Journal of Experimental Psychology: General

ISSN: 0096-3445

2024, Vol. 153, No. 12, 3130–3142
https://doi.org/10.1037/xge0001637

The Curve of Control: Nonmonotonic Effects of Task Difficulty on Cognitive Control

Miklos Bognar[1, 2], Mate Gyurkovics[3], Balazs Aczel[2], and Henk van Steenbergen[4, 5]
[1] Doctoral School of Psychology, ELTE Eotvos Lorand University
[2] Institute of Psychology, ELTE Eotvos Lorand University
[3] Institute of Neuroscience and Psychology, University of Glasgow
[4] Department of Cognitive Psychology, Institute of Psychology, Leiden University
[5] Leiden Institute for Brain and Cognition, Leiden University

The U-shaped curve has long been recognized as a fundamental concept in psychological science, particularly in theories about motivational accounts and cognitive control. In this study ($N = 330$), we empirically tested the prediction of a nonmonotonic, curvilinear relationship between task difficulty and control adaptation. Drawing from motivational intensity theory and the expected value of control framework, we hypothesized that control intensity would increase with task difficulty until a maximum tolerable level, after which it would decrease. To examine this hypothesis, we conducted two experiments utilizing Stroop-like conflict tasks, systematically manipulating the number of distractors to vary task difficulty. We assessed control adaptation and measured subjective task difficulty. Our results revealed a curvilinear pattern between perceived task difficulty and adaptation of control. The findings provide empirical support for the theoretical accounts of motivational intensity theory and expected value of control, highlighting the nonlinear nature of the relationship between task difficulty and cognitive control.

Public Significance Statement

Humans can improve their performance in certain situations when they are motivated or under some level of stress. However, there is a limit to how much we can handle, and pushing beyond that limit actually hinders our performance. In the study, we found that as we made the experimental task more challenging, people initially adapted better to conflicting signals. But there came a point where the tasks became too difficult, and their ability to adapt started to decline. Interestingly, this pattern of adaptation was influenced by how difficult people perceived the tasks to be.

Keywords: cognitive control, task difficulty, motivational intensity

Supplemental materials: https://doi.org/10.1037/xge0001637.supp

This article was published Online First September 19, 2024.

Joo-Hyun Song served as action editor.

Miklos Bognar https://orcid.org/0000-0003-2717-1769

Henk van Steenbergen https://orcid.org/0000-0003-1917-6412

Balazs Aczel and Henk van Steenbergen contributed equally as senior authors. Raw data, code for data management, and statistical analyses were written in R and are also available in the Open Science Framework and can be accessed at https://osf.io/cysx9/. Prior to peer review, a preprint version of this article was uploaded to https://psyarxiv.com/ywup9, and a poster presentation about the early findings of this study was held on the Psychonomic Society 63rd Annual Meeting. Eotvos Lorand University's institutional ethical review board approved these studies (2021/315). This study was conducted in accordance with the Declaration of Helsinki.

The authors have no conflicts of interest to disclose. This project (134918) has been implemented with the support provided by the Ministry of Innovation and Technology of Hungary from the National Research, Development and Innovation Fund (Grant FK_20) awarded to Balazs Aczel.

Miklos Bognar played a lead role in data curation, formal analysis, and investigation, a supporting role in conceptualization, funding acquisition, and writing–review and editing, and an equal role in methodology, project administration, resources, software, and writing–original draft. Mate Gyurkovics played a supporting role in conceptualization, methodology, and validation and an equal role in writing–review and editing. Balazs Aczel played a lead role in funding acquisition and an equal role in conceptualization, project administration, supervision, and writing–review and editing. Henk van Steenbergen played a lead role in validation, visualization, and writing–review and editing, a supporting role in formal analysis, and an equal role in conceptualization, investigation, methodology, project administration, software, supervision, and writing–original draft.

Correspondence concerning this article should be addressed to Henk van Steenbergen, Department of Cognitive Psychology, Institute of Psychology, Leiden University, Wassenaarseweg 52, 2333 AK, Leiden, The Netherlands. Email: HvanSteenbergen@fsw.leidenuniv.nl

U-shaped curves are ubiquitous in many theories of psychological science. Initially proposed by Yerkes and Dodson (1908) to describe the relationship between stimulus strength and speed of learning, it gained popularity when Hebb introduced the concept of U-shaped relationship between arousal and performance (Hebb, 1955). To this day, curvilinear functions remain a cornerstone in many psychological theories that aim to describe the relationship between motivational concepts such as arousal, conflict, and difficulty on the one hand and performance, effort, and cognitive control on the other hand. However, despite its popularity, the empirical evidence for these curvilinear relationships is often weak or simply not yet available (Teigen, 1994).

Influential theories on goal-directed behavior have proposed that the motivation to perform well on a task and the corresponding mental effort it involves responds to task difficulty in a nonlinear fashion. For example, in the classic motivational intensity theory (MIT), Brehm and Self (1989) have proposed that the level of motivation is a joint function of the perceived difficulty of the task and the maximum level of motivation possible or justified in a given context. Accordingly, motivation is assumed to increase with task difficulty, up to a certain level after which motivation drops. Indeed, numerous studies that have used cardiovascular measures of effort as an index of motivational intensity have consistently observed the U-shaped pattern in between-subject designs where participants perform a task with a given level of difficulty that can range from very easy to impossible (Richter et al., 2008).

At the same time, studies in the field of cognitive control have also observed adaptations in cognitive control that depend on the level of task difficulty, for example, using sequential trial-to-trial analyses (Gratton et al., 1992). The classic conflict monitoring account of cognitive control has proposed that these adaptations are driven by a neural conflict monitor that signals the need for additional control, assuming a monotonic relationship between conflict and subsequent control (Botvinick et al., 2001). However, given that task difficulty without conflict also triggers behavioral adaptation, it has been reasoned that these adaptations might reflect a response to a more generic signal that is also a response to disfluency or task difficulty (Dreisbach & Fischer, 2011), possibly triggering a transient negative affective state that facilitates the upregulation of control processes (Dignath et al., 2020; Dreisbach & Fischer, 2012; Saunders et al., 2017; van Steenbergen et al., 2009). This facilitation of cognitive control by emotional factors suggests a link between control processes and motivational states. Indeed, the expected value of control (EVC) framework (Shenhav et al., 2013) that has expanded the original conflict monitoring account does take into account these more general signals, as well as motivational factors. As such, the EVC framework has predicted a curvilinear relationship between task difficulty and cognitive effort intensity. According to EVC, when task difficulty increases, using the same level of control will reduce task success; therefore, to optimize the expected value of control, the level of control needs to increase, until a point where task difficulty is so high that the benefits of increasing control allocation no longer outweigh the costs associated with it. In line with EVC, neuropsychologically inspired models also have predicted U-shaped responses to task difficulty, which may be reflected by neurochemical boosting signals in the brain (Sarter et al., 2006; Silvetti et al., 2018).

There are several points where EVC is similar to the MIT framework, especially when considering the consensus that cognitive conflict resolution is effortful (Bouzidi & Gendolla, 2023). Recent work has indeed shown that MIT and EVC, although stemming from separate research traditions, can be integrated (Silvestrini et al., 2023) and share the prediction that the relationship between task difficulty and effort intensity is nonmonotonic. It is important to note, however, that there are several points where the definitions of motivation and effort are different in MIT and EVC. MIT predicts effects on mental effort, while EVC focuses only on cognitive control performance, and it does not consider effort intensity as a construct that is separate from control intensity. Critically, improvements in cognitive control performance may not always be directly proportional to increases in effort (Bouzidi & Gendolla, 2024; Silvestrini & Gendolla, 2019). For instance, due to the variability in individual cognitive control abilities, the same amount of control performance might reflect different levels of effort, as lower ability individuals might exert more effort to reach the same amount of cognitive control, as a compensatory mechanism (Hockey, 1997). There also might be a discrepancy between effort and cognitive performance in highly skilled individuals or those employing a cheating strategy, with low levels of effort still resulting in good performance (Silvestrini & Gendolla, 2019). MIT and EVC also posit different predictions on the relationship between effort intensity and motivation. In cases when the difficulty of the task is clear and fixed, MIT predicts effort intensity and potential motivation to be in a shark-fin-shaped relationship, as after the point of the maximal tolerable difficulty, effort allocation sharply drops to zero, rather than gradually decreasing as in an inverted U function. EVC on the other hand distinguishes between situations where an alternative worthwhile task is available or not. When it is available, EVC predicts a sharp decline in control allocation when the maximal tolerable difficulty is reached, similar to MIT. However, when no alternative task is available, EVC allows for a gradual decline in control allocation (Musslick et al., 2015). Critically, the prediction of a nonmonotonic relationship between task difficulty and effort is still shared by the two frameworks. To the best of our knowledge, this prediction has not been empirically tested on behavioral measures yet. In our model, we used cognitive control performance as a proxy to measure this effect on effort exertion.

To do so, we parametrically manipulated task difficulty in a given trial while measuring the effect on control adaptations in the next trial. We capitalized on the robust observation in cognitive control paradigms, such as the Stroop task, that conflict in a previous trial triggers increases in cognitive control in the next trial. This so-called conflict-adaptation effect can be observed in Stroop-like tasks where incongruent (conflict) and congruent (no-conflict) trials are randomly presented. The difference in response performance on incongruent and congruent trials is referred to as the congruency effect. This congruency effect is typically reduced if the preceding trial is incongruent, compared to when it is congruent, a phenomenon commonly referred to as the conflict-adaptation effect, or the congruency sequence effect (Egner, 2007; Gratton et al., 1992). There is a general, albeit not unanimous, consensus that this conflict-adaptation effect can be considered as a form of adaptive control that is triggered by the experienced difficulty level of the previous trial, provided that common confounds such as stimulus and responses repetitions are eliminated (Braem et al., 2019). However, to the best of our knowledge, only a few attempts have been made to systematically vary the level of trial difficulty to measure its impact on subsequent control. In one study, Forster et al. (2011) produced three levels of conflict by manipulating the number of distractors in the stimuli of an Eriksen flanker test. They observed that after conflict trials in comparison to no-conflict trials, control monotonically increased with the conflict level of the previous trial. More recently, Zhang et al. (2021, 2023) observed a similar pattern of findings in a confound-minimized design. However, these studies did not sample extreme levels of difficulty. Thus, the absence of a clear curvilinear pattern may be attributed to the limited range of difficulty levels sampled, which may reflect just one side of the alleged U shape.

In order to test the prediction that control intensity increases until a maximum tolerable task difficulty is reached and drops when the task becomes more difficult, we conducted two experiments with Stroop-like conflict tasks in which we systematically varied the levels number of distractors in inducer trials, using a wide range of difficulty levels, and tested the effect on control intensity in the subsequent diagnostic trials. In both experiments we measured perceived task difficulty using subjective ratings presented after the task. We hypothesized that increasing the number of distractors during incongruent inducer trials would result in a corresponding increase in subjective task difficulty and that control adaptation would exhibit a nonmonotonic, curvilinear pattern in response.

Method

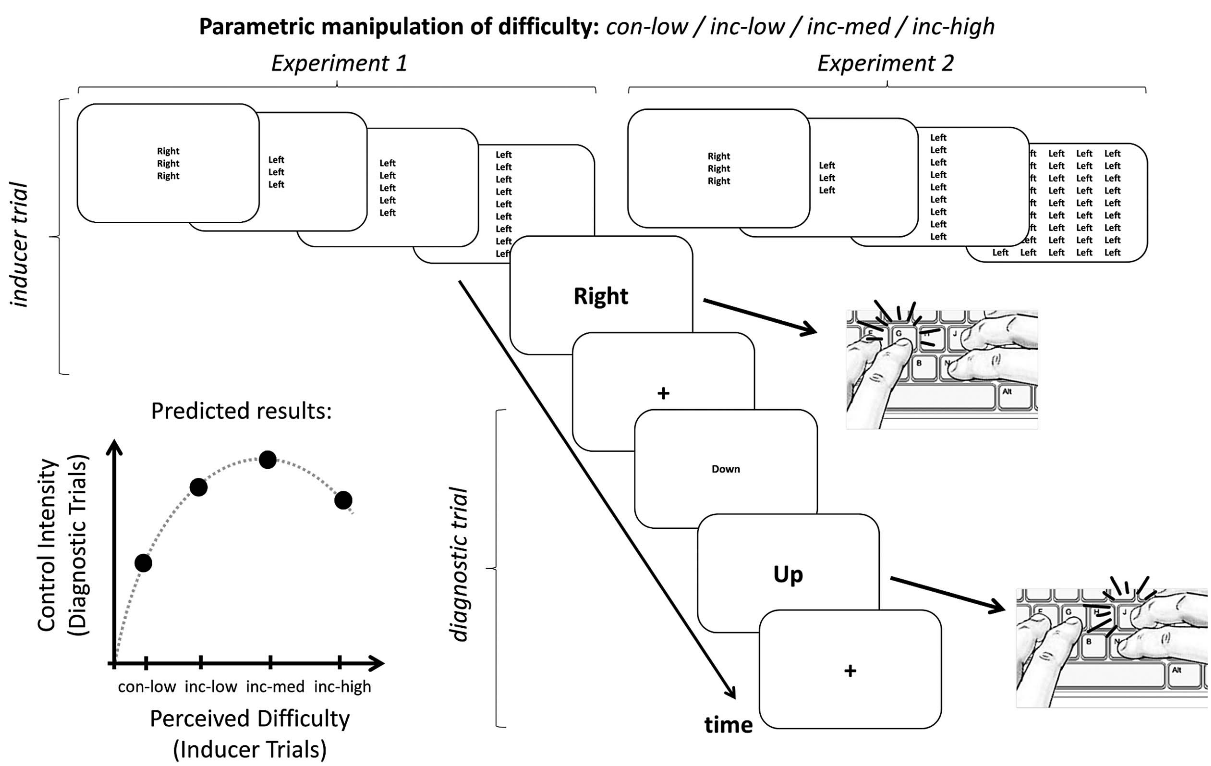
Figure 1 illustrates the key experimental design we employed. We used a Stroop-like prime-probe task, using congruent (con) and incongruent (inc) prime-probe pairs. In our design, we distinguished between inducer and diagnostic trials (Braem et al., 2019). To test the effect of trial difficulty on subsequent control, we parametrically manipulated the number of distractors during inducer trials and measured the impact on control in the subsequent diagnostic trial. These trials alternated throughout the experiment. Diagnostic trial difficulty was kept constant at a low level. We first conducted several small sample online pilot studies to determine an experimental manipulation that led to a monotonic increase in difficulty at the self-report level. In the two subsequent high-powered online experiments reported here, we tested two variants of our design using a different number of distractors.

Participants

We collected data for 330 participants in total in two experiments running at the same time from the local university participation pool, where participants received course credits for taking part in the

Figure 1

Experimental Design



Note. Schematic overview of trial setup and key prediction. In both experiments, we presented a prime-probe task with alternating inducer and diagnostic trials. In each trial, participants could ignore the prime and had to respond to the probe screen by pressing the corresponding button on the keyboard. In congruent trials, prime and probe directions were the same, whereas in incongruent trials these differed. Inducer trial difficulty was manipulated by increasing amounts of distractors in the prime stimulus. The number of distractors in the diagnostic trials' prime stimulus was fixed to one. Inset: We hypothesized that control intensity measured in the diagnostic trials would follow an inverted U-shaped response to the perceived difficulty of the diagnostic trials. con = congruent; inc = incongruent; med = medium.

experiment. Participants were randomly assigned to one of the two experiments when they signed up. They had 1 week to reach the online experiment page and download the task. At the end of the week, we analyzed all data received. In the beginning of the experiment, participants were asked about their sex and handedness. Every participant was a native Hungarian speaker, which was a participation criterion.

Note that we did not conduct precise sample size estimations before data collection because we aimed to reach as many available participants of our participant pool (600 participants) as possible. The online nature of the study allowed us to reach all potential participants; however, not everyone could participate given that E-Prime Go only runs on hardware with the Windows operating system. To detect small-to-medium within-subject effects with Cohen's $d \geq .41$ (reflecting the average effect size in psychology; Lakens & Evers, 2014; Richard et al., 2003) with 80% power requires 52 participants. A sample of 330 participants is more than enough, even when taking into account the very conservative rule of thumb (Barnhoorn et al., 2015) that online experiments may require samples that are four times larger than lab experiments. To test our crucial hypotheses, we used mixed models rather than repeated measures analyses of variance. Given that mixed-effect linear regressions are considered to be more effective in accounting for within-subject variability than analysis of variance measures (Bagiella et al., 2000), our statistical power estimations are probably underestimated. Our sample size also surpassed the recommended 1,600 measurements per conditions in mixed-effect model analysis of reaction time experiments (Brysbaert & Stevens, 2018).

In Experiment 1, a total of 169 participants took part in the task, and after exclusion (see Results section), data from 163 participants were analyzed ($Mdn_{\text{age}} = 21$ years, 70.4% female, 29.6% male, 84% right-handed). One hundred sixty-one participants took part in Experiment 2, and after exclusion, data from 155 participants were analyzed ($Mdn_{\text{age}} = 21$ years, 70.8% female, 29.2% male, 90.7% right-handed). Data were collected online in October 2021.

Apparatus and Stimuli

To collect data, we used an E-Prime Go experiment that was downloaded by the participants on their computer. The experiment was written in E-Prime 3.0 (Psychology Software Tools, Inc, 2020). After downloading the program, participants were instructed to close every other program on their computer and to start the experiment. The experiment tested the computer for performance issues. If no performance issue was found, the experiment started. Participants had to perform a prime-probe task during the experiment that was a modified version of a task developed earlier by Weissman et al. (2014). In the prime-probe task, two stimuli (a prime and a probe stimulus) follow each other in a short period of time on each trial. First a prime stimulus (a direction word, e.g., "left") is shown for 133 ms, then a short blank screen for 33 ms, and the probe stimulus (a direction word, e.g., "right") for 133 ms. The prime-probe trial is considered congruent if the prime and the probe directions are the same, and incongruent if they are different. After the probe stimulus, a fixation cross appeared until the end of the trial. Participants had to respond to the probe stimulus by pressing one of four response buttons ("f"—left; "g"—right; "n"—down; "j"—up). Participants had 1,500 ms in total to respond to the current trial (probe duration plus the fixation cross duration). Each trial lasted 2,000 ms. Please note that the actual timings varied slightly due to participants using displays with different refresh rates.

Procedure

In the prime-probe task, the vertical (up–down) and horizontal (left–right) dimensions alternated from trial to trial. Prime stimuli and probe stimuli could only contain directions from the given trial dimension (vertical or horizontal); there were no incongruent trials that combined, for example, left and down. Experiment 1 and Experiment 2 both manipulated the number of distractors of the prime stimulus of the inducer trials, but they differed in the specific levels used. In Experiment 1, the difficulty levels for low, medium (med), and high inducer trials were obtained by manipulating the number of vertically stacked distractor words using three, five, or nine words, whereas in Experiment 2, we used three words, nine words or a matrix of 9×5 words covering the entire display. We used a complete factorial design where the number of distractors was manipulated orthogonally to the level of congruency in the inducer trials. However, in the main analyses, we focused on the incongruent trials and used the low-difficulty congruent trial (with three distractors in both experiments) as the reference (easiest) condition. We refer to the other type of congruent trials (with more distractors) as filler trials. These trials were not of primary interest, but we included them to avoid associative learning between the number of distractors and trial congruency. Indeed, our earlier piloting work revealed that omitting filler trials made it actually easy to perform well on incongruent inducer trials, likely because the correct response could be predicted by the number of distractors presented. Trial difficulty was not manipulated for the diagnostic trials. Here, every prime stimulus consisted of only one direction word in the middle of the screen.

After the instructions, participants performed a practice block with 24 trials. In the practice block, an "error" message was shown after error or timeout trials. If the participant performed worse than 80% on the practice block, it was repeated until the 80% performance was reached. After the practice block, participants performed four test blocks, each including 120 trials. Each block started with a trial from the horizontal dimension, followed by a trial from the vertical dimension, in alternating order. In the first and the last block, we used the horizontal trials as inducer trials, whereas in the second and third block, we used the vertical trials as inducer trials. Error feedback was shown after every erroneous trial. Between every block, there was an information screen which allowed the participants to take a break or leave the experiment. At the end of the experiment, example trials from all difficulty types were listed to the participants who had to rate them by perceived difficulty on a scale from 1 to 9.

Data Preprocessing

We excluded six participants who performed below chance level (25%). For subjective difficulty diagnostics, we included only difficulty ratings data on inducer trials (six per participant) that were collected at the end of the experiments. In control adaptation analyses, only correct diagnostic trials were included. Overall, 6.8% of trials were excluded for being incorrect, 5.6% for being post-error trials, and 4.6% of trials were excluded for being outliers (2 *SD*s from the conditional means of the participant). In the inducer trial analyses, only inducer trials were included, with no additional exclusion criteria.

Statistics

All analyses and data filtering were performed using R Statistical Software (R Core Team, 2022), with the help of the "tidyverse" package (Wickham et al., 2019). In all of our analyses, we used mixed-effect regression models, using participants as random intercepts. To construct the random slope structure, first, we initially included all predictors as random effects in the model in a stepwise manner in the order of the term's contribution to the model and constructed the largest possible regression model that could still converge. Next, on the previously constructed largest possible model, we employed a backward elimination process to refine the model by excluding any fixed or random factors that were deemed unnecessary. If the second step eliminated any crucial fixed factors due to nonsignificant results, we reported the best fitted model output by the first step. We used the "buildmer" R package to conduct the above-described two-step process (Voeten, 2022). Both in the first and the second steps, we used the "bobyqa" optimizer to fit the models. By utilizing mixed-effect regression models with participants as random intercepts and employing a backward elimination approach for the random slope structure, we aimed to account for individual differences while identifying the most relevant predictors that contribute significantly to the model's overall performance. The experiments were entirely identical, except for inducer stimuli representing difficulty levels which were different in the two experiments. Relying on initial pilot data, we designed the stimuli with the focus on a measurable difference in perceived task difficulty across the four difficulty levels, while we kept all other parameters the same in the two experiments. Because the two experiments were nearly identical, to increase statistical power and to facilitate readability, we merged the data sets of the two experiments and added Experiment as a factorial predictor in all analyses. By this merge, we were able to create the inducer difficulty score which allowed us to measure the relative level of difficulty in both experiments. We created the inducer difficulty score by mapping the four inducer difficulty levels to numbers from 0 to 3, thus creating a numeric predictor. To measure the amount of adaptive control, we used the control intensity score, which was calculated on diagnostic trials by subtracting incongruent reaction times (RTs) from congruent RTs. By this method, we could calculate diagnostic congruency effects on different inducer conditions. We did not know a priori whether the highest level of task difficulty per experiment was sufficient to sample the right side of the proposed U shape. However, any differences between the experiments should become evident by the main effects or interaction effects with the factor Experiment in our statistical models. Nevertheless, on the request of one of the reviewers, we have conducted the same analyses on the two experiments separately and included their results in the Results section. These analyses confirmed the main findings reported in the main text, although the curvilinear pattern was not significant for some of the analyses. As we were hypothesizing curvilinear relationships in some of our analyses, we included both linear and quadratic terms in such models, so a significant quadratic term indicates evidence for a curvilinear pattern over and above potential linear effects. Figures were created using the "sjPlot" (Lüdecke, 2022) and "ggrain" (Allen et al., 2021) packages.

Transparency and Openness

Neither of the experiments were preregistered. Tasks and collected raw data are publicly shared on the Open Science Framework pages of the project. Code for data management and statistical analyses were written in R and are also available at https://osf.io/cysx9/. Prior to peer review, a preprint version of this article was published on PsyArXiv at https://osf.io/preprints/psyarxiv/ywup9/ (Bognar et al., 2024), and a poster presentation about the early findings of this study was held on the Psychonomic Society 63rd Annual Meeting.

Results

Our main hypothesis was that by parametrically increasing the level of inducer trial difficulty, control intensity in the diagnostic trial will increase, and thus the congruency effect will decrease, up to a certain inflection point from where the congruency effect may increase again. To do so, we first measured the perceived subjective difficulty of the inducer conditions. Then we analyzed the effect of difficulty conditions on control adaptation and tested whether subjective difficulty ratings predicted this adaptation of control. Finally, we ran a supplementary analysis in which we explored the behavioral effects during the inducer trials themselves.

Effect of Parametric Difficulty Manipulation on Perceived Difficulty

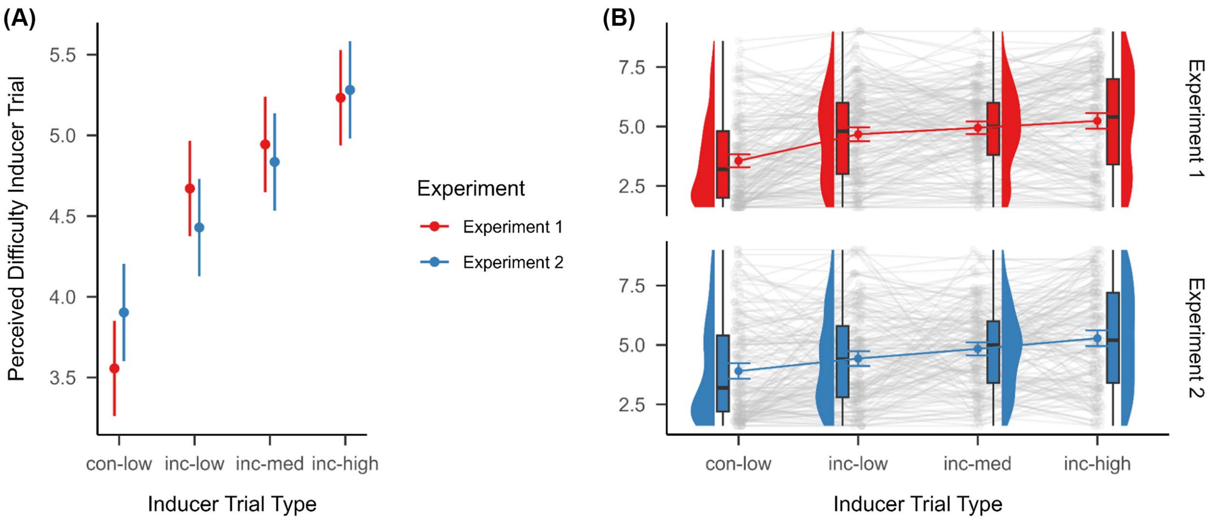
As shown in Table 1 and Figure 2, there was a monotonic increase in perceived difficulty as a function of inducer trial type, $p < .001$; $F(3, 960) = 36.78$, $\eta^2 = .10$. To see if the steps in the monotonic increase are equidistant, we analyzed the first derivate of difficulty level, and this revealed that the difference in mean ratings between subsequent difficulty levels are significant in Experiment 1, $p < .001$, $F(2, 326) = 11.94$, $\eta^2 = .068$, but not in Experiment 2, $p = .88$, $F(2, 315) = 0.12$, $\eta^2 = .001$. This suggests that we cannot simply assume that we manipulated perceived difficulty in equidistant steps, at least not in Experiment 1. Interestingly, we did not observe a significant difference in overall ratings $t(325) = -1.42$, $p = .15$,

Table 1
Summary for the ANOVA Analysis on the Mixed-Effect Polynomial Model for Inducer Trial Type on Perceived Difficulty

Term	Sum of squares	Mean of squares	Numerator degree of freedom	Denominator degree of freedom	F	p
Inducer trial type	264.6	88.2	3	960.0	36.8	<.001
Experiment	0.0	0.0	1	320.0	0.0	.93
Experiment × Inducer trial type	16.1	5.4	3	960.0	2.2	.08

Note. ANOVA = analysis of variance.

Figure 2
The Effect of Inducer Trial Type on Perceived Difficulty



Note. The effect of inducer trial type on perceived difficulty provided using a 9-point Likert scale after the task as a function of inducer trial type. Panel A: Predicted means and the 95% confidence intervals. Panel B: Individual rating data in gray and data distribution in color. con = congruent; inc = incongruent; med = medium. See the online article for the color version of this figure.

$d = -.16$ between Experiment 1 ($M = 4.35$, $SD = 1.98$) and Experiment 2 ($M = 4.55$, $SD = 2.01$), suggesting that the overall higher amount of distractors used in Experiment 2 did not increase overall subjective difficulty at the task level. Tables 2 and 3 show perceived difficulty across four difficulty levels of inducer trials.

Effect of Parametric Difficulty Manipulation on Adaptive Control

We then proceeded by verifying that the parametric difficulty manipulation affected adaptive control. To do so, we created the inducer difficulty score variable as a linear predictor of the RT in the diagnostic trials. A significant main effect of congruency was found in both Experiment 1, $t(162) = 37.1$, $p < .001$, and Experiment 2, $t(152) = 33.463$, $p < .001$. In Experiment 1 a significant curvilinear interaction was also found between Inducer Difficulty Score and Diagnostic Trial Congruency, $t(22{,}048) = 2.776$, $p = .005$; however, this interaction was not significant in Experiment 2, $t(20{,}508) = 1.776$, $p = .076$. Supplemental Tables S4 and S5 describe all fixed and random terms used in the models for the separate analyses.

Table 2
The Effect of Inducer Trial Type on Perceived Difficulty in Experiment 1

Inducer difficulty level	Mean rating	Rating *SD*
Con-low	3.53	0.73
Inc-low	4.63	1.89
Inc-med	4.90	1.71
Inc-high	5.20	2.11

Note. con = congruent; inc = incongruent; med = medium.

In the merged analysis, we used Experiment as a control predictor. Table 4 describes all the fixed and random terms that were used in the model for the merged analysis. The factor Experiment did not interact with the curvilinear interaction described above; Table 4 does not list this effect because it was eliminated from the initial model that included all main effects and interactions given that it did not explain substantial variance. As shown in Table 4, the main effect of Experiment was not significant in reaction time between the two experiments $t(319) = -.658$, $p = .511$. Figure 3A shows the corresponding estimated marginal means, showing both the typical main effect of congruency $t(309) = 33.245$, $p < .001$, and the curvilinear interaction between Inducer Difficulty Score and Diagnostic Trial Congruency, $t(42{,}583) = 3.188$, $p = .001$. All other significant effects are reported in Table 4. Figure 3B plots this effect using the control intensity score (inverse congruency effect). Consistent with our predictions, inducer trial difficulty on low levels first numerically increased control and then decreased. Descriptive statistics on reaction times can be found in Supplemental Table S11.

Effect of Perceived Difficulty on Adaptive Control

The initial analysis on perceived difficulty indicated that the variations in difficulty levels across conditions were not evenly spaced, at least not in Experiment 1. This observation might raise concerns about the validity of the analyses above that assumed a difficulty score with an interval level. We therefore created a new model that used the subjective difficulty reported as a predictor instead. This model used the average perceived difficulty calculated for each difficulty level and experiment separately as a predictor. When analyzing the experiments separately, the hypothesized curvilinear interaction between Subjective Difficulty and Diagnostic Trial Congruency did not yield significant results. Experiment 1: $t(22{,}047) = 1.906$, $p = .0566$. Experiment 2: $t(20{,}534) = 1.661$,

Table 3
The Effect of Inducer Trial Type on Perceived Difficulty in Experiment 2

Inducer difficulty level	Mean rating	Rating *SD*
Con-low	3.90	2.06
Inc-low	4.44	1.94
Inc-med	4.81	1.72
Inc-high	5.24	2.07

Note. con = congruent; inc = incongruent; med = medium.

p = .097. Supplemental Tables S6 and S7 describing the fixed and random factors used in the separate analysis models can be found in the Supplemental Materials.

In the merged analysis, we again used the averaged perceived difficulty calculated for each difficulty level and each experiment, yielding a predictor with eight unique values. Table 5 shows all fixed and random factors that were used in the merged analysis model as well as their significance levels. Critically, this model confirmed our main hypothesis of an inverted-U relationship, as reflected by a significant quadratic effect, $t(42{,}270) = 2.379$, $p = .017$, of perceived difficulty on the congruency effect. The corresponding effect on the control intensity score is visualized in Figure 4.

We also tested a similar model that included the individual ratings (i.e., without averaging), but this model did not reveal the hypothesized U shape ($t = 0.72$, $p = .47$). This could be due to the fact that the rating data was too noisy (one question per condition per participant) to predict behavior (~40 trials per condition per subject), or because the absolute value per participant cannot be directly used to predict effects at a population level.

Control Adaptation Effects Could Not Be Reliably Attributed to Conflict Strength in the Inducer Trials

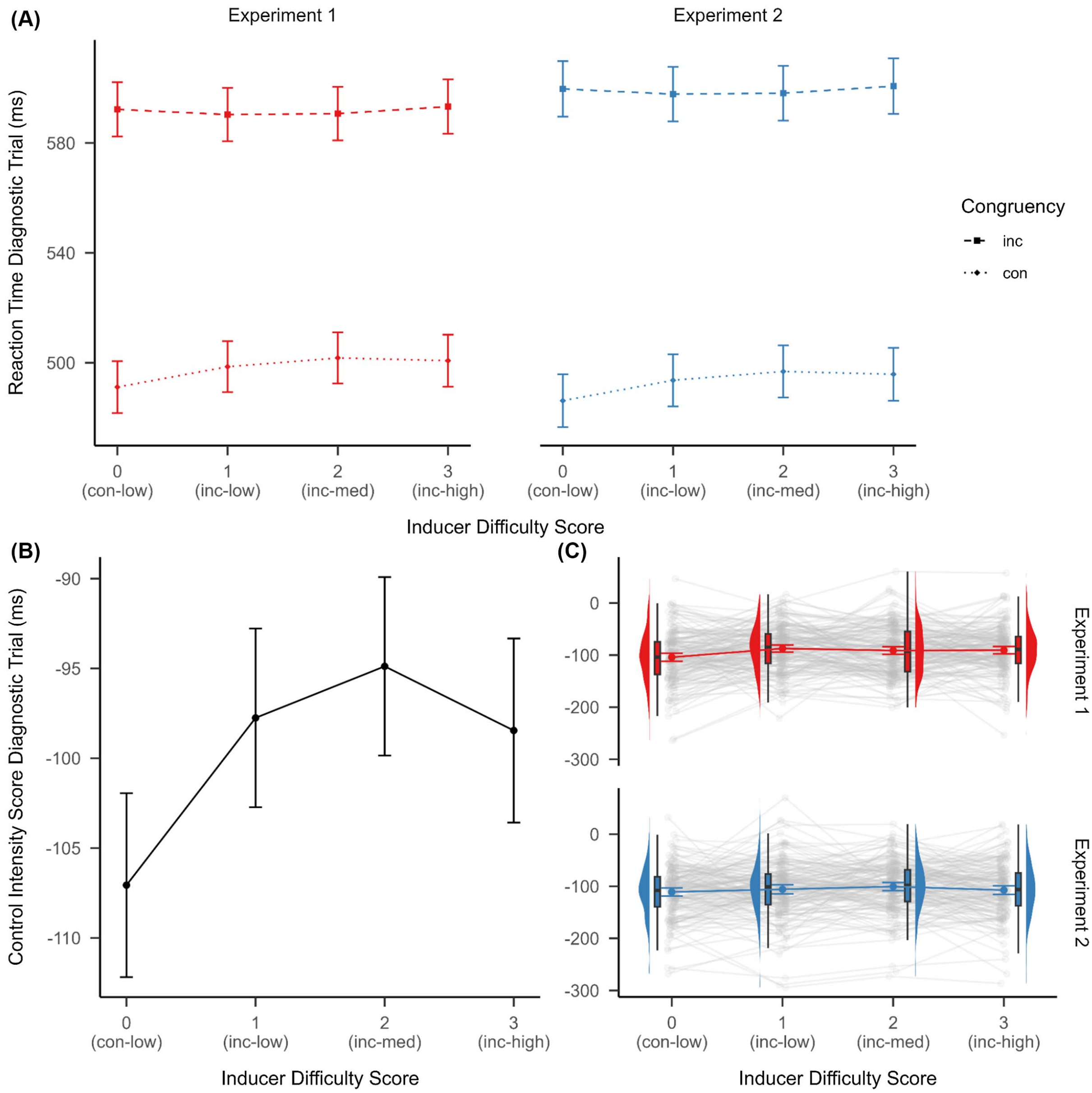
We performed a supplementary set of analyses that investigated the effect of difficulty level on performance during the inducer trials themselves. These were post hoc exploratory analyses aimed to reveal the processes that underlie the effect of increased perceived task difficulty and the associated adaptation effects reported above. In this analysis, we also included the filler trials (i.e., the con-med and con-high inducer trials; see Method section), thus employing a full factorial design. As Figure 5 shows, we run the same analyses on self-reported difficulty, reaction time, and error rates (see tables in Supplemental Materials). Reaction time, $t(77{,}143) = 51.407$, $p < .001$, error rate ($z = 19.081$, $p < .001$), and subjective difficulty rating, $t(1{,}612) = 12.745$, $p < .001$, all revealed a main effect of congruency, suggesting that conflict trials relative to no-conflict trials are associated with performance impairment and perceptions of increased task difficulty. However, this congruency effect did not increase with the number of distractors present. So this pattern of results suggests that increased task difficulty in our task was not due to an increase in conflict strength between primes and the probe. If anything, difficulty levels reduced the congruency effect in RT, $F(2, 77143) = 38.090$, $p < .001$, $\eta^2 = .0009$, for trials with high number of distractors (see Figure 5A). However, as it is visible in the figure, this effect was driven by a reaction time increase in the congruent trials. At the same time, participants made fewer errors (see Figure 5C) when the number of distractors increased (medium compared to low: $z = -3.313$, $p < .001$; high compared to low: $z = -5.458$, $p < .001$), an effect that was independent of congruency. This suggests that presenting a high number of distractors caused a shift in speed–accuracy trade-off, making participants more cautious, at least during the filler (con-med and con-high) trials. Perceived difficulty of congruent inducer trials as well as of incongruent trials increased with distractor numbers, $F(2, 1612) = 23.474$, $p < .001$, $\eta^2 = .03$, suggesting that an increase in the number of congruent distractors was also not associated with a perception of ease (see Figure 5E). It is interesting to note that the ratings of the con-med and con-high trials also showed great variability and suggest a bimodal distribution, suggesting that the one subset of participants perceived these trials as relatively easy whereas the others perceived them as relatively hard.

Table 4
Summary for the Mixed-Effect Polynomial Model for RT on Diagnostic Congruency × Difficulty

	Reaction time		
Predictor	Estimate	95% CI	p
Intercept	498.02	[488.83, 507.20]	**<.001**
Diagnostic trial congruency	93.60	[88.09, 99.12]	**<.001**
Inducer trial difficulty level (linear)	743.39	[461.22, 1025.57]	**<.001**
Inducer trial difficulty level (curvilinear)	−431.44	[−713.38, −149.50]	**.003**
Experiment	−4.40	[−17.52, 8.71]	.510
Diagnostic trial congruency × Trial difficulty level (linear)	−665.27	[−1070.57, −259.97]	**.001**
Diagnostic trial congruency × Trial difficulty level (curvilinear)	658.99	[253.81, 1064.18]	**.001**
Diagnostic trial congruency × Experiment	12.70	[4.80, 20.60]	**.002**
Random effects			
σ^2		10589.90	
$\tau_{00\text{participant}}$		3447.50	
$\tau_{11\text{participant. Diagnostic trial congruency}}$		984.42	
$\rho_{01\text{participant}}$		−0.16	
Intraclass correlation coefficient		0.26	
$N_{\text{participant}}$		322	
Observations		43,199	
Marginal R^2/conditional R^2		0.150/0.367	

Note. p values for statistically significant predictors are indicated in bold. RT = reaction time; CI = confidence interval.

Figure 3
Conflict Intensity as a Function of Inducer Difficulty Score



Note. Panel A: Predicted congruent (dotted line) and incongruent (dashed line) diagnostic trial means as a function of the four inducer difficulty scores with standard error as error bars. Panel B: Predicted diagnostic trial control intensity score as a function of the four inducer difficulty scores. Error bars represent pooled standard error. Panel C: Individual diagnostic trial control intensity scores in gray and data distribution in color. con = congruent; inc = incongruent; med = medium. See the online article for the color version of this figure.

Discussion

This study tested a critical prediction that can be inferred from both classic motivational intensity theory (Brehm & Self, 1989) and more recent neurocomputational and neurophysiological accounts (Sarter et al., 2006; Silvetti et al., 2018), as proposed in a recent unified framework (Silvestrini et al., 2023). Specifically, it examined whether perceived task difficulty is related to cognitive control recruitment in a curvilinear manner. In two experiments, we found evidence for this curvilinear relationship in control adaptation using sequential behavioral analyses. We observed that the parametric increase in distractors in conflict-inducer trials monotonically increased subjective difficulty but that the adaptation of cognitive control in the subsequent diagnostic trial increased with subjective task difficulty only to a certain point after which it plateaued or declined.

The general pattern of our findings was consistent across two experiments that used a different range of distractors, with a larger number of distractors being presented in some trials in Experiment 2

Table 5
Summary for the Mixed-Effect Polynomial Model for RT on Diagnostic Congruency × Subjective Difficulty

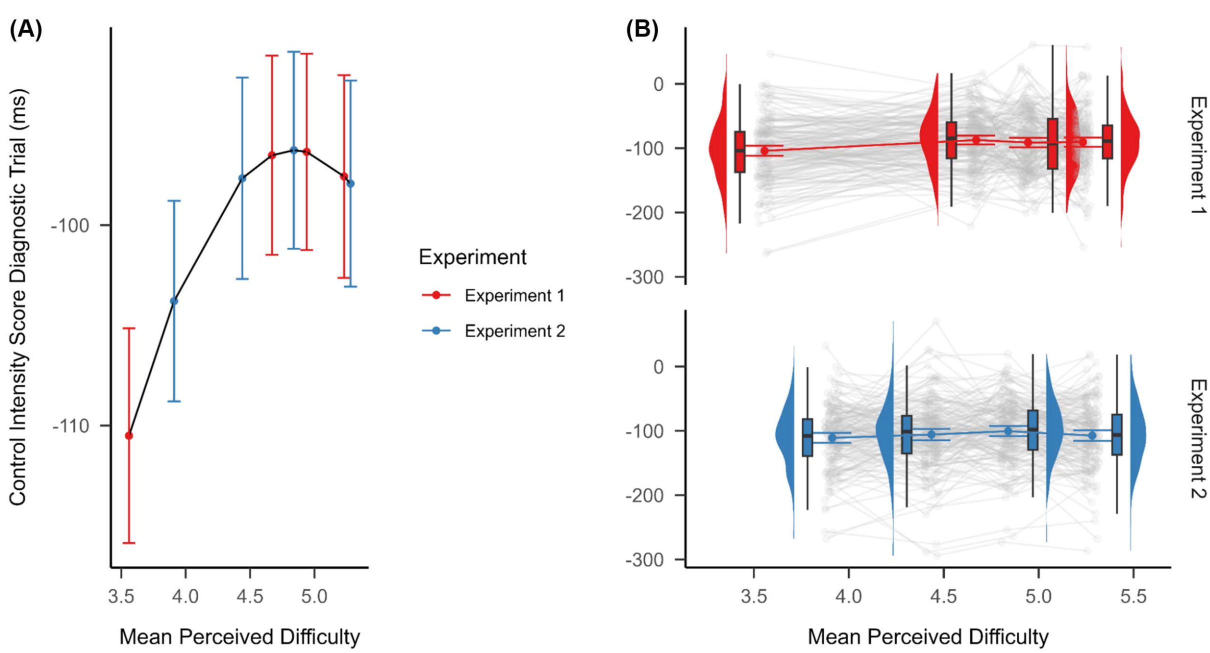
Predictor	Reaction time		
	Estimate	95% CI	p
Intercept	495.87	[489.31, 502.42]	**<.001**
Diagnostic trial congruency	99.81	[95.79, 103.83]	**<.001**
Subjective difficulty (linear)	848.38	[565.96, 1130.81]	**<.001**
Subjective difficulty (curvilinear)	−206.99	[−497.93, 83.95]	.163
Diagnostic trial congruency × Subjective difficulty (linear)	−850.69	[−1255.99, −445.39]	**<.001**
Diagnostic trial congruency × Subjective difficulty (curvilinear)	505.35	[88.98, 921.71]	**.017**
Random effects			
σ^2		10588.81	
$\tau_{00\text{participant}}$		3443.40	
$\tau_{11\text{participant. Diagnostic trial congruency}}$		1030.20	
$\rho_{01\text{participant}}$		−0.17	
Intraclass correlation coefficient		0.26	
$N_{\text{participant}}$		322	
Observations		43,199	
Marginal R^2/conditional R^2		0.149/0.367	

Note. p values for statistically significant predictors are indicated in bold. RT = reaction time; CI = confidence interval.

in comparison to Experiment 1. Despite these differences, subjective ratings were not higher in Experiment 2, nor was the inflection point of the U-shaped pattern in control intensity shifted. This suggests that experienced difficulty and the maximum amount of control justified are coded in a context-relative manner. This is consistent with the value normalization account (Rangel & Clithero, 2012), which proposes that signal values are computed using a normalized code relative to the signal value's position in the contextual distribution. It is also worth noting that the subjective ratings varied considerably between subjects, making it difficult to directly use self-report data to predict behavioral performance at an individual level. In our results, we therefore could only demonstrate U-shaped effects when using aggregated self-report data.

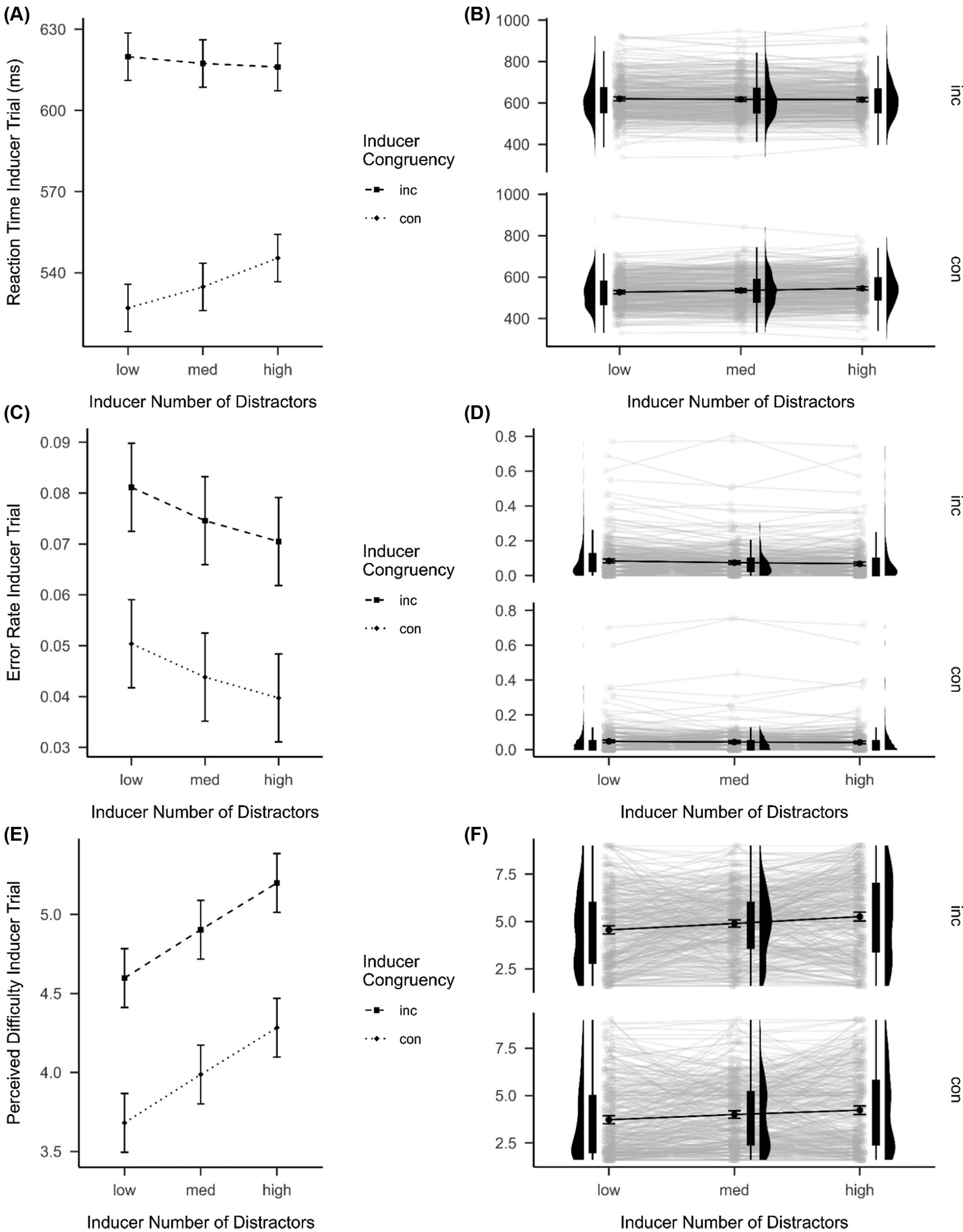
To the best of our knowledge, we are the first to report curvilinear, nonmonotonic effects of task difficulty, manipulated trial-wise, on cognitive control. Our findings thus go beyond earlier work that only

Figure 4
The Effect of Inducer Trial Perceived Difficulty on Control Intensity Score



Note. Panel A: Predicted diagnostic trial control intensity score as a function of inducer trial perceived difficulty. Colors represent the two experiments; error bars represent pooled standard error. Panel B: Individual diagnostic trial control intensity scores in gray, distribution in color. See the online article for the color version of this figure.

Figure 5
Inducer Trial Analysis



Note. Predicted reaction time (Panel A), error rate (Panel C), and perceived difficulty (Panel E) on inducer trials as a function of the different inducer trial types, in congruent (dotted line) and incongruent (dashed line) trials. Error bars represent standard error. (Panels B, D, F) Individual data of the variable plotted on the left are shown in gray and distribution in black. con = congruent; inc = incongruent; med = medium.

demonstrated monotonic effects of conflict strength on control adaptation (Forster et al., 2011; Zhang et al., 2021, 2023). Nevertheless, the curvilinear effects obtained here are consistent with early work that uses block-wise manipulations of task difficulty that reduced control adaptation effects (Fischer et al., 2010; van Steenbergen et al., 2015). The paradigm we developed here might also help to bridge literature that mainly used between-subject designs to show that cardiovascular measures of effort show a nonmonotonic response to task difficulty (Richter et al., 2008), studies in the neurochemistry field that have also suggested an inverted U-shaped relationship between adaptive control and dopamine levels (Bijleveld et al., 2023; Cools & D'Esposito, 2011; Westbrook et al., 2020), and recent work that has shown that cardiac effort can also be observed in conflict tasks (Bouzidi & Gendolla, 2023). While several studies suggested that the exertion of cognitive control is inherently effortful akin to physical labor (Kool & Botvinick, 2014; Kool et al., 2010), it is important to note that this study did not measure the amount of effort exerted by participants directly. Instead, we focused on the congruency effect as an inverse proxy for the intensity of cognitive control allocation. This way we assumed that more cognitive control allocation—and thus more effort—leads to smaller congruency effects, that is, to a range where floor and ceiling effects are absent (Norman & Bobrow, 1975). Note however that computational analyses have shown (e.g., Musslick et al., 2018) that congruency costs can reflect other variables, such as task automaticity; thus, we cannot be entirely sure that this study's curvilinear results represent pure effects on control intensity, and by extension, effort.

It is also worth noting that in contrast to the self-reported higher subjective difficulty of the inducer trials with more distractors during the conflict trials, we did not observe clear evidence that a higher number of distractors increased objective task difficulty as would be observed by slower and more error-prone behavior during conflicting inducer trials. To specify, although we did observe a congruency effect on inducer trials, this effect actually decreased (rather than increased) with the number of distractors presented, both in reaction time and accuracy. However, this effect was primarily due to slower responses on congruent trials. Note that we included congruent inducer trials to set congruency proportions on these trials to 50%. This prevented the participants from preparing for an incongruent probe after a difficulty-manipulated prime. In addition, participants also tended to make fewer errors when the number of distractors increased. Together, these findings suggest that our distractors display did not change conflict strength as such but rather shifted the participant's strategy to respond more cautiously. This change could be due to our decision to include diagnostic trials with very low task difficulty, which took up half of all trials in the experiment. As a result, more difficult trials were underrepresented in the overall experimental context. This decision may have caused higher difficulty trials to be less expected, making participants more cautious and perceiving these trials as more difficult. All in all, our findings point to the importance of taking into account subjective experience when investigating cognitive control. Although the causal role of subjective states in the adaptations of cognitive control is still a subject of dialogue (Questienne et al., 2021), some earlier work has suggested that adaptation effects in control only occur when participants report they experienced conflict (Desender et al., 2014). This suggests that adaptation of control can result from the meta-cognitive experience of difficulty. In addition, the experience of dysfluency often also involves negative integral (task-related) affect, which may play a causal role in conflict adaptation too (Dignath et al., 2020; Dreisbach & Fischer, 2015; Saunders et al., 2017; van Steenbergen et al., 2009).

This study also had a number of limitations. First, we assessed perceived difficulty only once through a questionnaire administered at the end of the study. Therefore, the usual caveats associated with self-report measures apply. To further test the role of integral affect on control adaptation, future work may probe subjective feelings within trials or combine our task with affective priming trials, similar to studies in regard to the affective modulation account (Dreisbach & Fischer, 2012; Fröber et al., 2017). Given the above approach, a posttest questionnaire would not be necessary. Second, we could not completely determine the origin of the increases in experienced task difficulty in high distractor conflict trials, which likely reflects the contribution of several cognitive processes, such as visual distraction and unexpectedness. Third, data was collected online, which may have increased noise and could explain why we failed to observe strong within-subject effects of perceived difficulty on performance. Fourth, as MIT and EVC predict slightly different shapes of the relationship between difficulty and effort, it would have been ideal to know which pattern would be supported more by the data. Analyses reported in Supplemental Material provided some evidence that our results at the individual level better resemble a gradual inverted U rather than a shark-fin shape, at least when compared to a linear shape. However, the overall variability did not allow us to conclude that the inverted U shape fit the data better than the shark-fin shape, so future studies are needed to investigate this issue in more detail, potentially with more observations per participant to allow for better estimation of individual-level effort curves. Fifth, as the results section showed, when analyzing the experiments separately, not all experimental effects were significant; however, patterns of effects were in line with the hypothesized nonmonotonic effect. At the same time, we considered pooling the two experiments—in order to increase statistical power—to be reasonable, because the different difficulty levels in the experiments were chosen arbitrarily and were controlled for in the pooled statistical models.

To conclude, using a novel design that parametrically manipulated a range of task difficulty levels, we were able to show that cognitive control measured in a Stroop-like task scales with perceived difficulty in a curvilinear fashion. These findings provide compelling evidence for a key prediction that can be derived from both classical accounts on motivation and modern accounts on cognitive control. Future studies may combine this paradigm with physiological and neural measures to determine the neurobiological mechanisms (Berger et al., 2020) underlying these quadratic effects. Moreover, given the assumed U-shaped effect that stress has on cognitive performance and recent evidence that control adjustments may predict the stress reactivity in daily life (Grueschow et al., 2021; Lin et al., 2023; van Steenbergen et al., 2021), our study also opens new avenues for investigating the complex interplay between subjective states and cognitive control, both in healthy and clinical samples.

Constraints on Generality

In this work, our goal was to address a fundamental scientific question regarding the relationship between perceived task difficulty and control adaptation, without a focus on lifespan or clinical aspects. As stated in the Method section, the experiments were conducted on healthy young adults due to the accessibility of this population in university settings and because it is a standard approach in the broader literature on conflict-adaptation research. There are findings indicating that conflict-adaptation effects might be influenced by age (Larson et al., 2016) or clinical conditions (e.g., Larson et al., 2011). Therefore, broad generalization of the reported findings beyond the current target population (healthy young adults) is not appropriate.

References

Allen, M., Poggiali, D., Whitaker, K., Marshall, T., van Lagen, J., & Kievit, R. A. (2021). Raincloud plots: A multi-platform tool for robust data visualization. *Wellcome Open Research*, *4*, Article 63. https://doi.org/10.12688/wellcomeopenres.15191.2

Bagiella, E., Sloan, R. P., & Heitjan, D. F. (2000). Mixed-effects models in psychophysiology. *Psychophysiology*, *37*(1), 13–20. https://doi.org/10.1111/1469-8986.3710013

Barnhoorn, J. S., Haasnoot, E., Bocanegra, B. R., & van Steenbergen, H. (2015). QRTEngine: An easy solution for running online reaction time experiments using Qualtrics. *Behavior Research Methods*, *47*(4), 918–929. https://doi.org/10.3758/s13428-014-0530-7

Berger, A., Mitschke, V., Dignath, D., Eder, A., & van Steenbergen, H. (2020). The face of control: Corrugator supercilii tracks aversive conflict signals in the service of adaptive cognitive control. *Psychophysiology*, *57*(4), Article e13524. https://doi.org/10.1111/psyp.13524

Bijleveld, E., van Breukelen, F. N., de Segovia Vicente, D., & Schutter, D. J. L. G. (2023). Mapping the dose–response relationship between monetary reward and cognitive performance. *Motivation Science*, *9*(3), 205–215. https://doi.org/10.1037/mot0000299

Bognar, M., Gyurkovics, M., Aczel, B., & van Steenbergen, H. (2024). *The curve of control: Non-monotonic effects of task difficulty on cognitive control*. https://osf.io/cysx9

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624–652. https://doi.org/10.1037/0033-295X.108.3.624

Bouzidi, Y. S., & Gendolla, G. H. E. (2023). Is cognitive conflict really effortful? Conflict priming and shielding effects on cardiac response. *Psychophysiology*, *60*(2), Article e14169. https://doi.org/10.1111/psyp.14169

Bouzidi, Y. S., & Gendolla, G. H. E. (2024). Cognitive conflict does not always mean high effort: Task difficulty's moderating effect on cardiac response. *Psychophysiology*, *61*(8), Article e14580. https://doi.org/10.1111/psyp.14580

Braem, S., Bugg, J. M., Schmidt, J. R., Crump, M. J. C., Weissman, D. H., Notebaert, W., & Egner, T. (2019). Measuring adaptive control in conflict tasks. *Trends in Cognitive Sciences*, *23*(9), 769–783. https://doi.org/10.1016/j.tics.2019.07.002

Brehm, J. W., & Self, E. A. (1989). The intensity of motivation. *Annual Review of Psychology*, *40*(1), 109–131. https://doi.org/10.1146/annurev.ps.40.020189.000545

Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition*, *1*(1), Article 9. https://doi.org/10.5334/joc.10

Cools, R., & D'Esposito, M. (2011). Inverted-U-shaped dopamine actions on human working memory and cognitive control. *Biological Psychiatry*, *69*(12), e113–e125. https://doi.org/10.1016/j.biopsych.2011.03.028

Desender, K., Van Opstal, F., & Van den Bussche, E. (2014). Feeling the conflict: The crucial role of conflict experience in adaptation. *Psychological Science*, *25*(3), 675–683. https://doi.org/10.1177/0956797613511468

Dignath, D., Eder, A. B., Steinhauser, M., & Kiesel, A. (2020). Conflict monitoring and the affective-signaling hypothesis-An integrative review. *Psychonomic Bulletin & Review*, *27*(2), 193–216. https://doi.org/10.3758/s13423-019-01668-9

Dreisbach, G., & Fischer, R. (2011). If it's hard to read… try harder! Processing fluency as signal for effort adjustments. *Psychological Research*, *75*(5), 376–383. https://doi.org/10.1007/s00426-010-0319-y

Dreisbach, G., & Fischer, R. (2012). Conflicts as aversive signals. *Brain and Cognition*, *78*(2), 94–98. https://doi.org/10.1016/j.bandc.2011.12.003

Dreisbach, G., & Fischer, R. (2015). Conflicts as aversive signals for control adaptation. *Current Directions in Psychological Science*, *24*(4), 255–260. https://doi.org/10.1177/0963721415569569

Egner, T. (2007). Congruency sequence effects and cognitive control. *Cognitive, Affective & Behavioral Neuroscience*, *7*(4), 380–390. https://doi.org/10.3758/CABN.7.4.380

Fischer, R., Plessow, F., Kunde, W., & Kiesel, A. (2010). Trial-to-trial modulations of the Simon effect in conditions of attentional limitations: Evidence from dual tasks. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(6), 1576–1594. https://doi.org/10.1037/a0019326

Forster, S. E., Carter, C. S., Cohen, J. D., & Cho, R. Y. (2011). Parametric manipulation of the conflict signal and control-state adaptation. *Journal of Cognitive Neuroscience*, *23*(4), 923–935. https://doi.org/10.1162/jocn.2010.21458

Fröber, K., Stürmer, B., Frömer, R., & Dreisbach, G. (2017). The role of affective evaluation in conflict adaptation: An LRP study. *Brain and Cognition*, *116*, 9–16. https://doi.org/10.1016/j.bandc.2017.05.003

Gratton, G., Coles, M. G., & Donchin, E. (1992). Optimizing the use of information: Strategic control of activation of responses. *Journal of Experimental Psychology: General*, *121*(4), 480–506. https://doi.org/10.1037/0096-3445.121.4.480

Grueschow, M., Stenz, N., Thörn, H., Ehlert, U., Breckwoldt, J., Brodmann Maeder, M., Exadaktylos, A. K., Bingisser, R., Ruff, C. C., & Kleim, B. (2021). Real-world stress resilience is associated with the responsivity of the locus coeruleus. *Nature Communications*, *12*(1), Article 2275. https://doi.org/10.1038/s41467-021-22509-1

Hebb, D. O. (1955). Drives and the C.N.S. (conceptual nervous system). *Psychological Review*, *62*(4), 243–254. https://doi.org/10.1037/h0041823

Hockey, G. R. J. (1997). Compensatory control in the regulation of human performance under stress and high workload; a cognitive-energetical framework. *Biological Psychology*, *45*(1–3), 73–93. https://doi.org/10.1016/S0301-0511(96)05223-4

Kool, W., & Botvinick, M. (2014). A labor/leisure tradeoff in cognitive control. *Journal of Experimental Psychology: General*, *143*(1), 131–141. https://doi.org/10.1037/a0031048

Kool, W., McGuire, J. T., Rosen, Z. B., & Botvinick, M. M. (2010). Decision making and the avoidance of cognitive demand. *Journal of Experimental Psychology: General*, *139*(4), 665–682. https://doi.org/10.1037/a0020198

Lakens, D., & Evers, E. R. K. (2014). Sailing from the seas of chaos into the corridor of stability: Practical recommendations to increase the informational value of studies. *Perspectives on Psychological Science*, *9*(3), 278–292. https://doi.org/10.1177/1745691614528520

Larson, M. J., Clayson, P. E., Keith, C. M., Hunt, I. J., Hedges, D. W., Nielsen, B. L., & Call, V. R. (2016). Cognitive control adjustments in healthy older and younger adults: Conflict adaptation, the error-related negativity (ERN), and evidence of generalized decline with age. *Biological Psychology*, *115*, 50–63. https://doi.org/10.1016/j.biopsycho.2016.01.008

Larson, M. J., Farrer, T. J., & Clayson, P. E. (2011). Cognitive control in mild traumatic brain injury: Conflict monitoring and conflict adaptation. *International Journal of Psychophysiology*, *82*(1), 69–78. https://doi.org/10.1016/j.ijpsycho.2011.02.018

Lin, L., Schwabe, L., Wang, X., Zhan, L., & Zhang, L. (2023). Neural correlates of conflict adaptation predict daily stress reactivity. *Psychophysiology*, *60*(8), Article e14279. https://doi.org/10.1111/psyp.14279

Lüdecke, D. (2022). *_sjPlot: Data visualization for statistics in social science_* [Computer software]. https://CRAN.R-project.org/package=sjPlot

Musslick, S., Cohen, J. D., & Shenhav, A. (2018). Estimating the costs of cognitive control from task performance: Theoretical validation and potential pitfalls. *40th annual meeting of the cognitive science society: Changing minds, CogSci 2018* (pp. 798–803). The Cognitive Science

Society. https://static1.squarespace.com/static/56e19a6d04426223be32a8b9/t/5b08209d352f533794ae90d3/1527259293998/MusslickCohenShenhav_CogSci2018.pdf

Musslick, S., Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2015). A computational model of control allocation based on the expected value of control. *Reinforcement learning and decision making conference* (Vol. *2015*). https://static1.squarespace.com/static/56e19a6d04426223be32a8b9/t/603ffcebfb86fd32de52b223/1614806254948/MusslickEtAl_RLDM2015.pdf

Norman, D. A., & Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology*, *7*(1), 44–64. https://doi.org/10.1016/0010-0285(75)90004-3

Psychology Software Tools, Inc. (2020). (E-Prime Go) [Computer software]. https://support.pstnet.com/

Questienne, L., van Dijck, J.-P., & Gevers, W. (2021). The Role of subjective experiences in conflict tasks: A review. *Psychologica Belgica*, *61*(1), 46–62. https://doi.org/10.5334/pb.508

R Core Team. (2022). *R: A language and environment for statistical computing* [Computer software]. https://www.R-project.org/

Rangel, A., & Clithero, J. A. (2012). Value normalization in decision making: Theory and evidence. *Current Opinion in Neurobiology*, *22*(6), 970–981. https://doi.org/10.1016/j.conb.2012.07.011

Richard, F. D., Bond, C. F., Jr., & Stokes-Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology*, *7*(4), 331–363. https://doi.org/10.1037/1089-2680.7.4.331

Richter, M., Friedrich, A., & Gendolla, G. H. E. (2008). Task difficulty effects on cardiac activity. *Psychophysiology*, *45*(5), 869–875. https://doi.org/10.1111/j.1469-8986.2008.00688.x

Sarter, M., Gehring, W. J., & Kozak, R. (2006). More attention must be paid: The neurobiology of attentional effort. *Brain Research Reviews*, *51*(2), 145–160. https://doi.org/10.1016/j.brainresrev.2005.11.002

Saunders, B., Lin, H., Milyavskaya, M., & Inzlicht, M. (2017). The emotive nature of conflict monitoring in the medial prefrontal cortex. *International Journal of Psychophysiology*, *119*, 31–40. https://doi.org/10.1016/j.ijpsycho.2017.01.004

Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The expected value of control: An integrative theory of anterior cingulate cortex function. *Neuron*, *79*(2), 217–240. https://doi.org/10.1016/j.neuron.2013.07.007

Silvestrini, N., & Gendolla, G. H. E. (2019). Affect and cognitive control: Insights from research on effort mobilization. *International Journal of Psychophysiology*, *143*, 116–125. https://doi.org/10.1016/j.ijpsycho.2019.07.003

Silvestrini, N., Musslick, S., Berry, A. S., & Vassena, E. (2023). An integrative effort: Bridging motivational intensity theory and recent neurocomputational and neuronal models of effort and control allocation. *Psychological Review*, *130*(4), 1081–1103. https://doi.org/10.1037/rev0000372

Silvetti, M., Vassena, E., Abrahamse, E., & Verguts, T. (2018). Dorsal anterior cingulate-brainstem ensemble as a reinforcement meta-learner. *PLOS Computational Biology*, *14*(8), Article e1006370. https://doi.org/10.1371/journal.pcbi.1006370

Teigen, K. H. (1994). Yerkes–Dodson: A law for all seasons. *Theory & Psychology*, *4*(4), 525–547. https://doi.org/10.1177/0959354394044004

van Steenbergen, H., Band, G. P. H., & Hommel, B. (2009). Reward counteracts conflict adaptation. Evidence for a role of affect in executive control. *Psychological Science*, *20*(12), 1473–1477. https://doi.org/10.1111/j.1467-9280.2009.02470.x

van Steenbergen, H., Band, G. P. H., & Hommel, B. (2015). Does conflict help or hurt cognitive control? Initial evidence for an inverted U-shape relationship between perceived task difficulty and conflict adaptation. *Frontiers in Psychology*, *6*, Article 974. https://doi.org/10.3389/fpsyg.2015.00974

van Steenbergen, H., de Bruijn, E. R., van Duijvenvoorde, A. C., & van Harmelen, A.-L. (2021). How positive affect buffers stress responses. *Current Opinion in Behavioral Sciences*, *39*, 153–160. https://doi.org/10.1016/j.cobeha.2021.03.014

Voeten, C. C. (2022). *_buildmer: Stepwise elimination and term reordering for mixed-effects regression_* [Computer software]. https://CRAN.R-project.org/package=buildmer

Weissman, D. H., Jiang, J., & Egner, T. (2014). Determinants of congruency sequence effects without learning and memory confounds. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(5), 2022–2037. https://doi.org/10.1037/a0037454

Westbrook, A., van den Bosch, R., Määttä, J. I., Hofmans, L., Papadopetraki, D., Cools, R., & Frank, M. J. (2020). Dopamine promotes cognitive effort by biasing the benefits versus costs of cognitive work. *Science*, *367*(6484), 1362–1366. https://doi.org/10.1126/science.aaz5891

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., Francois, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), Article 1686. https://doi.org/10.21105/joss.01686

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology & Psychology*, *18*(5), 459–482. https://doi.org/10.1002/cne.920180503

Zhang, M., Li, Q., Li, Y., Chen, Y., Gu, Y., Yin, S., & Chen, A. (2023). Temporal dynamics of conflict adaptation across different conflict strengths. *Psychophysiology*, *60*(1), Article e14160. https://doi.org/10.1111/psyp.14160

Zhang, M., Li, Q., Yin, S., & Chen, A. (2021). Changes in the level of conflict trigger conflict adaptation. *Acta Psychologica Sinica*, *53*(2), 128–138. https://doi.org/10.3724/SP.J.1041.2021.00128

Received June 26, 2023
Revision received May 20, 2024
Accepted June 12, 2024 ■