Consciousness and Cognition 143 (2026) 104095

Contents lists available at ScienceDirect

Consciousness and Cognition

journal homepage: www.elsevier.com/locate/yccog

Full Length Article

Introspection without execution: Evidence for introspective switch costs in NoGo trials

Jonathan Mendl [*], Gesine Dreisbach

University of Regensburg, Department of Psychology, Universitätsstraße 31, 93053 Regensburg, Germany

ARTICLE INFO

Keywords:
Introspection
Task switching
NoGo
Theory-based monitoring

ABSTRACT

In the task-switching paradigm, the most prominent measure are the switch costs indicating worse performance on task switches compared to task repetitions. Previous research showed that introspective RT estimations capture the switch costs. However, the meaning of these introspective switch costs is not yet well understood. They may simply reflect the immediate experience of an executed response or rely on abstract theory-based knowledge about task transitions ("task switches are difficult and take longer"). The present study investigated how introspective judgments are formed during task switching. In Experiment 1, we tested participants' awareness of task transitions as a prerequisite of theory-based knowledge. On a random subset of trials, participants had to indicate whether the current task repeated or switched relative to the previous trial. Judgments were collected with a short (1000 ms) and a long (5000 ms) time window to distinguish intuitive from deliberate decisions. As predicted, participants reliably recognized the transition even with a short time window. In Experiment 2, we reduced immediate experience-based knowledge by introducing NoGo trials where participants had to refrain from responding. Hence, required RT estimations most likely relied on theory-based processes. The results showed introspective switch costs even in NoGo trials suggesting that participants may use theory-based knowledge to inform their judgments although experience-based influences of covert responses may also play a role. Taken together, our findings provide novel insights into introspection during task switching contributing to a better understanding of how task-switching performance may relate to decision-making.

1. Introduction

How accurately can humans judge their own performance? This ability is known as introspection which refers to the metacognitive capacity to monitor one's own cognitive processes (Danziger, 1980). Previous studies have shown that self-evaluations can be distorted by countless biases (Guggenmos, 2022; Kruger & Dunning, 1999; Thompson et al., 1998). To remain accurate, individuals may therefore draw on abstract, theory-based knowledge (Koriat, 1997, 2012; Koriat et al., 2008) rather than relying solely on moment-moment impressions of how well they performed. When judging your performance in a tennis game, for example, you might combine your immediate sense of how the match felt with general knowledge, for instance, knowing that you usually play better in moderate weather, to reach a more reliable assessment. A central question is whether theory-based monitoring is also applied during more elementary cognitive operations, such as switching between tasks. Therefore, the present study directly tested the role of theory-based monitoring during task switching.

In typical task switching paradigms, participants have to perform two simple tasks, e.g., categorizing numbers as odd or even, and categorizing letters as consonants or vowels (Rogers & Monsell, 1995). The task can either repeat or switch from one trial to the next. As a prominent finding, participants perform worse on task switches compared to task repetitions. This performance difference is called switch costs (Koch & Kiesel, 2022; Monsell, 2003). They are often used as a measure of cognitive flexibility, with higher switch costs indicating lower flexibility (Dreisbach & Mendl, 2024). Switch costs can be explained by the need to activate the new task set in case of a task switch (Rogers & Monsell, 1995) and by passive decay of the previous task set (Allport & Wylie, 2000). In any case, switch costs reflect a fundamental limitation of human multitasking. This raises the question whether participants are aware of these costs, in other words, whether switch costs are captured by introspective judgments.

Previous studies investigated introspection in multitasking situations especially in the dual-task paradigm. Here, two tasks have to be performed simultaneously and a shorter interval between stimuli (stimulus

* Corresponding author.
E-mail addresses: JonathanMendl@gmx.de (J. Mendl), gesine.dreisbach@ur.de (G. Dreisbach).

https://doi.org/10.1016/j.concog.2026.104095
Received 9 February 2026; Received in revised form 8 July 2026; Accepted 8 July 2026
Available online 11 July 2026

onset asynchrony) typically leads to slower RTs to the second task (Fischer & Janczyk, 2022). Introspection can be measured using a visual analogue scale where participants must estimate their RT after every trial. Interestingly, estimations did not accurately reflect the strong influence of the SOA on performance (Bratzke et al., 2014; Corallo et al., 2008). Building on these findings, Bratzke and Bryce (2019) explored whether a similar introspective blind spot exists for the switch costs during task switching (i.e., with sequential, non-overlapping tasks). Participants had to estimate their RT after every other target. While introspection did not capture influences of the preparation time, switch costs were clearly evident in introspective estimations. These introspective switch costs have been replicated in multiple experiments (Bratzke & Bryce, 2019, 2022; Mendl & Dreisbach, 2022). Even subtle modulations of the objective switch cost due to varied switch proportions were evident in introspective estimations (Mendl & Dreisbach, 2022).

Together, these results demonstrate that people are able to report the performance costs associated with switching between tasks. However, the basis of these introspective switch costs is not yet well understood. One question, central to understanding the nature of introspection during task switching, is how participants form these estimates. Do they rely on their direct experience of how fast and easy their executed response felt (experience-based judgments), or do they use more abstract knowledge about the effort and difficulty of task switches (theory-based judgments)? This distinction between two potential sources for introspective monitoring was made by Koriat (1997). Generally, introspection can draw on multiple sources of information, with the accuracy of one's judgments depending on how reliable these cues are. Experience-based monitoring relies on intuitive, non-analytic access to immediate sensations whereas theory-based monitoring involves more analytic judgments grounded in beliefs or abstract knowledge (Koriat, 2012). Both sources can contribute to introspective estimates (Pavailler et al., 2025). The present study directly investigated whether introspective switch costs reflect more than simple estimations of the immediate experience. In other words, we examined whether theory-based monitoring can be applied during task switching.

In the present study, we conducted two experiments. Experiment 1 tested a prerequisite of theory-based monitoring during task switching. In order to apply theory-based introspection, participants need to be able to form and retrieve separate representations for task switches and task repetitions. They need to learn that a task switch is difficult and associated with worse performance, while a task repetition is easy and associated with better performance. This can only happen when participants are aware of the task transition (repetition or switch) in a given trial. Therefore, in Experiment 1, we tested this awareness by asking participants in a random subset of trials whether the just executed task repeated or switched relative to the previous trial. We additionally tested the extent of awareness by using a short (1000 ms) and long (5000 ms) time window for the transition judgments to distinguish intuitive and deliberate decisions. We expected that the transition judgment accuracy is overall better than chance and that judgments are more accurate with a longer time window. Additionally, we should find the typical objective switch costs. Last, we also considered the target-response transitions (response key repetition vs. response key switch) as a potentially relevant factor for the transition awareness. Previous task-switching research has consistently shown that task repetitions are typically associated with a response-repetition benefit, whereas task switches are often associated with a response-switch benefit (e.g., Benini et al., 2023; Dreisbach & Wenke, 2011; Hübner & Druey, 2006, 2008; Kleinsorge, 1999; Meiran et al., 2000; Schuch & Koch, 2004). Such response-transition effects can presumably influence subjective processing fluency and may therefore be introspectively accessible. Consequently, participants' transition judgments may partly rely on perceived response-transition fluency (being more accurate when the target response key and the task were repeated or when both switched).

In Experiment 2, we tested theory-based introspection during task switching by using NoGo trials, therefore reducing the influence of experience-based monitoring of the executed response. In NoGo trials, participants are instructed to withhold their response to the target. Introspection was measured on a visual analogue scale after every other target. In NoGo trials, participants had no moment-to-moment experience about their performance. Hence, the most likely source for introspection was knowledge about the present task transition. We expected to find objective and introspective switch costs when participants responded to the target (Go trials). Critically, we also expected to find introspective switch costs for NoGo trials in terms of longer RT estimates for (not executed) switches as compared to (not executed) repetitions.

2. Experiment 1[1]

2.1. Method

2.1.1. Participants

Experiment 1 was preregistered (https://aspredicted.org/wf8fj6.pdf). Data collection and analysis were performed according to the preregistration protocol. According to a power analysis in G*Power 3.1 to detect a medium effect (d = 0.5) in the paired *t*-test of judgment accuracy with a power of 0.80, at least 27 participants are required (Faul et al., 2009). Due to the novelty of the current approach, we aimed for at least 50 participants. Hence, 51 participants completed the experiment. One participant had to be excluded (for exclusion criteria, see 1.2.1. Data Preprocessing), resulting in a final sample of $N = 50$. The mean age was 21.08 years ($SD = 2.34$), ranging from 18 to 30 years. Forty-two were female (8 male) and 44 right-handed (6 left-handed). All participants provided informed consent following the ethical standards of the national research committee and with the 1964 Helsinki Declaration and its later amendments. Psychology students received course credit for their participation.

2.1.2. Apparatus and stimuli

The experiment was programmed in lab.js (Henninger et al., 2022) and hosted online on Open Lab (Shevchenko, 2022). The stimuli of the two tasks were adopted from Fröber and Dreisbach (2017). In the number task, participants had to categorize numbers (125, 132, 139, 146; 160, 167, 174, 181) as smaller or larger than 153. In the letter task, participants had to categorize letters (B, D, F, H; S, U, W, Y) as closer to A or Z in the alphabet. Stimuli were shown in black ink (5.33% of screen height) in the default sans-serif font. They were presented in the center of the screen on a white background. Participants had to respond to numbers smaller than 153 and letters closer to A using the S key with their left index finger and to numbers larger than 153 and letters closer to Z using the K key with their right index finger.[2]

On 25% of the test trials, participants had to indicate the current task transition, i.e., whether the task repeated or switched relative to the previous trial. In these judgment prompts, the German words for repetition (“Wiederholung”) and switch (“Wechsel”) were shown 25% of the

[1] All Experiments were preregistered: Experiment 1 https://aspredicted.org/wf8fj6.pdf; Experiment 2 https://aspredicted.org/qu4yd4.pdf.

Raw data files associated with this article can be found online under the following link:https://doi.org/10.5283/epub.79764. Additional study materials of all experiments (lab.js experiment files, SPSS analysis scripts) will be shared upon request.

[2] We deliberately chose univalent stimuli (i.e., the stimulus sets of the two tasks do not overlap) because our goal was to keep task transitions as implicit as possible. With bivalent stimuli (e.g., only using numbers as stimuli, two tasks consisting of judging parity or magnitude), participants necessarily rely on explicit task cues to determine the currently relevant task, making the (necessary) cue highly salient. In contrast, with univalent stimuli, the task is directly activated by the stimulus itself without requiring an additional cue interpretation process. Therefore, the present paradigm makes task transitions less salient compared to using bivalent stimuli.

screen height to the left and the right of the center with the same size and font as the task stimuli (see Fig. 1). Participants had to press the A key with their left middle finger to indicate the option presented on the left, and the L key with their right middle finger to indicate the option presented on the right. The locations of the task transitions in the prompt stayed consistent for the entire session but were counterbalanced across participants.

2.1.3. Procedure

After providing demographics, the experiment began with separate practice blocks for each task (8 trials each, task order counterbalanced across participants). Next, participants practiced switching between the two tasks (16 trials). In these first practice blocks, a trial always consisted of a fixation cross (500 ms), the target (until response), feedback (correct "Richtig!" or error "Fehler!"; 1000 ms), and an intertrial interval (ITI; 500 ms). In the last practice block, participants were familiarized with the judgment prompt (32 trials). Here, the participants had to indicate the present task transition (whether the task repeated or switched relative to the previous task) after the target in every second trial. The prompt stayed until a response was given.

The following test phase consisted of 8 blocks (64 trials each). The task switched in roughly 50% of the trials. In 25% of the trials (at random), the judgment prompt appeared. The prompt could not appear in the first trial or directly following another prompt. The judgment prompt was presented equally often after switch trials and repetition trials. Furthermore, we manipulated the time window of the judgment prompt. In alternating blocks, the time window was either short (1000 ms) or long (5000 ms). The time window condition in the first test block was counterbalanced across participants. Prior to each block, participants were instructed about the upcoming time window condition. They were told that they would only have very little time to judge the transition and that they should therefore decide quickly and spontaneously (short time window condition), or that they would have sufficient time to judge the transition and that they should therefore take the time to decide deliberately. We did not include performance feedback in the test phase, to leave the succession of trials unobstructed. Only after each block, participants were informed about the error rate in response to the task-switching targets (not about judgment errors) in the previous block (in %). During each block, participants only received feedback (1500 ms) in case they responded to slow to the judgment prompt (too slow "Zu langsam"). Each trial consisted of a fixation cross (500 ms), target (until response), judgment prompt (in 25% of the trials; block-wise 1000 ms or 5000 ms), feedback (only for too slow judgments; 1500 ms) and ITI (500 ms; see Fig. 1). At the end of the experiment, the participants were asked how they answered the transition question and what their answers were based on.

2.1.4. Design

The judgment accuracy (in %) was computed as the rate of correct judgments relative to the number of judgment prompts. The Time Window (short, long) of the judgment prompt served as an independent variable in a repeated-measures design. Additionally, the judgment accuracy was included in a 2 (Time Window) x 2 (Task Transition: repetition, switch; whether the task repeated or switched from one trial to the next) x 2 (Response Transition: repetition, switch; whether the correct response key for the target-based reaction repeated or switched from one trial to the next) repeated-measures design. The same 2x2x2 design was used for the judgment RT (in ms). One might expect that participants are more biased toward (and faster on) task-switch judgments when the target-response keys switched and vice versa (preregistered as secondary analysis). For the mean RT (in ms) and error rate (in %) in response to the target, we used a 2 (Time Window) x 2 (Task Transition) repeated-measures design to check whether the time window influenced task performance and thereby potentially the switch cost experience. All raw data files associated with this article are available online under the following link: https://doi.org/10.5283/epub.79764. Additional study materials of all experiments (lab.js experiment files, SPSS analysis scripts) will be shared upon request.

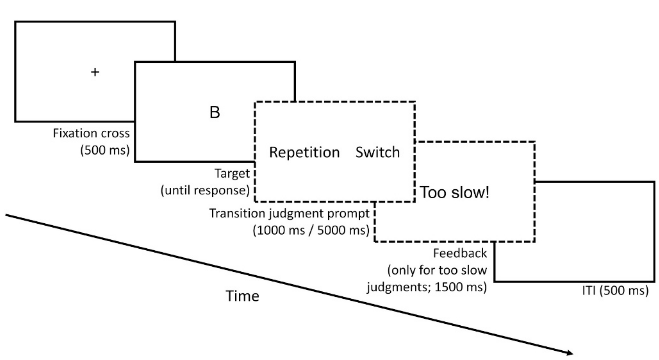

Fig. 1. Schematic Depiction of a Single Trial With the Judgment Prompt in Experiment 1. Note. The judgment prompt only appeared in random 25% of the trials.

2.2. Results

2.2.1. Data Preprocessing

All preprocessing steps and data analyses were performed according to the preregistration protocol unless stated otherwise. To investigate the transition judgment accuracy, we only included trials with responses within the time window (excluded too slow responses: 7.94% of all judgments).[3] As a last step, we excluded trials with erroneous responses to the present or previous target (3.57% of all judgments). To investigate the judgment RT, we additionally excluded erroneous judgments (17.92%) and judgment RTs faster than 150 ms or slower than 3000 ms (0.52%). To investigate the error rate in response to the target, we excluded the first trial of each block (1.56% of all trials) and trials following the judgment prompt (24.64%). For the analysis of RTs, we additionally excluded errors (3.05%), post errors (2.81%), and RTs faster than 150 ms or slower than 3000 ms (0.29%).

As stated in the preregistration, we excluded participants with an overall mean RT or error rate with more than three interquartile ranges above the third (or below the first) quartile. One participant had to be excluded due to the RT criterion ($M = 1407$ ms). This resulted in a final sample of $N = 50$.

2.2.2. Judgment accuracy

The overall judgment accuracy of 78.36% ($SE = 1.61$) was significantly higher than chance (50%), $t(49) = 17.64$, $p < .001$, $d = 2.50$. Furthermore, judgments were more accurate in blocks with a long judgment time window ($M = 84.27\%$, $SE = 1.91$) compared to blocks with a short judgment time window ($M = 70.71\%$, $SE = 1.85$; difference: $M = 13.54\%$, $SE = 1.88$), $t(49) = 7.19$, $p < .001$, $d = 1.02$. Critically, explorative one-sample t-tests showed that the judgment accuracy was significantly higher than chance in both conditions (all $ps < 0.001$).

In the 2 (Time Window) x 2 (Task Transition) x 2 (Response Transition) repeated-measures ANOVA of the judgment accuracy, we found significant main effects of Time Window, $F(1, 49) = 48.96$, $p < .001$, $\eta_p^2 = 0.50$ (see t-test), and Task Transition, $F(1, 49) = 11.68$, $p < .001$, $\eta_p^2 = 0.19$. Unexpectedly, participants were more accurate in judging task switches ($M = 82.17\%$, $SE = 1.95$) compared to task repetitions ($M = 72.91\%$, $SE = 2.27$; difference: $M = 9.27\%$, $SE = 2.71$). The main effects were further qualified by an interaction of Time Window and Transition, $F(1, 49) = 9.50$, $p = .003$, $\eta_p^2 = 0.16$. The difference in judgment accuracy between task switches and task repetitions was more pronounced in blocks with a short time window ($M_{switch} = 76.93\%$, $SE = 2.23$; $M_{repetition}$

[3] Including all judgment trials in the judgment accuracy analysis resulted in the same statistical pattern of results.

= 64.52%, SE = 2.46; difference: M = 12.41%, SE = 2.89, p < .001) compared to blocks with a long time window (M_{switch} = 87.42%, SE = 2.23; $M_{repetition}$ = 81.30%, SE = 2.61; difference: M = 6.12%, SE = 2.91, p = .040; see Fig. 2). The main effect and interactions of Response Transition were not significant (all Fs < 1.92, all ps > 0.172).

2.2.3. *Judgment RT*

The 2 (Time Window) x 2 (Task Transition) x 2 (Response Transition) repeated-measures ANOVA of the judgment RT revealed a significant main effect of Time Window, $F(1, 49) = 150.44$, $p < .001$, $\eta_p^2 = 0.75$. Participants gave quicker judgments in blocks with a short judgment time window (M = 685 ms, SE = 12) compared to blocks with a long judgment time window (M = 1115 ms, SE = 37; difference: M = 430 ms, SE = 35). No other effect or interaction was significant (all Fs < 3.25, all ps > 0.077).

2.2.4. *Target RT*

The 2 (Time Window) x 2 (Task Transition) repeated-measures ANOVA of the RT in response to the target resulted in significant main effects of Time Window, $F(1, 49) = 48.74$, $p < .001$, $\eta_p^2 = 0.50$, and Transition, $F(1, 49) = 136.65$, $p < .001$, $\eta_p^2 = 0.74$. Participants responded faster to the target in blocks with a short judgment time window (M = 681 ms, SE = 20) compared to blocks with a long judgment time window (M = 752 ms, SE = 23; difference: M = 71 ms, SE = 10), and faster on task repetitions (M = 676 ms, SE = 19) compared to task switches (M = 756 ms, SE = 24; switch cost: M = 80 ms, SE = 7). The interaction between Time Window and Transition was not significant, $F(1, 49) = 2.97$, $p = .091$, $\eta_p^2 = 0.06$.

2.2.5. *Target error rate*

In the 2 (Time Window) x 2 (Task Transition) repeated-measures ANOVA of the error rate in response to the target, we found a significant main effect of Time Window, $F(1, 49) = 44.41$, $p < .001$, $\eta_p^2 = 0.48$. Participants made more errors in response to the target in blocks with a short judgment time window (M = 5.35%, SE = 0.42) compared to blocks with a long judgment time window (M = 2.95, SE = 0.34; difference: M = 2.39%, SE = 0.36). No other effect was significant (all Fs < 2.46, all ps > 0.123).

2.2.6. *Exploratory analysis of the post-experiment question*

After the Experiment, participants were asked how they answered the transition question and what their answers were based on. Here, 21 (out of 50) participants mentioned that they constantly monitored the tasks or transitions (e.g., using inner speech to repeat the task in every trial and compare the tasks in case of a judgment prompt). To investigate the effects of such constant monitoring processes, we divided the sample in two groups (no constant monitoring: n = 29, constant monitoring: n = 21) for the key analyses of judgment accuracy and Target RT. This control analysis was not preregistered.

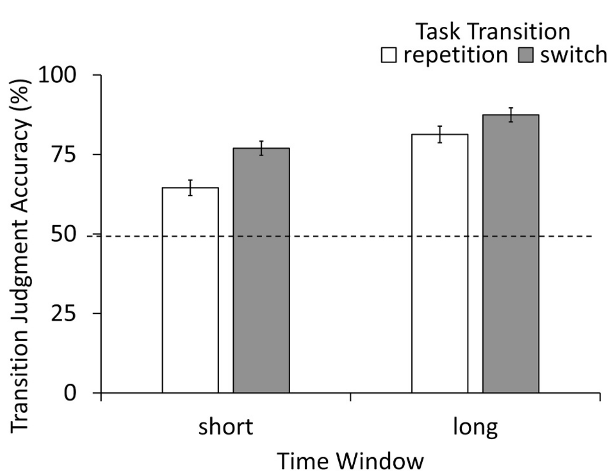

Fig. 2. Judgment Accuracy (in %) as a Function of Time Window (Short, Long) and Task Transition (Repetition, Switch). Note. Error bars represent ± one standard error of the mean. The dashed line indicates chance level (50%).

The 2(Time Window; within-subjects) x 2 (Monitoring; between-subjects) mixed ANOVA revealed a significant main effect of Monitoring, $F(1, 48) = 8.40$, $p = .006$, $\eta_p^2 = 0.15$. Participants in the constant monitoring group were more accurate in judging the transition (M = 82.66%, SE = 2.34) compared to the group without constant monitoring (M = 73.74%, SE = 1.99). There was a significant main effect of Time Window, $F(1, 48) = 48.65$, $p < .001$, $\eta_p^2 = 0.50$ (see main analysis). The interaction of Time Window and Monitoring was not significant, $F(1, 48) = 0.15$, $p = .699$, $\eta_p^2 < 0.01$.

The 2 (Time Window; within-subjects) x 2 (Task Transition; within-subjects) x 2 (Monitoring; between-subjects) mixed ANOVA of the Target RT showed a significant main effect of Monitoring, $F(1, 48) = 4.72$, $p = .035$, $\eta_p^2 = 0.09$. Participants in the constant monitoring group responded slower to the target (M = 768 ms, SE = 31) compared to the group without constant monitoring (M = 679 ms, SE = 27). The main effects of Time Window, $F(1, 48) = 47.01$, $p < .001$, $\eta_p^2 = 0.50$, and Task Transition, $F(1, 48) = 140.48$, $p < .001$, $\eta_p^2 = 0.75$, were also significant (see main analysis). No other effect was significant (all Fs < 2.83, all ps > 0.099).

2.3. *Discussion*

Taken together, the results confirm our expectation that participants can judge the task transition significantly better than chance. Deliberate judgments (with a long time window of 5000 ms) were more accurate than intuitive ones (with a short time window of 1000 ms). Still, even when participants only had a short judgment time window, they indicated the correct transition above chance level. This suggests that participants acquire knowledge about task transitions and may form separate representations for task switches and task repetitions.

The transition judgments were not influenced by target-response key repetitions or switches. Thus, the observed metacognitive judgments cannot primarily be explained by well-established response-transition effects, such as response-repetition benefits in task repetitions and response-switch benefits in task switches (e.g., Benini et al., 2023; Dreisbach & Wenke, 2011; Hübner & Druey, 2006, 2008; Kleinsorge, 1999; Meiran et al., 2000; Schuch & Koch, 2004). Unexpectedly, participants were more accurate in judging task switches rather than task repetitions. A potential reason might be that task switches and their associated higher difficulty and effort (Dreisbach & Jurczyk, 2022; Kool et al., 2010) are more salient than task repetitions.

The performance data in response to the target showed the typical RT switch costs (Koch & Kiesel, 2022). Interestingly, the time pressure of the judgment time window did not only influence the RT of the judgment response but also carried over to the target response. Participants responded faster and made more errors in response to the target in blocks with a shorter judgment time window.

Some participants mentioned that they had constantly monitored the current tasks or transitions. This appeared to increase the judgment accuracy but also slowed the RT in response to the target. These costs of constant monitoring are in line with research about prospective memory costs where performance is typically slower with concurrent prospective memory demand (Anderson et al., 2019). Interestingly, in the present study, the constant monitoring costs are self-inflicted by a subset of participants.

One limitation of Experiment 1 is that the present design used univalent stimuli. Hence, participants may not only have relied on representations of task transitions themselves but potentially also on changes in stimulus features such as stimulus category (letters vs. numbers) or stimulus length.[4] However, we deliberately chose univalent stimuli because our goal was to keep task transitions as implicit as possible. Using bivalent stimuli would have required explicit task cues to indicate the currently relevant task, making cue transitions more salient and introducing additional cue-interpretation processes. In contrast, with univalent stimuli, the task is directly activated by the stimulus itself without requiring explicit task cues. That is, the usage of univalent stimuli makes task transitions less salient compared to bivalent stimuli. Still, we cannot completely rule out that participants base their transition judgments on changes in stimulus features.

In sum, Experiment 1 showed that participants are able to indicate whether the current trial involved a task switch or a task repetition. This knowledge may be used to form separate mental representations for task switches and task repetitions, providing a basis for theory-based monitoring during task switching. In order to directly test for theory-based introspection in Experiment 2, we reduced the possibility to rely on experience-based monitoring of executed responses by introducing NoGo trials.

3. Experiment 2

In Experiment 2, we tested introspection by instructing participants to estimate their own RT after every other target in a task-switching paradigm (Bratzke & Bryce, 2019, 2022; Mendl & Dreisbach, 2022). Critically, we included NoGo trials, where participants had to avoid responding to the target and just had to wait for a short duration. This should reduce immediate experience-based knowledge. If we still find significant introspective switch costs, this would most likely be the result of theory-based introspection where participants use abstract knowledge about the task transition to inform their estimation.

3.1. Method

3.1.1. Participants

Experiment 2 was also preregistered (https://aspredicted.org/qu4yd4.pdf). Data collection and analysis were performed according to the preregistration protocol. According to the same logic as in Experiment 1, we collected the data of 50 participants. No participant had to be excluded. The mean age was 21.28 years ($SD = 2.25$), ranging from 18 to 30 years. Thirty-seven were female (11 male, 2 other) and 46 right-handed (4 left-handed). All participants provided informed consent following the ethical standards of the national research committee and with the 1964 Helsinki Declaration and its later amendments. Psychology students received course credit for their participation.

3.1.2. Apparatus and stimuli

We used the same software and tasks as in Experiment 1. As a first change in Experiment 2, the number and the letter task were mapped to separate fingers. One task was mapped to both index fingers (keys D and K). The other task was mapped to both middle fingers (keys S and L).[5] Instead of a fixation cross, we used a box (length: 13.33% of screen height; height: 8.33% of screen height). This box was presented in black in Go trials and in red (rgb: 255, 0, 0) to indicate NoGo trials.

Introspection was measured using a visual analogue scale with labels at the extremes (0 ms, 1500 ms). The default position of the slider was the center of the scale. Participants had to use their mouse to click on the scale to reposition the slider and estimate their RT. Only after clicking somewhere on the scale, participants could confirm their estimation by pressing the spacebar (see Fig. 3).

3.1.3. Procedure

In Experiment 2, participants first practiced each task (8 trials each). Next, they practiced switching between the two tasks (16 trials). Each practice trial consisted of a fixation box (500 ms) which stayed on screen during target presentation, the target (until response), feedback (correct "Richtig!", error "Fehler!"; 1000 ms), and an ITI (500 ms). In the next practice block, participants were familiarized with NoGo trials (16 trials; 50% Go, 50% NoGo). Participants had to withhold responses when the fixation box was red. In NoGo trials the target inside the red fixation box stayed on screen for 1000 ms. Any response during this interval resulted in error feedback. In the last practice block, we introduced the introspection method (8 trials). In each trial, we presented two targets. At the end of each trial, participants had to estimate their own response to the second target on the visual analogue scale. Each trial consisted of a fixation box (500 ms), the first target (until response), a blank screen (500 ms), a fixation box (500 ms) the second target (until response), a blank screen (500 ms), introspection on the visual analogue scale (until response), and an ITI (500 ms). During introspection, we also showed feedback in case of an error in the first and/or the second target (error on the first target "Fehler bei erstem Reiz!", error on the second target "Fehler bei zweitem Reiz!").

In the test phase (4 blocks with 64 trials each), we used the introspection method while including 25% NoGo as the second target (equally often in switch and repetition trials). Participants were instructed to estimate their RT to the second target at the end of each trial. In case of a NoGo target, participants were instructed to estimate how fast they would have responded to the target. The trial structure is depicted in Fig. 3. At the end of the experiment, the participants were asked what their RT estimations were based on after Go targets and after NoGo targets.

3.1.4. Design

First, the objective RT (in ms) on the second target in the Go condition was used as a dependent variable. The Task Transition (repetition, switch; whether the task repeated or switched from the first to the second target in the current trial) served as an independent variable in a repeated-measures design. Second, the introspective RT (in ms) measured on the visual analogue scale was used in 2 (Task Execution: NoGo, Go) x 2 (Task Transition) repeated-measures design. For completeness, we will also analyze the error rate (in %) on the second target in a 2 (Task Execution) x 2 (Task Transition) repeated-measures design.

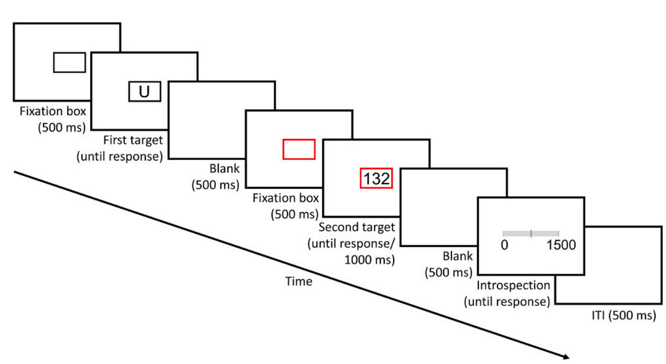

Fig. 3. Schematic Depiction of a Single Trial With NoGo on the Second Target in Experiment 2.

[4] We thank an anonymous reviewer for bringing this to our attention.

[5] The task design of Experiment 2, with the number and letter task being mapped to different fingers, is typically used to avoid response crosstalk between the two tasks. For Experiment 1, we deviated from this design because participants also had to perform transition judgments on each trial. To avoid overload from too many response keys/fingers, we limited the response set to two keys for the transition judgment and two keys for the task responses.

3.2. Results

3.2.1. Data Preprocessing

All preprocessing steps and data analyses were performed according to the preregistration protocol unless stated otherwise. In all analyses, only the second target of each trial was of interest. To investigate the error rate, we excluded the first trial of each block (1.56% of all trials). To investigate objective and introspective RT, we additionally excluded errors on the first target (9.08%), errors on the second target (5.84%), and objective RTs faster than 150 ms or slower than 3000 ms (0.21%). Last, for objective RTs, we additionally excluded all NoGo trials (21.97%). We excluded participants with an overall mean RT or error rate with more than three interquartile ranges above the third (or below the first) quartile. No participant had to be excluded resulting in a final sample of $N = 50$.

3.2.2. Objective RT

The objective RT significantly varied as a function of Task Transition, $t(49) = 7.99$, $p < .001$, $d = 1.13$. We found the typical RT switch costs ($M = 83$ ms, $SE = 10$). Participants responded faster in task repetitions ($M = 616$ ms, $SE = 15$) compared to task switches ($M = 699$ ms, $SE = 20$; see Fig. 4).

3.2.3. Introspective RT

The 2 (Task Execution) x 2 (Task Transition) repeated-measures ANOVA of the introspective RT showed a significant main effect of Task Transition, $F(1, 49) = 52.36$, $p < .001$, $\eta_p^2 = 0.51$. Participants showed introspective switch costs by estimating faster RTs in task repetitions ($M = 623$ ms, $SE = 20$) compared to task switches ($M = 691$ ms, $SE = 21$; switch costs: $M = 68$ ms, $SE = 9$). The main effect of Task Execution and the interaction of Task Execution and Task Transition were not significant (all $Fs < 3.04$, all $ps > 0.088$). Critically, participants estimated significant introspective switch costs in both NoGo trials ($M_{repetition} = 624$ ms, $SE = 21$; $M_{switch} = 701$, $SE = 23$; switch costs: $M = 78$ ms, $SE = 13$, $p < .001$) and Go trials ($M_{repetition} = 622$ ms, $SE = 21$; $M_{switch} = 682$, $SE = 21$; switch costs: $M = 59$ ms, $SE = 8$, $p < .001$; see Fig. 4).

3.2.4. Error rate

In the 2 (Task Execution) x 2 (Task Transition) repeated-measures ANOVA of the error rates, we found a significant main effect of Task Execution, $F(1, 49) = 39.69$, $p < .001$, $\eta_p^2 = 0.45$. Participants made more errors in Go trials ($M = 8.32\%$, $SE = 0.69$) where an error meant pressing the wrong key in response to the target, compared to NoGo trials ($M = 3.71\%$, $SE = 0.46$; difference: $M = 4.61\%$, $SE = 0.73$) where an error represented a false alarm. There was a significant interaction between Task Execution and Task Transition, $F(1, 49) = 13.18$, $p < .001$, $\eta_p^2 = 0.21$. Post-hoc comparisons showed that the switch costs were only significant for error rates in Go trials ($M_{repetition} = 7.31\%$, $SE = 0.62$; $M_{switch} = 9.33\%$, $SE = 0.85$; switch costs: $M = 2.02\%$, $SE = 0.55$, $p < .001$) and not for false alarms in NoGo trials ($M_{repetition} = 4.15\%$, $SE = 0.58$; $M_{switch} = 3.28\%$, $SE = 0.53$; switch costs: $M = -0.87$, $SE = 0.62$, $p = .161$; see Fig. 5).

3.2.5. Exploratory analysis of the post-experiment question

In the post-experiment question, the participants were asked what their RT estimations were based on after Go targets and after NoGo targets. Here, 13 (out of 50) participants mentioned that they in some way simulated their response in NoGo trials (e.g., estimating the time until their response finger twitched). To explore the effects of this simulation strategy, we divided the sample in two groups (no simulated response: $n = 37$, simulated response: $n = 13$) for the key analyses of objective and introspective RT. One may argue that introspective switch costs in NoGo are only the result of these simulated responses. This analysis was not preregistered.

The 2 (Task Transition; within-subjects) x 2 (Simulation Strategy; between-subjects) mixed ANOVA of the objective RT in Go trials showed a significant main effect of Task Transition, $F(1, 48) = 41.97$, $p < .001$, $\eta_p^2 = 0.47$ (see main analysis). The main effect of Simulation Strategy and the interaction of Simulation Strategy and Task Transition were not significant (all $Fs < 1.36$, all $ps > 0.249$). So, there were objective switch costs irrespective of the Simulation Strategy.

The 2 (Task Execution; within-subjects) x 2 (Task Transition; within-subjects) x 2 (Simulation Strategy; between-subjects) mixed ANOVA of the introspective RT resulted in a significant main effect of Task Transition, $F(1, 48) = 30.61$, $p < .001$, $\eta_p^2 = 0.39$ (see main analysis). The interaction of Task Transition and Simulation Strategy was significant, $F(1, 48) = 4.90$, $p = .032$, $\eta_p^2 = 0.09$, which was further qualified by a significant three-way interaction between Task Execution, Task Transition, and Simulation Strategy, $F(1, 48) = 5.15$, $p = .028$, $\eta_p^2 = 0.10$. The group without simulated responses showed significant introspective switch costs in NoGo and Go trials (all $ps < 0.001$) whereas the group with simulated responses showed significant introspective switch costs only in Go trials ($p = .007$) but not in NoGo trials ($p = .320$; see Table 1).

3.2.6. Exploratory correlational analyses of objective and introspective switch costs

We also ran exploratory correlational analyses between the objective switch costs in Go, introspective switch costs in Go, and introspective switch costs in NoGo trials. Switch costs were computed as the difference between the RT on task switches minus the RT on task repetitions. All three correlations were significant. Participants with higher objective switch costs estimated higher introspective switch costs in Go trials ($r =$

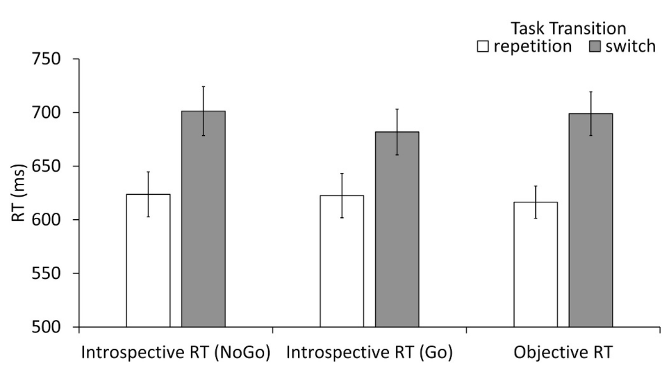

Fig. 4. Objective and Introspective RT (in ms) as a Function of Task Execution (NoGo, Go; Only for Introspective RT) and Task Transition (Repetition, Switch). Note. Error bars represent ± one standard error of the mean.

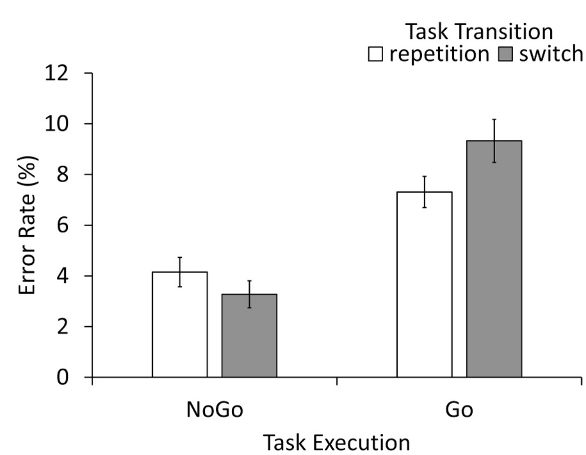

Fig. 5. Error Rate (in %) as a Function of Task Execution (NoGo, Go) and Task Transition (Repetition, Switch). Note. Error bars represent ± one standard error of the mean. Error Rates in NoGo trials denote false alarms.

Table 1
Introspective RT Estimations (in ms) as a Function of Simulation Strategy (No Simulated Responses, Simulated Responses), Task Execution (NoGo, Go), and Task Transition (Repetition, Switch).

	NoGo			Go		
	Repetition	Switch	SC	Repetition	Switch	SC
No Simulated Responses	620 (*25*)	716 (*26*)	96 (*14*)	631 (*24*)	696 (*25*)	65 (*9*)
Simulated Responses	634 (*42*)	659 (*45*)	24 (*24*)	598 (*41*)	642 (*42*)	45 (*16*)

Note. SC stands for the mean introspective switch costs (in ms) per condition. SE are presented in parentheses.

0.493, $p < .001$). Participants with higher objective switch costs estimated higher introspective switch costs in NoGo trials ($r = 0.285$, $p = .045$). Last, Participants with higher estimated introspective switch costs in Go trials also estimated higher introspective switch costs in NoGo trials ($r = 0.591$, $p < .001$).

3.2.7. Exploratory comparison of objective and introspective switch costs

Last, we compared the magnitude of the objective switch costs and the introspective switch costs in Go and NoGo trials. We thank an anonymous reviewer for suggesting this analysis. Introspective switch costs on Go trials ($M = 59$ ms, $SE = 8$) were significantly smaller than the objective switch costs ($M = 83$ ms, $SE = 10$), $t(49) = 2.45$, $p = .018$, $d = 0.35$. In contrast, introspective switch costs on NoGo trials ($M = 78$ ms, $SE = 13$) did not significantly differ from the objective switch costs, $t(49) = 0.35$, $p = .731$, $d = 0.05$. Please note that the absolute magnitude of introspective switch costs should be interpreted with caution because it is highly influenced by the range of the visual analogue scale and therefore does not directly map onto millisecond-based objective RT measures.

3.3. Discussion

The results of Experiment 2 demonstrated that, even without experiencing task performance in the current trial (NoGo), participants were able to estimate that they would have been slower on a task switch compared to a task repetition. In other words, we found introspective switch costs even when experience-based monitoring of an executed response on the current trial was impossible. This suggests that participants likely applied theory-based monitoring (Koriat, 1997, 2012; Koriat et al., 2008) using abstract knowledge about task transitions to guide their estimations. Furthermore, we replicated established effects of prior task-switching research. Participants showed objective switch costs (Koch & Kiesel, 2022; Monsell, 2003) and were able to correctly estimate switch costs in Go trials (Bratzke & Bryce, 2019, 2022; Mendl & Dreisbach, 2022). Additionally, we found that participants showed switch costs in the error rates in Go trials. NoGo error rates (false alarms) were not influenced by the task transition.

The post-experiment question indicated that 13 participants simulated their response in NoGo trials (e.g., by estimating the time until their response finger twitched). These simulated responses may allow for experience-based monitoring even on NoGo trials. However, exploratory analyses including the between-subjects factor Simulation Strategy pointed in a different direction. Introspective switch costs in NoGo were only evident in the group without simulated responses. When interpreting these results, caution is advised due to the analysis being underpowered especially for the group who simulated responses. Still, there were no hints in favor of the alternative explanation that experience-based monitoring was applied in NoGo via simulated responses. Taken together, the analysis of the post-experiment question does not challenge the interpretation that the introspective switch costs in NoGo are most likely the result of theory-based monitoring. However, we cannot exclude that participants may have had some awareness of their readiness to respond. That is, experience-based knowledge might have contributed to the effect. We come back to this in the General Discussion.

Theory-based knowledge does not arise out of thin air. To gain abstract knowledge, *prior* experiences are necessary (not to be confused with the moment-to-moment experience of experience-based monitoring). In the present paradigm, this meant that participants first experienced how their performance in Go trials is influenced by the task transition to later use this information for their RT estimations in NoGo trials. This rationale is supported by the exploratory correlational analyses showing that objective switch costs in Go trials and introspective switch costs in NoGo trials were positively related. This suggests that participants may have formed theory-based knowledge about task transitions during their RT estimations in Go trials which then informed their estimations in NoGo trials.

4. General discussion

The present study investigated whether theory-based introspection can be applied during task switching. This would mean that participants may possess abstract knowledge about task transitions (e.g., that task switches are more difficult and take longer than task repetitions) and may use this knowledge to inform their estimations. In Experiment 1, participants had to indicate the current task transition on a subset of trials. The results suggested that participants were generally able to recognize whether the task had just repeated or switched. However, results of Experiment 1 do not unequivocally speak to the idea that participants had acquired abstract knowledge about task transitions (repetition, switch) that then informed their judgments. In Experiment 2, we therefore tested whether abstract knowledge about task transitions is applied when participants estimate their RT on a visual analogue scale. To this end, we reduced the contribution of immediate experience-based monitoring by introducing NoGo trials. Results confirmed out predictions: We still found significant introspective switch costs. In NoGo trials participants had to refrain from responding to the target. Hence, theory-based monitoring likely contributed to estimations of how fast participants would have responded on the current NoGo trial, although we cannot exclude that experience-based knowledge in terms of subtle moment-to-moment familiarities contributed to the effect. Together, these findings provide novel insights into introspection during task switching.

Previous research on introspection during task switching already showed that participants are generally able to report introspective switch costs (Bratzke & Bryce, 2019, 2022; Mendl & Dreisbach, 2022). The participants' estimations of the switch costs were rather accurate and even reflected subtle effects of the switch proportion (Mendl & Dreisbach, 2022). However, prior to the present study the meaning of these introspective switch costs remained unclear. Here, we showed that introspective switch costs may not only reflect experience-based estimations of immediate task performance but may also be influenced by theory-based, abstract knowledge about task transitions. This finding is critical when making assumptions about the role of the (introspective) switch costs for decision making (Kool et al., 2010; Mendl & Dreisbach, 2022; Mittelstädt et al., 2018, 2019; Monno et al., 2021). It is often assumed that participants take their switch costs into account when deciding to switch or repeat. In the often-used voluntary task switching paradigm, participants can decide for themselves which task to perform on a given trial (Arrington & Logan, 2004). Participants with higher switch costs typically switch tasks less often voluntarily (Mayr & Bell, 2006; Mendl & Dreisbach, 2022; Mittelstädt et al., 2018, 2019). This finding was often attributed to an economic cost-benefit analysis. Participants avoid the cost of switching. The present study provides evidence that participants possess some awareness of the costs associated with task switching. Hence, our findings provide a plausible mechanism through which switch costs may inform voluntary task choices.

Furthermore, the present findings shed light on how theory-based knowledge is formed during task switching. First, participants appear to notice the current transition (as suggested by Experiment 1) and may form separate representations for task repetitions and task switches based on their performance. These representations may be updated through repeated experience and can later be accessed (even in NoGo trials as shown in Experiment 2). The correlational analyses in Experiment 2 suggest that both objective performance and introspective estimations of this performance contribute to the formation of these representations. Specifically, participants with larger objective and introspective switch costs in Go trials also showed larger introspective switch costs in NoGo trials. Together, this suggests that prior experience may contribute to the formation of theory-based representations. Hence, experience-based introspection and theory-based introspection can interact and influence each other. One open question is whether introspective switch costs in NoGo trials would emerge in the absence of prior task-switching experience. One possibility is that individuals already hold pre-existing beliefs about task switching, such as a general aversion to change or an expectation that switching tasks is effortful and costly (Dreisbach & Jurczyk, 2022; Kool et al., 2010; Vermeylen et al., 2022). In this case, we would expect to find introspective switch costs in NoGo trials even without any prior task performance.

The present findings support the broader idea that introspection about one's own performance most likely relies on a mixture of theory-based and experience-based monitoring (Koriat et al., 2004). The present study suggests that theory-based monitoring may play a role during introspection in task switching. Future research may investigate the accuracy of such theory-based introspection processes and determine potential mismatches between objective performance (closely linked to experience-based monitoring) and theory-based monitoring (Koriat et al., 2008). To that end, it would be interesting whether typical (subtle) influences on the objective switch costs, e.g., the proportion of task switches in a given context, are reflected in theory-based introspective estimations. One may argue that only more salient experiences are encoded in theory-based representations (Tversky & Kahneman, 1973).

4.1. Limitations

One important limitation is that the present findings do not unequivocally demonstrate that introspective switch costs in NoGo trials of Experiment 2 were based exclusively on theory-based monitoring. Participants may still have relied on subtle moment-to-moment experiences during NoGo trials even in the absence of an overt response. For example, participants may have had some awareness of their readiness to respond, differences in the effort required to inhibit a response, or covert processing of the target stimulus up to the stage of response selection. Additionally, participants may have learned associations between changes in perceptual stimulus features and differences in processing difficulty across trials. Such processes could also contribute to introspective RT estimations in NoGo trials without requiring fully abstract representations of task transitions. At the same time, the present design attempted to reduce the contribution of such covert trial-specific processing. Specifically, the 500 ms NoGo fixation box allowed participants to prepare the NoGo response itself before target onset. This likely interrupted or reduced typical task processing toward overt response preparation, as participants first had to prepare withholding a response before processing the upcoming target. While this manipulation cannot fully exclude covert processing or subjective experiences related to response readiness and inhibition, it likely reduced the extent to which participants could simultaneously engage in such processes. Moreover, the need for response inhibition should be stronger for task repetitions than task switches (due to the activation advantage of task repetitions, cf. Dreisbach et al., 2002) and thus more effortful, which would rather translate in higher RTs estimations for task repetitions which was not the case. However, participants may also have interpreted greater inhibition effort as reflecting a stronger tendency to respond and, consequently, inferred that their response would have been faster had it been executed. Such an inference could contribute to the observed introspective switch costs. However, this interpretation would require an additional inferential step. Therefore, although alternative explanations remain possible, we consider theory-based monitoring the most likely source of the introspective switch costs observed in NoGo trials.

5. Conclusion

To conclude, the present study provides new insights into introspection during task switching. Experiment 1 showed that participants can reliably identify task transitions even within a short time window. In Experiment 2, by reducing the possibility to rely on the immediate experience of an executed response, we found that participants still showed introspective switch costs. Therefore, the present study provides evidence that theory-based monitoring (on top of experienced-based knowledge of covert responses) may contribute to introspection during task switching and offers a better understanding of how task-switching performance relates to decision-making.

Declaration***

No generative AI tools were used beyond standard grammar and spell checking tools.

CRediT authorship contribution statement

Jonathan Mendl: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization. **Gesine Dreisbach:** Writing – review & editing, Supervision, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

All raw data files associated with this article are available online under the provided link.

References

Allport, A., & Wylie, G. (2000). Task switching, stimulus-response bindings, and negative priming. In S. Monsell, & J. Driver (Eds.), *Control of cognitive processes: Attention and performance XVIII* (pp. 35–70). MIT Press.

Anderson, F. T., Strube, M. J., & McDaniel, M. A. (2019). Toward a better understanding of costs in prospective memory: A meta-analytic review. *Psychological Bulletin, 145* (11), 1053–1081. https://doi.org/10.1037/bul0000208

Arrington, C. M., & Logan, G. D. (2004). The cost of a voluntary task switch. *Psychological Science, 15*(9), 610–615. https://doi.org/10.1111/j.0956-7976.2004.00728.x

Benini, E., Möller, M., Koch, I., Philipp, A. M., Qiu, R., & Mayr, S. (2023). Evidence of task-triggered retrieval of the previous response: A binding perspective on response-repetition benefits in task switching. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-023-02409-9

Bratzke, D., & Bryce, D. (2019). Introspection is not always blind to the costs of multitasking: The case of task switching. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 45*(6), 980–992. https://doi.org/10.1037/xlm0000635

Bratzke, D., & Bryce, D. (2022). Timing of internal processes: Investigating introspection about the costs of task switching and memory search. *Attention, Perception, & Psychophysics, 84*(5), 1501–1508. https://doi.org/10.3758/s13414-022-02510-6

Bratzke, D., Bryce, D., & Seifried-Dübon, T. (2014). Distorted subjective reports of stimulus onsets under dual-task conditions: Delayed conscious perception or estimation bias? *Consciousness and Cognition, 30*, 36–47. https://doi.org/10.1016/j.concog.2014.07.016

Corallo, G., Sackur, J., Dehaene, S., & Sigman, M. (2008). Limits on introspection: Distorted subjective time during the dual-task bottleneck. *Psychological Science, 19* (11), 1110–1117. https://doi.org/10.1111/j.1467-9280.2008.02211.x

Danziger, K. (1980). The history of introspection reconsidered. *Journal of the History of the Behavioral Sciences, 16*(3), 241–262. https://doi.org/10.1002/1520-6696(198007)16:3<241::AID-JHBS2300160306>3.0.CO;2-O

Dreisbach, G., Haider, H., & Kluwe, R. H. (2002). Preparatory processes in the task-switching paradigm: Evidence from the use of probability cues. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(3), 468–483. https://doi.org/10.1037/0278-7393.28.3.468

Dreisbach, G., & Jurczyk, V. (2022). The role of objective and subjective effort costs in voluntary task choice. *Psychological Research, 86*(5), 1366–1381. https://doi.org/10.1007/s00426-021-01587-2

Dreisbach, G., & Mendl, J. (2024). Flexibility as a matter of context, effort, and ability: Evidence from the task-switching paradigm. *Current Opinion in Behavioral Sciences, 55*, Article 101348. https://doi.org/10.1016/j.cobeha.2023.101348

Dreisbach, G., & Wenke, D. (2011). The shielding function of task sets and its relaxation during task switching. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 37*(6), 1540–1546. https://doi.org/10.1037/a0024077

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Fischer, R., & Janczyk, M. (2022). Dual-task performance with simple tasks. In A. Kiesel, L. Johannsen, I. Koch, & H. Müller (Eds.), *Handbook of human multitasking* (pp. 3–36). Springer International Publishing. https://doi.org/10.1007/978-3-031-04760-2_1.

Fröber, K., & Dreisbach, G. (2017). Keep flexible - keep switching! The influence of forced task switching on voluntary task switching. *Cognition, 162*, 48–53. https://doi.org/10.1016/j.cognition.2017.01.024

Guggenmos, M. (2022). Reverse engineering of metacognition. *ELife, 11*. https://doi.org/10.7554/eLife.75420

Henninger, F., Shevchenko, Y., Mertens, U. K., Kieslich, P. J., & Hilbig, B. E. (2022). Lab. Js: A free, open, online study builder. *Behavior Research Methods, 54*(2), 556–573. https://doi.org/10.3758/s13428-019-01283-5

Hübner, R., & Druey, M. D. (2006). Response execution, selection, or activation: What is sufficient for response-related repetition effects under task shifting? *Psychological Research, 70*(4), 245–261. https://doi.org/10.1007/s00426-005-0219-8

Hübner, R., & Druey, M. D. (2008). Multiple response codes play specific roles in response selection and inhibition under task switching. *Psychological Research, 72*(4), 415–424. https://doi.org/10.1007/s00426-007-0118-2

Kleinsorge, T. (1999). Response repetition benefits and costs. *Acta Psychologica, 103*(3), 295–310. https://doi.org/10.1016/S0001-6918(99)00047-5

Koch, I., & Kiesel, A. (2022). Task switching: Cognitive control in sequential multitasking. In A. Kiesel, L. Johannsen, I. Koch, & H. Müller (Eds.), *Handbook of human multitasking* (pp. 85–143). Springer International Publishing. https://doi.org/10.1007/978-3-031-04760-2_3.

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*(4), 665–682. https://doi.org/10.1037/a0020198

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General, 126*(4), 349–370. https://doi.org/10.1037/0096-3445.126.4.349

Koriat, A. (2012). Metacognition and consciousness. In P. D. Zelazo, M. Moscovitch, & E. Thompson (Eds.), *The Cambridge handbook of consciousness*. Cambridge University Press. https://doi.org/10.1017/CBO9780511816789.012.

Koriat, A., Bjork, R. A., Sheffer, L., & Bar, S. K. (2004). Predicting one's own forgetting: The role of experience-based and theory-based processes. *Journal of Experimental Psychology: General, 133*(4), 643–656. https://doi.org/10.1037/0096-3445.133.4.643

Koriat, A., Nussinson, R., Bless, H., & Shaked, N. (2008). Information-based and experience-based metacognitive judgments: Evidence from subjective confidence. In J. Dunlosky, & R. A. Bjork (Eds.), *Handbook of metamemory and memory* (pp. 117–135). Psychology Press.

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77*(6), 1121–1134. https://doi.org/10.1037/0022-3514.77.6.1121

Mayr, U., & Bell, T. (2006). On how to be unpredictable: Evidence from the voluntary task-switching paradigm. *Psychological Science, 17*(9), 774–780. https://doi.org/10.1111/j.1467-9280.2006.01781.x

Meiran, N., Chorev, Z., & Sapir, A. (2000). Component processes in task switching. *Cognitive Psychology, 41*(3), 211–253. https://doi.org/10.1006/cogp.2000.0736

Mendl, J., & Dreisbach, G. (2022). The role of objective and introspective switch costs in voluntary task choice. *Journal of Experimental Psychology: Human Perception and Performance, 48*(12), 1373–1389. https://doi.org/10.1037/xhp0001051

Mittelstädt, V., Miller, J., & Kiesel, A. (2018). Trading off switch costs and stimulus availability benefits: An investigation of voluntary task-switching behavior in a predictable dynamic multitasking environment. *Memory & Cognition, 46*(5), 699–715. https://doi.org/10.3758/s13421-018-0802-z

Mittelstädt, V., Miller, J., & Kiesel, A. (2019). Linking task selection to task performance: Internal and predictable external processing constraints jointly influence voluntary task switching behavior. *Journal of Experimental Psychology. Human Perception and Performance, 45*(12), 1529–1548. https://doi.org/10.1037/xhp0000690

Monno, I., Spitzer, M., Miller, J., Dignath, D., & Kiesel, A. (2021). Scaling of the parameters for cost balancing in self-organized task switching. *Journal of Cognition, 4*(1), 8. https://doi.org/10.5334/joc.137

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences, 7*(3), 134–140. https://doi.org/10.1016/S1364-6613(03)00028-7

Pavailler, N., Gevers, W., & Burle, B. (2025). Temporal metacognition: Direct readout or mental construct? The case of introspective reaction time. *Journal of Experimental Psychology: General, 154*(4), 1122–1148. https://doi.org/10.1037/xge0001708

Rogers, R. D., & Monsell, S. (1995). Costs of a predictible switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124*(2), 207–231. https://doi.org/10.1037/0096-3445.124.2.207

Schuch, S., & Koch, I. (2004). The costs of changing the representation of action: Response repetition and response-response compatibility in dual tasks. *Journal of Experimental Psychology: Human Perception and Performance, 30*(3), 566–582. https://doi.org/10.1037/0096-1523.30.3.566

Shevchenko, Y. (2022). Open lab: A web application for running and sharing online experiments. *Behavior Research Methods*, 1–8. https://doi.org/10.3758/s13428-021-01776-2

Thompson, T., Davis, H., & Davidson, J. (1998). Attributional and affective responses of impostors to academic success and failure outcomes. *Personality and Individual Differences, 25*(2), 381–396. https://doi.org/10.1016/S0191-8869(98)00065-8

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5*(2), 207–232. https://doi.org/10.1016/0010-0285(73)90033-9

Vermeylen, L., Braem, S., Notebaert, W., & Ruitenberg, M. F. L. (2022). The subjective evaluation of task switch cues is related to voluntary task switching. *Cognition, 224*, Article 105063. https://doi.org/10.1016/j.cognition.2022.105063