question, was this something you were explicitly aware of during the experiment?

Note. Participants were asked to answer each question as elaborate as possible, not using very short or yes/no answers. The original questionnaire was written in Dutch.

low-demand context as easy (14 out of 14, $\chi^2(1, N = 14) = 14, p < .001$). Several metacognitive unaware participants did not report any differences in difficulty (10 out of 18), and for those who reported a difference in difficulty, about half of them indicated the high-demand context to be easier (5 out of 8, $\chi^2(1, N = 8) = 0.5, p = .50$). In this experiment, participants who were classified as metacognitive aware noted that in one of both boxes the task simply felt easier ($n = 10$), they made less mistakes ($n = 4$), or it was easier to concentrate and keep focus ($n = 2$). Participants who were classified as metacognitive unaware noted that they preferred one of both boxes because of its position on the screen ($n = 12$), or they didn't notice any difference at all ($n = 6$).

The selection of low-demand contexts. To examine the preference for low-demand contexts, it was computed how frequently participants selected the low-demand box when they had to decide whether to perform the next block in the upper or the lower box. In Figure 2, it can be seen that both in Experiment 1 (Figure 2A) and Experiment 2 (Figure 2C), only participants who noticed a difference in difficulty between both boxes (i.e., metacognitive aware participants) developed a clear preference for the low-demand context. In Experiment 1, all but one metacognitive aware participant preferred the low-demand context (see Figure 2B). In Experiment 2, this was the case for all metacognitive aware participants (Figure 2D). Contrary to this, the distribution of low-demand selections was more uniform for the metacognitive unaware participants in both experiments (see Figure 2B and 2D).

A univariate analysis of variance (ANOVA) predicting mean percentage of low-demand choices based on metacognitive aware-

ness (2 levels: metacognitive aware or unaware), showed a reliable main effect of metacognitive awareness, Experiment 1: $F(1, 30) = 10.97, p = .002, \eta_p^2 = .27$, Experiment 2: $F(1, 30) = 4.99, p = .033, \eta_p^2 = .14$. To interpret this main effect, we compared the mean percentage of low-demand choices to chance level selection (i.e., 50%), for both awareness groups separately. This analysis showed that in both experiments the metacognitive aware group strongly deviated from chance level selection (Experiment 1: on average, 77% low-demand selections; $t(16) = 5.70, p < .001, d = 1.38, BF = 767.47$; Experiment 2: on average, 81% low-demand selections; $t(13) = 9.74, p < .001, d = 2.60, BF = 61,378.08$) whereas the metacognitive unaware group did not (Experiment 1: on average 45% low-demand selections; $t(14) = -0.55, p = .59, d = 0.14, BF = 0.30$; Experiment 2: on average 60% low-demand selections; $t(17) = 1.33, p = .20, d = 0.31, BF = 0.50$).

Task performance. An important requirement is that behavioral performance on the main task (i.e., responding to the target arrow) is the same for both metacognitive awareness groups. We want both groups to differ in metacognitive awareness, but not in performance. Therefore, participants' performance on the main task (i.e., their responses to the target arrow) was analyzed using a 2 (demand context: high or low) \times 2 (congruency: congruent or incongruent) repeated measures ANOVA on the median RTs of correct trials and mean error rates, with metacognitive awareness (aware or unaware) as a between-subjects factor. Note that these results should be treated with caution because the number of trials for some cells are highly variable and fairly low for some subjects (i.e., the number of congruent trials in a high-demand block ranged

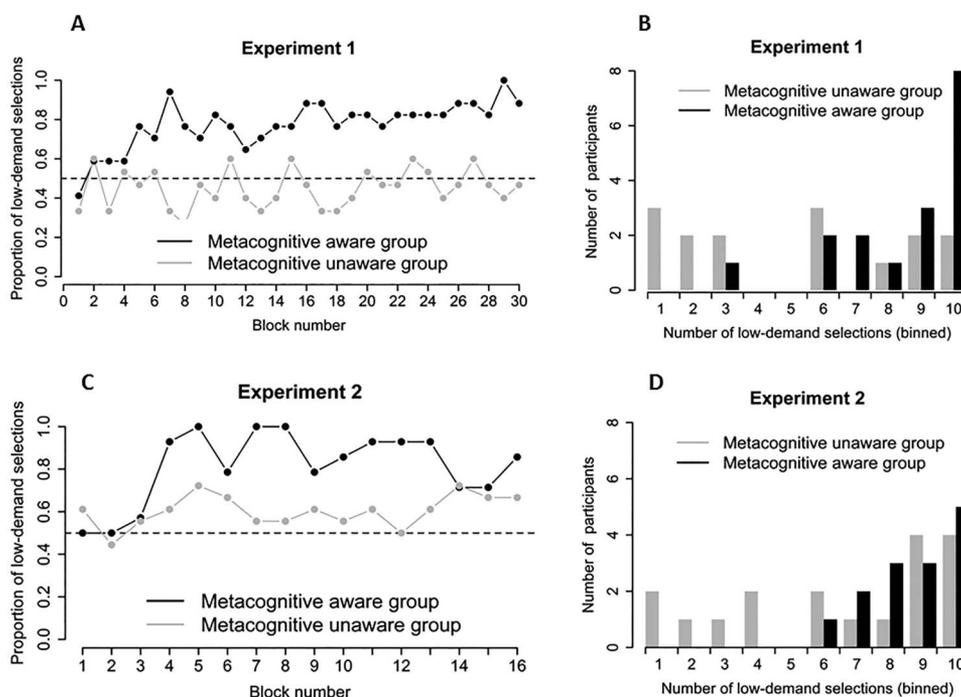


Figure 2. Results of Experiments 1 and 2. Figures on the left reflect the development of preferences for group averages for (A) Experiment 1 and (C) Experiment 2 across the blocks of trials. Figures on the right reflect mean preference on the individual level grouped in 10 bins, for (B) Experiment 1 and (D) Experiment 2. Bin 1 corresponds to the lowest proportion of low-demand choices (between 0% and 10%) and bin 10 the highest proportion (between 90% and 100%). Note that all participants performed the task at least once in each box.

between participants from four to 112 observations in Experiment 1 and from three to 60 in Experiment 2; the number of incongruent trials in the low-demand block ranged from five to 110 observations in Experiment 1 and from two to 59 in Experiment 2).

In Experiment 1, RTs only showed a main effect of congruency, $F(1, 30) = 250.23, p < .001, \eta_p^2 = .89$. Responses were faster on congruent (312 ms) compared with incongruent trials (382 ms). There were no other significant effects (all $ps > .28$). A similar effect of congruency was observed in the error rates, $F(1, 30) = 26.87, p < .001, \eta_p^2 = .47$. Fewer errors were made on congruent (2.7%) compared with incongruent trials (14.9%). There was also a main effect of demand context, $F(1, 30) = 10.41, p < .001, \eta_p^2 = .26$, reflecting the fact that more errors were committed in the low-demand context (10.3%) compared with the high-demand context (7.4%). Interestingly, the interaction between congruency and demand context reached significance, $F(1, 30) = 10.10, p < .001, \eta_p^2 = .25$. The congruency effect in the error rates was larger in the low-demand context (14.8%) compared with the high-demand context (9.6%). This interaction was not significantly modulated by the between-subjects factor awareness, $p > .09, \eta_p^2 = .09$. If anything, this three-way interaction showed that the modulation of the congruency effect by the demand context was larger for the metacognitive unaware group (a difference of 8.5%, $t(14) = 2.95, p = .01, d = 0.76$) than for the metacognitive aware group (a difference of 2.12%, $t(16) = 0.91, p = .38, d = 0.22$).

Highly similar results were obtained in Experiment 2. Responses were faster on congruent (316 ms) compared with incongruent trials (379 ms), $F(1, 30) = 114.71, p < .001, \eta_p^2 = .79$, and there were no other significant effects (all $ps > .23$). Error rates showed a main effect of congruency, $F(1, 30) = 50.71, p < .001, \eta_p^2 = .62$. Fewer errors were made on congruent (1.7%) compared with incongruent trials (15.7%). There was also a main effect of demand context, $F(1, 30) = 15.40, p < .001, \eta_p^2 = .34$, reflecting the fact that more errors were made in the low-demand context (11.2%) compared with the high-demand context (6.2%). The interaction between congruency and demand context also reached significance, $F(1, 30) = 11.34, p = .002, \eta_p^2 = .27$. The congruency effect in the error rates was larger in the low-demand context (18.4%) compared with the high-demand context (9.4%). This interaction was not significantly modulated by the between-subjects factor awareness, $p > .41, \eta_p^2 = .02$.

In sum, these results tentatively suggest that there was no evidence for differences in performance between the metacognitive aware and unaware group in Experiments 1 and 2.

Prime visibility. To assess prime visibility, signal detection theory was used to compute d' (Green & Swets, 1966). Left-pointing primes were treated as signal, right-pointing primes as noise. A left response to a left-pointing prime was considered a hit; a left response to a right-pointing prime was considered a false alarm. Hit and false alarm proportions were computed by dividing the total number of hits and false alarms by the number of signals. One false alarm proportion of 0 in Experiment 1 was adjusted to .05. In Experiment 1, mean d' was 0.17, which was slightly above chance level performance, $t(31) = 2.12, p = .042, d = 0.37, BF = 1.33$. However, and crucially, there was no significant difference in d' between both awareness groups, $F < 1, \eta_p^2 = .03, BF = 0.47$. In fact, d' was numerically higher for the metacognitive unaware group ($d' = 0.25$) compared with the metacognitive aware group

($d' = 0.10$), ruling out the possibility that metacognitive awareness was caused by prime visibility. In Experiment 2, d' was 0.11 on average, which was not significantly different from chance level performance, $t(31) = 1.59, p = .12, d = 0.28, BF = 0.59$. As expected, there was no significant difference in d' between the aware ($M = 0.16$) and the unaware ($M = 0.07$) group, $F < 1, \eta_p^2 = .01, BF = 0.40$.

Interim Discussion

Experiments 1 and 2 showed that only participants who subjectively experienced a difference in difficulty between performing a task in a high-demand and a low-demand context developed a preference for the low-demand context. However, some issues remain to be clarified before we can conclude on the role of metacognitive awareness in the selection of low-demand contexts. A first issue is our assumption that higher processing costs are induced in the high-demand context compared with the low-demand context. While these context-dependent processing costs should elicit different behavioral patterns for both demand contexts, they should be identical for both metacognitive awareness groups. Considering, for example, a hypothetical situation in which half of the participants do not show a difference in processing costs between the high- and low-demand context. Because of the absence of differing processing costs, these participants are unlikely to develop a specific preference or to report a difference in the experience of difficulty. To ensure that absence of a choice preference can be uniquely attributed to the absence of metacognitive experiences, we should therefore rule out the possibility that the two metacognitive awareness groups differ in performance. Although both Experiments 1 and 2 suggested that task performance was not modulated by metacognitive awareness, the free-choice nature of the design does not allow drawing strong conclusions from them. Participants with a strong preference for the low-demand context have performed most of the trials in this context, and barely any trials in the other context. These data are thus highly biased, and cannot be reliably interpreted. In Experiment 3, we therefore added baseline blocks at the beginning of the experiment that can provide an unbiased measure of performance to contrast both metacognitive awareness groups. A second issue is that differences in subjective experiences might be present, even in conditions with matched behavioral performance (e.g., aversiveness; McGuire & Botvinick, 2010). Therefore, two additional questions were added to the questionnaire to explicitly probe participants' capacity to be introspective regarding their performance. Following the questions assessing their metacognitive experiences, participants were also asked to indicate which context took them longest to perform, and which context they thought others would find more difficult. Chance level responses to these questions in the metacognitive unaware group, but above chance level responses in the metacognitive aware group, would further indicate that metacognitive awareness is the crucial factor differing between both groups. Third, a final issue is the possibility that metacognitive awareness is not an antecedent but rather a consequence of choice behavior. Under this scenario, preference is actually driven by performance, and metacognitive awareness results from participants observing their choice behavior. To exclude this possibility, the data of the baseline blocks can be used in a multiple regression analysis to examine whether metacognitive

awareness, but not performance, is the main factor driving choice behavior.

Experiment 3

Method

Participants. Thirty-five participants (6 males) participated in return for monetary compensation (7 €), and provided written informed consent. One participant was excluded because she switched to the wrong response buttons halfway through the task. Two more participants were excluded because they made more than 20% errors during the baseline blocks or during the main task. Mean age of the final sample ($n = 32$) was 20.9 years ($SD = 1.62$, range 19–25). All reported normal or corrected to normal vision and were naive with respect to the hypothesis.

Stimuli and apparatus. Stimuli and apparatus were identical to Experiment 2.

Procedure. The procedure was identical to Experiment 2, except for the following. The experiment started with three baseline blocks. In the first baseline block, participants performed 60 experimental trials in a medium-demand context (i.e., equal proportion of congruent and incongruent trials) in a box appearing in the middle of the screen. This block provides a measure of participants' sensitivity to congruency. After this first baseline block, half of the participants performed 150 trials in a low-demand context (i.e., 80% congruent trials) in a box appearing at the top or the bottom of the screen, followed by 150 trials in a high-demand context (i.e., 20% congruent trials) in the other box. For the other half of the participants, this order was reversed. After these baseline blocks, the free-choice part of the experiment started, in which participants could freely choose in which box they wanted to perform the task. Because the high- and low-demand baseline blocks were performed in the same boxes that also constituted the high- and low-demand blocks in the subsequent free-choice part, it could be expected that the metacognitive aware group had already developed a preference for the low-demand context right from the onset of the free-choice part. Therefore, and to keep the duration of the experiment the same as in Experiment 2, only 10 blocks of 20 trials were included in the free-choice part of this experiment. Together, the baseline blocks and the free-choice part took participants about 20 min to perform.

Finally, unlike Experiments 1 and 2, the questions in the post-experiment questionnaire were presented sequentially on the screen to prevent participants from reading ahead. Also, two additional questions were added at the end (translated from Dutch): "Which box took you longer to perform the task?" and "Which of both boxes do you think other participants found the easiest?"

Results

Metacognitive awareness. Fifteen participants were classified as metacognitive aware and 17 as metacognitive unaware. The same three coders as in Experiments 1 and 2 unanimously agreed on the classification of 24 participants. Of the remaining eight, four were classified as metacognitive aware. The interrater reliability (Fleiss' kappa for three raters) was .66, indicating substantial agreement among raters (Landis & Koch, 1977).¹ Similar to

the previous experiments, all metacognitive aware participants correctly identified the low-demand context as the easiest (18 out of 18, $\chi^2(1, N = 18) = 18, p < .001$), whereas metacognitive unaware participants either did not report any differences in difficulty (6 out of 17), or for those who did report a difference, about half of them indicated the high-demand context to be easier (6 out of 11, $\chi^2(1, N = 11) = 0.09, p = .76$). Participants who were classified as metacognitive aware responded that in one of both boxes the task simply felt more difficult ($n = 3$), they made less mistakes ($n = 7$), or it was easier to process the stimuli ($n = 5$). Participants who were classified as metacognitive unaware responded that it was easier to view the screen at one of both locations ($n = 15$), that they simply liked one box more ($n = 1$), or that they didn't notice any difference at all ($n = 1$).

When asked which box they thought took them longer to perform the task, most metacognitive aware participants guessed correctly (12/15, $\chi^2(1, N = 15) = 5.4, p = .020$), whereas metacognitive unaware participants guessed at chance level (8/15, $\chi^2(1, N = 15) = 0.07, p = .80$). Note that, despite being asked to guess, two metacognitive unaware participants nevertheless refused to guess, stating that there really was no difference. When asked which box they thought others found the easiest, most metacognitive aware participants guessed correctly (11/13, $\chi^2(1, N = 13) = 6.23, p = .012$), whereas metacognitive unaware participants guessed at chance level (7/13, $\chi^2(1, N = 13) = 0.08, p = .78$). Again, four metacognitive unaware and two metacognitive aware participants did not make a guess here, despite being explicitly asked to do so.

Baseline measures. To examine whether the difference in metacognitive experience between the two awareness groups was related to differences in behavioral performance, median RTs of correct trials and mean error rates of the baseline blocks were submitted to a 2 (congruency: congruent or incongruent) \times 3 (demand context: low, medium or high) repeated measures ANOVA.

RTs showed a main effect of congruency, $F(1, 30) = 406.55, p < .001, \eta_p^2 = .93$, which was significantly modulated by demand context, $F(2, 29) = 5.27, p = .011, \eta_p^2 = .27$. The congruency effect was 71 ms in the medium-demand context, becoming significantly smaller in the high-demand context, 55 ms, $t(31) = 4.03, p < .001, d = .90$, and only numerically larger in the low-demand context, 74 ms, $p > .47$. Congruency effects were 18 ms smaller in the high-demand context compared with the low-demand context, a difference which we will refer to as the *context-specific*

¹ In all three experiments, the three raters were blind to the data, but not to the hypothesis. Therefore, in Experiment 3, three new raters who were both blind to the data and the hypothesis also coded the data. They were told that participants performed a masked priming task with a congruency proportion manipulation, and asked to judge for each participant whether he or she noted that the task was more difficult to perform in one of both boxes. Their interrater reliability was comparable, Fleiss' kappa for three raters = .61, and 30 out of 32 participants were coded identically as done by the first group of raters. When using these metacognitive awareness labels to analyze the data, results remained similar, showing an effect of metacognitive awareness, $t(28.16) = 1.88, p_{\text{one-sided}} = .035$, indicating that the metacognitive aware participants developed a preference for the low-demand context, 69%, $t(14) = 3.32, p = .005$, whereas the metacognitive unaware participants did not develop a preference, 50%, $t(16) = 0.07, p > .90$

congruency effect, $t(31) = 4.31, p < .001, d = 0.76$. Crucially, the Congruency \times Demand Context interaction was not modulated by metacognitive awareness, $F < 1, \eta_p^2 < .01, BF = 0.15$, showing that both metacognitive awareness groups produced the same results (Figure 3). Finally, there was a marginally significant main effect of demand context, $F(2, 29) = 3.29, p = .051, \eta_p^2 = .18$, showing that responses were slightly faster in the low-demand context ($M = 341$ ms) than in the high-demand ($M = 348$ ms, $t(31) = -2.47, p = .019$) and the medium-demand context ($M = 350$ ms, $t(31) = -2.57, p = .015$). The latter two did not differ from each other, $p > .50$. All other effects did not reach significance, all $ps > .14$.

In the error rates, we also found a main effect of congruency, $F(1, 30) = 12.67, p = .001, \eta_p^2 = .30$, which was significantly modulated by demand context, $F(2, 29) = 8.96, p < .001, \eta_p^2 = .38$. The congruency effect was 8.1% in the medium-demand context, becoming significantly smaller in the high-demand context, 5.07%, $t(31) = 2.52, p = .017, d = 0.37$, and significantly larger in the low-demand context, 11.20%, $t(31) = -2.24, p = .032, d = 0.31$. As can be seen in Figure 3C and 3D, more errors

were committed on incongruent trials in the high-demand context, compared with the low-demand context. To explain this pattern, one should take into account that in the low-demand context only 20% of the trials are incongruent, and participants are hypothesized to adopt a liberal response criterion. As a result, many mistakes occur on the occasional incongruent trials. In the high-demand context, 80% of the trials are incongruent and participants are hypothesized to respond more conservatively, which explains why fewer errors were made on the frequently occurring incongruent trials. Confirming the pattern in RTs, we observed a highly significant context-specific congruency effect, 6.1%, $t(31) = 5.92, p < .001, d = 1.05$. Importantly, the Congruency \times Demand Context interaction was not modulated by metacognitive awareness, $F(2, 29) = 2.20, p = .13, \eta_p^2 = .13, BF = .032$, indicating that both metacognitive awareness groups produced the same results. Finally, there also was a significant main effect of demand context, $F(2, 29) = 8.96, p < .001, \eta_p^2 = .47$, reflecting that less errors were made in the high-demand context ($M = 2.85\%$) than in the medium-demand context ($M = 5.52\%$, $t(31) = -3.85, p < .001, d = 0.64$) and the low-demand context ($M = 6.84\%$,

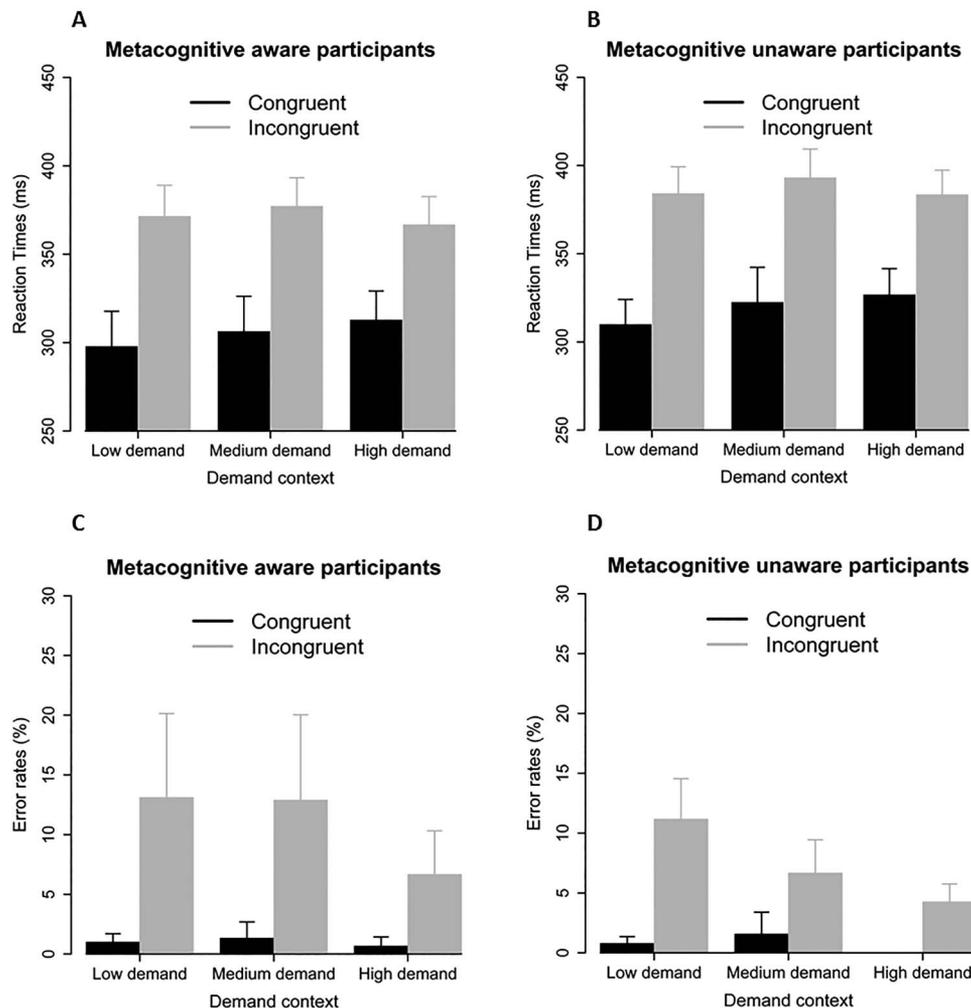


Figure 3. Reaction times and error rates of the baseline blocks of Experiment 3, separately for metacognitive aware (A and C) and unaware (B and D) participants. Error bars reflect 95% within-subjects confidence intervals.

$t(31) = -6.69, p < .001, d = 0.91$). The latter two did not differ from each other, $p > .13$. All other effects did not reach significance, all $ps > .18$.

The selection of low-demand contexts. In line with Experiments 1 and 2, Figure 4A shows that metacognitive awareness significantly influenced the selection of low-demand choices, $F(1, 30) = 5.80, p = .022, \eta_p^2 = .16$: only metacognitive aware participants showed a preference for the low-demand context (on average, 72% low-demand selections; $t(14) = 5.14, p < .001, d = 1.33, BF = 202.05$), whereas the metacognitive unaware group did not (on average, 48% low-demand selections; $t(16) = -0.21, p = .84, d = 0.05, BF = 0.25$). Figure 4B shows that all metacognitive aware participants selected the low-demand context more frequently, whereas this distribution was more uniform for metacognitive unaware participants.

The main purpose of Experiment 3 was to examine whether metacognitive awareness, but not performance, is the main factor driving preference. To investigate this, the mean-centered congruency effect and the context-specific congruency effect (in both RTs and errors) were extracted from the baseline trials and used as indices of performance. The congruency effect was used as a measure of overall sensitivity for response conflict, because participants with higher congruency effects are those who have the most to gain from reducing the cost of that experience. The context-specific congruency effect provides us with a good approximation of the difference in processing costs between the high- and low-demand contexts. A multiple regression analysis was performed with the mean percentage of low-demand choices as the dependent measure, and metacognitive awareness (2 levels: metacognitive aware or unaware) and the indices of performance and all possible interactions as regressors.

For RTs, results showed that only metacognitive awareness significantly predicted the proportion of low-demand choices, $t(24) = 2.54, p = .018, \eta_p^2 = 0.20$. None of the other variables or interactions reached significance, all $ps > .16$.

For the error rates, this analysis could not be performed because the congruency effect correlated positively with the context-specific congruency effect, $r = .43, t(30) = 2.64, p = .012$, causing problems with multicollinearity (i.e., variance inflation factor scores for their interaction and the three-way interaction were 31 and 33, respectively). Therefore, we added these different

measures to separate analyses. When the congruency effect was added to the regression, only a marginally significant main effect of metacognitive awareness was found, $t(28) = 1.98, p = .057, \eta_p^2 = 0.14$. Neither the congruency effect, nor its interaction with metacognitive awareness reached significance (both $ps > .68$). When adding the context-specific congruency effect in the regression, we observed both a main effect of metacognitive awareness, $t(28) = 2.61, p = .014, \eta_p^2 = 0.19$, a main effect of context-specific congruency effect, $t(28) = 2.65, p = .013, \eta_p^2 = 0.10$, and a marginally significant interaction between both, $t(28) = -2.00, p = .055, \eta_p^2 = 0.12$. This interaction reflected the fact that the size of the context-specific congruency effect was predictive of the proportion of low-demand choices in the metacognitive unaware group, $t(15) = 2.18, p = .046, \eta_p^2 = 0.24$, but not in the metacognitive aware group, $p = .62$. In sum, it can be concluded that there was only a minor hint for a role of performance in driving preferences, where metacognitive awareness consistently predicted the avoidance of cognitive demand.

Task performance in the free-choice part. For completeness, we also report the data of the free-choice part of the experiment. Note that four participants (one metacognitive aware) always performed the task in the same box, likely because they already developed a preference during the baseline blocks. The data of these participants are omitted here. RTs showed a main effect of congruency, $F(1, 25) = 187.27, p < .001, \eta_p^2 = .88$, reflecting faster responses on congruent (305 ms) than on incongruent trials (368 ms). Also the main effect of context was significant, $F(1, 25) = 5.24, p = .031, \eta_p^2 = .17$, showing slightly faster RTs in the low-demand (332 ms) compared with the high-demand (341 ms) context. There were no other significant effects (all $ps > .17$).

Error rates showed a main effect of congruency, $F(1, 25) = 14.48, p < .001, \eta_p^2 = .37$. Fewer errors were made on congruent (1.2%) compared with incongruent trials (10.3%). There were no other significant effects (all $ps > .10$).

Prime visibility. Our index of prime visibility (d') was 0.12, on average, which was not significantly different from chance level performance, $t(31) = 1.53, p = .13, d = 0.27, BF = 0.54$. As expected, there was no significant difference in d' between the aware ($M = 0.10$) and the unaware ($M = 0.13$) group, $F < 1, \eta_p^2 < .01, BF = 0.34$.

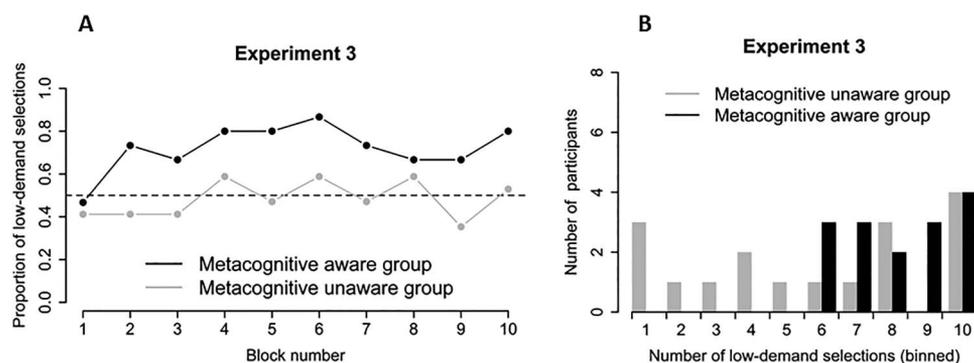


Figure 4. Main results of Experiment 3. (A) Development of preferences for group averages across blocks of trials. (B) Mean preference for the individual level.

Experiment 4

Given the importance of the questionnaire in classifying participants as either metacognitive aware or metacognitive unaware, it is important to demonstrate that it indeed measures metacognitive awareness. One way to test this is to compare the classification based on the questionnaire, to a forced-choice trial-by-trial measurement of metacognition. If the questionnaire is indeed sensitive to differences in metacognition, this should become evident in this forced-choice measure. Specifically, participants who are classified as metacognitive aware should be better at distinguishing difficult (i.e., incongruent) from easy (i.e., congruent) trials, compared with participants who are classified as metacognitive unaware. Therefore, in Experiment 4 participants provided trial-by-trial forced-choice ratings of their metacognitive awareness, and participants' ability to dissociate congruent from incongruent trials was compared with their classification using the questionnaire.

Method

Participants. Thirty participants (7 males) participated in return for course credit and provided written informed consent. Mean age of the sample was 19.4 years ($SD = 1.47$, range 18–25). All reported normal or corrected-to-normal vision and were naive with respect to the hypothesis.

Stimuli and apparatus. Stimuli and apparatus were identical to Experiments 2 and 3 except that responses were collected using a standard keyboard.

Procedure. Participants performed the same priming task that was used in Experiments 1–3, but now they additionally provided a metacognitive response on a trial-by-trial basis by rating how much difficulty they experienced while responding to the target arrow (e.g., Desender, Van Opstal, Hughes, & Van den Bussche, 2016). On each trial, participants were instructed to press “d” in response to a left-pointing target arrow and “k” in response to a right-pointing target arrow, with the middle finger of each hand. After this response, a blank screen was presented for 500 ms, followed by the question: “How much difficulty did you experience when responding to the arrow?” (translated from Dutch). They could answer either by pressing the “o” key with the ring finger of their right hand (“Rather more difficulty”) or by pressing the “m” key with the index finger of their right hand (“Rather less difficulty”). There was no time limit to answer this question. The intertrial interval was set to 800 ms.

The experiment started with 20 practice trials identical to the baseline block presented in Experiment 3 (i.e., without the metacognitive question). Afterward, participants performed 20 practice trials with the metacognition question. After these two practice phases, participants performed 20 trials in a box at the top of the screen, followed by 20 trials in a box at the bottom of the screen (reversed for half of the participants). This was repeated 10 times (400 trials in total). Boxes in this experiment were identical to those used in Experiments 2 and 3 (see Figure 1). Similar to the previous experiments, one of the boxes constituted a high-demand context (i.e., 80% congruent trials) while the other box constituted a low-demand context (i.e., 20% congruent trials; counterbalanced). Contrary to Experiments 1–3, there was no free-choice part; participants simply alternated between the boxes.

After the experiment, participants answered the metacognitive questionnaire. Note that Question 2 was omitted, because there

was no free-choice part in this experiment. Finally, participants took part in a detection task, identical to Experiments 2 and 3.

Results

Metacognitive awareness. Seventeen participants were classified as metacognitive aware and 13 as metacognitive unaware. The same three coders as in Experiments 1–3 unanimously agreed on the classification of 26 participants. Of the remaining four, one was classified as metacognitive aware. The interrater reliability (Fleiss' kappa for three raters) was .816, indicating almost perfect agreement among raters (Landis & Koch, 1977).

In contrast to the previous experiments, however, there were only 10 metacognitive aware participants that correctly identified the low-demand context as the easiest, 10 out of 17, $\chi^2(1, N = 17) = 0.53, p = .467$. Thus, seven participants who were classified as metacognitive aware indicated that they felt the high-demand box was the easiest. Because these participants clearly did not have accurate metacognitive awareness, they were reclassified as metacognitive unaware. This left a total of 20 metacognitive unaware participants, from which seven did not report a difference in difficulty, and for those who did report a difference, about half then indicated the high-demand context to be easier, eight out of 13, $\chi^2(1, N = 13) = 0.69, p = .405$.

When asked which box they thought took them longer to perform the task, most metacognitive aware participants guessed correctly, nine out of 10, $\chi^2(1, N = 10) = 6.4, p = .011$, whereas metacognitive unaware participants guessed around chance level, 11 out of 17, $\chi^2(1, N = 17) = 1.5, p = .225$. Note that, despite being asked to guess, three metacognitive unaware participants nevertheless refused to guess, stating that there really was no difference. When asked which box they thought others found the easiest, most metacognitive aware participants guessed correctly, seven out of eight, $\chi^2(1, N = 8) = 4.5, p = .033$, whereas metacognitive unaware participants guessed at chance level, eight out of 16, $\chi^2(1, N = 16) = 0, p = 1$. Again, four metacognitive unaware and two metacognitive aware participants did not make a guess here, despite being explicitly asked to do so.

Trial-by-trial metacognitive awareness. Trial-by-trial judgments of difficulty were analyzed using signal detection theory (Green & Swets, 1966). This method can separate sensitivity from response bias. Incongruent trials were treated as signal, congruent trials as noise. A “difficult” judgment on an incongruent trial was considered a hit; a “difficult” judgment on a congruent trial was considered a false alarm. We calculated d' (sensitivity) by subtracting the standardized proportion of false alarms from the standardized proportion of hits. In order not to confuse this index with our measure of prime visibility, we will refer to it as conflict- d' , a measure of how good participants' judgments track the actual congruency. Only trials where the response to the target arrow was correct were included in this calculation, in order not to conflate this metacognitive awareness of difficulty with error awareness. Note that in previous reports (e.g., Desender et al., 2014; Reuss, Desender, Kiesel, & Kunde, 2014), participants were asked to indicate on each trial whether they thought there was a conflict between prime and target. Although this question closely resembles the question used in the current experiment, it remains to be seen whether or not participants introspect on the same experience when asked slightly different questions. More importantly for the

present purpose, however, is that our current question measures participants' metacognitive awareness of difficulty, which is expected to be related to the metacognitive awareness of a difference in difficulty. Finally, by adding the proportion of hits and the proportion of false alarms, multiplied by -0.5 , we computed c' (response bias), which quantifies the bias toward "easy" responses.

Both metacognitive awareness groups displayed above chance level trial-by-trial metacognitive awareness (metacognitive aware: conflict- $d' = 0.66$, $t(9) = 2.65$, $p = .026$; metacognitive unaware, conflict- $d' = 0.26$, $t(19) = 4.43$, $p < .001$). Interestingly, as predicted, conflict- d' values were higher, on average, in the metacognitive aware group than in the metacognitive unaware group, $F(1, 28) = 4.33$, $p = .047$. Thus, participants who noted a difference in difficulty between both contexts (classified using the questionnaire), were also better in noting differences between congruent and incongruent trials on the trial level, compared with participants who were classified as metacognitive unaware. There were no differences in response bias between participants classified as metacognitive aware ($M = 0.67$) and metacognitive unaware ($M = 0.78$), $F < 1$.

Confirming that the seven participants initially classified as metacognitive aware indeed display inaccurate metacognitive awareness, they also display poor trial-by-trial metacognitive awareness, conflict- $d' = .014$, $t(6) = 4.73$, $p < .001$. However, because it is debatable whether they are indeed metacognitive unaware, we additionally report the results without these seven participants. Numerically, the results remain highly similar (metacognitive aware: conflict- $d' = 0.66$, $t(9) = 2.65$, $p = .026$; metacognitive unaware, conflict- $d' = 0.32$, $t(12) = 3.80$, $p = .002$), however, due to the drop in power the difference between both groups no longer reaches statistical significance, $F(1, 21) = 2.01$, $p = .171$.

Task performance. RTs showed a main effect of congruency, $F(1, 28) = 57.82$, $p < .001$, $\eta_p^2 = .67$, reflecting faster responses on congruent (407 ms) than on incongruent trials (464 ms). There were no other significant effects (all $ps > .097$).

Error rates showed a main effect of congruency, $F(1, 28) = 13.17$, $p = .001$, $\eta_p^2 = .32$. Fewer errors were made on congruent (1.9%) compared with incongruent trials (8.9%). There was a significant main effect of context, $F(1, 28) = 4.73$, $p = .038$, $\eta_p^2 = .14$, reflecting higher error rates in the mainly congruent context (6.2%) compared with the mainly incongruent context (4.6%). There was an interaction between congruency and context that just failed to reach significance, $F(1, 28) = 3.69$, $p = .065$, $\eta_p^2 = .12$. The congruency effect in the error rates was larger in the low-demand context (8.3%) compared with the high-demand context (5.6%). There were no other significant effects, all $ps > .466$.

Prime visibility. Our index of prime visibility (d') was 0.07 on average, which was not significantly different from chance level performance, $t(29) = 1.19$, $p = .24$, $d = 0.21$, $BF = 0.37$. As expected, there was no significant difference in d' between the aware ($M = 0.06$) and the unaware ($M = 0.07$) group, $F < 1$, $\eta_p^2 < .01$, $BF = 0.36$.

Discussion

This study reports three experiments in which participants could freely choose to perform a task in either of two boxes. One of the boxes created a low-demand context (i.e., low proportion of con-

flict trials) and the other a high-demand context (i.e., high proportion of conflict trials). Because we used subliminal priming to create these differences in response conflict, task performance and perceived difficulty are influenced without participants being aware of the visual stimuli driving these changes. In all three experiments, only participants who had the subjective experience that the task was more difficult in one of the two boxes developed a clear preference for the low-demand context, whereas participants who did not have this experience did not. Crucially, baseline blocks in Experiment 3 showed that both metacognitive awareness groups had identical task performance, while the avoidance of cognitive demand was again observed only for metacognitive aware participants. Finally, a multiple regression approach confirmed that awareness was the main factor driving the preference for low-demand contexts. We conclude that the avoidance of cognitive demand depends on the metacognitive appreciation of the difference in processing costs.

Avoiding the Conflict

Conflicts in information processing can interfere with ongoing actions, and call for remedial measures to be taken in order to adapt to this conflict. Depending on the situation, two viable strategies are to increase cognitive control and deal with the conflict, or to avoid the conflicting situation (Botvinick, 2007). The mechanisms behind the strategic increase of control in response to conflict have been studied extensively during the last decade (for reviews, see Bugg & Crump, 2012; Egner, 2007). However, only recently the avoidance of cognitive conflict has become a topic of empirical research (Kool et al., 2010; Schouppe, Demanet, Boehler, Ridderinkhof, & Notebaert, 2014). Therefore, the exact cognitive mechanisms behind this process remain unclear. Here, in three studies, we found that only participants who have the subjective experience that the task was more difficult in the high-demand context prefer the low-demand context. Although in Experiment 1 participants received explicit instructions about the difference in demand between contexts, similar results were obtained in Experiments 2 and 3 where this instruction was not provided. Combining the data of all three studies to predict low-demand choices showed a clear effect of metacognitive awareness, $F(1, 92) = 20.14$, $p < .001$, $\eta_p^2 = 0.18$, $BF = 1,044.46$, but no main effect of Experiment, $F < 1$, $\eta_p^2 < .01$, $BF = 0.23$, nor an interaction between both, $F < 1$, $\eta_p^2 < .01$, $BF = 0.32$. This additional analysis shows that choice behavior depends on metacognitive awareness, and not on the exact instructions that were provided. Overall, our results suggest that the metacognitive appreciation of the difference in difficulty acts as a cue for participants to decide that it is not worth to invest more effort in a task and instead switch to another, less demanding, task context when available.

The most crucial findings were those of Experiment 3. There, it was observed that for participants with similar task performance, metacognition was the decisive feature to develop a preference for the easiest context. It might seem odd that groups with identical behavioral performance can differ in their metacognitive experience of the difference in difficulty. However, work in the field of metacognition has a rich tradition in explaining subjective experiences as reflecting a *noisy* read-out of performance. As such, even with matched performance, substantial individual variability

can exist in the capacity to introspect on this performance (Fleming, Weil, Nagy, Dolan, & Rees, 2010), a dissociation which can also be induced using noninvasive brain stimulation (Rounis, Maniscalco, Rothwell, Passingham, & Lau, 2010). In line with these observations, the current findings form another prime example showing how groups with similar task performance can still show differences in metacognitive experiences. The findings of Experiment 3 are difficult to explain within a framework in which processing costs directly influence the allocation of control. Without taking metacognition into account, it is hard to grasp how participants with identical task performance can produce such different patterns of demand avoidance. However, this result can elegantly be explained by assuming that not the actual difference in demand between both contexts but rather the metacognitive appreciation of this difference acts as a cue for the allocation of control. This latter proposal is in line with recent theoretical work arguing that cost/benefit computations give rise to subjective experiences of effort, which, in turn, influence the allocation of future effort (Kurzman et al., 2013). So far, several comprehensive accounts of cognitive control have already been proposed, which motivated a large and diverse body of research (Alexander & Brown, 2011; Botvinick, 2007; Shenhav, Botvinick, & Cohen, 2013). However, these different accounts mostly remain silent on the role of metacognition in the avoidance of cognitive conflict. Therefore, the challenge for future models of cognition will be to design a framework which captures these dynamic influences of both objective performance and subjective experiences on behavior.

The Multifaceted Nature of Demand Selection

It is important to stress that choice behavior in a demand selection task, and cognitive control more generally, is likely to be multifaceted in nature (e.g., Gold et al., 2015; McGuire & Botvinick, 2010). Preferences are determined by various factors, only one of them being the metacognitive experience of difficulty. Using a multiple regression approach, we were able to exclude the possibility that performance was the actual factor driving preferences for the low-demand context. However, this analysis does not directly prove that metacognitive awareness is the underlying cause. In theory, it could still be that another, currently unknown, factor drives both preferences and metacognitive awareness. To maximally exclude this possibility, in the experiments reported here, special care was taken to minimize visual differences between the low- and high-demand contexts (i.e., primes were subliminal, boxes had the same color in Experiments 2–4, etc.), in order to assure that the only difference was the subjective experience of difficulty. Furthermore, although proponents of implicit learning effects might argue that the metacognitive unaware participants might develop a preference when another experimental design is used, the spatial nature of our task is an excellent candidate for implicit learning effects to take place (Beck, Angelone, Peterson, & Varakin, 2008). The observation that these participants nevertheless did not, on average, develop any preference, strengthens our conclusion that metacognitive awareness is crucial for the avoidance of high-demand.

However, in many other situations, contexts or objects typically differ in various ways, and any of these differences might act as a cue determining preference. The importance of this point becomes

particularly clear when considering the results of the metacognitive unaware group. In all three experiments, the majority of participants showed a clear preference for one of both contexts. For the metacognitive aware participants this was based on a genuine experience of difficulty, whereas for the metacognitive unaware participants this was probably based on external variables unrelated to cognitive demand, such as color or location. Metacognitive unaware participants did not necessarily have bad metacognitive abilities, but might simply have used different cues upon which to base their choices. Therefore, one should be cautious not to interpret the data of participants selecting the high-demand option more often as a reflection of a general preference for a cognitive challenge (see, e.g., Experiment 3 of Kool et al., 2010).

The importance of carefully minimizing visual differences between the low- and high-demand contexts, as was done in the current report, becomes clear when considering studies that do not make these efforts. For example, a recent report documented that patients with schizophrenia show reduced avoidance of cognitive demand compared with matched healthy controls, a finding which the authors attribute to a reduced ability in patients to detect the effort demands associated with each context (Gold et al., 2015). Because the degree of demand avoidance in this study was also positively related to IQ scores, this finding was attributed to individual differences in general intellectual ability. Interestingly, these findings were only obtained when *explicitly* informing participants about the crucial task-switching manipulation that was used to create differences in cognitive demand. When this information was not provided, neither patients nor controls showed demand avoidance. This finding contrasts with the current observation that low-demand preferences were identical regardless of whether explicit instructions were provided (Experiment 1) or not (Experiments 2 and 3). An important difference with the current report is that Gold et al. (2015) used a task-switching manipulation to create different levels of cognitive demand and all stimuli were always clearly presented on the screen. Therefore, differences in demand context could, in principle, also be noted by mentally keeping track of the number of task switches. Indeed, this was also what Gold et al. (2015) proposed. Contrary to this, when carefully matching both demand contexts for visual differences, metacognition might be the only possible cue informative about cognitive demand, and therefore found to play an important role in avoiding high-demand contexts. An alternative possibility is, of course, that metacognition is related to intellectual ability. However, at present, this hypothesis is not supported by the literature (Fleming, Huijgen, & Dolan, 2012; Weil et al., 2013). Future studies will have to shed light on this prediction that different mechanisms underlie the avoidance of cognitive demand depending on the information available to the participants.

In sum, we carefully conclude here that in the absence of major visual differences, the metacognitive appreciation of response conflict is the main cue for people to base their choices on. Whether this also holds in more complex forms of decision making, remains to be answered in future studies.

Investing Control Versus Avoiding Control

In line with the current results, we recently also showed that deciding to invest more effort in a high-demand task (rather than avoiding it) likewise seems to depend on metacognitive experi-

ences (Desender et al., 2014). In that study, the strategic processing of conflicting information after a conflict trial (i.e., conflict adaptation; Gratton, Coles, & Donchin, 1992) was present only after trials on which participants had the subjective experience of difficulty. When the subjective experience of difficulty did not coincide with actual conflict, an adaptation effect in the error rates was observed after the experience of conflict, but not after actual response conflict. Together with the present results, this suggests that the subjective experience of task difficulty is crucial to evaluate the cost of performing that task. Without subjectively experiencing this cost, the benefit of exerting more cognitive control or switching to another task cannot appropriately be appraised. Given this proposed link between metacognition and cognitive control, one might wonder why task performance did not differ between both metacognitive awareness groups in the baseline blocks of Experiment 3. This question is particularly important, given that in Experiment 4 it was shown that there is a relationship between metacognitive awareness about the difference in difficulty between both contexts and metacognition measured on a trial-by-trial basis. More specifically, participants who were classified as metacognitive aware (i.e., based on the questionnaire) were also better at distinguishing congruent (i.e., easy) from incongruent (i.e., difficult) trials, when judging the difficulty of a trial, compared with participants classified as metacognitive unaware.

To explain why both metacognitive awareness groups in Experiment 3 nevertheless displayed similar behavioral performance on the priming task, we note that it is still highly debated whether the difference in performance between blocks with mainly congruent and mainly incongruent trials actually reflects an expression of cognitive control (Schmidt & Besner, 2008). A straightforward interpretation of our data would be that the context-specific congruency effect is the result of participants automatically changing their response threshold dependent on the task difficulty (Kinoshita, Forster, & Mozer, 2008), and some of them become aware of the difference in task difficulty and others do not. Given the empirical dissociation between trial-by-trial adaptation effects and block-wise adaptation effects (Funes, Lupiáñez, & Humphreys, 2010), future research is needed to shed light on the necessary involvement of metacognition in each of these processes.

Conclusion

In three experiments, we provided evidence that the subjective experience that a task is difficult to perform in a high-demand context plays a role in shaping the preference for a low-demand context. Thus, we conclude that the avoidance of cognitive demand depends on the metacognitive appreciation of the difference in cognitive demand.

References

- Alexander, W. H., & Brown, J. W. (2011). Medial prefrontal cortex as an action-outcome predictor. *Nature Neuroscience*, *14*, 1338–1344. <http://dx.doi.org/10.1038/nn.2921>
- Beck, M. R., Angelone, B., Peterson, M., & Varakin, D. A. (2008). Implicit learning for probable changes in a visual change detection task. *Consciousness and Cognition*, *17*, 1192–1201. <http://doi.org/10.1016/j.concog.2008.06.011>
- Blais, C., Harris, M. B., Guerrero, J. V., & Bunge, S. A. (2012). Rethinking the role of automaticity in cognitive control. *The Quarterly Journal of Experimental Psychology*, *65*, 268–276. <http://dx.doi.org/10.1080/17470211003775234>
- Botvinick, M. M. (2007). Conflict monitoring and decision making: Reconciling two perspectives on anterior cingulate function. *Cognitive, Affective & Behavioral Neuroscience*, *7*, 356–366. <http://dx.doi.org/10.3758/CABN.7.4.356>
- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*, 624–652. <http://dx.doi.org/10.1037/0033-295X.108.3.624>
- Bugg, J. M., & Crump, M. J. C. (2012). In support of a distinction between voluntary and stimulus-driven control: a review of the literature on proportion congruent effects. *Frontiers in Psychology*, *3*, 367.
- Chambon, V., & Haggard, P. (2012). Sense of control depends on fluency of action selection, not motor performance. *Cognition*, *125*, 441–451. <http://dx.doi.org/10.1016/j.cognition.2012.07.011>
- Desender, K., Van Opstal, F., Hughes, G., & Van den Bussche, E. (2016). The temporal dynamics of metacognition: Dissociating task-related activity from later metacognitive processes. *Neuropsychologia*, *82*, 54–64. <http://dx.doi.org/10.1016/j.neuropsychologia.2016.01.003>
- Desender, K., Van Opstal, F., & Van den Bussche, E. (2014). Feeling the conflict: The crucial role of conflict experience in adaptation. *Psychological Science*, *6*, 375–383. <http://dx.doi.org/10.1177/0956797613511468>
- Egner, T. (2007). Congruency sequence effects and cognitive control. *Cognitive, Affective & Behavioral Neuroscience*, *7*, 380–390. <http://dx.doi.org/10.3758/CABN.7.4.380>
- Fleming, S. M., Huijgen, J., & Dolan, R. J. (2012). Prefrontal contributions to metacognition in perceptual decision making. *The Journal of Neuroscience*, *32*, 6117–6125. <http://dx.doi.org/10.1523/JNEUROSCI.6489-11.2012>
- Fleming, S. M., Weil, R. S., Nagy, Z., Dolan, R. J., & Rees, G. (2010). Relating introspective accuracy to individual differences in brain structure. *Science*, *329*, 1541–1543. <http://dx.doi.org/10.1126/science.1191883>
- Funes, M. J., Lupiáñez, J., & Humphreys, G. (2010). Sustained vs. transient cognitive control: Evidence of a behavioral dissociation. *Cognition*, *114*, 338–347. <http://dx.doi.org/10.1016/j.cognition.2009.10.007>
- Gold, J. M., Kool, W., Botvinick, M. M., Hubzin, L., August, S., & Waltz, J. A. (2015). Cognitive effort avoidance and detection in people with schizophrenia. *Cognitive, Affective & Behavioral Neuroscience*, *15*, 145–154. <http://dx.doi.org/10.3758/s13415-014-0308-5>
- Gotler, A., Meiran, N., & Tzelgov, J. (2003). Nonintentional task set activation: Evidence from implicit task sequence learning. *Psychonomic Bulletin & Review*, *10*, 890–896. <http://dx.doi.org/10.3758/BF03196549>
- Gratton, G., Coles, M. G. H., & Donchin, E. (1992). Optimizing the use of information: Strategic control of activation of responses. *Journal of Experimental Psychology: General*, *121*, 480–506. <http://dx.doi.org/10.1037/0096-3445.121.4.480>
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York, NY: Wiley.
- Hassin, R. R. (2013). Yes it can: On the functional abilities of the human unconscious. *Perspectives on Psychological Science*, *8*, 195–207. <http://dx.doi.org/10.1177/1745691612460684>
- Hockey, G. R. J. (2011). A motivational control theory of cognitive fatigue. In P. L. Ackerman (Ed.), *Cognitive fatigue: Multidisciplinary perspectives on current research and future applications* (pp. 167–187). Washington, DC: American Psychological Association. <http://dx.doi.org/10.1037/12343-008>
- Hull, C. L. (1943). *Principles of behavior*. New York, NY: Appleton-Century.
- Hurlburt, R. T., & Heavey, C. L. (2001). Telling what we know: Describing inner experience. *Trends in Cognitive Sciences*, *5*, 400–403. [http://dx.doi.org/10.1016/S1364-6613\(00\)01724-1](http://dx.doi.org/10.1016/S1364-6613(00)01724-1)

- Inzlicht, M., & Schmeichel, B. J. (2012). What is ego depletion? Toward a mechanistic revision of the resource model of self-control. *Perspectives on Psychological Science*, 7, 450–463. <http://dx.doi.org/10.1177/1745691612454134>
- Jeffreys, H. (1961). *The theory of probability* (3rd ed.). New York, NY: Oxford University Press.
- Kahneman, D. (1973). *Attention and effort*. Upper Saddle River, NJ: Prentice Hall.
- Kinoshita, S., Forster, K. I., & Mozer, M. C. (2008). Unconscious cognition isn't that smart: Modulation of masked repetition priming effect in the word naming task. *Cognition*, 107, 623–649. <http://dx.doi.org/10.1016/j.cognition.2007.11.011>
- Klotz, W., & Neumann, O. (1999). Motor activation without conscious discrimination in metacontrast masking. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 976–992. <http://dx.doi.org/10.1037/0096-1523.25.4.976>
- Koch, I. (2001). Automatic and intentional activation of task sets. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 1474–1486. <http://dx.doi.org/10.1037/0278-7393.27.6.1474>
- Kool, W., McGuire, J. T., Rosen, Z. B., & Botvinick, M. M. (2010). Decision making and the avoidance of cognitive demand. *Journal of Experimental Psychology: General*, 139, 665–682. <http://dx.doi.org/10.1037/a0020198>
- Kurzban, R., Duckworth, A., Kable, J. W., & Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *Behavioral and Brain Sciences*, 36, 661–679. <http://dx.doi.org/10.1017/S0140525X12003196>
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159–174. <http://dx.doi.org/10.2307/2529310>
- Lau, H. C., & Passingham, R. E. (2007). Unconscious activation of the cognitive control system in the human prefrontal cortex. *The Journal of Neuroscience*, 27, 5805–5811. <http://dx.doi.org/10.1523/JNEUROSCI.4335-06.2007>
- Mattler, U. (2003). Priming of mental operations by masked stimuli. *Perception & Psychophysics*, 65, 167–187. <http://dx.doi.org/10.3758/BF03194793>
- McGuire, J. T., & Botvinick, M. M. (2010). Prefrontal cortex, cognitive control, and the registration of decision costs. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, 107, 7922–7926. <http://dx.doi.org/10.1073/pnas.0910662107>
- Metcalfe, J., & Finn, B. (2008). Evidence that judgments of learning are causally related to study choice. *Psychonomic Bulletin & Review*, 15, 174–179. <http://dx.doi.org/10.3758/PBR.15.1.174>
- Morey, R. D., & Rouder, J. N. (2014). BayesFactor: Computation of Bayes factors for common designs (R package version 0.9.11-1) [Computer software]. Retrieved from <http://CRAN.R-project.org/package=BayesFactor>
- Morsella, E., Gray, J. R., Krieger, S. C., & Bargh, J. A. (2009). The essence of conscious conflict: Subjective effects of sustaining incompatible intentions. *Emotion*, 9, 717–728. <http://dx.doi.org/10.1037/a0017121>
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, 231–259. <http://dx.doi.org/10.1037/0033-295X.84.3.231>
- Reder, L. M. (1988). Strategic control of retrieval strategies. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22, pp. 227–259). New York, NY: Academic Press. [http://dx.doi.org/10.1016/S0079-7421\(08\)60042-0](http://dx.doi.org/10.1016/S0079-7421(08)60042-0)
- Reder, L. M., & Schunn, C. D. (1996). Metacognition does not imply awareness: Strategy Choice is governed by implicit learning and memory. In L. M. Reder (Ed.), *Implicit memory and metacognition* (pp. 45–77). Mahwah, NJ: Erlbaum.
- Reuss, H., Desender, K., Kiesel, A., & Kunde, W. (2014). Unconscious conflicts in unconscious contexts: The role of awareness and timing in flexible conflict adaptation. *Journal of Experimental Psychology: General*, 143, 1701–1718. <http://dx.doi.org/10.1037/a0036437>
- Reuss, H., Kiesel, A., Kunde, W., & Hommel, B. (2011). Unconscious activation of task sets. *Consciousness and Cognition: An International Journal*, 20, 556–567. <http://dx.doi.org/10.1016/j.concog.2011.02.014>
- Rounis, E., Maniscalco, B., Rothwell, J. C., Passingham, R. E., & Lau, H. (2010). Theta-burst transcranial magnetic stimulation to the prefrontal cortex impairs metacognitive visual awareness. *Cognitive Neuroscience*, 1, 165–175. <http://dx.doi.org/10.1080/17588921003632529>
- Schmidt, J. R., & Besner, D. (2008). The Stroop effect: Why proportion congruent has nothing to do with congruency and everything to do with contingency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34, 514–523. <http://dx.doi.org/10.1037/0278-7393.34.3.514>
- Schmidt, J. R., Crump, M. J. C., Cheesman, J., & Besner, D. (2007). Contingency learning without awareness: Evidence for implicit control. *Consciousness and Cognition: An International Journal*, 16, 421–435. <http://dx.doi.org/10.1016/j.concog.2006.06.010>
- Schouppe, N., Demanet, J., Boehler, C. N., Ridderinkhof, K. R., & Notebaert, W. (2014). The role of the striatum in effort-based decision-making in the absence of reward. *The Journal of Neuroscience*, 34, 2148–2154. <http://dx.doi.org/10.1523/JNEUROSCI.1214-13.2014>
- Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The expected value of control: An integrative theory of anterior cingulate cortex function. *Neuron*, 79, 217–240. <http://dx.doi.org/10.1016/j.neuron.2013.07.007>
- Song, H., & Schwarz, N. (2008). If it's hard to read, it's hard to do: Processing fluency affects effort prediction and motivation. *Psychological Science*, 19, 986–988. <http://dx.doi.org/10.1111/j.1467-9280.2008.02189.x>
- Song, H., & Schwarz, N. (2009). If it's difficult to pronounce, it must be risky. *Psychological Science*, 20, 135–138. <http://dx.doi.org/10.1111/j.1467-9280.2009.02267.x>
- Thompson, V., Evans, J. S. B. T., & Campbell, J. I. D. (2013). Matching bias on the selection task: It's fast and feels good. *Thinking & Reasoning*, 19(3–4), 431–452. <http://dx.doi.org/10.1080/13546783.2013.820220>
- Thompson, V. A., Turner, J. A., Pennycook, G., Ball, L. J., Brack, H., Ophir, Y., & Ackerman, R. (2013). The role of answer fluency and perceptual fluency as metacognitive cues for initiating analytic thinking. *Cognition*, 128, 237–251. <http://dx.doi.org/10.1016/j.cognition.2012.09.012>
- Verguts, T., & Notebaert, W. (2009). Adaptation by binding: A learning account of cognitive control. *Trends in Cognitive Sciences*, 13, 252–257. <http://dx.doi.org/10.1016/j.tics.2009.02.007>
- Vermeiren, A., & Cleeremans, A. (2012). The validity of d' measures. *PLoS ONE*, 7, e31595. <http://dx.doi.org/10.1371/journal.pone.0031595>
- Weil, L. G., Fleming, S. M., Dumontheil, I., Kilford, E. J., Weil, R. S., Rees, G., . . . Blakemore, S.-J. (2013). The development of metacognitive ability in adolescence. *Consciousness and Cognition: An International Journal*, 22, 264–271. <http://dx.doi.org/10.1016/j.concog.2013.01.004>
- Wenke, D., Fleming, S. M., & Haggard, P. (2010). Subliminal priming of actions influences sense of control over effects of action. *Cognition*, 115, 26–38. <http://dx.doi.org/10.1016/j.cognition.2009.10.016>
- Westbrook, A., Kester, D., & Braver, T. S. (2013). What is the subjective cost of cognitive effort? Load, trait, and aging effects revealed by economic preference. *PLoS ONE*, 8, e68210. <http://dx.doi.org/10.1371/journal.pone.0068210>

Received February 12, 2016

Revision received December 19, 2016

Accepted January 2, 2017 ■