Attention, Perception, & Psychophysics (2020) 82:1858–1871
https://doi.org/10.3758/s13414-019-01887-1

Declarative and procedural working memory updating processes are mutually facilitative

Anthony W. Sali[1,2,3] · Tobias Egner[2,3]

Published online: 24 December 2019
© The Psychonomic Society, Inc. 2019

Abstract

Executive function, or cognitive control, describes the ability to guide information processing in line with internal goals, but the nature of—and relationship between—the component processes supporting this ability remains poorly understood. Two key components of cognitive control are thought to be the regulation of the declarative contents of working memory (WM) and the selection of task sets, or procedural rules that determine how declarative items are employed. Factor-analytic studies have suggested that updating the items held in WM and updating task sets are cognitively distinct, but interrelated, core domains of executive function. However, the precise relationship between these processes remains unknown, since they have rarely been tested simultaneously in a single task. In the present study, we devised a novel method of independently manipulating declarative item-updating and procedural task-updating processes in WM. Across two experiments, we found that the updating of declarative and procedural WM representations interacted subadditively, suggesting they are not constrained by a common processing bottleneck. Moreover, in a third experiment, we found that updating two declarative items in WM simultaneously did not incur a behavioral cost in response time above and beyond the cost of one item alone. Taken together, our results provide new evidence that the updating of information in declarative and procedural WM is mutually facilitative, such that opening the gate for updating declarative content reduces the time needed to update procedural content, and vice versa.

Keywords Working memory · Task switching · Executive control · Cognitive and attentional control

Working memory (WM) refers to a limited short-term store of active memory representations that are readily accessible to guide behavior. The regulation of the contents of WM is considered a key function of cognitive control. Due to the limited capacity of WM (e.g., Cowan, 2001; Luck & Vogel, 1997), individuals must flexibly replace outdated representations with those that become relevant as the demands of the environment change (e.g., Frank, Loughry, & O'Reilly, 2001). Two well-studied components of WM are the storage of individual items and the maintenance of task sets that determine the rules by which these item representations are manipulated or employed to influence behavior. Thus, the contents of WM not only reflect behaviorally relevant items (declarative WM content), but also the goals and rules (procedural WM content) that determine how these items ultimately guide behavior (Cole, Ito, & Braver, 2015; Desimone & Duncan, 1995; Rogers & Monsell, 1995).

Individuals must frequently update the declarative contents and procedural task rules that are stored in WM. Although both of these processes incur robust behavioral costs in performance (e.g., Risse & Oberauer, 2010; Rogers & Monsell, 1995), the degree to which executing one operation interferes or benefits the other remains poorly understood. A growing body of factor-analytic work has suggested that updating the declarative and procedural contents of WM might reflect dissociable components of executive functioning (e.g., Friedman et al., 2006; Miyake & Friedman, 2012; Miyake et al., 2000), and an influential model of WM posits that separable stores of declarative and procedural information exist (Oberauer, 2003, 2009; Risse & Oberauer, 2010; Souza, Oberauer, Gade, & Druey, 2012). Here we present three experiments that characterize the behavioral costs associated with simultaneous updates of declarative and procedural representations in WM, to test experimentally the relationship between these vital control processes.

✉ Anthony W. Sali
saliaw@wfu.edu

[1] Department of Psychology, Wake Forest University, Winston-Salem, NC, USA

[2] Center for Cognitive Neuroscience, Duke University, Durham, NC, USA

[3] Department of Psychology and Neuroscience, Duke University, Durham, NC, USA

Dissociating declarative and procedural WM

In Oberauer's (2009) proposed architecture of WM, a declarative system maintains items, such as the digits of a phone number, and a procedural system maintains the tasks that are applied to the items. Within each subsystem, several representations can be activated (accessible) simultaneously but only one representation can be selected by the "internal focus of attention" for declarative memories or the "response focus" for procedural memories at any one time. In line with this proposal, a growing body of research has suggested that the attentional selection of specific task and item representations within WM present individuals with similar limitations, as both task switching and switching the selected item in declarative WM are associated with robust and reliable costs in behavioral performance (Garavan, 1998; Oberauer, 2003; Rogers & Monsell, 1995). When the number of declarative item lists (e.g., multiple lists of digits) is manipulated, participants demonstrate list-switch costs, mixing costs, and, in the face of additional preparation time, residual switch costs, mirroring effects frequently found in studies of task switching (Souza et al., 2012). Moreover, declarative and procedural WM are both associated with *n*–2 repetition costs when participants switch among lists of three or more tasks or items, and interference between list items is high (Gade, Souza, Druey, & Oberauer, 2017; Mayr & Keele, 2000). Thus, there is ample evidence for analogous selection processes of objects and tasks in WM.

Factor-analytic studies of executive functioning have provided additional evidence that declarative and procedural WM representations are cognitively dissociable. Miyake et al. (2000) applied factor analysis to a large set of cognitive control tasks, which produced support for cognitively distinct domains, termed "updating" (changing the declarative contents of WM), "shifting" (flexibly switching between tasks), and "inhibition" (goal-directed stopping of a prepotent response; Friedman, Miyake, Robinson, & Hewitt, 2011; Friedman et al., 2008; Miyake & Friedman, 2012; Miyake et al., 2000). Moreover, subsequent work has shown that these three putative executive functions are stable over the course of development (Mischel et al., 2011; Moffitt et al., 2011) and are differentially related to individual differences in a variety of classic neuropsychological tests such as the Wisconsin Card Sorting Task (Miyake et al., 2000) and to individual differences in IQ (Friedman et al., 2006), thus providing additional credence to the diversity of executive functions. Direct empirical tests of the relationship between declarative and procedural updating processes are lacking, however. Therefore, in the present study, we used a variant of the one-back paradigm to test for possible interaction effects of simultaneous declarative and procedural updating processes in WM.

Declarative and procedural WM gating

In light of existing theories regarding WM updating, the question we sought to address in the present study translates into how the "gating" of information into declarative and procedural WM is related. In addition to regularly updating WM representations, individuals must also maintain these WM representations in a fashion that is resistant to interference from other cognitive processes or external stimuli. A commonly proposed solution to this challenge is a gating mechanism that prevents conflicting information from interfering with WM representations when it is closed, but that opens whenever an update is needed (e.g., Frank et al., 2001; Hazy, Frank, & O'Reilly, 2006; Kessler & Oberauer, 2014, 2015). In declarative WM, it has been proposed that a behavioral cost is associated with opening and closing the gate, with the magnitude of the cost scaling positively with the number of gate opening/closing operations needed on a given trial, and the state of the gate remaining open or closed until the task demands require a change (Kessler & Oberauer, 2015; Rac-Lubashevsky & Kessler, 2016). An equivalent gating mechanism has also been proposed to operate in relation to the updating and shielding of task sets—that is, procedural WM representations (Braver & Cohen, 2000; D'Ardenne et al., 2012; Dreisbach, 2012; Dreisbach, & Wenke, 2011; Kessler, 2017; Waszak, Hommel, & Allport, 2003).

Although the relationship between declarative and procedural gating is currently not well understood, work regarding the attentional selection of declarative and procedural representations in WM provides some indication of the interaction of gating processes. Risse and Oberauer (2010) cued participants on a trial-by-trial basis to select a digit and an arithmetic task to apply to the digit from memory. When both the object–cue and task–cue mappings were variable across time, forcing participants to rely on WM representations of the cue mappings, object- and task-switching costs were additive, suggesting a serial selection bottleneck. However, when at least one of the object–cue or task–cue mappings was consistent across trials, and long-term memory representations could thus aid the recall of at least one dimension, there was a subadditive interaction of object and task switching. The authors interpreted this pattern of results as evidence of parallel processing since the combined cost in response times (RT) associated with switching the selected object and task was less than the sum of the costs associated with switching either dimension alone (Risse & Oberauer, 2010; see also Souza et al., 2012). Thus, the bottleneck that was observed when selecting objects and tasks from WM with variable cues can be overcome when at least one selection is primed by well-learned cue–target associations.

The present study

In the present study, we addressed the novel question of whether and how procedural updating interacts with the need to simultaneously encode new information (a picture of a face) into declarative WM. In line with prior theoretical proposals (e.g., Braver & Cohen, 2000; Dreisbach, 2012) and empirical evidence (e.g., Dreisbach & Wenke, 2011), we assume that cued task switching is synonymous to an updating of a procedural WM representation, even if there are two well-learned task possibilities that participants switch among. We here combine this manipulation of procedural updating with an independent, declarative updating manipulation that requires the maintaining or updating of stimulus category information in WM. Specifically, participants completed a modified version of the one-back task in which they updated a declarative item held in WM, occasionally requiring a change in the categorical nature of the representation, as well as periodically switched a classification task rule, thus requiring them also to update the procedural information held in WM. Our task design allowed us to test whether these costs interacted when categorical updating and task-set shifting were performed simultaneously. The need to update declarative WM information was manipulated at the categorical level (based on face gender and age) and was unpredictable from the participant's perspective, and thus did not represent a cued selection of information already in WM (as had been done in Risse & Oberauer, 2010), but a true updating of WM content with new information. In particular, on some trials the to-be-remembered stimulus category had to be maintained, and on other trials it had to be updated (see the Method sections).

Three possible behavioral patterns would indicate differing underlying processing structures (cf. Risse & Oberauer, 2010; see also Souza et al., 2012): Purely additive costs of declarative and procedural updating would suggest that the two actions are constrained by a shared processing bottleneck, and thus have to be carried out serially (e.g., if only one gating process could take place at a given moment). Conversely, supra-additive costs would indicate that the execution of one operation interferes with an individual's ability to execute the other beyond sharing a common processing bottleneck. This could be the case if, for example, declarative and procedural gating processes were both serial and also require a shared supervisory process that needs to be shifted from one gating process to another. Finally, a subadditive interaction allows for multiple interpretations. As has been outlined in previous studies (e.g., Kessler, 2017; Souza et al., 2012), two processes that are executed, at least partially, in parallel will produce a subadditive interaction, as the total time needed to execute both processes is less than the sum of the times needed to execute each of the processes in isolation. Thus, one interpretation of subadditive effects would be that declarative and procedural gating can take place in parallel—for instance, via two independently operating gating mechanisms. Relatedly, it is possible that executing one updating process (e.g., declarative) facilitates the execution of the other updating process (procedural or a second declarative), referred to here as the *facilitation hypothesis*. For instance, there may be a general cost for initiating any updating/gating process, but that cost is shared among declarative and procedural WM. In the present study, Experiments 1 and 2 probed the interaction of declarative and procedural WM, whereas Experiment 3 provided a comparison with the case in which two declarative updates, rather than one procedural and one declarative one, are required simultaneously.

Experiment 1

In Experiment 1 we sought to test the interaction of declarative and procedural updating processes in WM. To this end, participants completed a variant of a one-back task in which they responded whether a face presented on trial *n* was a categorical match to that presented on trial *n*–1, according to either an age or a gender rule. Critically, we factorially manipulated whether participants needed to update the categorical information in declarative WM (age and gender) and whether they needed to update procedural information to determine whether the current face matched the previous one according to either an age or gender rule.

Method

Participants The sample sizes in similar previous studies (Risse & Oberauer, 2010; Souza et al., 2012) had ranged from 16 to 20 individuals per experiment. Given the potential of slightly noisier data from online collection, we approximately doubled the low end of that range and implemented a stopping rule of recruitment, such that we ran additional participants until we had achieved a useable sample size of 30 participants. Forty-two individuals (24 men, 18 women) ranging in age from 21 to 60 years (M = 35.0, SD = 10.36) completed the experiment and successfully submitted their data on Amazon Mechanical Turk in exchange for monetary compensation. All participants had an Amazon Mechanical Turk approval rating greater than 85% and had successfully completed more than 50 previous assignments. Twelve participants were excluded for having overall behavioral accuracies less than 80%, resulting in a final sample of 30 individuals. Participants agreed to the terms of a Duke University institutional review board (IRB)-approved consent form and received $3.00 as compensation.

Stimuli We selected 412 faces from the 10K Face Database (Bainbridge, Isola, & Oliva, 2013) to serve as the stimuli. These faces, which were taken from the internet and are

intended to represent the general population, had previously been rated along numerous dimensions in an earlier study (Bainbridge et al., 2013). On the basis of these ratings, we selected faces that were easily distinguishable according to age and gender. The images selected for the "younger face" group were all categorized as 20–30 years of age in the Bainbridge ratings. To select images for the "older face" group, we first selected all available images that were rated over 60 years of age. We then added faces that were ranked in the next oldest age group (45–60 years) to complete the stimulus set. Equal numbers of faces that were rated male and female were selected for the younger and older face groups. To simplify the instructions to participants, we asked them to base the young/old judgment on whether the individual appeared younger or older than 40 years of age. The high performance accuracy in Experiment 1 (see below) indicates that the participants in our study judged the stimuli similarly to those who had done the initial ratings. We selected equal numbers of faces (n = 103) falling into each combination of the age and gender categories (i.e., young male, young female, old male, old female). Each face was presented inside an oval aperture and was only presented once over the course of the experiment.

Design and procedure Participants completed a modified one-back paradigm in which they viewed a stream of trial-unique faces and reported the presence or absence of a match between the trial n and trial n–1 stimuli according to their age or gender (see Fig. 1A). Stimulus presentation and response polling were controlled by code written in JavaScript to run in a web browser. With the exception of the first face presented in each run, participants were tasked with determining whether the current stimulus (stimulus n) matched the previous stimulus (stimulus n–1) according to one of two potential rules. Each stimulus was presented for 1,000 ms, followed by a variable intertrial interval that ranged from 2,500 to 3,500 ms. The response window was set to 3,500 ms. A colored border surrounding each face cued the to-be-applied classification rule. Specifically, if the border was red, participants judged whether the current and previous faces were both the same age (where age was defined as less than or greater than 40 years). If the border was blue, participants instead reported whether a match occurred according to the gender of the faces (male vs. female).

Critically, although the identity of the stimulus changed on each trial, consecutive items could match along one or both of the categorical dimensions (age and gender). Thus, our design allowed us to independently manipulate whether the stimulus category held in WM needed to be updated from the previous trial (on category change trials) and, independently, whether the procedural task held in WM must also be updated (on trials in which the task cue changed from the previous trial), creating four conditions: (1) a categorical repeat (e.g., older man/older man) while repeating task set, (2) a categorical update (e.g., older man/younger man) while repeating task set, (3) a categorical repeat (e.g., younger woman/younger woman) while updating task set, or (4) a categorical update (e.g., younger woman/older man) with a simultaneous update of task set. Thus, our design allowed us to test whether the behavioral cost associated with a category update in declarative WM interacted with the cost associated with procedural updating.

The use of trial-unique stimuli ensured that there were never any exact stimulus repetitions across consecutive trials, thus controlling for any influence of stimulus repetition priming. For the purposes of classification, a category update trial was defined as any trial in which there was at least one categorical change across consecutive stimuli. For example, the presentation of a young female face followed by a different young female face served as a category repeat trial. Conversely, a young female face followed by any other combination of age and gender would be classified as a category update. Each factorial combination of category updating and task-set shifting occurred with equal frequency. Participants completed a total of ten blocks consisting of 41 trials each and received behavioral accuracy feedback during self-timed breaks between blocks. The first trial was excluded from all analyses. The total duration of the experiment was approximately 40 min.

Data analysis We subjected the RT data for trials in which participants made an accurate response to a repeated measures analysis of variance (ANOVA) with factors of category operation (repeat vs. update) and task operation (repeat vs. update). Any RTs that were greater or less than three *SD*s from the mean of each condition for each participant were excluded from the analysis, as well as any anticipatory RTs less than 300 ms. In this experiment and in subsequent experiments, the *SD* cutoff was computed prior to removing anticipatory RTs. In total, this procedure resulted in a loss of less than 2% of all trials in Experiment 1 in which an accurate response was made. All data and code for running the analysis and generating figures for this experiment and for all subsequent experiments are available at https://osf.io/5u829/.

Results and discussion

We observed significant main effects of category and task operations on RTs. Participants were slower to respond when there was a categorical change in WM than when the category repeated across consecutive trials, $F(1, 29) = 79.30$, $p < .001$, $\eta_p^2 = .732$ (see Fig. 1B). Furthermore, there was a reliable cost associated with updating task sets, $F(1, 29) = 55.50$, $p < .001$, $\eta_p^2 = .657$. Importantly, we also found a significant subadditive interaction of category updating and task-set updating, $F(1, 29) = 27.91$, p

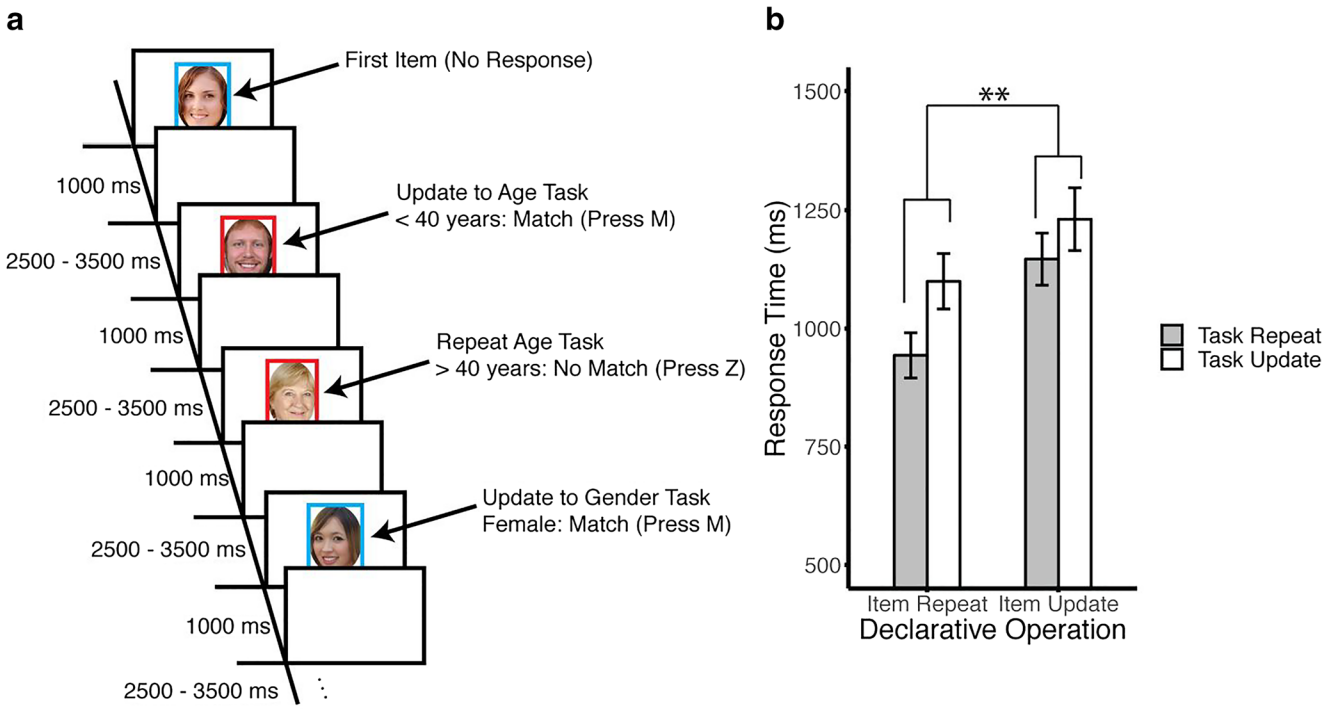

Fig. 1 (A) Behavioral paradigm for Experiment 1. The participants reported whether each stimulus categorically matched the previous stimulus according to one of two potential rules by making a button press (Z vs. M). Red frames cued participants to match according to face age (less than 40 years or greater than 40 years), and blue frames cued participants to match according to face gender (male or female). (B) Behavioral response times in Experiment 1. Error bars denote one between-subjects standard error of the mean. $^{**}p < .001$.

$< .001$, $\eta_p^2 = .490$. The RT cost of updating a declarative category was smaller when a concurrent procedural update was required than when the procedural task set was repeated; or, put the other way around, the cost of updating task sets was reduced when the stimulus category had to be updated, as compared to when the category stayed the same.

When testing performance accuracy, significant main effects of category operation, $F(1, 29) = 40.03$, $p < .001$, $\eta_p^2 = .580$, and task operation, $F(1, 29) = 20.24$, $p < .001$, $\eta_p^2 = .411$, emerged, such that participants were more accurate when categories repeated than when they changed, and also more accurate when tasks repeated than on task update trials (see Table 1). There was no significant interaction of the two control operations, $F(1, 29) = 0.79$, $p = .382$, $\eta_p^2 = .026$.

Table 1 Descriptive statistics for Experiment 1

Condition	Response Times *M* (*SD*)	Accuracies *M* (*SD*)
Repeat category/Repeat task	943.29 (261.14)	96.57 (3.14)
Repeat category/Update task	1,099.87 (320.46)	93.33 (3.36)
Update category/Repeat task	1,146.71 (300.81)	87.47 (7.03)
Update category/Update task	1,231.04 (364.49)	82.87 (11.02)

Taken together, the results of Experiment 1 indicate that executing an updating operation in either declarative or procedural WM did not increase the time needed to execute an updating operation in the other domain, but rather decreased RTs relative to the sum of each control operation in isolation. As we described above, the significant subadditive RT interaction could be attributed to two potential underlying cognitive architectures. First, it is possible that declarative and procedural updating processes in WM are able to proceed, at least partially, in parallel—for instance, via two independent gating mechanisms. This finding would be consistent with models of separable declarative and procedural stores, as well as with factor-analytic work suggesting that the latent variables of updating and shifting are cognitively dissociable (Miyake & Friedman, 2012; Miyake et al., 2000; Oberauer, 2009; Risse & Oberauer, 2010). Alternatively, it is possible that the two operations are not independent but mutually facilitative—for instance, due to reliance on a shared control process such as the opening of a gating mechanism that, once activated, can benefit both types of operations, or due to one updating operation "priming" the other. Before trying to tease apart these two alternative possibilities (in Exp. 3), in Experiment 2 we first sought to replicate the subadditive interaction observed in Experiment 1, while ruling out the possibility that the interaction could have been driven by differences in response uncertainty across conditions.

Experiment 2

In the first experiment, we factorially manipulated declarative category and procedural task updating in WM. We observed robust evidence in favor of a subadditive RT interaction of the two control processes, suggesting that either the updating processes are independent and can be carried out, at least partially, in parallel, and/or that completing one control operation facilitates the completion of the other, concurrent operation. However, in Experiment 1 the button mapping for "match" and "no match" responses remained constant throughout the course of the experiment. Consequently, category repeat trials, regardless of whether or not there was a task switch, were always associated with the same "match" button, whereas category update trials could be associated with either of the response buttons. Importantly, task-switching costs are reduced when motor responses change from one trial to the next, relative to when the response remains the same (e.g., Kleinsorge & Heuer, 1999; Rogers & Monsell, 1995). Since there were equal numbers of category repeat trials and category update trials in Experiment 1, and category repeat trials were always associated with the "match" response, it was more likely that a category repeat trial would require the same button response as the previous trial than would a category update trial. With fewer response changes than in the category update trials, it is possible that task switch costs in the category repeat condition were artificially elevated, thus potentially leading to the observed subadditive interaction. To rule out any possible influence of this imbalance in response uncertainty across trial types, in Experiment 2 we removed the consistent button–response mapping and instead manipulated whether the response mapping needed to be updated on each trial, in addition to whether participants updated the category held in WM and updated task set. Crucially, the design of Experiment 2 allowed us to approximately equate the numbers of instances of each button mapping per condition, remove any potential differences in response uncertainty across conditions, and minimize differences in response repetitions across conditions.

Method

Participants Thirty-five individuals (13 men, 22 women), ranging in age from 19 to 62 years (M = 34.8, SD = 8.54), completed the study and successfully submitted their data in exchange for monetary compensation. Six additional participants were excluded for having previously completed Experiment 1 or an earlier version of this experiment. As in Experiment 1, all participants had an Amazon Mechanical Turk approval rating greater than 85% and had successfully completed more than 50 previous assignments. Of these participants, five were excluded for having overall behavioral accuracies less than 80%, thus resulting in a final sample of 30 individuals. No participants in the final sample had previously completed Experiment 1. All participants agreed to a consent form that was approved by the Duke University IRB and received $4 for participation.

Stimuli We added more faces to those used in Experiment 1, from the same database according to the same criteria described for Experiment 1 (Bainbridge et al., 2013). The resulting set of 880 faces was again balanced, such that 220 faces apiece were classified as "young/male," "young/female," "old/male," and "old/female." The face stimuli presented for each participant were randomly sampled without replacement from this set of images, such that each face could only be presented once in a single session.

Design and procedure All aspects of Experiment 2 were identical to those of Experiment 1, except where noted below. In Experiment 1, participants had always pressed the M key if there was an item match according to the cued rule, or the Z key if there was no match. Conversely, in Experiment 2, we varied this response mapping on a trial-by-trial basis. Specifically, on each trial, a verbal cue was presented at the bottom of the screen along with the stimulus to indicate the relevant response mapping, such that participants either read "Z = No Match, M = Match" or "Z = Match, M = No Match" (see Fig. 2A). Due to the increase in difficulty associated with the variable response mapping, we lengthened the stimulus presentation to 1,500 ms, and the response window was set to 4,000 ms. Since response-mapping updating might interact with category updating and task-set updating, we approximately equated the numbers of trials in which the response mapping repeated or updated for each possible combination of the other factors. Although there were still more response repetitions across consecutive trials for category repeat trials when the response mapping repeated across trials, there were more response repetitions for category update trials when the response mapping also updated. Thus, any residual influence of the number of response repetitions across conditions should manifest as an interaction with the response mapping operation (e.g., repeat vs. update response mapping).

Participants completed ten blocks of 41 trials each. The first trial of each block was excluded from the analysis as in Experiment 1. Furthermore, since the first trial had no button mapping, and thus the second trial's mapping could not repeat or update, the second trial was also excluded. Participants received accuracy feedback and were given a self-paced break between blocks. The experiment lasted approximately 45 min.

Data analysis We again trimmed RTs for all trials with an accurate response that were more than three SDs above or below the mean for each condition for each participant, as well as those that were shorter than 300 ms, resulting in a loss of less than 1% of all trials in which an accurate response was

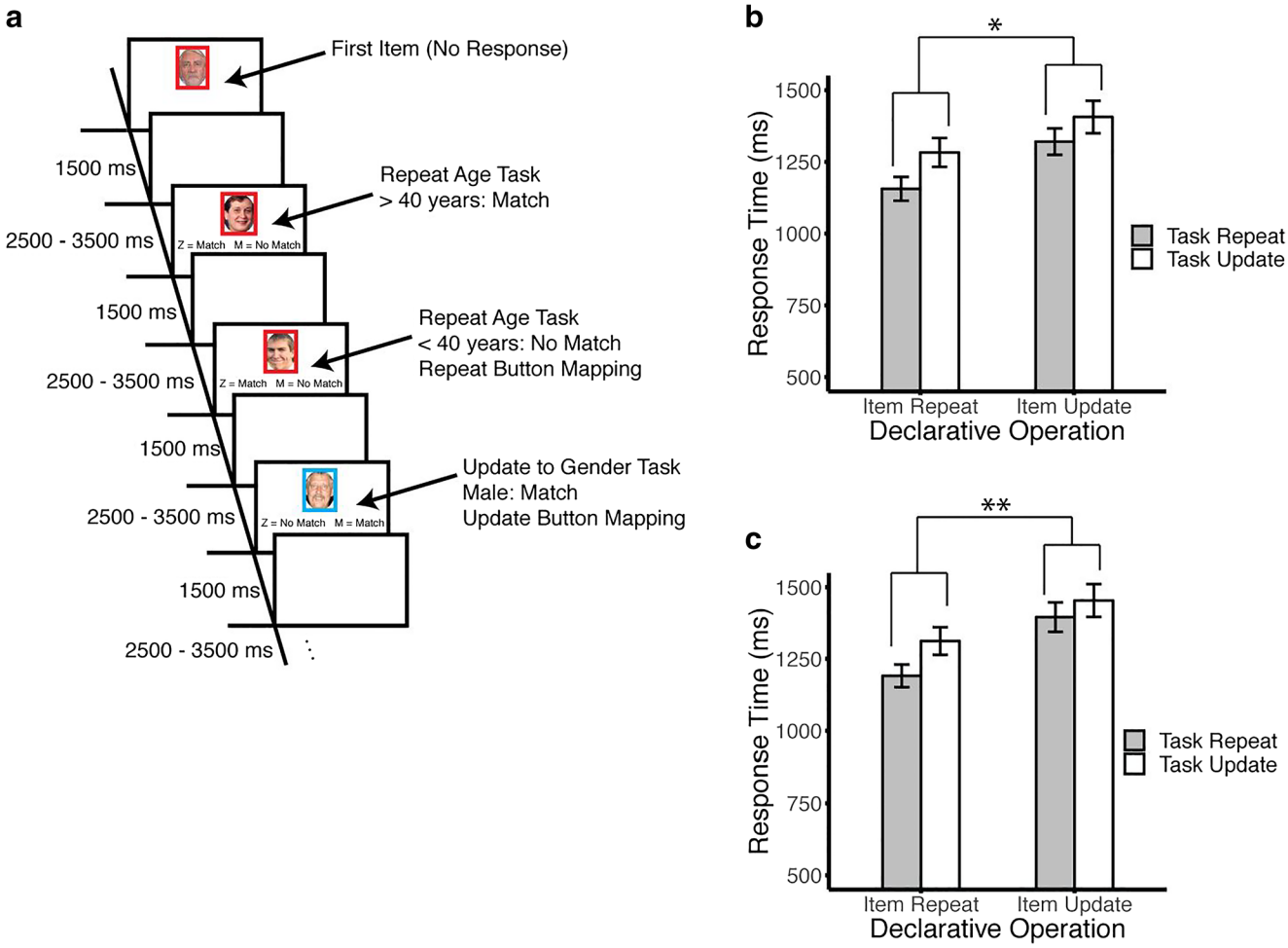

Fig. 2 (A) Behavioral paradigm for Experiment 2. Participants reported whether each stimulus categorically matched the previous stimulus according to one of two potential rules by making a button press (Z vs. M). Red frames cued participants to match according to face age (less than 40 years or greater than 40 years), and blue frames cued participants to match according to face gender (male or female). (B) Behavioral response times for button-mapping repeat trials. (C) Behavioral response times for button-mapping update trials. Error bars denote one between-subjects standard error of the mean. $^{*}p < .05$; $^{**}p < .001$.

made. To evaluate whether individuals can engage in category-updating and task-set-updating processes in parallel when button–response mappings are variable, we subjected the RT data to three-way repeated measures ANOVAs with the factors category operation (repeat vs. update), task operation (repeat vs. update), and response-mapping operation (repeat vs. update). Performance accuracy was not of primary interest for this study's purpose, but we report equivalent ANOVA results on accuracy for completeness' sake.

Results and discussion

We found significant main effects on RTs of the category operation, $F(1, 29) = 72.56$, $p < .001$, $\eta_p^2 = .714$, and the task operation, $F(1, 29) = 65.72$, $p < .001$, $\eta_p^2 = .694$. As in Experiment 1, participants were slower on trials in which the categorical classification of declarative information in working memory updated relative to when the category repeated, and were slower when they updated task set relative to when they repeated the same task set across consecutive trials (see Table 2). Furthermore, there was a significant main effect of response-mapping operation, $F(1, 29) = 54.03$, $p < .001$, $\eta_p^2 = .651$, as participants were slower on trials in which the button mapping changed than on those in which it repeated. Critically, we again found a significant subadditive interaction of the category-updating and task-set-updating operations, $F(1, 29) = 12.69$, $p = .001$, $\eta_p^2 = .304$, such that procedural-updating costs were smaller in the case of a category update than in the case of a category repeat (see Figs. 2b–2c), and this interaction did not vary as a function of whether the button mapping was repeated or updated (i.e., the three-way interaction failed to reach statistical significance), $F(1, 29) = 1.51$, $p = .229$, $\eta_p^2 = .049$. Interestingly, there was also a significant interaction of button-mapping operation and category-updating operation, $F(1, 29) = 5.29$, $p = .029$, $\eta_p^2 = .154$, such that the cost in RTs associated with updating the button-

Table 2 Descriptive statistics for Experiment 2

Condition	Response Times *M* (*SD*)	Accuracies *M* (*SD*)
Repeat category/Repeat task/Button repeat	1,155.98 (227.14)	95.52 (3.66)
Repeat category/Update task/Button repeat	1,282.70 (276.49)	91.58 (6.35)
Update category/Repeat task/Button repeat	1,320.39 (251.68)	89.03 (6.99)
Update category/Update task/Button repeat	1,406.50 (309.72)	75.03 (19.16)
Repeat category/Repeat task/Button update	1,190.69 (215.83)	93.00 (4.06)
Repeat category/Update task/Button update	1,312.42 (264.63)	91.04 (5.79)
Update category/Repeat task/Button update	1,395.79 (281.06)	84.30 (8.74)
Update category/Update task/button update	1,453.59 (314.65)	76.29 (15.19)

mapping was larger when declarative categories updated than when they repeated. The interaction of button mapping and task operations was not significant, $F(1, 29) = 1.37$, $p = .251$, $\eta_p^2 = .045$. Finally, the interaction of category and task operations was statistically significant for button repeat trials, $F(1, 29) = 4.22$, $p = .049$, $\eta_p^2 = .127$, and for button update trials, $F(1, 29) = 18.61$, $p < .001$, $\eta_p^2 = .391$, when tested in two separate ANOVAs.

To ensure that the subadditive interaction of the category and task operations was not due to any remaining differences in response repetitions across conditions, we reran the ANOVA with an added factor accounting for whether or not there was a direct response repetition (e.g., the correct response on two consecutive trials was the Z key). Retrimming the RT data according to the criteria above with the added response repetition factor again resulted in a reduction of less than 1% of trials with an accurate response. When accounting for response repetitions, there was still a significant subadditive interaction of category and task operations, $F(1, 29) = 10.00$, $p = .004$, $\eta_p^2 = .256$, and, importantly, the three-way interaction of response repetition with category operation and task operation, $F(1, 29) = 1.83$, $p = .187$, $\eta_p^2 = .059$, as well as the four-way interaction of response repetition with category operation, task operation, and button-mapping operation, failed to reach statistical significance, $F(1, 29) = 0.05$, $p = .831$, $\eta_p^2 = .002$. Given the relatively low number of observations per cell and the unequal numbers of response repetitions per condition, we also ran the ANOVA with the added factor of response repetition without any outlier trimming, to guard against any influence of unequal trimming across conditions. The outcomes of the three tests noted above remained the same. In sum, we found clear evidence in favor of a subadditive relationship between category and task-set updating/switching in WM, regardless of whether the response mapping was repeated or updated.

An identical analysis of behavioral accuracies also yielded significant main effects of category operation, $F(1, 29) = 41.25$, $p < .001$, $\eta_p^2 = .587$; task operation, $F(1, 29) = 17.30$, $p < .001$, $\eta_p^2 = .374$; and button-mapping operation, $F(1, 29) = 5.53$, $p = .026$, $\eta_p^2 = .160$. Overall, participants were less accurate when performing any of the updating operations. Moreover, there was a significant button-mapping operation by task operation interaction, $F(1, 29) = 13.94$, $p = .001$, $\eta_p^2 = .325$, such that participants demonstrated a greater cost in accuracy when updating task representations when the button mapping repeated across consecutive trials than when it updated. We also observed a supra-additive interaction of category and task operations, $F(1, 29) = 4.72$, $p = .038$, $\eta_p^2 = .140$, such that task-updating costs were larger when there was a categorical change of the declarative information in WM than when the category repeated (see Table 2). Given the subadditive interaction in RTs but the supra-additive interaction in behavioral accuracies, we computed inverse efficiency scores by dividing the mean RT for each condition for each participant by the corresponding proportion of correct responses (Townsend & Ashby, 1978, 1983). When testing these inverse efficiency scores, neither the interaction of category and task operations, $F(1, 29) = 3.27$, $p = .081$, $\eta_p^2 = .101$, nor the interaction of category operation and button operation, $F(1, 29) = 0.06$, $p = .809$, $\eta_p^2 = .002$, was statistically significant, thus not providing strong support for the possibility that the RT effects were mediated by a speed–accuracy trade-off (see Table 3). The remaining two-way and three-way interactions of the accuracy ANOVA failed to reach statistical significance, Fs < 3.09, ps > .089.

When probing RTs for trials in which participants made a correct behavioral response, we again found robust evidence in favor of the parallel processing and/or facilitative accounts of declarative and procedural memory in Experiment 2. Importantly, by varying the button–response mapping on a trial-by-trial basis, Experiment 2 ruled out the possibility that differences in response certainty across category repeat and category update trials could account for the findings of Experiment 1. Furthermore, the lack of a significant three-way interaction of the item, task, and button operations in Experiment 2 suggests that the act of updating button operations did not influence the degree to which declarative and procedural updating interacted. Surprisingly, we found a supra-additive relationship between category updating and

Table 3 Inverse efficiency scores for Experiment 2

Condition	Inverse Efficiency Scores *M* (*SD*)
Repeat category/Repeat task/Button repeat	1,213.63 (258.46)
Repeat category/Update task/Button repeat	1,409.17 (326.77)
Update category/Repeat task/Button repeat	1,492.08 (311.70)
Update category/Update task/Button repeat	2,006.87 (668.89)
Repeat category/Repeat task/Button update	1,281.85 (231.02)
Repeat category/Update task/Button update	1,448.30 (308.21)
Update category/Repeat task/Button update	1,663.29 (311.53)
Update category/Update task/Button update	1,962.91 (479.75)

button-mapping updating. Although it was not the focus of the present study, this finding suggests that the button-mapping instructions on each trial, unlike the task rules themselves, may have been encoded into declarative WM and consequently interfered with participants' abilities to update the categorical information held in WM. Differences in the mnemonic representations of response-mapping instructions as opposed to cognitive task rules, such as the age and gender judgments used here, is beyond the scope of the present study, but it poses an interesting question for future research. In the cognitive neuroscience literature, there is some evidence for distinct brain loci mediating task-goal versus response-set representation (Muhle-Karbe, Andres, & Brass, 2014).

Unlike in Experiment 1, we found a supra-additive interaction of declarative and procedural updating in behavioral accuracies in Experiment 2, such that task-updating costs were magnified when participants simultaneously updated the category held in WM. In other words, the risk of an error was disproportionally higher in the condition in which both declarative and procedural processes were most likely to be erroneous, namely when category and task updating co-occurred. Furthermore, there was no significant supra-additive interaction of declarative and procedural updating when considering inverse efficiency scores, thus suggesting that the RT and accuracy effects stem from differing underlying mechanisms. The supra-additive effect in accuracy might have arisen in Experiment 2 because the added task of reading the trial-by-trial button mapping increased the attentional load. Combined with the subadditive processing time effects, the accuracy data suggest that although simultaneous declarative and procedural updating operations proceed faster than the sum of the processing times needed for the two operations in isolation, the simultaneous application of these operations may nevertheless enhance the likelihood of an erroneous response. If it is replicable, this is an intriguing finding, as it suggests some interdependence between declarative and procedural updating operations in their ultimate impact on the response selection stage. Notably, such an interdependence is consistent with a locus of control that is shared across declarative and procedural WM, as is implied in the facilitative account.

The subadditive RT effects we observed in Experiments 1 and 2 could reflect either independent or mutually facilitative declarative and procedural updating processes. In particular, one WM updating operation might prime the system for another update, regardless of whether the to-be-updated information is declarative or procedural in nature. A somewhat analogous finding has been obtained in studies that have documented smaller costs of switching a motor response when simultaneously updating the task set (Kleinsorge & Heuer, 1999; Korb, Jiang, King, & Egner, 2017). This form of priming is consistent with the possibility of a shared gating mechanism for declarative and procedural WM, such that executing one form of updating opens the gate, which in turn allows the other to proceed with a smaller cost than if the gate needed to be opened a second time. In Experiment 3, we tested this facilitation hypothesis more directly.

Experiment 3

In Experiments 1 and 2, we found that the behavioral cost in RTs associated with simultaneously updating declarative and procedural information in WM interacted subadditively. These results suggest that at the level of processing times, there is some benefit of engaging in both control processes simultaneously. As we stated above, the presence of a subadditive interaction in RT could be attributed to (at least) two plausible underlying cognitive architectures. First, declarative and procedural updating might proceed, at least partially, in parallel—for instance, via two independent gating mechanisms. However, another possibility is that there is a generalized cost the first time a new declarative item or procedural task is brought into WM that is then spared for the immediately following operations, such that the two control operations are mutually facilitative. For instance, if declarative and procedural WM were to share an input gate, once the gate had been opened by the necessity to update declarative content, there would be no additional opening cost for also updating procedural content. In Experiment 3, we explicitly tested this hypothesis by factorially manipulating the number of

simultaneous categorical updates in a given trial while keeping the task rule constant. In particular, if updating categorical information in WM or updating procedural task sets has a generalized benefit for upcoming cognitive switches, we would expect to see a similar behavioral pattern when individuals must perform two categorical updates in WM simultaneously.

Method

Participants Forty-nine participants (28 men, 17 women, four of whom failed to complete the demographic survey), ranging in age from 21 to 60 years ($M = 32.7$, $SD = 9.74$), all of whom had an Amazon Mechanical Turk approval rating that exceeded 85% and had previously completed more than 50 assignments, completed the study on Amazon Mechanical Turk and successfully submitted their data in exchange for monetary compensation. No participants had previously completed Experiment 1 or 2. Two participants were excluded for technical difficulties. We adopted the same accuracy cutoff of 80% as in the previous experiments, resulting in the exclusion of 17 additional participants and yielding a final sample of 30 individuals. As in Experiment 2, all participants agreed to a consent form that was approved by the Duke University IRB and received $4 for participation.

Stimuli The stimulus set was identical to that used in Experiment 2.

Design and procedure All aspects of Experiment 3 were identical to those in Experiment 2, except where noted below. Unlike in Experiments 1 and 2, participants viewed two faces simultaneously for 1,500 ms (see Fig. 3A). On each trial, participants made a button press to indicate whether the gender category of the left and/or right face matched the category of the face presented at each of those locations on the previous trial. Importantly, participants always matched the left and right faces on trial n to the left and right faces, respectively, on trial n–1. Thus, there were four possible conditions: (1) no categorical update for left or right locations, (2) a categorical update for the left location alone, (3) a categorical update for the right location alone, and (4) a categorical update for both locations. Participants used their index, middle, ring, and pinky fingers to press the V key if there were no updates for left or right, the B key if there was a left update only, the N key if there was a right update only, and the M key if both face locations updated. As in the earlier experiments, no response was associated with the first trial. All faces were surrounded by a blue border, in order to match the previous two experiments. Following the presentation of the faces, there was an intertrial interval ranging from 2,500 to 3,500 ms. Participants needed to respond within a 4,000-ms response window.

Participants again completed ten blocks of 41 trials each, and the first trial was thrown out from each block. Participants received accuracy feedback during a self-paced break between each of the blocks. The entire experiment lasted approximately 45 min.

Data analysis As in the previous two experiments, we focused our analysis on RTs for trials in which the participant made an accurate response. We again trimmed RTs that were more than three *SD*s above and below the mean of each condition for each participant, as well as anticipatory RTs that were shorter than 300 ms. This procedure resulted in a loss of less than 2% of all trials in which an accurate response was made. We subjected the trimmed RTs to an ANOVA with the single factor of control operation (zero WM updates, one WM update, two WM updates). We then subjected the accuracy data to an identical ANOVA.

Results and discussion

When testing RTs, we found a significant effect of updating operation, $F(2, 58) = 28.01$, $p < .001$ (Greenhouse–Geisser-corrected for violation of the sphericity assumption), $\eta_p^2 = .491$ (see Fig. 3B). To follow up this significant effect, we ran a series of pairwise comparisons to adjudicate which conditions differed significantly. To account for running multiple tests, we applied a Bonferroni correction, which yielded a corrected critical alpha of .017 for the following three tests. Participants were significantly slowed in both the one-update, $t(29) = 6.25$, $p < .001$, $d = 1.140$, and the two-update, $t(29) = 6.59$, $p < .001$, $d = 1.203$, conditions, relative to the zero-update condition. Critically, after correcting for multiple comparisons, there was no significant difference in RTs between the one-update and two-update conditions, $t(29) = 2.46$, $p = .020$, $d = 0.448$. In fact, RTs were actually numerically faster in the two-update than in the one-update condition, providing no evidence that adding a second declarative update added a significant cost to processing time.

An equivalent ANOVA on accuracy also reached statistical significance, $F(2, 58) = 7.65$, $p = .001$, $\eta_p^2 = .209$. Follow-up comparisons revealed that participants were less accurate for one update than for zero updates, $t(29) = 3.52$, $p = .001$, $d = 0.642$, and for two than for zero updates, $t(29) = 3.01$, $p = .005$, $d = 0.549$. However, there was no difference in behavioral accuracies between the one-update and two-update conditions, $t(29) = 0.79$, $p = .434$, $d = 0.145$ (see Table 4). In sum, the results of Experiment 3 showed a subadditive effect of multi-item updating in declarative WM, as there was a significant performance cost for updating per se, but no difference in cost between the one-update and two-update conditions. This pattern resembles that of the subadditive interaction between declarative and procedural updating operations in Experiments 1 and 2, providing suggestive evidence for a shared gating mechanism.

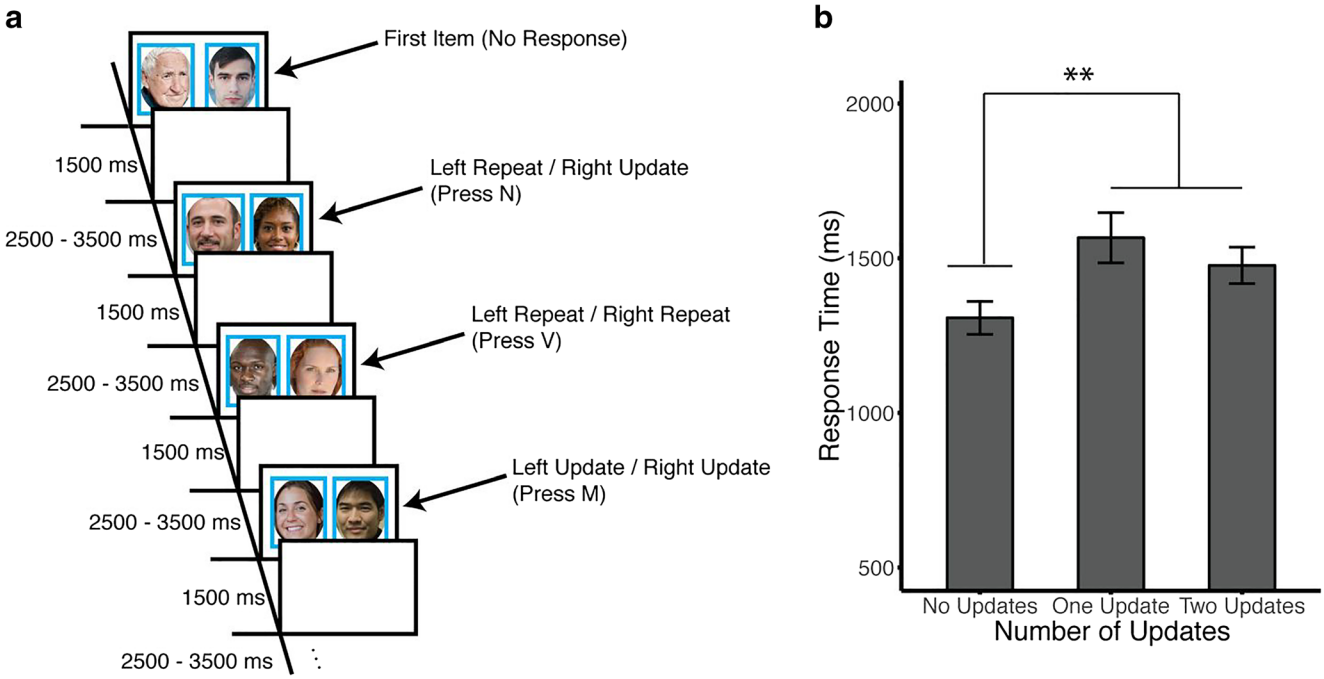

Fig. 3 (A) Behavioral paradigm for Experiment 3. Participants reported whether there were no stimulus updates, a left stimulus update, a right stimulus update, or two stimulus updates with respect to each face's gender. (B) Behavioral response times. Error bars denote one between-subjects standard error of the mean. $^{**}p < .001$.

General discussion

In the present study, we interrogated the relationship between declarative category and procedural task updating processes in WM. Across the first two experiments, we found no evidence of compounding behavioral costs in RTs for category and task-set updating. Instead, we found robust evidence in favor of subadditive costs that was not related to a speed–accuracy trade-off. In particular, the cost associated with updating procedural task sets in WM was smaller when there was a categorical update of declarative information in WM than would have been expected if the category and procedural updating costs were merely summed together. Relatedly, we again found a subadditive relationship in Experiment 3, suggesting that updating a declarative item in WM temporarily reduces the cost of executing a second declarative update.

The present study demonstrates that just as attentional selections from the declarative and procedural subsystems can occur in parallel (e.g., Risse & Oberauer, 2010), updating information in one (putative) subsystem does not interfere with updating information in the other. In particular, the costs associated with updating task-relevant information in the declarative store, such as what occurs when there is a change in the category of a face held in memory, did not compound the costs associated with updating the task held in the procedural store. Thus, the regulation of information in declarative and procedural stores does not interfere with each other (e.g., Montojo & Courtney, 2008; Souza et al., 2012).

A gating mechanism, potentially mediated by the basal ganglia, is thought to regulate the maintenance and updating of WM, shielding WM representations from unimportant and potentially distracting information during closed-gate states, but opening to allow behaviorally relevant information to enter WM from either perception or from long-term memory when needed (Braver & Cohen, 2000; Frank et al., 2001; O'Reilly, 2006; O'Reilly & Frank, 2006). One potential interpretation of our results is that since procedural updating required the selection of representations that were already maintained in long-term memory, they could proceed in parallel with the opening of the gate to allow new information into declarative WM. This interpretation is consistent with Souza et al.'s (2012) model of WM, in which there are activated regions of declarative and procedural WM, termed the "region of direct access" and the "bridge," respectively, from which selections are ultimately made. However, the accuracy results of Experiment 2 also suggest some form of interdependent detrimental impact on appropriate response selection in the case in which both declarative and procedural subprocesses are more error-prone (i.e., when both domains require an

Table 4 Descriptive statistics for Experiment 3

Condition	Response Times *M* (*SD*)	Accuracies *M* (*SD*)
Zero updates	1,306.91 (291.12)	91.47 (4.77)
One update	1,566.41 (445.94)	87.85 (6.44)
Two updates	1,476.86 (324.44)	88.60 (6.07)

update). The source of the latter effect—if it proves reliable—represents an interesting target for future investigations, perhaps employing neuroimaging, which would facilitate the teasing apart of the impact of our manipulations on perceptual versus central versus motor processing stages.

Although it is possible that we observed an underadditive interaction because the gate to declarative WM was free to open and close independently of participants' selection of tasks from procedural WM, another possibility is that a shared gate regulates the encoding of information in WM and shields those representations from external as well as internal interference. Although there were only two tasks that participants repeatedly switched between in the present study, a growing body of work suggests that switching tasks involves opening a gate to procedural WM, as distractors have been shown to interfere substantially more with performance during task switch than repeat trials (e.g., Dreisbach, 2012; Dreisbach & Wenke, 2011). To prevent interference from irrelevant dimensions of some stimulus (e.g., the gender of a face when age is the relevant task rule), participants may engage in a process of task shielding in which a (closed) gate blocks out associations between the stimulus and other potential responses (Dreisbach & Haider, 2008). The importance of such a mechanism is apparent from studies showing that participants develop stimulus-task bindings that—when weak—insufficiently prevent interference from unwanted stimulus associations (Waszak et al., 2003). Given this resilience to interference, switching between tasks requires a momentary relaxation of task shielding, which may contribute to the behavioral cost typically associated with updating tasks (Dreisbach, & Wenke, 2011; Kessler, 2017; Kessler, Baruchin, & Bouhsira-Sabag, 2017). An additional explanation of the present findings is therefore that opening the gate to declarative or procedural WM allows new information to pass into the other WM subsystem momentarily without incurring the cost of reopening a gate. However, further research will be needed to determine whether there is a shared single gate for declarative and procedural WM or whether there are separate, but parallel, procedural and declarative gates that, once opened, allow multiple updates in a single domain (as in Exp. 3) to occur without additional behavioral costs.

Given that the results of the present study provide evidence in favor of independent declarative and procedural WM subsystems or a facilitative relationship between WM updating across domains, an important question for future research is the degree to which intrinsic and external factors modulate the efficacy of control operations in each system. There is variability in the degree of cognitive flexibility, both across individuals (e.g., Bertolino et al., 2006; Cools, 2008; Cools & D'Esposito, 2011; Heatherton & Wagner, 2011; Nolan, Bilder, Lachman, & Volavka, 2004) and within individuals (Leber, Turk-Browne, & Chun, 2008; Sali, Courtney, & Yantis, 2016). However, despite research on the internal factors that determine cognitive flexibility, it is presently unclear whether flexibility over manipulating declarative and procedural information in WM vary together in a predictable fashion. Likewise, an important capacity of cognitive control is to flexibly adapt to the properties of the environment. Previous research has demonstrated that individuals display increased cognitive flexibility in contexts associated with frequent updating than in contexts associated with infrequent updating (Chiu & Egner, 2017; Crump & Logan, 2010; Dreisbach & Haider, 2006; Leboe, Wong, Crump, & Stobbe, 2008; Monsell & Mizon, 2006; Sali, Anderson, & Yantis, 2015). An interesting topic for future research is thus whether the flexibility of declarative and procedural control processes vary independently according to environmental demands. Independence in this kind of control learning would argue that, regardless of any mutual facilitation, a dissociation exists between declarative and procedural WM updating operations.

The participants in the present study updated procedural WM by selecting one of two potential task sets from long-term memory during each trial on the basis of a visual cue. In contrast, individuals updated and stored properties of a novel and unpredictable stimulus in declarative memory on each trial. Given that recurring task sets were used throughout each experiment, we cannot rule out the possibility that participants, at least partially, relied on long-term memory representations of the procedural rules. The constancy of the task rules may have allowed participants to move a procedural rule from long-term memory into the focus of WM and encode a new item into declarative WM in parallel. However, individuals are able to rapidly learn novel task rules (e.g., Cole, Bagic, Kass, & Schneider, 2010; Cole, Laurent, & Stocco, 2013). An important consideration for future research is thus whether category updating and task-set updating processes still proceed in parallel when individuals must periodically store a novel task set in WM rather than retrieve a set that has already been prepared. By broadening the potential set of tasks to include rules that are novel to the participant, future research may better define the boundary conditions in which declarative and procedural updating are facilitative.

The carving of executive functioning into its composite parts has important implications for understanding healthy as well as disordered variability in behavior across individuals. For example, individual differences in updating the declarative contents of WM are associated with both intelligence (Friedman et al., 2006) and childhood deficits in attentional control (Friedman et al., 2007). Moreover, hallmark symptoms of attention deficit hyperactivity disorder (ADHD) include inattention and impulsivity (e.g., Barkley, 1997), both of which may be viewed as the consequence of extreme and inappropriate cognitive flexibility (e.g., Cools, 2008). In the present study we investigated the relationship between two fundamental dimensions of control that are central to many everyday tasks. A better understanding of the relationship between updating and shifting processes may therefore aid in determining the specific deficits associated with common disorders.

Conclusions

In the present study, we tested a manipulation of declarative and procedural updating simultaneously, in the same paradigm. Across two experiments, we found no evidence of compounding behavioral costs in processing time when declarative and procedural control processes were carried out on the same trial, suggesting that completing one operation does not delay an individual's ability to simultaneously complete the other. In a subsequent experiment, we found that completing one declarative update in WM facilitates a second, simultaneous declarative update, replicating the pattern for simultaneous declarative and procedural updates. Taken together, the results add to our understanding of two key components of executive functioning by suggesting that declarative updating and procedural set shifting are not constrained by a common serial-processing bottleneck. Instead, our results suggest that the gating of information into declarative and procedural WM is mutually beneficial.

Open Practices Statement The data and analysis code for all of the experiments are available at https://osf.io/5u829/, and none of the experiments were preregistered.

Author Note This research was supported in part by National Institutes of Health Grant R01-MH087610 to T.E and in part by US-Israel Binational Science Foundation (BSF) grant 2016234 to TE. We thank Yoav Kessler for helpful comments on an earlier draft of this manuscript.

References

Bainbridge, W. A., Isola, P., & Oliva, A. (2013). The intrinsic memorability of face photographs. *Journal of Experimental Psychology: General*, *142*, 1323–1334. https://doi.org/10.1037/a0033872

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin*, *121*, 65–94. https://doi.org/10.1037/0033-2909.121.1.65

Bertolino, A., Blasi, G., Latorre, V., Rubino, V., Rampino, A., Sinibaldi, L., … Dallapiccola, B. (2006). Additive effects of genetic variation in dopamine regulating genes on working memory cortical activity in human brain. *Journal of Neuroscience*, *26*, 3918–3922. https://doi.org/10.1523/jneurosci.4975-05.2006

Braver, T. S., & Cohen, J. D. (2000). On the control of control: The role of dopamine in regulating prefrontal function and working memory. In S. Monsell & J. Driver (Eds.), *Control of cognitive processes: Attention and performance XVIII* (pp. 713–737). Cambridge, MA: MIT Press.

Chiu, Y.-C., & Egner, T. (2017). Cueing cognitive flexibility: Item-specific learning of switch readiness. *Journal of Experimental Psychology: Human Perception and Performance*, *43*, 1950–1960. https://doi.org/10.1037/xhp0000420

Cole, M. W., Bagic, A., Kass, R., & Schneider, W. (2010). Prefrontal dynamics underlying rapid instructed task learning reverse with practice. *Journal of Neuroscience*, *30*, 14245–14254. https://doi.org/10.1523/jneurosci.1662-10.2010

Cole, M. W., Ito, T., & Braver, T. S. (2015). The behavioral relevance of task information in human prefrontal cortex. *Cerebral Cortex*, *26*, 2497–2505. https://doi.org/10.1093/cercor/bhv072

Cole, M. W., Laurent, P., & Stocco, A. (2013). Rapid instructed task learning: A new window into the human brain's unique capacity for flexible cognitive control. *Cognitive, Affective, & Behavioral Neuroscience*, *13*, 1–22. https://doi.org/10.3758/s13415-012-0125-7

Cools, R. (2008). Role of dopamine in the motivational and cognitive control of behavior. *Neuroscientist*, *14*, 381–395. https://doi.org/10.1177/1073858408317009

Cools, R., & D'Esposito, M. (2011). Inverted-U-shaped dopamine actions on human working memory and cognitive control. *Biological Psychiatry*, *69*, e113–e125. https://doi.org/10.1016/j.biopsych.2011.03.028

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*, 87–185. https://doi.org/10.1017/S0140525X01003922

Crump, M. J. C., & Logan, G. D. (2010). Contextual control over task-set retrieval. *Attention, Perception, & Psychophysics*, *72*, 2047–2053. https://doi.org/10.3758/BF03196681

D'Ardenne, K., Eshel, N., Luka, J., Lenartowicz, A., Nystrom, L. E., & Cohen, J. D. (2012). Role of prefrontal cortex and the midbrain dopamine system in working memory updating. *Proceedings of the National Academy of Sciences*, *109*, 1900–1909.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222. https://doi.org/10.1146/annurev.ne.18.030195.001205

Dreisbach, G. (2012). Mechanisms of cognition control: The functional role of task rules. *Current Directions in Psychological Science*, *21*, 227–231. https://doi.org/10.1177/0963721412449830

Dreisbach G., & Haider, H. (2006). Preparatory adjustment of cognitive control in the task switching paradigm. *Psychonomic Bulletin & Review*, *13*, 334-338

Dreisbach, G., & Haider, H. (2008). That's what task sets are for: Shielding against irrelevant information. *Psychological Research*, *72*, 355–361.

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*, 1540–1546.

Frank, M. J., Loughry, B., & O'Reilly, R. C. (2001). Interactions between frontal cortex and basal ganglia in working memory: A computational model. *Cognitive, Affective, & Behavioral Neuroscience*, *1*, 137–160. https://doi.org/10.3758/CABN.1.2.137

Friedman, N. P., Haberstick, B. C., Willcutt, E. G., Miyake, A., Young, S. E., Corley, R. P., & Hewitt, J. K. (2007). Greater attention problems during childhood predict poorer executive functioning in late adolescence. *Psychological Science*, *18*, 893–900. https://doi.org/10.1111/j.1467-9280.2007.01997.x

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., Defries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science*, *17*, 172–179. https://doi.org/10.1111/j.1467-9280.2006.01681.x

Friedman, N. P., Miyake, A., Robinson, J. L., & Hewitt, J. K. (2011). Developmental trajectories in toddlers' self-restraint predict individual differences in executive functions 14 years later: A behavioral genetic analysis. *Developmental Psychology*, *47*, 1410–1430. https://doi.org/10.1037/a0023750

Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology: General*, *137*, 201–225. https://doi.org/10.1037/0096-3445.137.2.201

Gade, M., Souza, A. S., Druey, M. D., & Oberauer, K. (2017). Analogous selection processes in declarative and procedural working memory: *N*–2 list-repetition and task-repetition costs. *Memory & Cognition*, *45*, 26–39. https://doi.org/10.3758/s13421-016-0645-4

Garavan, H. (1998). Serial attention within working memory. *Memory & Cognition*, *26*, 263–276. https://doi.org/10.3758/BF03201138

Hazy, T. E., Frank, M. J., & O'Reilly, R. C. (2006). Banishing the homunculus: Making working memory work. *Neuroscience*, *139*, 105–118.

Heatherton, T. F., & Wagner, D. D. (2011). Cognitive neuroscience of self-regulation failure. *Trends in Cognitive Sciences*, *15*, 132–139. https://doi.org/10.1016/j.tics.2010.12.005.

Kessler, Y. (2017). The role of working memory gating in task switching: A procedural version of the reference-back paradigm. *Frontiers in Psychology*, *8*, 2260. https://doi.org/10.3389/fpsyg.2017.02260

Kessler, Y., Baruchin, L. J., & Bouhsira-Sabag, A. (2017). Working memory updating occurs independently of the need to maintain task-context: Accounting for triggering updating in the AX-CPT paradigm. *Psychological Research*, *81*, 191–203. https://doi.org/10.1007/s00426-015-0717-2

Kessler, Y., & Oberauer, K. (2014). Working memory updating latency reflects the cost of switching between maintenance and updating modes of operation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 738–754. https://doi.org/10.1037/a0035545

Kessler, Y., & Oberauer, K. (2015). Forward scanning in verbal working memory updating. *Psychonomic Bulletin & Review*, *22*, 1770–1776. https://doi.org/10.3758/s13423-015-0853-0

Kleinsorge, T., & Heuer, H. (1999). Hierarchical switching in a multidimensional task space. *Psychological Research*, *62*, 300–312. https://doi.org/10.1007/s004260050060

Korb, F. M., Jiang, J., King, J. A., & Egner, T. (2017). Hierarchically organized medial frontal cortex-basal ganglia loops selectively control task-and response-selection. *Journal of Neuroscience*, *37*, 7893–7905.

Leber, A. B., Turk-Browne, N. B., & Chun, M. M. (2008). Neural predictors of moment-to-moment fluctuations in cognitive flexibility. *Proceedings of the National Academy of Sciences*, *105*, 13592–13597. https://doi.org/10.1073/pnas.0805423105

Leboe, J. P., Wong, J., Crump, M., & Stobbe, K. (2008). Probe-specific proportion task repetition effects on switching costs. *Perception & Psychophysics*, *70*, 935–945. https://doi.org/10.3758/pp.70.6.935

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*, 279–281. https://doi.org/10.1038/36846

Mayr, U., & Keele, S. W. (2000). Changing internal constraints on action: The role of backward inhibition. *Journal of Experimental Psychology: General*, *129*, 4–26. https://doi.org/10.1037/0096-3445.129.1.4

Mischel, W., Ayduk, O., Berman, M. G., Casey, B. J., Gotlib, I. H., Jonides, J., … Shoda, Y. (2011). "Willpower" over the life span: Decomposing self-regulation. *Social Cognitive and Affective Neuroscience*, *6*, 252–256. https://doi.org/10.1093/scan/nsq081

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: four general conclusions. *Current Directions in Psychological Science*, *21*, 8–14. https://doi.org/10.1177/0963721411429458

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49–100. https://doi.org/10.1006/cogp.1999.0734

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., . . . Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences*, *108*, 2693–2698. https://doi.org/10.1073/pnas.1010076108

Monsell, S., & Mizon, G. A. (2006). Can the task-cuing paradigm measure an endogenous task-set reconfiguration process? *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 493–516. https://doi.org/10.1037/0096-1523.32.3.493

Montojo, C. A., & Courtney, S. M. (2008). Differential neural activation for updating rule versus stimulus information in working memory. *Neuron*, *59*, 173–182. https://doi.org/10.1016/j.neuron.2008.05.012

Muhle-Karbe, P. S., Andres, M., & Brass, M. (2014). Transcranial magnetic stimulation dissociates prefrontal and parietal contributions to task preparation. *Journal of Neuroscience*, *34*, 12481–12489. https://doi.org/10.1523/jneurosci.4931-13.2014

Nolan, K. A., Bilder, R. M., Lachman, H. M., & Volavka, J. (2004). Catechol *O*-methyltransferase Val^{158}Met polymorphism in schizophrenia: Differential effects of Val and Met alleles on cognitive stability and flexibility. *American Journal of Psychiatry*, *161*, 359–361. https://doi.org/10.1176/appi.ajp.161.2.359

O'Reilly, R. C. (2006). Biologically based computational models of high-level cognition. *Science*, *314*, 91–94.

O'Reilly, R. C., & Frank, M. J. (2006). Making working memory work: A computational model of learning in the prefrontal cortex and basal ganglia. *Neural Computation*, *18*, 283–328. https://doi.org/10.1162/089976606775093909

Oberauer, K. (2003). Selective attention to elements in working memory. *Experimental Psychology*, *50*, 257–269. https://doi.org/10.1026//1618-3169.50.4.257

Oberauer, K. (2009). Design for a working memory. In B. H. Ross (Eds.), *The psychology of learning and motivation* (Vol. 51, pp. 45–100). San Diego, CA, US: Elsevier Academic Press.

Rac-Lubashevsky, R., & Kessler, Y. (2016). Dissociating working memory updating and automatic updating: The reference-back paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*, 951–969.

Risse, S., & Oberauer, K. (2010). Selection of objects and tasks in working memory. *Quarterly Journal of Experimental Psychology*, *63*, 784–804. https://doi.org/10.1080/17470210903147486

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

Sali, A. W., Anderson, B. A., & Yantis, S. (2015). Learned states of preparatory attentional control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*, 1790–1805. https://doi.org/10.1037/xlm0000146

Sali, A. W., Courtney, S. M., & Yantis, S. (2016). Spontaneous fluctuations in the flexible control of covert attention. *Journal of Neuroscience*, *36*, 445–454. https://doi.org/10.1523/jneurosci.2323-15.2016

Souza, A. S., Oberauer, K., Gade, M., & Druey, M. D. (2012). Processing of representations in declarative and procedural working memory. *Quarterly Journal of Experimental Psychology*, *65*, 1006–1033. https://doi.org/10.1080/17470218.2011.640403

Townsend, J. T., & Ashby, F. G. (1978). Methods of modeling capacity in simple processing systems. In N. J. Castellan Jr. & F. Restle (Eds.), *Cognitive theory* (Vol. 3, pp. 200–239). Hillsdale, NJ: Erlbaum.

Townsend J. T., & Ashby F. G. (1983). The stochastic modeling of elementary psychological processes. Cambridge: Cambridge University Press.

Waszak, F., Hommel, B., & Allport, A. (2003). Task-switching and long-term priming: Role of episodic stimulus-task bindings in task-shift costs. *Cognitive Psychology*, *46*, 361–413. https://doi.org/10.1016/S0010-0285(02)00520-0

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.