Journal of Experimental Psychology:
Learning, Memory, and Cognition

ISSN: 0278-7393

2024, Vol. 50, No. 2, 287–305
https://doi.org/10.1037/xlm0001195

Working Memory Capacity Preferentially Enhances Implementation of Proactive Control

Yanli Lin, Rachel E. Brough, Allison Tay, Joshua J. Jackson, and Todd S. Braver
Department of Psychological and Brain Sciences, Washington University in St. Louis

Previous research has linked working memory capacity (WMC) with enhanced proactive control. However, it remains unclear the extent to which this relationship reflects the influence of WMC on the tendency to engage proactive control, or rather, the ability to implement it. The current study sought to clarify this ambiguity by leveraging the Dual Mechanisms of Cognitive Control (DMCC) version of the AX-CPT task, in which the mode of cognitive control is experimentally manipulated across distinct testing sessions. To adjudicate between competing hypotheses, Bayesian mixed modeling was used to conduct sequential analyses involving two separate data sets. Posterior parameter estimates obtained from the initial analysis were entered as informed priors during the replication analysis to evaluate the influence of new data on previous estimates. Results yielded strong evidence demonstrating that the influence of WMC on proactive control is most robust under experimentally controlled conditions, during which use of proactive control is standardized across participants via explicit training and instruction. Critically, the observed pattern of findings suggests that the relationship between WMC and proactive control may be better characterized as individual differences in the ability to implement proactive control, rather than a more generalized tendency to engage it.

Keywords: cognition, cognitive control, dual mechanisms of control, individual differences

Cognitive control is a fundamental ability that enables coordination and adaptive execution of goal-directed behavior (Egner, 2017). Given the central role of cognitive control in navigating wide domains of human functioning, it is unsurprising that control is considered to be a dynamical process sensitive to changing demands and circumstances related to the environmental context. One theory which aims to systematically parse this variability is the Dual Mechanisms of Cognitive Control (DMC) framework (Braver, 2012; Braver et al., 2021). The DMC proposes that cognitive control can be deployed in two distinct modes: proactive and reactive. Proactive control involves preparatory, sustained activation of task rules and goal representations; whereas reactive control involves transient, stimulus-driven task/goal activation.

Importantly, each mode of control is theorized to confer unique costs and advantages, such that successful use is likely to require flexible adoption of both control strategies. Conversely, disruption of this adaptability has been shown to underlie psychological impairment and behavioral dysfunction. For example, psychological and neurocognitive disorders such as schizophrenia and dementia have been associated with diminished use of proactive control (Barch & Ceaser, 2012; Braver et al., 2005), whereas developmental risk for anxiety has been linked to overreliance on reactive control (Troller-Renfree et al., 2019). Furthermore, individual differences in personality and cognitive ability are likewise thought to exert systemic influence on control dynamics (Braver, 2012; Braver et al., 2007). Within this domain, one of the most well studied individual difference constructs is working memory capacity (WMC), commonly defined as the ability to temporarily store, manipulate, and retrieve goal-relevant information (Unsworth & Engle, 2008).

Investigating WMC and Cognitive Control Using the AX-CPT

In particular, a growing number of studies have leveraged the AX version of the continuous performance test (AX-CPT; Braver et al., 2001; Servan-Schreiber et al., 1996), to investigate the influence of WMC on cognitive control strategy use. Briefly, the AX-CPT is a widely used task of context processing, during which participants respond to trials of sequential cue-probe letter pairs. A target response is required only when an A cue is followed by an X (AX trials); conversely, a nontarget response is required for all other cue-probe pairs including AY, BX, and BY trials such that B and Y represent any letters except A and X. Proactive control is indexed by preparatory response patterns influenced by the A cue (e.g., committing more correct nontarget responses, following a

This article was published Online First November 3, 2022.

Yanli Lin https://orcid.org/0000-0002-2931-9096
Rachel E. Brough is now at Department of Psychology, University of Denver.

We thank Erin Gourley for her contributions to task programming and data collection for the DMCC project and Rongxiang Tang for her many contributions to data collection and initial analyses of this dataset. This research was supported by National Institutes of Health Grants R37 MH066078, F32 AG069499 and through funds provided by the McDonnell Center for Systems Neuroscience at Washington University in St. Louis.

Online Data Sharing & Repositories: https://osf.io/zuvry/ (Lin et al., 2022).

Correspondence concerning this article should be addressed to Yanli Lin, Department of Psychological and Brain Sciences, Washington University in St. Louis, Somers Family Hall St. Louis, MO 63105. Email: lyanli@wustl.edu

non-A cue, such as on BX trials, but more incorrect target responses on AY trials), whereas reactive control is indexed by "just-in-time" response patterns that are influenced by the X probe (e.g., committing more correct nontarget responses on AY trials, but slower and potentially more error-prone responses on BX trials).

Using this paradigm, investigators examining the relationship between WMC and cognitive control have consistently reported that adult individuals with higher WMC exhibit greater use of proactive control relative to lower WMC participants (Belletier et al., 2019; Boudewyn et al., 2015; Redick, 2014; Redick & Engle, 2011; Richmond et al., 2015; Stawarczyk et al., 2014; Wiemers & Redick, 2018). There is also emerging evidence that this relationship extends to young children and may begin to develop as early as age 5 (Gonthier et al., 2019; Troller-Renfree et al., 2020; Wang et al., 2021). Importantly, these findings provide empirical support for the theoretical implications of WMC in the context of the DMC framework—namely, that WMC enables active maintenance of task and goal relevant information (see also Engle et al., 1999; Engle & Kane, 2004; Kane et al., 2001; Kane & Engle, 2003; for other related theoretical frameworks); and because proactive control essentially relies on this capacity, it stands to reason that higher WMC should facilitate the use of proactive control.

Conceptual Ambiguities and Methodological Challenges

With that said, the studies to date have primarily relied on extreme-group comparison (Redick, 2014; Redick & Engle, 2011; Wiemers & Redick, 2018) or correlational (Belletier et al., 2019; Boudewyn et al., 2015; Gonthier et al., 2019; Richmond et al., 2015; Stawarczyk et al., 2014; Troller-Renfree et al., 2020; Wang et al., 2021) designs. Although it is instructive to compare AX-CPT performance across low versus high WMC individuals, the salient limitation of these methods is that it precludes the ability to discriminate between the *tendency* to engage in proactive control and the *ability* to implement it. Without stringent experimental controls to standardize control strategy use and establish appropriate comparisons, the admittedly porous but potentially important boundary conditions between the tendency and ability to implement different modes of control is likely to remain analytically indistinguishable. Specifically, experimental manipulations that involve explicit instruction of proactive/reactive control are critical for interpreting observed individual differences in ability. This is because failure to specify and standardize cognitive control strategy use leaves open the unwanted possibility that participants may spontaneously vary their mode of control during task performance. In other words, under conditions for which cognitive control strategy is not explicitly instructed and/or constrained, it is possible that observed individual differences may reflect the tendency to spontaneously adopt a control mode, or to flexibly switch among them, rather than the ability to implement a singular mode of control when the task requires it.

In addition to resolving these ambiguities, a related but broader challenge involves circumventing the psychometric and analytic limitations commonly associated with using cognitive experimental behavioral paradigms when conducting individual differences research. Here, we briefly highlight four pervasive and well-documented problems. First, use of conventional summary scores derived from averaging across aggregated trial performance overlooks trial-level variability, and as a consequence, can result in poor reliability and underestimated effect sizes (Rouder & Haaf, 2019; Snijder et al., 2022). Second, the equally common practice of using subtraction-based difference scores can increase measurement error, constrain between-subjects variance, and attenuate reliability (Caruso, 2004; Cronbach & Furby, 1970; Draheim et al., 2019; Hedge et al., 2018). Third, classic experimental psychology statistical approaches such as ANOVA and ANCOVA often violate assumptions of independence, and do not model subject-level variability as a unique source of variance, potentially leading to overestimated effect sizes and increased risk of type I error (Judd et al., 2012; Singmann & Kellen, 2019). Fourth, the practice of null hypothesis testing (NHST), which has been subject to increasing scrutiny (Cumming, 2014; Halsey et al., 2015; Ioannidis, 2005; Simmons et al., 2011), only weighs evidence against the plausibility of the null hypothesis but not evidence in favor the alternative hypothesis, precluding inferential evaluation of both the existence and magnitude of the hypothesized effect (Wagenmakers, 2007; Wagenmakers et al., 2018).

Study Rationale

With these challenges in mind, the primary aims of the present study are twofold: (a) to clarify the conceptual ambiguities regarding the nature of the relationship between WMC and proactive control; and (b) to remediate the common methodological issues pertaining to measurement and analysis described above. As alluded to above, the key to achieving the first aim lies in deriving methods to place cognitive control mode under direct experimental manipulation. Toward this end, one of the most comprehensive efforts to date involves the development of the DMCC task battery (see Braver et al., 2021; Snijder et al., 2022). In particular, the DMCC version of the AX-CPT was designed to elicit selective use of proactive, reactive, and baseline control across three separate namesake testing sessions.

Briefly, the proactive session leverages prior work by explicitly instructing participants to use contextual cue information in preparing their responses (Gonthier et al., 2016), effectively training participants to implement a proactive strategy during task performance. On the other hand, consistent with other approaches to reactive control, the reactive session uses an implicit, item-specific cueing manipulation (Braver et al., 2021; Bugg & Crump, 2012; Tang et al., 2021). Specifically, in this session, the probe is presented in a distinct location with unique borders on AY, BX, and no-go trials, serving as an accentuated "just-in-time" stimulus signal to implement reactive control during high demand conflict trials. Finally, the baseline session excludes the prior manipulations but retains no-go trials (presented across all sessions), during which the probe is replaced by a numerical digit. Importantly, implementation of no-go trials decreases the predictive utility of cue information, and is designed to reduce the proactive bias observed in healthy young adults (Gonthier et al., 2016)—rendering the baseline session an ideal "active control" condition from which intended shifts of cognitive control can be evaluated via cross-session comparison (e.g., contrasting proactive vs. baseline performance).

Importantly, the multisession, within-subject design of the DMCC AX-CPT affords a powerful experimental-correlational approach to clarify the influence of WMC on cognitive control (Cronbach, 1957). By standardizing cognitive control strategy use via experimental manipulation, variability associated with the natural tendency to adopt a preferred control mode is minimized, enabling any observed associations between WMC and performance within a given session to be interpreted as between-subjects differences in

the ability or effectiveness of using a given control strategy instead. Moreover, the addition of multiple sessions enables comparative analyses to ascertain the relative *specificity* of the purported influence of WMC on control. For example, data can be aggregated across sessions to explicitly test whether a hypothesized relationship between higher WMC and enhanced proactive control is observed preferentially within the proactive session, over and above any relationships present in the baseline and reactive sessions.

Together, these design features provide the inferential ability to parse the extent to which observed correlations between WMC and proactive control reflect the influence of WMC on the ability to implement proactive control or the tendency to engage it. For example, finding a specific association between higher WMC and enhanced proactive control preferentially in the proactive session (during which all participants are trained and instructed to engage in proactive control) would link higher WMC with superior proactive control ability. On the other hand, observing an association between higher WMC and enhanced proactive control during the baseline but *not* the proactive session would signal that WMC influences the tendency to use proactive control during unconstrained conditions where strategy use is not actively manipulated. Lastly, a third possibility involves finding that higher WMC is associated with enhanced proactive control in both the proactive and baseline sessions, suggesting that WMC may influence both the tendency and ability to use proactive control.

To adjudicate between these competing possibilities, we preregistered a set of initial analyses aimed at broadly replicating past findings linking WMC with enhanced proactive control (see https://osf.io/n9mqw). For full transparency, this initial preregistration also included hypotheses aimed at testing the relationship between trait anxiety and reactive control, which were not broadly supported by preliminary analyses. Consequently, we elected to focus investigative efforts around WMC and proactive control, while prospectively committing to a second follow-up preregistration that was explicitly aimed at replication, as well as minimization of researcher's degrees of freedom, through the use of a holdout dataset (described further below). Importantly, the holdout dataset remained unanalyzed until after all variables, data inclusion criteria, and statistical analytic procedures were specified in the submission of the 2nd preregistration. The primary rationale behind this "two-step" preregistration process was to balance transparency and rigor with the flexibility to derive the most robust yet tractable method for analyzing the data.

We hypothesized that higher WMC scores would be associated with enhanced proactive control metrics in the baseline and or proactive sessions. Specifically, analyses focused on three indices of proactive control: (a) A-cue bias, (b) BX interference, and (c) the *d*-prime context effect. Briefly, A-cue bias assesses the propensity to commit target responses across "A" trials, including both AX (hits) and AY (false alarms) trials—rendering it a collective measure of response bias based on contextual cue use and response preparation. Because WMC is needed to store and actively maintain contextual cue information, whereas proactive control enables preparatory cue use (i.e., activation of the expected target response), we predicted that higher WMC would be associated with stronger A-cue bias (i.e., more target responding across "A" trials). The BX interference effect contrasts errors and reaction time (RT) across BX and BY trials, enabling measurement of the interference associated with the presentation of an "X" probe following a *nontarget* ("B") cue. Here, enhanced B-cue use afforded from proactive control should elicit greater preparation of correct nontarget responses and attenuate "X" probe interference. Higher WMC is therefore expected to be associated with reduced BX error (i.e., fewer errors) and RT interference (i.e., faster RTs) on high conflict BX trials, referenced to the low conflict BY trials. Last, the *d*-prime context effect compares target responding across AX and BX trials, measuring the degree to which the response to "X" probes discriminates target ("A") and nontarget ("B") cues. Similarly, greater cue use should elicit *more* target responses on AX trials (i.e., correct hits) relative to BX trials (i.e., false alarms)—a metric referred to as *d*-prime sensitivity. Extending the rationale developed above, higher WMC is expected to be associated with enhanced *d*-prime sensitivity. Taken together, these three indices provide a strong signature of proactive control from which to observe sensitivity to individual differences in WMC, and how these effects are impacted by the session manipulation of cognitive control mode. Specifically, the presence of session specific patterns that emerged to support these predictions (e.g., WMC positively correlated with A-cue bias in only the proactive session but not the baseline session) were identified in the initial dataset. As mentioned above, these were then treated as fixed predictions to be retested in a (second) preregistered follow-up replication analysis, using a separately collected holdout dataset (https://osf.io/x96wn).

Regarding the second aim, we endeavored to address the aforementioned methodological and analytic issues by adopting a trial-level Bayesian mixed modeling approach. First, we used complete trial-level data across all task conditions to appropriately capture trial-level variability and obviate use of difference scores. Second, we leveraged mixed modeling with random slopes and intercepts to circumvent nonindependence and account for subject-level variability in task performance. Third, we used a Bayesian regression approach to quantify and assess evidence in favor of the predicted effects against the null. Moreover, as mentioned above, we used this approach to showcase how results from a preregistered analysis can be subject to sequential replication with Bayesian updating procedures (Ly et al., 2019; Verhagen & Wagenmakers, 2014). Although Bayesian statistical approaches are far from new, and have become increasingly popular within various areas of psychological science (van de Schoot et al., 2017), to our knowledge, the application of Bayesian approaches to investigate cognitive control remains relatively underused relative to traditional frequentist approaches. Consequently, a final motivation for the current study was to provide a practical demonstration of the advantages and implementational features of this analytic and inferential approach, as applied to the research prerogatives of the DMC framework and of cognitive control studies more generally. Below, we separately detail the specific methods and results of the initial and replication analyses before synthesizing the collective implications of the findings in the general discussion section at the end.

Initial Analysis

Method

Participants

Two-hundred seventy-eight participants were recruited via the Amazon Mechanical Turk (MTurk) online platform. Data collection

was completed in March 2018 and included two rounds of testing, during which participants completed a retest assessment several weeks after initial testing (see Tang et al., 2021, for details of the full study protocol). Of this recruited set, 140 participants failed to complete the full study protocol because they either did not complete the sessions on time, experienced prohibitive technical issues, or failed to comply with task instructions. Participants with overall accuracy below 50% and or 3 *SD*s below the sample mean were excluded (one participant), resulting in a final sample of 137 participants. Participants were not restricted with regard to age (22–64, $M = 36.70$, $SD = 10.03$; 85 females, 52 males). The TurkPrime interface was used for all aspects of participant recruitment and management (e.g., advertisement, communication, and payment). Prospective participants were given a link to review and sign the consent form. On consenting, the web-links for the tasks were made available over MTurk. Participants were compensated a total of $122 for full completion (includes test and retest). The study protocol (IRB #201608003) was approved by the institutional review board of Washington University, St. Louis.

Design and Procedure

The study protocol consisted of thirty testing sessions lasting 20–40 minutes (15 sessions for test, and 15 sessions for retest), including the baseline, proactive, and reactive sessions of the AX-CPT among other components of the DMCC battery. For the purposes of the current study (i.e., to maintain continuity with the replication analysis described below), analyses focused exclusively on the test data; the retest data were excluded from all analyses and not examined. The DMCC task battery was originally developed for the neuroimaging environment (Braver et al., 2021); the online behavioral protocol and task variants (Tang et al., 2021), which included additional self-report/demographic questionnaires, are fully described elsewhere (see Etzel et al., 2021). Participants were asked to complete the sessions at a rate of five per week in fixed sequential order (with baseline conditions completed first, followed by reactive, then proactive), taking about three weeks to complete the full protocol. Each five-session set was posted at the beginning of the week and two reminder emails were sent to remind subjects to complete the set by the end of the week. Completed sessions were checked for accuracy and compliance (see Tang et al., 2021). Subjects who did not complete the weekly set or failed to comply with instructions were discontinued from future sessions and received a prorated payment for sessions completed.

Tasks

Ax-CPT. Participants were instructed to make either target ("z") or nontarget ("m") button press responses to visually presented cue-probe pairs. Consistent with previous versions of the AX-CPT, a target response was required to the probe on AX trials, whereas a nontarget response was required to the probe on all other trial types (AX, BX, BY), as well as to the cue on all trials. As mentioned above, this version of the task also included no-go trials, which required withholding response to the probe; no-go trials were indicated by a digit (1–9) rather than letter probe (Gonthier et al., 2016). There were a total of 216 trials, encompassing 72 AX trials, 72 BY trials, 18 AY trials, 18 BX trials and 36 no-go trials (18 following an A cue, 18 following a B cue). The task was performed in three 72 trial blocks with all trials presented in random order. Subjects were instructed to take a minimum one-minute rest break between blocks to mitigate fatigue. Across all trials, the cue was presented at the center of a white screen for 500 milliseconds (ms). After a fixed blank duration of 4,000 ms, the target probe was presented for 500 ms which was preceded by a bounding box presented 250 ms earlier. Each trial concluded with a 1,500 ms intertrial interval during which a triangle arrangement of fixation crosses was presented at the center of the screen.

Baseline Session. The baseline session was identical to the description above. Participants performed a 12-trial practice block before beginning the actual session.

Proactive Session. Participants completed two phases of strategy training prior to beginning the session (Gonthier et al., 2016). In the first phase, an audio clip instructed participants how to prepare their button presses in response to the cue across six hypothetical trials (i.e., to prepare a target response following A cues and nontarget otherwise). In the second phase, participants completed six practice trials by typing out "left" or "right" to indicate the button they were preparing to press. Feedback was provided after incorrect responses, reminding participants of the cue letter and requesting them to try again. Finally, participants were prompted with the visual message "Use the strategy!" during the intertrial interval periods across the actual testing session. All other task components were identical to the baseline session.

Reactive Session. The reactive session featured a new AX-CPT variant that was adopted to preferentially encourage and enhance use of reactive control (for further discussion of this version see Braver et al., 2021; Tang et al., 2021). High conflict trials (AY, BX, no-go) were preceded by a unique border color and further accentuated by placing the probe in a distinct spatial location. Specifically, on low-conflict AX and BY trials, the probe was presented on the upper half of the screen, whereas the probe was presented on the lower half of the screen during the high-conflict AY, BX, and no-go trials. Furthermore, a black border preceded the probe on AX and BY trials, whereas a red border preceded the probe on AY, BX, and no-go trials. Cues were presented at the center of the screen. All other trial parameters were identical to the baseline and proactive sessions.

Working Memory Tasks.

Operation Span Task. An automated online version of the Operation Span Task (OSPAN; Turner and Engle, 1989; Unsworth et al., 2005) was used to assess WMC (https://www.millisecond.com/download/library/ospan). During each trial, participants were required to verify the accuracy of a mathematical equation before being presented with a random letter to remember. The number of math-letter sequences (i.e., set size) varied from three to seven per trial. At the end of the trial, participants selected the presented letters in the order that they had appeared. The task consisted of three trials of each set size for a total of 15 trials and imposed a response deadline for math problems based on average performance during practice trials. Processing task accuracy was not an exclusionary criterion. Performance was scored by summing the total number of correctly recalled letters (i.e., partial span score).

Symmetry Span Task. Likewise, an automated online version of the Symmetry Span Task (SYMSPAN; Unsworth et al., 2009) was implemented to derive another measure of WMC (https://www.millisecond.com/download/library/symmspan). Similar to the structure of the OSPAN, participants were required to judge whether a displayed shape is symmetrical along its vertical axis (a

response deadline was again imposed based on average performance during practice trials) before being presented with a red square in a 4 × 4 grid of potential locations to remember. At the end of each trial, participants selected the location of the red squares in the order of presentation. The number of symmetry-location sequences (i.e., set size) ranged from two to five per trial. The task consisted of three trials of each set size for a total of 12 trials. Again, processing task accuracy was not an exclusionary criterion. Performance was scored by summing the total number of correctly recalled square locations (partial span score).

Statistical Analyses and Predictions

Bayesian linear regression models were fit using the brms package in the R software environment (Bürkner, 2017). WMC was quantified as a composite score and mean-centered by way of *z*-scoring the average of OSPAN and SYMSPAN scores (each span task was also z-scored first). Categorical variables (trial type, session) were effect coded. As further detailed below, the data were filtered so that trial type was rendered as a binary variable contingent on the metric of interest. Logistic regression on trial-level data was then used to model interference effects on response type (target vs. nontarget) and accuracy (correct vs. incorrect). Similarly, trial-level RTs were modeled using shifted lognormal functions (Haines et al., 2020), with mixed-effects linear regression used to estimate interference effects. Trial type, session (baseline, proactive, reactive), and WMC score were predictors in the model. Subject-level variability was modeled by entering the intercept and trial type as random effects nested within subject. Owing to the novel nature of the analyses, we conservatively elected to use uniform priors on the fixed effects across all models. Random effects were weakly informative based on brms defaults.

Below, we describe how each control metric was modeled and tested. Models involving response type and accuracy was run with 4 Monte Carlo chains, each containing 2,000 sample iterations and 1,000 warm-up iterations, with the warmup iterations discarded. The modeling of RT was likewise run with four chains, but each containing 4,000 sample iterations and 2,000 warm-up iterations—sample iterations were doubled to ensure sufficient effective sample sizes (ESS) across model parameters. For every parameter estimate, we report the mean, standard deviation, and the 95% credible interval (CI; quantile-based equal tailed interval) of the posterior distribution, as well as the R-hat and ESS values. Where applicable, log-odds were exponentiated to odds and presented in the model summary tables. The preregistration for this initial analysis is accessible at https://osf.io/n9mqw. Moreover, all data, materials, and analysis code are openly available at https://osf.io/zuvry/ (Lin et al., 2022).

A-Cue Bias. A-cue bias was modeled as the log-likelihood of committing a target, relative to nontarget, probe response on AX/AY trials. To test the general prediction that higher WMC would be associated with stronger A-cue bias, we ran the following model on AX/AY trial data aggregated across all sessions: *Probe Response* ∼ *Trial Type* × *WMC* × *Session* + *(1* + *trial type* | *subject)*. Specifically, we expected to observe that higher WMC would be associated with higher log-odds of committing target responses in either the baseline session, proactive session, or both (i.e., WMC × session interactions). Although finding a main effect of WMC was also a possibility (i.e., that higher WMC is associated with higher log-odds of target responding irrespective of session manipulations), we nonetheless expected that the effect of WMC on A-cue bias would be most salient in the baseline or proactive sessions based on prior work and theoretical grounds (Richmond et al., 2015). Failure to obtain evidence for either effect would serve as grounds for falsification that WMC is related to A-Cue bias.

BX Interference. BX interference on trial accuracy was modeled as the log-likelihood of committing a correct, relative to incorrect, response on BX/BY trials, whereas interference on RT was modeled using lognormal linear regression on BX/BY trial RTs. To test the prediction that higher WMC would be associated with reduced BX interference, we applied the same predictors specified above to estimate trial accuracy and RT. Here, we expected that higher WMC would be associated with higher log-odds of correct responses and faster RTs on BX trials relative to BY trials (i.e., reduced BX error and RT interference, respectively) in either the baseline session, proactive session, or both. Note that in contrast to the A-cue bias, the BX interference predictions are three-way interactions that involve trial type (i.e., Session × WMC × Trial Type). Again, it was also possible, but not predicted, that higher WMC would be associated with reduced BX error and RT interference collapsing across session (i.e., WMC × Trial Type). The absence of these three-way interaction effects would fail to support the hypothesized relationship between WMC and BX interference.

D-Prime Context Effect. Lastly, the *d*-prime context effect was modeled as the log-likelihood of committing a target, relative to nontarget, probe response on AX/BX trials. Similar to above, we leveraged the same predictors to test the hypothesis that higher WMC would be associated with enhanced *d*-prime sensitivity. Specifically, we expected that higher WMC would be associated with higher log-odds of target responding on AX relative to BX trials (i.e., more correct hits relative to fewer false alarms) in either the baseline session, proactive session, or both. Similar to the BX interference predictions (but distinct from A-cue bias predictions), the *d*-prime context predictions are three-way interactions that involve trial type (i.e., Session × WMC × Trial Type). Once again, we acknowledged the possibility, but did not predict, that higher WMC would be associated with enhanced *d*-prime sensitivity collapsing across all sessions (i.e., WMC × Trial Type). Failure to obtain evidence for these three-way interaction effects would fail to support the hypothesis that WMC is related to *d*-prime sensitivity.

Results

Descriptive statistics of all working memory measures are provided in Table 1. Full model summaries involving all parameter estimates are provided in Tables 2–5. To maintain focus and tractability, we circumscribe descriptive reporting to only the predicted effects below. In particular, we do not provide descriptive or psychometric characteristics of the full AX-CPT task data here, because these are the focus of two additional reports on the DMCC battery (Snijder et al., 2022; Tang et al., 2021).

In addition to the primary analyses reported below, we also ran all models including age as a fixed effect covariate. In all of these supplementary analyses, age did not alter the pattern or magnitude of the WMC effects reported below. Consistent with prior findings (Bopp & Verhaeghen, 2005; Borella et al., 2008; Park et al., 2002), increasing age was associated with a decrease in WMC ($r = -.23$,

Table 1
Descriptive Statistics of Working Memory Measures Separated by Study Wave

Measure	Study wave	*M*	*SD*	Skew	Kurtosis	α	Correlation
OSPAN	Initial	60.92	12.6	−1.87	4.36	0.85	.34**
SYMSPAN		29.96	7.12	−0.42	−0.35	0.73	—
WMC Composite		0.01	0.99	−0.86	1.17	—	—
OSPAN	Replication	60.69	13.00	−1.85	4.35	0.85	.32**
SYMSPAN		28.85	6.99	−0.47	−0.13	0.71	—
WMC Composite		0	1.00	−0.99	1.10	—	—

Note. OSPAN = Operation Span Task; SYMSPAN = Symmetry Span Task; Correlation = Pearson correlation between OSPAN and SYMSPAN for each study wave.
** $p < .001$.

$p < .01$). Likewise, age was associated with a small but statistically significant increase in both BX/BY trial accuracy ($b = .02$, $SD = .01$, 95% CI [.01, .04]) and RT ($b = .005$, $SD = .001$, 95% CI [.003, .007]), suggesting that age corresponded to a more cautious responding style that involved trading speed for accuracy, again consistent with prior reports (cf. Braver et al., 2005). All code and full output summaries for the age analyses are available online (https://osf.io/zuvry/).

A-Cue Bias

Consistent with predictions, higher WMC was not significantly associated with a general tendency to make a target response (i.e., nonsignificant main effect; $b = .08$, $SD = .06$, 95% CI [−.04, .19]), but instead was uniquely associated with higher log-odds of target responding in the proactive session ($b = .16$, $SD = .03$, 95% CI [.09, .23]). Conversely, WMC was associated with relatively lower log-odds of target responding in the reactive ($b = -.08$, $SD = .04$, 95% CI [−.15, −.01]) and baseline session ($b = -.08$, $SD = .04$, 95% CI [−.15, −.01]).

BX Interference

In analyzing trial accuracy, higher WMC was associated with relatively higher log-odds of correct responses on BX trials relative to BY trials (i.e., reduced BX error interference, 3-way interaction) in the baseline session ($b = .14$, $SD = .05$, 95% CI [.06, .23]), but not in the proactive ($b = -.09$, $SD = .05$, 95% CI [−.19, .00]) or reactive session ($b = -.05$, $SD = .04$, 95% CI [−.13, .04]). In contrast for RT however, higher WMC was associated with faster RTs on BX relative to BY trials in the proactive session (i.e., reduced BX RT interference, 3-way interaction; $b = -.006$, $SD = .003$, 95% CI [−.011, −.001]),[1] but not the baseline ($b = .001$, $SD = .003$, 95% CI [−.004, .007]) or reactive session ($b = .005$, $SD = .003$, 95% CI [.000, .010]).

D-Prime Context Effect

Contrary to expectations, WMC was unrelated to target responding on *d*-prime sensitivity across all sessions (i.e., no significant WMC × Trial or three-way interactions; bs $< |.05|$, all CIs contain 0).

Summary

Briefly, the initial analyses broadly supported the hypothesized relationship between higher WMC and enhanced proactive control. In particular, higher WMC was associated with stronger A-cue bias and reduced BX RT interference in the proactive session. Interestingly and somewhat inconsistently, higher WMC was associated with reduced BX error interference in the *baseline* session, with no significant effect in the proactive session. Finally, WMC was unrelated to *d*-prime sensitivity. Although in need of further testing, this constellation of findings suggested that the relationship between WMC and enhanced proactive control may be relatively more specific to the proactive session as opposed to the baseline session. Moreover, there is limited evidence in favor of a "session general" effect of WMC on control metrics (i.e., that WMC is related to enhanced indices of proactive control irrespective of session).

With these possibilities in mind, we aimed to replicate the results prior to extrapolating the significance and implications of the pattern of findings. Toward this end, all observed effects were subject to retesting in a replication analysis, with the posterior parameter estimates obtained above entered as informed priors to fully use the advantages of Bayesian updating and hypothesis testing (Ly et al., 2019; Verhagen & Wagenmakers, 2014). Critically, the specific predictions involving WMC and control metrics were narrowed to mirror the results obtained in the initial analyses. Below, we detail the model specifications and statistical analytic procedures associated with the replicatory testing of each prediction, including use of Bayes factors and probability of direction to quantify the strength of evidence for the expected effects.

Replication Analysis

Method

Participants

Two-hundred forty-five participants enrolled in the study in October 2020. Of this recruited set, 135 participants completed

[1] Because RTs were fit to a lognormal distribution, effect estimates involved small values. Consequently, we report parameter estimates to the third decimal place. It is important to acknowledge that this effect differed from what was described in the second preregistration (i.e., that we initially obtained and expected to replicate null effects for BX RT interference). The preregistered prediction was informed by a model that inadvertently retained the assumption of a normal distribution, which failed to properly account for the skewness of RT data. We remediated this issue by updating the model to specify a shifted log-normal distribution. We acknowledge that this updating of the statistical model did occur after the 2nd preregistration, thus adding additional researcher degrees of freedom, but we believe this was outweighed by the improved statistical estimation obtained.

Table 2
Model Output From Initial A-Cue Bias Logistic Regression Analysis

Model	Fixed effects	Estimate	95% CI	Odds (*SD*)	R-hat	Bulk ESS	Tail ESS
A-Cue Bias	Intercept	0.23 (0.06)	[0.11, 0.35][a]	1.26 (0.08)	1.00	1,918	2,895
	Base	−0.37 (0.04)	[−0.45, −0.30][a]	0.69 (0.03)	1.00	6,016	3,417
	Pro	0.61 (0.03)	[0.55, 0.68][a]	1.85 (0.06)	1.00	5,707	3,062
	Rea	−0.24 (0.04)	[−0.32, −0.17][a]	0.79 (0.03)	1.00	6,740	3,256
	WMC	0.08 (0.06)	[−0.04, 0.19]	1.08 (0.06)	1.00	1,995	2,486
	Trial	−2.89 (0.07)	[−3.04, −2.76][a]	0.06 (0.00)	1.00	1,268	2,010
	Base:WMC	−0.08 (0.04)	[−0.15, −0.01][a]	0.92 (0.04)	1.00	5,157	3,114
	Pro:WMC	0.16 (0.03)	[0.09, 0.23][a]	1.17 (0.04)	1.00	6,173	3,238
	Rea:WMC	−0.08 (0.04)	[−0.15, −0.01][a]	0.92 (0.03)	1.00	6,359	3,325
	Base:Trial	−0.24 (0.04)	[−0.32, −0.16][a]	0.79 (0.03)	1.00	5,223	3,385
	Pro:Trial	0.36 (0.03)	[0.30, 0.43][a]	1.44 (0.05)	1.00	6,690	3,196
	Rea:Trial	−0.12 (0.04)	[−0.20, −0.04][a]	0.89 (0.03)	1.00	6,402	3,213
	WMC:Trial	0.06 (0.07)	[−0.09, 0.20]	1.06 (0.08)	1.01	1,394	2,198
	Base:WMC:Trial	0.00 (0.04)	[0.04, −0.07]	1.00 (0.04)	1.00	5,696	3,615
	Pro:WMC:Trial	−0.01 (0.03)	[−0.07, 0.06]	0.99 (0.03)	1.00	6,126	3,451
	Rea:WMC:Trial	0.00 (0.04)	[−0.07, 0.08]	1.00 (0.04)	1.00	6,423	3,446

Note. WMC = working memory capacity; Base = baseline session; Pro = proactive session; Rea = reactive session; Trial = trial type.
[a] 95% CI does not contain 0.

the full MTurk task battery; the remainder failed to complete the study because they either did not finish sessions on time, experienced prohibitive technical issues, or failed to comply with task instructions. Once again, participants were also excluded for overall accuracy below 50% and or 3 *SD*s below the sample mean (two participants excluded), resulting in a final sample of 133 participants. The TurkPrime interface was again used for all aspects of participant recruitment and management. All participants reviewed and signed the consent form prior to study enrollment. Similarly, participants were not restricted to age (18–77, M = 39.30, SD = 11.28; 73 females, 59 males, one prefer not to answer), and were paid $51 for full completion of the study (which did not include a retest component). The study protocol (IRB #201608003) was approved by the institutional review board of Washington University, St. Louis.

Design and Procedure

The second wave of data collection did not include a retest phase and therefore comprised 15 sessions. Likewise, in this wave of data collection, a fixed session order was used, but this was slightly different from the 2018 wave in that although the baseline sessions were first, the proactive sessions came before reactive. This change was made to counterbalance against the possibility of order-related carryover effects. If the effects of the first study could be replicated with a different order, then the findings were unlikely influenced by session order. The baseline session was always presented first because it did not contain strategic/contextual manipulations that would be prone to carryover (also because it reduced the number of possible counterbalance orders from 6 to 2). All other aspects of the tasks, materials, and procedures were otherwise identical to what was previously described regarding the first wave.

Table 3
Model Output From Initial BX Error Interference Logistic Regression Analysis

Model	Fixed effects	Estimate (*SD*)	95% CI	Odds (*SD*)	R-hat	Bulk ESS	Tail ESS
BX Error Interference	Intercept	3.62 (0.10)	[3.24, 3.82][a]	37.4 (3.80)	1.00	1,319	1,858
	Base	−0.31 (0.05)	[−0.39, −0.21][a]	0.74 (0.03)	1.00	5,241	3,466
	Pro	0.30 (0.05)	[0.21, 0.40][a]	1.36 (0.07)	1.00	6,133	3,419
	Rea	0.00 (0.05)	[−0.09, 0.10]	1.01 (0.05)	1.00	6,441	3,549
	WMC	0.06 (0.10)	[−0.14, 0.24]	1.06 (0.10)	1.00	1,299	2,118
	Trial	−1.51 (0.07)	[−1.66, −1.38][a]	0.22 (0.02)	1.00	2,604	2,520
	Base:WMC	−0.12 (0.05)	[−0.21, −0.03][a]	0.89 (0.04)	1.00	5,360	3,385
	Pro:WMC	0.02 (0.05)	[−0.08, 0.11]	1.02 (0.05)	1.00	5,827	3,616
	Rea:WMC	0.10 (0.05)	[0.01, 0.19][a]	1.11 (0.05)	1.00	6,588	3,362
	Base:Trial	−0.13 (0.05)	[−0.21, −0.03][a]	0.88 (0.04)	1.00	5,741	2,882
	Pro:Trial	0.05 (0.05)	[−0.05, 0.15]	1.06 (0.05)	1.00	5,603	3,275
	Rea:Trial	0.07 (0.05)	[−0.02, 0.16]	1.08 (0.05)	1.00	6,425	3,394
	WMC:Trial	0.08 (0.06)	[−0.03, 0.21]	1.09 (0.07)	1.00	2,230	2,744
	Base:WMC:Trial	0.14 (0.05)	[0.06, 0.23][a]	1.15 (0.05)	1.00	4,982	3,488
	Pro:WMC:Trial	−0.09 (0.05)	[−0.19, 0.00]	0.91 (0.05)	1.00	6,007	3,424
	Rea:WMC:Trial	−0.05 (0.04)	[−0.13, 0.04]	0.96 (0.04)	1.00	5,703	3,571

Note. WMC = working memory capacity; Base = baseline session; Pro = proactive session; Rea = reactive session; Trial = trial type.
[a] 95% CI does not contain 0.

Table 4
Model Output From Initial BX RT Interference Lognormal Regression Analysis

Model	Fixed effects	Estimate (*SD*)	95% CI	R-hat	Bulk ESS	Tail ESS
BX RT Interference	Intercept	6.107 (0.014)	[6.079, 6.135][a]	1.00	410	804
	Base	0.060 (0.003)	[0.055, 0.065][a]	1.00	12,119	6,662
	Pro	−0.107 (0.003)	[−0.112, −0.102][a]	1.00	11,393	7,121
	Rea	0.048 (0.003)	[0.043, 0.053][a]	1.00	11,705	6,661
	WMC	−0.040 (0.015)	[−0.069, −0.010][a]	1.00	509	1,557
	Trial	0.078 (0.004)	[0.069, 0.085][a]	1.00	1,598	3,684
	Base:WMC	0.008 (0.003)	[0.003, 0.013][a]	1.00	12,995	6,254
	Pro:WMC	−0.008 (0.003)	[−0.014, −0.003][a]	1.00	11,512	6,380
	Rea:WMC	0.000 (0.003)	[−0.005, 0.005]	1.00	12,489	6,641
	Base:Trial	−0.012 (0.003)	[−0.017, −0.007][a]	1.00	12,269	6,723
	Pro:Trial	−0.036 (0.003)	[−0.041, −0.031][a]	1.00	12,586	6,798
	Rea:Trial	0.048 (0.003)	[0.043, 0.053][a]	1.00	11,946	6,420
	WMC:Trial	−0.001 (0.004)	[−0.009, 0.008]	1.00	1,737	3,776
	Base:WMC:Trial	0.001 (0.003)	[−0.004, 0.007]	1.00	13,278	7,375
	Pro:WMC:Trial	−0.006 (0.003)	[−0.011, −0.001][a]	1.00	11,650	6,620
	Rea:WMC:Trial	0.005 (0.003)	[0.000, 0.010]	1.00	12,653	6,804

Note. WMC = working memory capacity; Base = baseline session; Pro = proactive session; Rea = reactive session; Trial = trial type.
[a] 95% CI does not contain 0.

Statistical Analyses and Predictions

Again, all Bayesian linear regression models were fit using the brms package in the R software environment (Bürkner, 2017). Coding of variables and mixed effect modeling procedures were identical to what was specified in the initial analyses. The key difference is that posterior parameter estimates obtained from the first dataset were modeled as informed Gaussian priors (specified as the mean and *SD* of the distribution) to evaluate the influence of new data on previous estimates. Indeed, the central aim of the analysis was to replicate the key findings involving WMC and proactive control. Models involving response type and accuracy were again run with four Monte Carlo chains, each containing 2,000 sample iterations and 1,000 warm-up iterations, whereas modeling of RT used 4,000 sample iterations and 2,000 warm-up iterations. Similarly, for every parameter estimate, we again report the mean, standard deviation, and the 95% CI of the posterior distribution.

Novel to the replication analysis, however, is the use of Bayes factors (BFs)—specifically, the Savage-Dickey density ratio (SDR) and evidence ratio (ER)—and the probability of direction (PD) metric to evaluate the strength of replication. For each predicted effect outlined below, the SDR was computed to evaluate the extent to which the mean parameter estimates obtained from the initial dataset differed as a function of incorporating the replication sample. Briefly, the SDR is formalized as the posterior density divided by the prior density at the specified point value, which for our purposes will be the mean value of the *prior* distribution (i.e., mean parameter estimate obtained from the initial analysis). Moreover, we calculate the ER to evaluate the overall amount of evidence favoring the presence of the predicted effect against the null. Contextualized within the scope of the research question, this involves testing the extent to which effects involving WMC and proactive control metrics are greater or less than 0, contingent on the specific directionality of the original finding. For example, when attempting

Table 5
Model Output From Initial d-Prime Sensitivity Logistic Regression Analysis

Model	Fixed effects	Estimate (*SD*)	95% CI	Odds (*SD*)	R-hat	Bulk ESS	Tail ESS
D-Prime Sensitivity	Intercept	0.47 (0.05)	[0.36, 0.57][a]	1.60 (0.08)	1.00	1,476	2,231
	Base	0.15 (0.03)	[0.08, 0.21][a]	1.16 (0.04)	1.00	5,853	3,421
	Pro	−0.04 (0.03)	[−0.11, 0.02]	0.96 (0.03)	1.00	6,355	3,382
	Rea	−0.10 (0.03)	[−0.17, −0.04][a]	0.90 (0.03)	1.00	5,199	3,565
	WMC	−0.05 (0.05)	[−0.14, 0.05]	0.96 (0.05)	1.00	1,911	2,477
	Trial	2.66 (0.08)	[2.51, 2.81][a]	14.3 (1.13)	1.00	902	1,657
	Base:WMC	−0.05 (0.03)	[−0.11, 0.01]	0.95 (0.03)	1.00	6,758	3,303
	Pro:WMC	0.11 (0.03)	[0.05, 0.18][a]	1.12 (0.04)	1.00	6,501	2,955
	Rea:WMC	−0.06 (0.03)	[−0.13, 0.00]	0.94 (0.03)	1.00	6,580	3,280
	Base:Trial	−0.28 (0.03)	[−0.34, −0.22][a]	0.76 (0.02)	1.00	5,412	3,213
	Pro:Trial	0.30 (0.03)	[0.23, 0.36][a]	1.12 (0.04)	1.00	5,867	3,809
	Rea:Trial	−0.02 (0.03)	[−0.08, 0.04]	0.98 (0.03)	1.00	5,874	3,320
	WMC:Trial	0.08 (0.08)	[−0.07, 0.23]	1.08 (0.08)	1.01	982	2,015
	Base:WMC:Trial	−0.03 (0.03)	[−0.09, 0.03]	0.97 (0.03)	1.00	6,100	3,221
	Pro:WMC:Trial	0.05 (0.03)	[−0.01, 0.12]	1.05 (0.04)	1.00	6,081	3,246
	Rea:WMC:Trial	−0.02 (0.03)	[−0.08, 0.04]	0.98 (0.03)	1.00	7,043	3,548

Note. WMC = working memory capacity; Base = baseline session; Pro = proactive session; Rea = reactive session; Trial = trial type.
[a] 95% CI does not contain 0.

to replicate a hypothetical main effect of WMC on A-cue bias (i.e., higher WMC associated with stronger A-cue bias), the ER is the ratio of the posterior probability that WMC > 0 (i.e., alternative hypothesis) relative to the posterior probability of WMC < 0 (i.e., null hypothesis). Lastly, we present the PD, computed as the posterior probability that a parameter estimate is positive or negative, to evaluate the amount of evidence that the effect falls within the expected direction.

Together, the SDR, ER, and PD provide a quantification of point estimate replicability (i.e., the degree to which the parameter means changed after incorporating replication data), as well as the overall strength of evidence favoring the presence of the predicted effect (i.e., the relative proportion of the posterior distribution falling above or below 0 in the predicted direction). When the SDR is greater than one, it indicates that the weight of evidence is in favor of the prior mean; conversely, SDRs less than one indicate that the prior and posterior mean may differ. When the ER is greater than one, it indicates that the evidence is in favor of the tested hypothesis; conversely, ER values less than one indicate evidence against the tested hypothesis. Per Lee and Wagenmakers (2014), values between 1–3 (1–1/3) suggest anecdotal evidence, 3–10 (1/3–1/10) moderate evidence, 10–30 (1/10–1/30) strong evidence, greater than 30 (less than 1/30) very strong evidence. The PD has a direct correspondence to the frequentist p-value, such that a PD of 95%, 97.5%, and 99.5% is approximately equivalent to a two-sided *p* value of .10, .05, and .01, respectively. To further aid interpretation of these metrics, we provide plots of the prior and posterior distributions as a visual supplement.

Lastly, we plotted predicted probability figures of each model, graphed as a function of WMC and session, and separated by trial type using the *interactions* package in R (Long, 2019). Importantly, it needs to be noted that the transformation of logits to predicted probabilities is nonlinear, and also that probability slope estimates are derived as an aggregated function of *all* model parameters (Ai & Norton, 2003; Osborne, 2019). Consequently, some of the hypothesized effects, which are based linearly in logits and involve higher order interactions, are not always visualizable in the probability domain. Nonetheless, we present these graphs to provide a more comprehensive portrayal of our models and to solicit engagement of other researchers whom might be interested in extending this work more formally into the probability domain (e.g., via the Stata methods outlined in Mitchell & Chen, 2005). The preregistration for this replication analysis is available at https://osf.io/x96wn, whereas all data and code used to conduct the analyses are openly available at https://osf.io/zuvry/ (Lin et al., 2022).

Results

Full model summaries involving all parameter estimates are provided in Tables 6–9. We again circumscribe descriptive reporting to only the effects of interest. Visualization of the prior and posterior distribution for each predicted effect is presented on Figure 1. Predicted probability graphs for each model are presented sequentially on Figures 2–5. Finally, Table 10 provides a summarized output for all control indices and relevant effect estimates collated across both initial and replication analyses.

As was done for the initial dataset, in addition to the primary analyses, we conducted supplementary analyses of all models that included age as a fixed effect covariate. Again, age did not alter the pattern of the WMC effects reported below. Replicating the initial dataset, increasing age was associated with a decrease in WMC ($r = -.27$, $p < .01$). Likewise, we replicated the small but significant effect of age on both BX/BY trial accuracy ($b = .01$, $SD = .01$, 95% CI [.01, .02]) and RT ($b = .006$, $SD = .001$, 95% CI [.004, .007]), consistent with the interpretation that increasing age was associated with a more cautious responding style, resulting in higher accuracy but slower responses on BX trials. All code and full output summaries for the age analyses are available online (https://osf.io/zuvry/).

A-Cue Bias

Confirming the original finding, higher WMC was uniquely associated with higher log-odds of target responding in the proactive session ($b = .12$, $SD = .02$, 95% CI [.08, .16], ER = Inf[2], PD = 100%), but not in the baseline condition ($b = -.02$, $SD = .03$, 95% CI [−.06, .04]). Interestingly, higher WMC was associated with relatively *lower* log-odds of target responding in the reactive session ($b = -.12$, $SD = .03$, 95% CI [−.17, −.06]). The SDR comparing the posterior and prior density at the prior mean ($b = .16$) for the predicted effect that WMC would be associated with stronger A-cue bias in the proactive session was .25, indicating that the prior mean estimate may have been overestimated in the original dataset. Nonetheless, given that the ER approached infinity and the PD was 100%, there was decisive evidence supporting the presence of the predicted effect (see Figure 1A).

BX Interference

Replicating the initial finding, higher WMC was associated with higher log-odds of correct responses on BX trials relative to BY trials (i.e., reduced BX error interference) in the baseline session ($b = .09$, $SD = .03$, 95% CI [.03, .16], ER = 443.44, PD = 99.78%), but not in the proactive ($b = -.02$, $SD = .03$, 95% CI [−.08, .04]) and reactive sessions ($b = -.05$, $SD = .03$, 95% CI [−.11, .00]). The SDR comparing the posterior and prior density at the prior mean ($b = .14$) for the predicted effect that higher WMC would be associated with reduced BX interference in the baseline session was .64, suggesting that the prior mean may have been slightly overestimated. The ER of 443.44 and PD of 99.78% provided very strong evidence in favor of the predicted effect, indicating that the vast majority of the distribution fell above 0 (see Figure 1B).

The RT interference effect was also replicated. Consistent with the finding from the initial analysis, higher WMC was associated with faster RTs on BX relative to BY trials (i.e., reduced BX RT interference) in only the proactive session ($b = -.006$, $SD = .002$, 95% CI [−.010, −.003], ER = 799, PD = 99.99%). In contrast, higher WMC was associated with comparatively slower RTs on BX relative to BY trials in the reactive session ($b = .005$, $SD = .002$, 95% CI [.001, .008]). WMC did not influence BX-BY RTs in the baseline session ($b = .001$, $SD = .002$, 95% CI [−.002, .005]). The SDR comparing the posterior and prior density at the prior mean ($b = -.006$) was 1.42, indicating increased evidence in

[2] An ER approaching infinity indicates that the entire posterior distribution exceeded the test value in the predicted direction.

Table 6
Model Output From Replication A-Cue Bias Logistic Regression Analysis

Model	Fixed effects	Estimate (*SD*)	95% CI	Odds (*SD*)	R-hat	Bulk ESS	Tail ESS
A-Cue Bias	Intercept	0.24 (0.04)	[0.16, 0.32][a]	1.27 (0.05)	1.00	2,842	2,989
	Base	−0.36 (0.03)	[−0.41, −0.31][a]	0.70 (0.02)	1.00	6,231	2,683
	Pro	0.51 (0.02)	[0.47, 0.55][a]	1.66 (0.04)	1.00	7,774	2,906
	Rea	−0.20 (0.03)	[−0.25, −0.14][a]	0.82 (0.02)	1.00	7,548	3,102
	WMC	0.16 (0.04)	[0.07, 0.24][a]	1.17 (0.05)	1.00	3,503	3,416
	Trial	−2.89 (0.05)	[−2.99, −2.80][a]	0.06 (0.01)	1.00	2,677	2,821
	Base:WMC	−0.02 (0.03)	[−0.06, 0.04]	0.99 (0.03)	1.00	6,950	2,467
	Pro:WMC	0.12 (0.02)	[0.08, 0.16][a]	1.13 (0.02)	1.00	6,080	3,203
	Rea:WMC	−0.12 (0.03)	[−0.17, −0.06][a]	0.89 (0.02)	1.00	7,543	2,930
	Base:Trial	−0.28 (0.03)	[−0.33, −0.23][a]	0.76 (0.02)	1.00	7,602	3,369
	Pro:Trial	0.34 (0.02)	[0.29, 0.38][a]	1.40 (0.03)	1.00	7,895	3,090
	Rea:Trial	−0.08 (0.03)	[−0.14, −0.03][a]	0.92 (0.02)	1.00	7,978	2,615
	WMC:Trial	−0.03 (0.05)	[−0.12, 0.07]	0.97 (0.05)	1.00	3,054	2,643
	Base:WMC:Trial	−0.08 (0.03)	[−0.13, −0.03][a]	0.92 (0.02)	1.00	7,344	3,282
	Pro:WMC:Trial	0.00 (0.02)	[−0.04, 0.04]	1.00 (0.02)	1.00	6,596	3,052
	Rea:WMC:Trial	0.07 (0.03)	[0.02, 0.12][a]	1.07 (0.03)	1.00	6,875	3,136

Note. WMC = working memory capacity; Base = baseline session; Pro = proactive session; Rea = reactive session; Trial = trial type.
[a] 95% CI does not contain 0.

support of the prior mean estimate. The ER of 799 and PD of 99.99% decisively confirmed the predicted effect (see Figure 1C).

D-Prime Context Effect

Notably, the null effects involving *d*-prime sensitivity were not replicated. Instead, higher WMC was associated with higher log-odds of target responding on AX trials relative to BX trials (i.e., enhanced *d*-prime sensitivity) in the proactive session (b = .06, *SD* = .02, 95% CI [.02, .10], ER = 499, PD = 99.80%). The emergence of this finding is not necessarily surprising given that there was already trending evidence in support of this relationship in the initial analysis (see Figure 1D for comparison of the prior and posterior distributions). Interestingly, higher WMC was associated with relatively lower log-odds of AX-BX target responding in the reactive session (i.e., reduced *d*-prime sensitivity; b = −.08, *SD* = .02, 95% CI [−.12, −.04]), whereas WMC did not influence *d*-prime sensitivity in the baseline session (b = .04, *SD* = .02, 95% CI [−.01, .08]). As shown in Figure 1D, the SDR comparing the posterior and prior density at the prior mean (b = .05) for the effect that higher WMC would be associated with enhanced *d*-prime sensitivity in the proactive session was 1.32, demonstrating that incorporation of additional data narrowed the posterior distribution closer to the prior mean estimate. Importantly, the ER of 499 and PD of 99.80% provided decisive evidence for the presence of the effect.

General Discussion

The current study sought to advance understanding of the nature of the relationship between WMC and proactive control using the theoretically optimized DMCC versions of the AX-CPT task. Importantly, the DMCC AX-CPT places cognitive control mode under experimental manipulation across three testing sessions, each designed to elicit selective engagement of proactive and reactive control in

Table 7
Model Output From Replication BX Error Interference Logistic Regression Analysis

Model	Fixed effects	Estimate (*SD*)	95% CI	Odds (*SD*)	R-hat	Bulk ESS	Tail ESS
BX Error Interference	Intercept	3.64 (0.07)	[3.50, 3.78][a]	38.2 (2.76)	1.00	1,177	1,968
	Base	−0.29 (0.03)	[−0.35, −0.22][a]	0.75 (0.02)	1.00	4,473	3,030
	Pro	0.14 (0.03)	[0.07, 0.20][a]	1.15 (0.04)	1.00	5,091	3,192
	Rea	0.09 (0.03)	[0.02, 0.15][a]	1.09 (0.04)	1.00	4,795	3,433
	WMC	0.18 (0.07)	[0.05, 0.32][a]	1.20 (0.08)	1.00	1,111	2,032
	Trial	−1.46 (0.05)	[−1.56, −1.37][a]	0.23 (0.01)	1.00	2,417	2,801
	Base:WMC	−0.06 (0.03)	[−0.12, 0.00]	0.94 (0.03)	1.00	4,836	3,301
	Pro:WMC	0.02 (0.03)	[−0.04, 0.09]	1.02 (0.03)	1.00	5,041	3,070
	Rea:WMC	0.03 (0.03)	[−0.03, 0.09]	1.04 (0.03)	1.00	5,300	3,021
	Base:Trial	−0.16 (0.03)	[−0.22, −0.10][a]	0.85 (0.03)	1.00	4,978	3,266
	Pro:Trial	0.02 (0.03)	[−0.05, 0.08]	1.02 (0.03)	1.00	5,372	2,828
	Rea:Trial	0.13 (0.03)	[0.07, 0.20][a]	1.14 (0.04)	1.00	5,352	3,020
	WMC:Trial	0.05 (0.04)	[−0.03, 0.13]	1.05 (0.04)	1.00	2,619	2,797
	Base:WMC:Trial	0.09 (0.03)	[0.03, 0.16][a]	1.10 (0.04)	1.00	4,645	3,236
	Pro:WMC:Trial	−0.02 (0.03)	[−0.08, 0.04]	0.98 (0.03)	1.00	5,107	3,106
	Rea:WMC:Trial	−0.05 (0.03)	[−0.11, 0.00]	0.95 (0.03)	1.00	4,971	2,801

Note. WMC = working memory capacity; Base = baseline session; Pro = proactive session; Rea = reactive session; Trial = trial type.
[a] 95% CI does not contain 0.

Table 8
Model Output From Replication BX RT Interference Lognormal Regression Analysis

Model	Fixed effects	Estimate (*SD*)	95% CI	R-hat	Bulk ESS	Tail ESS
BX RT Interference	Intercept	6.106 (0.010)	[6.085, 6.126][a]	1.00	590	1,037
	Base	0.065 (0.002)	[0.062, 0.068][a]	1.00	14,614	6,412
	Pro	−0.082 (0.002)	[−0.086, −0.079][a]	1.00	15,302	6,154
	Rea	0.028 (0.002)	[0.025, 0.032][a]	1.00	16,177	6,493
	WMC	−0.043 (0.011)	[−0.065, −0.022][a]	1.00	678	1,514
	Trial	0.079 (0.003)	[0.073, 0.084][a]	1.00	3,099	4,688
	Base:WMC	0.011 (0.002)	[0.008, 0.015][a]	1.00	16,724	6,402
	Pro:WMC	−0.012 (0.002)	[−0.015, −0.008][a]	1.00	14,934	6,591
	Rea:WMC	−0.001 (0.002)	[−0.005, 0.002]	1.00	14,234	6,162
	Base:Trial	−0.013 (0.002)	[−0.017, −0.010][a]	1.00	15,536	6,168
	Pro:Trial	−0.041 (0.002)	[−0.044, −0.037][a]	1.00	16,074	6,138
	Rea:Trial	0.056 (0.002)	[0.053, 0.060][a]	1.00	16,418	6,613
	WMC:Trial	−0.002 (0.003)	[−0.008, 0.004]	1.00	3,760	5,265
	Base:WMC:Trial	0.001 (0.002)	[−0.002, 0.005]	1.00	15,483	6,575
	Pro:WMC:Trial	−0.006 (0.002)	[−0.010, −0.003][a]	1.00	14,834	6,715
	Rea:WMC:Trial	0.005 (0.002)	[0.001, 0.008][a]	1.00	16,286	5,892

Note. WMC = working memory capacity; Base = baseline session; Pro = proactive session; Rea = reactive session; Trial = trial type.
[a] 95% CI does not contain 0.

addition to a baseline reference condition. Capitalizing on this core task feature, we adopted a unique experimental-correlational investigative approach to parse the extent to which WMC influences the *tendency* to engage proactive control relative to the *ability* to implement it. Moreover, to address broader relevant methodological issues pertaining to the ambiguities of using summary/subtraction-based difference scores, statistical nonindependence of repeated subject-level observations, and the inferential limitations of NHST, we leveraged trial-level Bayesian mixed modeling to conduct initial hypothesis testing and subsequent replication analysis across two separate data sets.

Conceptual and Methodological Implications

Overall, our findings were consistent with prior work (Belletier et al., 2019; Boudewyn et al., 2015; Redick, 2014; Redick & Engle, 2011; Richmond et al., 2015; Stawarczyk et al., 2014; Wiemers & Redick, 2018)—higher WMC was indeed broadly associated with enhanced proactive control, as indicated by stronger A-cue bias, reduced BX RT interference, and to a lesser extent enhanced *d*-prime context sensitivity (i.e., although overall Bayesian evidence for the *d*-prime effect was strong, it was only statistically significant in the replication study). Central to the novel aim of the investigation, we observed that most of these associations were specific to the proactive session relative to the baseline and reactive sessions. Consequently, our findings demonstrate that the influence of WMC on proactive control is most robust under regularized conditions during which participants are explicitly trained and instructed to use proactive control across task performance. Although the tendency to use proactive control is undoubtedly intertwined with the ability to implement it, this pattern suggests that within the current experimental approach, the relationship between WMC and proactive control may be more influenced by between-subjects variation in the ability to implement proactive control than the preferential tendency to engage it.

With that said, the notable exception is that WMC did not influence BX error interference in the proactive session but was rather

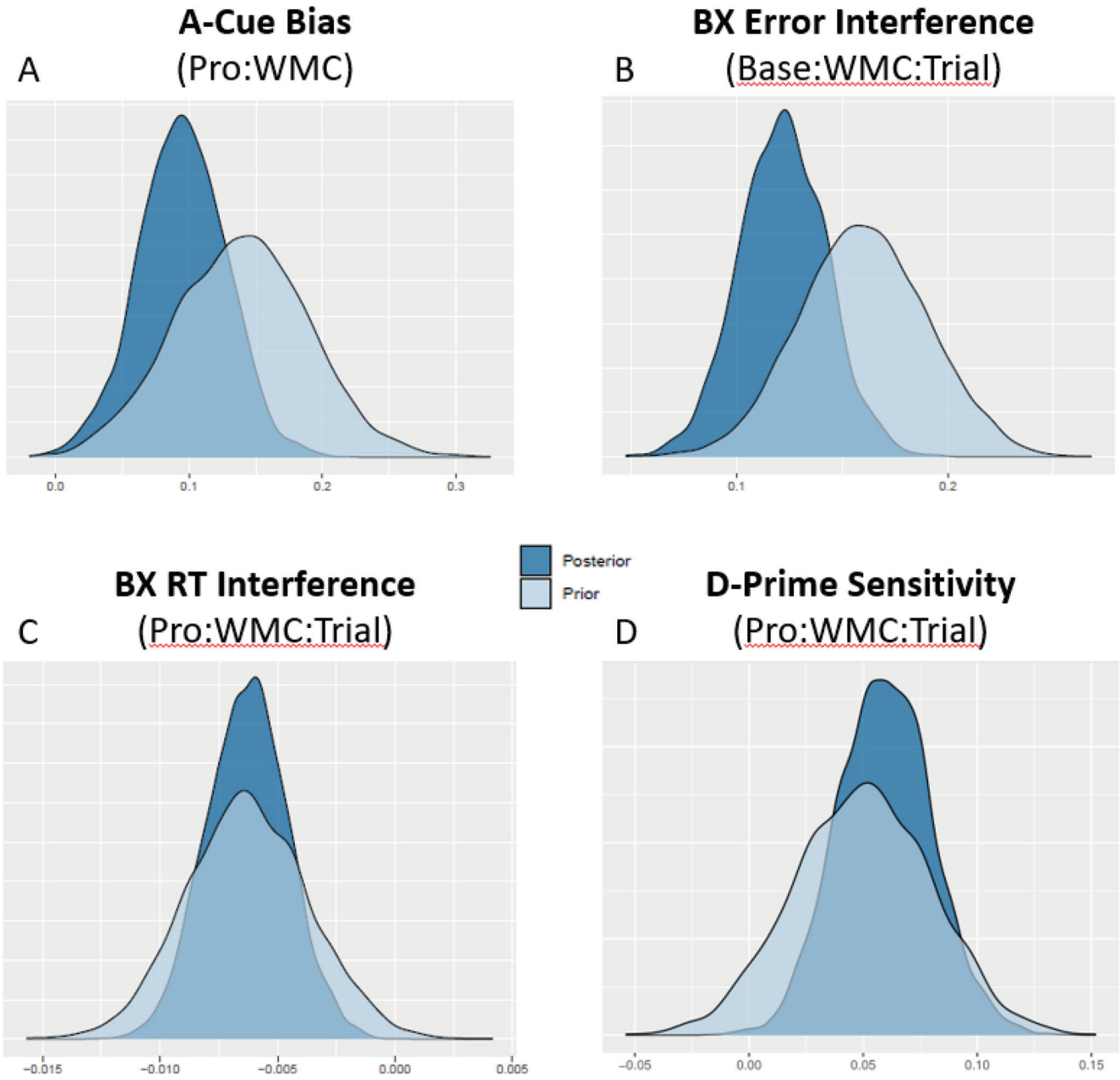
Table 9
Model Output From Replication d-Prime Sensitivity Logistic Regression Analysis

Model	Fixed effects	Estimate (*SD*)	95% CI	Odds (*SD*)	R-hat	Bulk ESS	Tail ESS
D-Prime Sensitivity	Intercept	0.45 (0.04)	[0.37, 0.52][a]	1.57 (0.06)	1.00	1,942	2,755
	Base	0.19 (0.02)	[0.15, 0.23][a]	1.21 (0.03)	1.00	7,194	3,071
	Pro	−0.02 (0.02)	[−0.06, 0.02]	0.98 (0.02)	1.00	7,464	3,254
	Rea	−0.14 (0.02)	[−0.18, −0.10][a]	0.87 (0.02)	1.00	5,820	2,748
	WMC	−0.02 (0.04)	[−0.09, 0.06]	0.99 (0.04)	1.00	2,023	2,323
	Trial	2.69 (0.05)	[2.59, 2.80][a]	14.80 (0.80)	1.00	1,545	2,460
	Base:WMC	0.01 (0.02)	[−0.03, 0.05]	1.01 (0.02)	1.00	6,905	2,925
	Pro:WMC	0.07 (0.02)	[0.02, 0.11][a]	1.07 (0.02)	1.00	7,043	3,113
	Rea:WMC	−0.08 (0.02)	[−0.12, −0.04][a]	0.92 (0.02)	1.00	7,339	2,919
	Base:Trial	−0.28 (0.02)	[−0.32, −0.24][a]	0.76 (0.02)	1.00	7,868	3,077
	Pro:Trial	0.20 (0.02)	[0.16, 0.24][a]	1.22 (0.03)	1.00	7,911	3,166
	Rea:Trial	0.03 (0.02)	[−0.01, 0.07]	1.03 (0.02)	1.00	6,507	2,984
	WMC:Trial	0.21 (0.05)	[0.11, 0.31][a]	1.24 (0.07)	1.00	1,414	2,175
	Base:WMC:Trial	0.04 (0.02)	[−0.01, 0.08]	1.04 (0.02)	1.00	5,887	2,520
	Pro:WMC:Trial	0.06 (0.02)	[0.02, 0.10][a]	1.06 (0.02)	1.00	7,166	3,171
	Rea:WMC:Trial	−0.08 (0.02)	[−0.12, −0.04][a]	0.92 (0.02)	1.00	5,450	3,228

Note. WMC = working memory capacity; Base = baseline session; Pro = proactive session; Rea = reactive session; Trial = trial type.
[a] 95% CI does not contain 0.

Figure 1

Posterior and Prior Density Distributions of A-Cue Bias, BX Error, and RT Interference, and d-Prime Sensitivity Estimates



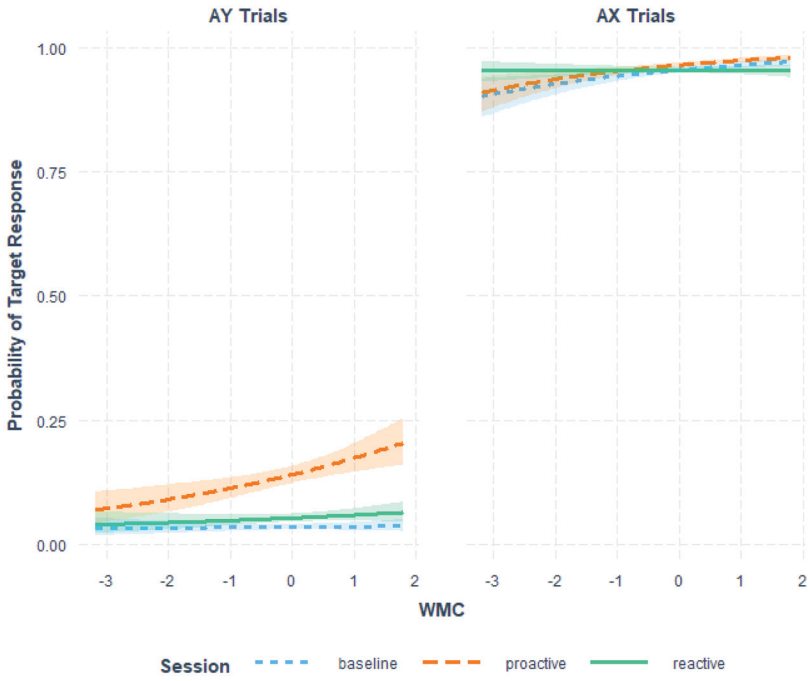
Note. Prior distributions reflect parameter estimates obtained during initial analysis, which were entered as informed priors to derive the posterior distributions of the replication analysis. WMC = working memory capacity; Base = baseline session; Pro = proactive session. See the online article for the color version of this figure.

associated with reduced error interference in the baseline session. To contextualize this unexpected finding, it is important to note that relative to the baseline session, the proactive session yielded higher overall likelihood of committing correct responses collapsing across trial type, but did not preferentially increase the log-odds of BX trial accuracy (no interactive effect of the proactive session on trial type; see Table 6). Together, this introduces the possibility that the proactive manipulation may have instituted a ceiling effect on trial accuracy, thereby restricting the variability needed to detect the expected relationship between WMC and BX error interference.

It is further worth mentioning that the RT interference analysis showed that higher WMC was associated with *reduced* BX RT interference in the *proactive* session, suggesting that although the potency of the proactive manipulation may have instituted a ceiling effect on trial accuracy, participants with

Figure 2

Predicted Probability of Target Responding Plotted as a Function of WMC and Session Separated Across AY and AX Trials



Note. WMC = working memory capacity. See the online article for the color version of this figure.

higher WMC were nonetheless able to respond more quickly when using proactive control. On the other hand, not only did the baseline session produce slower overall RTs, WMC was associated with slower RTs in the baseline session relative to the other sessions. Rather than reflecting increased tendency to engage in proactive control per se, the possible presence of a speed-accuracy tradeoff suggests that that individuals with higher WMC may have adopted a more cautious style of responding during the baseline session, slowing down to the benefit of increased accuracy. Given these considerations, the overall pattern of results appears consistent with the emerging picture that the relationship between WMC and proactive control is more circumscribed to the proactive session.

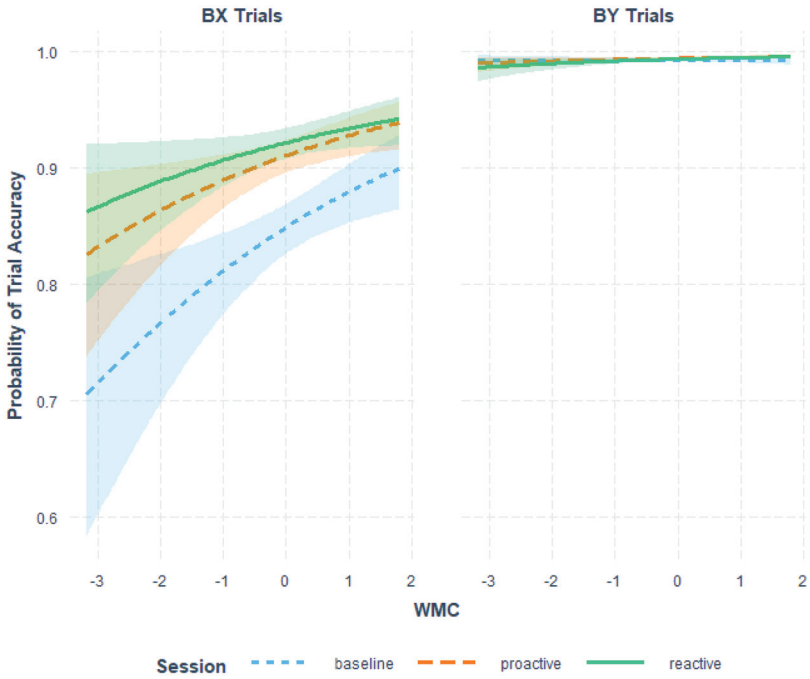
From a broader perspective, it is also worth noting that the relationship between WMC and BX interference effects seems variable in the literature across RT and accuracy. For example, some investigations, including the current study, report associations between WMC and BX RT interference (e.g., Redick & Engle, 2011), whereas others primarily find WMC differences in relation to accuracy but not RTs (e.g., Redick, 2014; Richmond et al., 2015). Fundamentally, deriving separate measures of RT and accuracy creates two possible indices of variability, which although related, may exhibit differential relationships to a common individual difference measure, such as WMC. This is particularly evident when considering performance from a decision-making framework, insofar that the amount of time and evidence needed to prepare and execute a response varies across individuals; thus, task related factors are likely to exert a systematic influence on these parameters, including speed-accuracy tradeoff dynamics. In other words, natural interindividual variation in evidence response thresholds could contribute to mixed findings, with apparent inconsistencies across studies being likely to occur if the task and sample related characteristics differ across them.

To better account for this variability, promising solutions might involve improved experimental manipulation and/or measurement of response thresholds. For example, adapted response deadline approaches have been used to constrain RT windows so that meaningful performance variability is circumscribed to accuracy (see Draheim et al., 2021, for a detailed discussion and task examples). Likewise, accuracy can be constrained (e.g., demanding a criterion level of trial accuracy while elongating response windows) so that performance variability is shifted toward RTs. Lastly, an analytical way to address this issue involves the use of generative models (e.g., drift diffusion modeling; Ratcliff et al., 2016) that integrate both accuracy and RT into the analysis, partitioning the observable data into evidence accumulation and response threshold parameters. Collectively, these approaches represent promising options for future studies to adopt to further clarify the relationship between BX interference and WMC, and strengthen investigation of individual differences in cognitive control more generally.

Interestingly, WMC was demonstrably associated with *less* A-cue bias, *d*-prime sensitivity, and BX interference attenuation in the *reactive* session relative to the other sessions. Considering that one of the major aims of the DMCC task battery is to empirically dissociate between proactive and reactive control, the opposing directionality of effects consistently observed in the reactive

Figure 3

Predicted Probability of Trial Accuracy Plotted as a Function of WMC and Session Separated Across BX and BY Trials



Note. WMC = working memory capacity. See the online article for the color version of this figure.

session introduces the intriguing possibility that between-session dissociations may be evidenced in relation to individual differences. One specific interpretation of this pattern is that individuals with higher WMC possess more flexibility in shifting toward or away from a prospective cue-based strategy (i.e., proactive control) to meet the contextual demands of the situation.

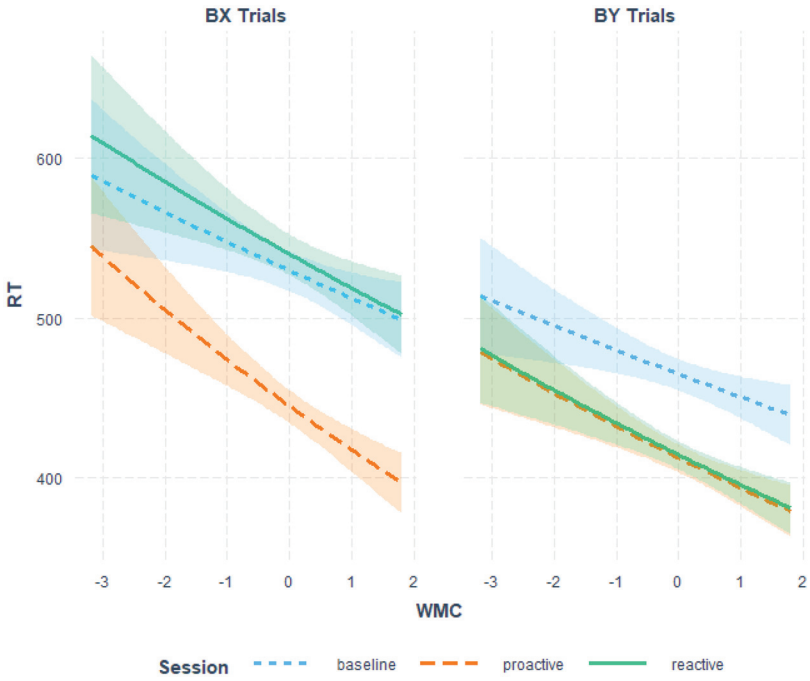
Indeed, this possibility is an extension of the emergent notion that higher WMC does not necessarily translate to a ubiquitously increased tendency to engage proactive control. Even though higher WMC may enable relatively more effective use of proactive control, it does not appear to lead to overreliance of proactive control during situational conditions that call for the use of reactive control. Put another way, higher WMC is not associated with more proactive control at the expense of less reactive control per se. Although the current data cannot fully resolve whether proactive and reactive control are independent constructs or polar ends of a dimensional continuum (see also Gonthier et al., 2016), it does demonstrate that experimentally induced shifts in proactive and reactive control can yield differential relationships between performance and a common individual difference variable. From this broader perspective, the results support and further illustrate the extent of orthogonality between these two modes of control.

Methodologically speaking, the findings reflect the strengths and unique advantages of our analytic approach, enabling us to conduct sequential replication analyses to rigorously test open-ended hypotheses while mitigating common statistical concerns prevalent in the field. First, aggregated trial-level mixed modeling allowed us to leverage the entire dataset while appropriately accounting for nonindependence and subject-level performance variability. Consequently, the results obtained here make full use of the data without relying on the computation of summary and difference scores, and do not involve arbitrary partitioning of the data, or imposing multiple sets of analyses unnecessarily (e.g., separating across session). Second, we applied Bayesian regression to derive posterior distribution estimates for each model parameter. Furthermore, we showcased how Bayesian updating procedures can be used to conduct confirmatory replication analyses by modeling the posterior parameter distributions obtained from the initial dataset as informed priors in analysis of separate holdout data (Ly et al., 2019; Verhagen & Wagenmakers, 2014). Collectively, this allowed us to visually and quantitatively assess how incorporation of new data influences the likelihood or amount of evidence favoring the predicted effect relative to the null as opposed to traditional NHST.

With all that said, it must be acknowledged that our conclusions are not unequivocal. In particular, a very recent study using a similar experimental-correlational approach reported null findings, observing that the relationship between WMC and proactive control was not moderated by experimental condition (Rosales et al., 2022). In fact, WMC was unrelated to proactive control performance across baseline, proactive, or reactive conditions. Instead, higher WMC was associated with generally faster RTs across all trial types, leading to the contradictory conclusion that WMC may have a domain general influence on performance but is not related to the specific tendency (or ability) to use proactive control. In attempting to reconcile the mixed nature

Figure 4

Predicted Trial RT Plotted as a Function of WMC and Session Separated Across BX and BY Trials



Note. WMC = working memory capacity. See the online article for the color version of this figure.

of these findings, it may be valuable to consider several key methodological factors that might differentiate the current study from Rosales et al. (2022).

First, although the proactive control manipulation was essentially identical across the studies, the baseline and reactive conditions differed substantially. Specifically, Rosales et al. (2022) used no-go trials to induce reactive control, whereas here we implemented an item-specific cueing approach for reactive control, while including no-go trials across *all* conditions. As detailed elsewhere (Braver et al., 2021; Tang et al., 2021), the purpose of this change was to explicitly distinguish selective *enhancement* of reactive control from reduced engagement of proactive control (i.e., the effect of no-go trials)—our baseline condition was therefore equivalent to the reactive condition in Rosales and colleague's study. Consequently, the experimental distinctions across the control conditions, in combination with our aggregated trial-level mixed modeling approach (designed specifically to assess for session-level interactions) may have increased performance variability and enhanced sensitivity to detect relationships between WMC and proactive control. Second, the mode of study recruitment and task administration (i.e., online MTurk vs. in-person university study) may have also played an influential role, specifically by increasing sample heterogeneity in the current study and introducing risk of participant cheating (though the automated response deadlines imposed during the WMC tasks ward against this possibility, but see Hicks et al., 2016, regarding threat of cheating). Finally, the current study was not age restricted and as a consequence had a relatively older participant sample. To address this issue, we conducted supplementary analyses that included age as a covariate and found that it did not alter any of the pattern of findings we reported. Nevertheless, the use of an older as opposed to younger sample might have resulted in greater performance variability, which may have buffered against issues related to range restriction.

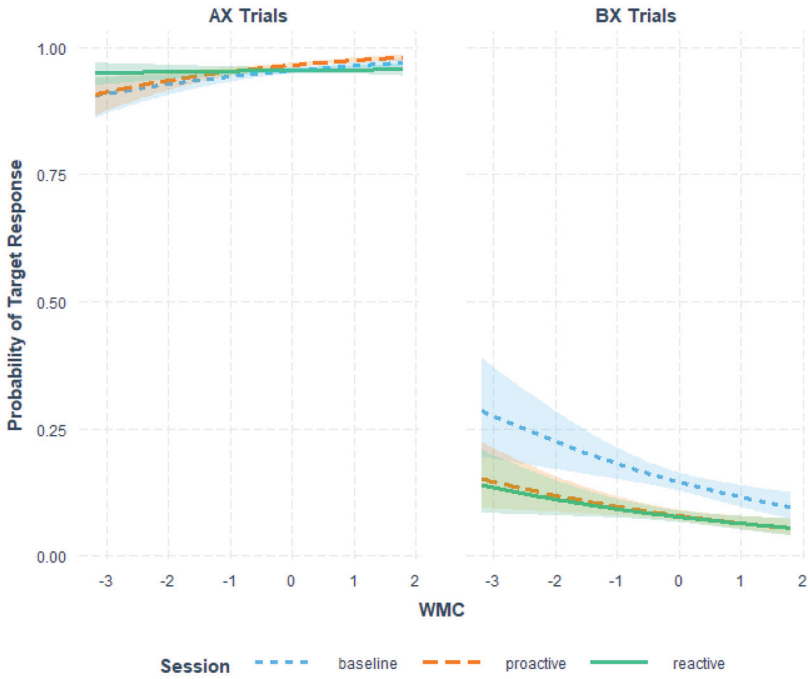
Although further research is needed to adjudicate between these different possibilities, we ultimately view these studies as complementary efforts toward elucidating the nature of the relationship between WMC and cognitive control. In fact, our findings strongly support Rosales and colleagues' astute postulations that: (a) individuals with higher WMC may be better at adapting/shifting mode of control in response to changing task demands; and (b) stronger and more explicit manipulations, such as item-specific cueing and strategy training, may be better suited to induce respective reactive and proactive control strategy shifts, than manipulations such as the inclusion of no-go trials. Indeed, it appears reasonable that the ability to detect relationships between WMC and proactive/reactive control (and the strength of the relationship itself) may be contingent on the potency of the manipulations. Taken together, there is considerable promise in conducting follow-up investigations that seek to validate and build off our approach in service of testing and extending these possibilities. Toward this end, we outline some specific directions for future research below.

Conclusion and Future Directions

The current study leveraged the design innovations of the DMCC task battery to show that the influence of WMC on

Figure 5

Predicted Probability of Target Responding Plotted as a Function of WMC and Session Separated Across AX and BX Trials



Note. WMC = working memory capacity. See the online article for the color version of this figure.

proactive control may be better characterized as individual differences in the ability to implement control, rather than the tendency to spontaneously engage it. Furthermore, we also found that although individuals with higher WMC exhibited more proactive control in the proactive session, they also exhibited greater shifts away from proactive control in the reactive session—suggesting that proactive and reactive control may be partially dissociable in relation to WMC. Indeed, from the perspective that proactive and reactive control represent orthogonal dimensions of control, it stands to reason that neither the ability nor tendency to engage in proactive control should *necessarily* have to come at the expense of reactive control (and vice versa), even though some individual differences may very well confer specific influence on only one mode of control.

With these considerations in mind, three immediate and potentially fruitful future directions involve: (a) leveraging the retest component of the initial dataset to determine the extent to which the reported associations are reliable across repeated testing over time within the same individuals; (b) examining whether the current pattern of findings can generalize to other DMCC tasks such as the Stroop or Cued Task-Switching; and (c) investigating how other individual differences, ideally ones that are both theoretically relevant to the DMC framework and fall within close nomological proximity to WMC, might be related to control using the DMCC task battery. Indeed, the two waves of MTurk data from which the current article is based are well equipped to address all of these questions, and we ourselves intend to pursue many of these avenues in the near future. With that said, we strongly encourage other investigators to make full use of our shared data to collaboratively accelerate progress along these domains. Critically, these directions synergize to enable incremental evaluation of the construct validity of the DMC framework across repeated assessment, different tasks, and other individual difference constructs, all the while providing a natural avenue through which to adjudicate between the mixed findings mentioned above. Obtaining convergent evidence across these domains would not only strengthen confidence in the validity and generalizability of the conclusions we make here, but would also further substantiate the explanatory power of the DMC framework and methodological sensitivity of the DMCC task battery in predicting and testing within-individual and between-individual variability in cognitive control (Braver et al., 2021; Snijder et al., 2022; Tang et al., 2021).

Finally, to bolster these efforts, we encourage future researchers to consider adopting our Bayesian mixed modeling approach, because this makes full use of trial-level data while also enabling estimation of the evidence for both the presence and magnitude of predicted effects. For example, the posterior parameter distributions reported here can be used to develop reasonable priors for future studies, including direct replication attempts or new investigations involving different DMCC tasks or individual difference measures. As we have shown here, this enables an accretive approach to research, continuously quantifying and reassessing evidence for the alternative hypothesis or effect of interest across new studies and accumulated data. In conclusion, we hope that other investigators will take interest in using the DMCC task

Table 10
Summarized Output of All Proactive Control Indices and Relevant Fixed Effects Separated by Study Wave

Model	Study wave	Fixed effects	Estimate (*SD*)	95% CI	Odds (*SD*)
A-Cue Bias	Initial	Base:WMC	−0.08 (0.04)	[−0.15, −0.01][a]	0.92 (0.04)
		Pro:WMC	0.16 (0.03)	[0.09, 0.23][a]	1.17 (0.04)
		Rea:WMC	−0.08 (0.04)	[−0.15, −0.01][a]	0.92 (0.03)
	Replication	Base:WMC	−0.02 (0.03)	[−0.06, 0.04]	0.99 (0.03)
		Pro:WMC	0.12 (0.02)	[0.08, 0.16][a]	1.13 (0.02)
		Rea:WMC	−0.12 (0.03)	[−0.17, −0.06][a]	0.89 (0.02)
BX Error interference	Initial	Base:WMC:Trial	0.14 (0.05)	[0.06, 0.23][a]	1.15 (0.05)
		Pro:WMC:Trial	−0.09 (0.05)	[−0.19, 0.00]	0.91 (0.05)
		Rea:WMC:Trial	−0.05 (0.04)	[−0.13, 0.04]	0.96 (0.04)
	Replication	Base:WMC:Trial	0.09 (0.03)	[0.03, 0.16][a]	1.10 (0.04)
		Pro:WMC:Trial	−0.02 (0.03)	[−0.08, 0.04]	0.98 (0.03)
		Rea:WMC:Trial	−0.05 (0.03)	[−0.11, 0.00]	0.95 (0.03)
BX RT interference	Initial	Base:WMC:Trial	0.001 (0.003)	[−0.004, 0.007]	—
		Pro:WMC:Trial	−0.006 (0.003)	[−0.011, −0.001][a]	—
		Rea:WMC:Trial	0.005 (0.003)	[0.000, 0.010]	—
	Replication	Base:WMC:Trial	0.001 (0.002)	[−0.002, 0.005]	—
		Pro:WMC:Trial	−0.006 (0.002)	[−0.010, −0.003][a]	—
		Rea:WMC:Trial	0.005 (0.002)	[0.001, 0.008][a]	—
D-Prime sensitivity	Initial	Base:WMC:Trial	−0.03 (0.03)	[−0.09, 0.03]	0.97 (0.03)
		Pro:WMC:Trial	0.05 (0.03)	[−0.01, 0.12]	1.05 (0.04)
		Rea:WMC:Trial	−0.02 (0.03)	[−0.08, 0.04]	0.98 (0.03)
	Replication	Base:WMC:Trial	0.04 (0.02)	[−0.01, 0.08]	1.04 (0.02)
		Pro:WMC:Trial	0.06 (0.02)	[0.02, 0.10][a]	1.06 (0.02)
		Rea:WMC:Trial	−0.08 (0.02)	[−0.12, −0.04][a]	0.92 (0.02)

Note. WMC = working memory capacity; Base = baseline session; Pro = proactive session; Rea = reactive session; Trial = trial type.
[a] 95% CI does not contain 0.

battery and applying these methods to their own work. Together, we look forward to furthering understanding of the role of individual differences in shaping the development and use of cognitive control.

References

Ai, C., & Norton, E. C. (2003). Interaction terms in logit and probit models. *Economics Letters*, *80*(1), 123–129. https://doi.org/10.1016/S0165-1765(03)00032-6

Barch, D. M., & Ceaser, A. (2012). Cognition in schizophrenia: Core psychological and neural mechanisms. *Trends in Cognitive Sciences*, *16*(1), 27–34. https://doi.org/10.1016/j.tics.2011.11.015

Belletier, C., Normand, A., Camos, V., Barrouillet, P., & Huguet, P. (2019). Choking under experimenter's presence: Impact on proactive control and practical consequences for psychological science. *Cognition*, *189*, 60–64. https://doi.org/10.1016/j.cognition.2019.03.018

Bopp, K. L., & Verhaeghen, P. (2005). Aging and verbal memory span: A meta-analysis. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *60*(5), 223–233. https://doi.org/10.1093/geronb/60.5.P223

Borella, E., Carretti, B., & De Beni, R. (2008). Working memory and inhibition across the adult life-span. *Acta Psychologica*, *128*(1), 33–44. https://doi.org/10.1016/j.actpsy.2007.09.008

Boudewyn, M. A., Long, D. L., Traxler, M. J., Lesh, T. A., Dave, S., Mangun, G. R., Carter, C. S., & Swaab, T. Y. (2015). Sensitivity to referential ambiguity in discourse: The role of attention, working memory, and verbal ability. *Journal of Cognitive Neuroscience*, *27*(12), 2309–2323. https://doi.org/10.1162/jocn_a_00837

Braver, T. S. (2012). The variable nature of cognitive control: A dual mechanisms framework. *Trends in Cognitive Sciences*, *16*(2), 106–113. https://doi.org/10.1016/j.tics.2011.12.010

Braver, T. S., Barch, D. M., Keys, B. A., Carter, C. S., Cohen, J. D., Kaye, J. A., Janowsky, J. S., Taylor, S. F., Yesavage, J. A., Mumenthaler, M. S., Jagust, W. J., & Reed, B. R. (2001). Context processing in older adults: Evidence for a theory relating cognitive control to neurobiology in healthy aging. *Journal of Experimental Psychology: General*, *130*(4), 746–763. https://doi.org/10.1037/0096-3445.130.4.746

Braver, T. S., Gray, J. R., & Burgess, G. C. (2007). Explaining the many varieties of working memory variation: Dual mechanisms of cognitive control. In *Variation in working memory* (pp. 76–106). Oxford University Press.

Braver, T. S., Kizhner, A., Tang, R., Freund, M. C., & Etzel, J. A. (2021). The dual mechanisms of cognitive control project. *Journal of Cognitive Neuroscience*, *33*(9), 1–26. https://doi.org/10.1162/jocn_a_01768

Braver, T. S., Satpute, A. B., Rush, B. K., Racine, C. A., & Barch, D. M. (2005). Context processing and context maintenance in healthy aging and early stage dementia of the Alzheimer's type. *Psychology and Aging*, *20*(1), 33–46. https://doi.org/10.1037/0882-7974.20.1.33

Bugg, J. M., & Crump, M. J. (2012). In support of a distinction between voluntary and stimulus-driven control: A review of the literature on proportion congruent effects. *Frontiers in Psychology*, *3*, 367. https://doi.org/10.3389/fpsyg.2012.00367

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01

Caruso, J. C. (2004). A comparison of the reliabilities of four types of difference scores for five cognitive assessment batteries. *European Journal of Psychological Assessment*, *20*(3), 166–171. https://doi.org/10.1027/1015-5759.20.3.166

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, *12*(11), 671–684. https://doi.org/10.1037/h0043943

Cronbach, L. J., & Furby, L. (1970). How we should measure "change": Or should we? *Psychological Bulletin*, *74*(1), 68–80. https://doi.org/10.1037/h0029382

Cumming, G. (2014). The new statistics: Why and how. *Psychological Science*, *25*(1), 7–29. https://doi.org/10.1177/0956797613504966

Draheim, C., Mashburn, C. A., Martin, J. D., & Engle, R. W. (2019). Reaction time in differential and developmental research: A review and commentary on the problems and alternatives. *Psychological Bulletin*, *145*(5), 508–535. https://doi.org/10.1037/bul0000192

Draheim, C., Tsukahara, J. S., Martin, J. D., Mashburn, C. A., & Engle, R. W. (2021). A toolbox approach to improving the measurement of attention control. *Journal of Experimental Psychology: General*, *150*(2), 242–275. https://doi.org/10.1037/xge0000783

Egner, T. (2017). *The Wiley handbook of cognitive control*. Wiley. https://doi.org/10.1002/9781118920497

Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In *The psychology of learning and motivation: Advances in research and theory* (Vol. 44, pp. 145–199). Elsevier Science.

Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102–134). Cambridge University Press. https://doi.org/10.1017/CBO9781139174909.007

Etzel, J. A., Brough, R. E., Freund, M. C., Kizhner, A., Lin, Y., Singh, M. F., Tang, R., Tay, A., Wang, A., & Braver, T. S. (2021). The dual mechanisms of cognitive control dataset: A theoretically-guided within-subject task fMRI battery. *Scientific Data*, *9*(1), 1–14. https://doi.org/10.1101/2021.05.28.446178

Gonthier, C., Macnamara, B. N., Chow, M., Conway, A. R. A., & Braver, T. S. (2016). Inducing Proactive Control Shifts in the AX-CPT. *Frontiers in Psychology*, *7*, 1822. https://doi.org/10.3389/fpsyg.2016.01822

Gonthier, C., Zira, M., Colé, P., & Blaye, A. (2019). Evidencing the developmental shift from reactive to proactive control in early childhood and its relationship to working memory. *Journal of Experimental Child Psychology*, *177*, 1–16. https://doi.org/10.1016/j.jecp.2018.07.001

Haines, N., Kvam, P. D., Irving, L. H., Smith, C., Beauchaine, T. P., Pitt, M. A., Ahn, W.-Y., & Turner, B. M. (2020). Theoretically informed generative models can advance the psychological and brain sciences: lessons from the reliability paradox. *PsyArXiv*. https://doi.org/10.31234/osf.io/xr7y3

Halsey, L. G., Curran-Everett, D., Vowler, S. L., & Drummond, G. B. (2015). The fickle P value generates irreproducible results. *Nature Methods*, *12*(3), 179–185. https://doi.org/10.1038/nmeth.3288

Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods*, *50*(3), 1166–1186. https://doi.org/10.3758/s13428-017-0935-1

Hicks, K. L., Foster, J. L., & Engle, R. W. (2016). Measuring working memory capacity on the web with the online working memory lab (the OWL). *Journal of Applied Research in Memory & Cognition*, *5*(4), 478–489. https://doi.org/10.1016/j.jarmac.2016.07.010

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, *2*(8), e124. https://doi.org/10.1371/journal.pmed.0020124

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, *103*(1), 54–69. https://doi.org/10.1037/a0028347

Kane, M. J., Bleckley, M. K., Conway, A. R. A., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General*, *130*(2), 169–183. https://doi.org/10.1037/0096-3445.130.2.169

Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General*, *132*(1), 47–70. https://doi.org/10.1037/0096-3445.132.1.47

Lee, M. D., & Wagenmakers, E.-J. (2014). *Bayesian cognitive modeling: A practical course*. Cambridge University Press. https://doi.org/10.1017/CBO9781139087759

Lin, Y., Braver, T. S., Tay, A., Brough, R. E., Etzel, J. A., & Tang, R. (2022, September 6). DMCC individual differences. Retrieved from osf.io/zuvry

Long, J. A. (2019). Interactions: Comprehensive, User-Friendly Toolkit for Probing Interactions. R package version 1.1.0. https://cran.r-project.org/package=interactions

Ly, A., Etz, A., Marsman, M., & Wagenmakers, E.-J. (2019). Replication Bayes factors from evidence updating. *Behavior Research Methods*, *51*(6), 2498–2508. https://doi.org/10.3758/s13428-018-1092-x

Mitchell, M. N., & Chen, X. (2005). Visualizing main effects and interactions for binary logit models. *The Stata Journal*, *5*(1), 64–82. https://doi.org/10.1177/1536867X0500500111

Osborne, J. (2019). Logits and tigers and bears, oh my! A brief look at the simple math of logistic regression and how it can improve dissemination of results. *Practical Assessment, Research & Evaluation*. Advance online publication. https://doi.org/10.7275/39h8-n858

Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N. S., Smith, A. D., & Smith, P. K. (2002). Models of visuospatial and verbal memory across the adult life span. *Psychology and Aging*, *17*(2), 299–320. https://doi.org/10.1037/0882-7974.17.2.299

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion Decision Model: Current Issues and History. *Trends in Cognitive Sciences*, *20*(4), 260–281. https://doi.org/10.1016/j.tics.2016.01.007

Redick, T. S. (2014). Cognitive control in context: Working memory capacity and proactive control. *Acta Psychologica*, *145*, 1–9. https://doi.org/10.1016/j.actpsy.2013.10.010

Redick, T. S., & Engle, R. W. (2011). Integrating working memory capacity and context-processing views of cognitive control. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *64*(6), 1048–1055. https://doi.org/10.1080/17470218.2011.577226

Richmond, L. L., Redick, T. S., & Braver, T. S. (2015). Remembering to prepare: The benefits (and costs) of high working memory capacity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(6), 1764–1777. https://doi.org/10.1037/xlm0000122

Rosales, K. P., Snijder, J.-P., Conway, A. R., & Gonthier, C. (2022). Working memory capacity and dual mechanisms of cognitive control: An experimental-correlational approach. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*. Advance online publication. https://doi.org/10.1177/17470218211066410

Rouder, J. N., & Haaf, J. M. (2019). A psychometrics of individual differences in experimental tasks. *Psychonomic Bulletin & Review*, *26*(2), 452–467. https://doi.org/10.3758/s13423-018-1558-y

Servan-Schreiber, D., Cohen, J. D., & Steingard, S. (1996). Schizophrenic deficits in the processing of context. A test of a theoretical model. *Archives of General Psychiatry*, *53*(12), 1105–1112. https://doi.org/10.1001/archpsyc.1996.01830120037008

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*(11), 1359–1366. https://doi.org/10.1177/0956797611417632

Singmann, H., & Kellen, D. (2019). An introduction to mixed models for experimental psychology. In D. H. Spieler & E. Schumacher (Eds.), *New Methods in Cognitive Psychology* (pp. 4–31). Routledge. https://doi.org/10.4324/9780429318405-2

Snijder, J.-P., Tang, R., Bugg, J., Conway, A. R. A., & Braver, T. (2022). On the psychometric evaluation of cognitive control tasks: An investigation with the dual mechanisms of cognitive control (DMCC) battery. *PsyArXiv*. https://doi.org/10.31234/osf.io/z7ury

Stawarczyk, D., Majerus, S., Catale, C., & D'Argembeau, A. (2014). Relationships between mind-wandering and attentional control abilities in

young adults and adolescents. *Acta Psychologica*, *148*, 25–36. https://doi.org/10.1016/j.actpsy.2014.01.007

Tang, R., Bugg, J., Snijder, J.-P., Conway, A. R. A., & Braver, T. (2021). The dual mechanisms of cognitive control (DMCC) project: Validation of an on-line behavioral task battery. *PsyArXiv*. https://doi.org/10.31234/osf.io/ngwqc

Troller-Renfree, S. V., Buzzell, G. A., & Fox, N. A. (2020). Changes in working memory influence the transition from reactive to proactive cognitive control during childhood. *Developmental Science*, *23*(6), e12959. https://doi.org/10.1111/desc.12959

Troller-Renfree, S. V., Buzzell, G. A., Pine, D. S., Henderson, H. A., & Fox, N. A. (2019). Consequences of not planning ahead: Reduced proactive control moderates longitudinal relations between behavioral inhibition and anxiety. *Journal of the American Academy of Child & Adolescent Psychiatry*, *58*(8), 768–775.e1. https://doi.org/10.1016/j.jaac.2018.06.040

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent. *Journal of Memory and Language*, *28*(2), 127–154.

Unsworth, N., Redick, T. S., Heitz, R. P., Broadway, J. M., & Engle, R. W. (2009). Complex working memory span tasks and higher-order cognition: A latent-variable analysis of the relationship between processing and storage. *Memory*, *17*(6), 635–654. https://doi.org/10.1080/09658210902998047

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, *37*(3), 498–505.

Unsworth, N., & Engle, R. W. (2008). Speed and accuracy of accessing information in working memory: An individual differences investigation of focus switching. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(3), 616–630. https://doi.org/10.1037/0278-7393.34.3.616

van de Schoot, R., Winter, S. D., Ryan, O., Zondervan-Zwijnenburg, M., & Depaoli, S. (2017). A systematic review of Bayesian articles in psychology: The last 25 years. *Psychological Methods*, *22*(2), 217–239. https://doi.org/10.1037/met0000100

Verhagen, J., & Wagenmakers, E.-J. (2014). Bayesian tests to quantify the result of a replication attempt. *Journal of Experimental Psychology: General*, *143*(4), 1457–1475. https://doi.org/10.1037/a0036731

Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of p values. *Psychonomic Bulletin & Review*, *14*(5), 779–804. https://doi.org/10.3758/BF03194105

Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., Selker, R., Gronau, Q. F., Šmíra, M., Epskamp, S., Matzke, D., Rouder, J. N., & Morey, R. D. (2018). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review*, *25*(1), 35–57. https://doi.org/10.3758/s13423-017-1343-3

Wang, C., Li, B., & Yao, Y. (2021). Proactive control mediates the relationship between working memory and math ability in early childhood. *Frontiers in Psychology*, *12*, 611429. https://doi.org/10.3389/fpsyg.2021.611429

Wiemers, E. A., & Redick, T. S. (2018). Working memory capacity and intra-individual variability of proactive control. *Acta Psychologica*, *182*, 21–31. https://doi.org/10.1016/j.actpsy.2017.11.002

Received February 8, 2022
Revision received September 8, 2022
Accepted September 9, 2022 ■