Received: 11 February 2022 | Accepted: 1 October 2022

DOI: 10.1002/ejsp.2921

EASP WILEY

RESEARCH ARTICLE

Expecting tasks to help or hurt subsequent cognitive performance: Variability, accuracy, and bias in forecasted after-effects

Zoë Francis[1] | **Michael Inzlicht**[2]

[1]University of the Fraser Valley, British Columbia, Canada

[2]University of Toronto, Ontario, Canada

Correspondence
Zoë Francis, University of the Fraser Valley, British Columbia, Canada.
Email: zoe.francis@ufv.ca

Abstract

After-effects on cognition—where a prior activity either benefits or hinders subsequent cognitive performance—are empirically inconsistent. Do people have insight into when their subjective energy and cognition will be helped or hurt by engaging in prior activities? Studies 1a and 1b (combined $N = 316$) find that people expect more demanding and unenjoyable tasks to hinder their subsequent energy and cognitive performance, regardless of their willpower lay theory. Study 2 ($N = 167$) examines the accuracy of these forecasts using a within-subject design. Participants' forecasts of their future subjective states did predict their actual experienced subjective states, but participants were not able to accurately forecast their subsequent maths performance. Additionally, they significantly overestimated the detrimental effects of demanding prior activities on both subjective state and performance. Study 3 ($N = 210$) found that participants' overestimation of detrimental after-effects could result in unnecessary financial costs, suggesting these biased forecasts can have consequences.

KEYWORDS
cognition, depletion, forecasting, lay theories, self-control

1 | INTRODUCTION

How would you feel after babysitting a neighbour's children? Or after exercising? Or writing an exam? You might predict that these activities would affect how you felt afterwards, your feelings of mental energy, and even your ability to succeed on a subsequent cognitive task. If you personally imagine babysitting or exercise to be demanding and unenjoyable, you might expect that these activities would make you feel mentally drained. If you imagine these activities as enjoyable, you might expect them to leave you feeling sharp and ready to tackle another task. Or you may believe that you would be similarly energetic and focused on a cognitive task, regardless of what activity you engaged in beforehand.

After-effects on cognition, where engaging in a prior task affects one's subsequent cognitive performance, have been extensively researched in fields of both ego depletion and exercise psychology (Inzlicht & Friese, 2019; Pontifex et al., 2019). Empirically, these literatures find mixed results—both positive and negative after-effects can occur—but people's personal forecasts of mental energy and after-effects have been relatively neglected. Personal forecasts, however, are what people rely on to make decisions and plan for the future. Here, we examine people's personal forecasts of after-effects, how they relate to willpower lay theories, and the extent to which these forecasts are accurate or biased.

1.1 | After-effects on cognitive performance

In everyday life, our mood and subjective energy levels fluctuate and can be influenced by the activities that we engage in (Zhang et al., 2018). Can activities also affect our subsequent cognition? After-effects on cognition have been of interest since the early days of psychology (Franz, 1897), but have recently been most studied in the fields of *ego depletion* and *exercise after-effects*.

The theory of ego depletion states that prior mental exertion (especially involving self-control) has detrimental effects on subsequent performance on a self-control task (Baumeister et al., 1998). While hundreds of depletion studies have been published, depletion effects do not consistently occur—high-powered replication attempts have found that prior effortful cognitive tasks often have no effect on subsequent performance at the group level (Hagger et al., 2016; Vohs et al., 2021). Some research suggests that detrimental after-effects may be most likely to occur if the initial task is sufficiently effortful (Dang et al., 2017; Lin et al., 2020) and may be less likely to occur if the initial task is rewarding (Francis & Inzlicht, 2016; Kurzban et al., 2013) or creates a feeling of self-efficacy (Chow et al., 2015).

Faciliatory effects of prior effort on subsequent self-control are also sometimes reported; these beneficial after-effects have been referred to as reverse-depletion, faciliatory effects, or learned industriousness, depending on the proposed underlying mechanism (Converse & Deshon, 2009; Primoceri et al., 2021; Savani & Job, 2017). Beneficial effects of short periods of physical activity on subsequent cognition, especially, have been reported in the field of exercise cognition (Pontifex et al., 2019). These after-effects of physical activity may likewise be moderated by perceptions of exercise as fun or effortful (Werle et al., 2015), such that more enjoyable forms of physical activity have more beneficial after-effects. After-effects on cognition are thus heterogenous and are likely to depend on specific characteristics of the preceding task, including effortfulness of the task, perceptions of the task enjoyment, similarity between the preceding and subsequent task, and additional yet unknown factors.

Even when examining the after-effects of one particular preceding task on a particular subsequent cognitive task, researchers find substantial variability across samples and across people (Carter et al., 2015; Dang, 2018; Wenzel et al., 2019). Most studies are interested only in the *average* after-effect on cognition—and often, studies find no overall effect. However, an absence of an overall after-effect does not mean that the activity did not affect different people in different ways; between-person effects cannot be generalized to within-person conclusions (Fisher et al., 2018). Indeed, hidden moderators have been proposed as one explanation for null effects in studies of depletion (Dang, 2016), although no known individual difference variables have been shown to reliably explain between-subject heterogeneity (Wenzel et al., 2019). Potentially, however, different people could experience either beneficial or detrimental after-effects after a given task, depending on how they personally perceived and experienced the task.

1.2 | Forecasting after-effects

If people respond differently to effortful tasks—in ways that are systematic, rather than random—people themselves may know and be able to forecast what types of tasks will help or hurt their own subsequent cognitive performance. Can people accurately state whether they will experience a hindering after-effect (depletion), a faciliatory after-effect (reverse-depletion), or neither after a given task? If so, participants' meta-cognitive knowledge could act as a promising moderator for heterogeneity in the depletion effect.

It could also be that people's personal perceptions of a task—how enjoyable, difficult, or rewarding they find the activity—are related to their forecasted after-effect of that task. For example, if someone expects physical exertion to be unenjoyable, they might also expect that physical exertion will hurt their subsequent cognitive performance. By asking participants about their perceptions of tasks—how much the task is difficult (Dang et al., 2017; Lin et al., 2020), enjoyable (Kurzban et al., 2013; Werle et al., 2015) or leads to a sense of accomplishment (Chow et al., 2015)—we can aim to better explain potential heterogeneity in forecasted after-effects. Note that these potential associations between perceived task characteristics and forecasted depletion could exist regardless of whether people themselves consciously believe that task enjoyment (or difficulty, etc.) is related to after-effects on cognition—for example, the association could be merely a consequence of someone's past experiences with exercise as unenjoyable and their past experiences with exercise making it more difficult to later concentrate on their homework.

Aim 1: *Characterize people's forecasts of after-effects. To what extent is there consensus in forecasts across and within participants? What tasks—with what characteristics—are expected to cause detrimental or beneficial after-effects (e.g., depletion or reverse-depletion)? (Study 1 and 2)*

1.3 | Willpower lay theories

General beliefs or expectations of after-effects on cognition are known to vary across people. Most research on people's after-effect beliefs has focused on the individual difference of willpower theory (Francis & Job, 2018; Job et al., 2010; Mukhopadhyay & Johar, 2005). People who have a limited willpower theory report that they must take a break after strenuous mental exertion, that they need to re-charge after an effortful task. People who have a non-limited willpower theory report that prior mental exertion has no effect on them, or even that prior exertion prepares them to perform even better—a belief in *faciliatory* after-effects. Beliefs about physical exertion's effect on cognitive performance are also likely to vary across people, although these have been less extensively studied.

Although an initially promising moderator of heterogenous after-effects (Job et al., 2010; Savani & Job, 2017), willpower theory has not reliably predicted whether people actually perform more poorly on a cognitive self-control task after prior mental exertion. Replication studies of depletion have most often failed to find a moderation by willpower theory (Carruth et al., 2018; Dang et al., 2021; Vohs et al., 2021; Wenzel et al., 2019). The lack of explanatory power of willpower theory may be because the measure asks about 'strenuous mental exertion', in general terms. Two people may both believe that 'strenuous mental exertion' will hurt their subsequent performance, but differ in their perceptions about whether a particular task is actually strenuous: for example, controlling one's emotions might seem strenuous for one person, but not for another (Bernecker & Job, 2017; Sun et al., 2019). In this current work, we thus ask people to forecast how *specific* tasks will affect their own personal state and subsequent cognitive performance.

We do expect associations between willpower theory and people's forecasts about specific tasks. We hypothesize that people with more limited willpower theories will forecast detrimental after-effects more than will people with non-limited willpower theories. We also expect that limited willpower theorists will specifically forecast negative after-effects (i.e., depletion) to occur after activities that they perceive as demanding. In other words, willpower theory should moderate the relationship between forecasted demand of the preceding activity and forecasted after-effects, such that the relationship between forecasted demand and negative after-effect is stronger for limited willpower theorists.

Aim 2: *Investigate whether participants with more limited willpower theories forecast and/or experience more substantial detrimental after-effects than do participants with more non-limited willpower theories. (Study 1, 2, and 3)*

1.4 | Forecasts: Biases and accuracy

Forecasts of our future states and behaviours are fundamental to humans' ability to engage in future-oriented behaviour, including all types of planning and self-control decisions (Wilson & Gilbert, 2003). Forecasts are known to affect people's self-regulation and are integral for prospective decision-making (Fayard et al., 2012; Kotabe et al., 2019). While much forecasting research focuses on anticipated emotions, understanding forecasts of cognitive performance are also critical. For example, people use forecasts of their future self-control when deciding when to schedule their exercise (Delose et al., 2015), and when deciding whether or not to put themselves in a tempting situation (Nordgren et al., 2009). Despite their potential importance, little is known about the accuracy or biases for forecasts of after-effects, specifically.

In general, people's forecasts are not always accurate. Both impact bias and optimism bias frequently influence people's expectations of the future—these biases may also be involved in after-effect forecasts, with potentially opposite effects.

1.4.1 | Optimism bias

Optimism bias describes how people are often overly and unrealistically optimistic when forecasting their own futures (Armor & Taylor, 1998; Weinstein, 1980). People believe that negative outcomes are less likely to happen to them (Clarke et al., 2000; Shepperd et al., 2017), expect to complete tasks more quickly than is realistic (Roger et al., 1994), and overestimate the likelihood of positive outcomes occurring (Weinstein, 1980). Optimism bias is thought to occur due to self-enhancement needs, where people tend to attribute more positive characteristics to themselves (Brown, 1986; Regan et al., 1995). While accuracy in one's forecasts or expectations is important, maintaining a positive self-perception is also important, and optimism biases are thus regularly observed (Armor & Taylor, 1998).

In the context of after-effect forecasts, optimism biases could result in people enhancing their expectations for their self-control and ability to succeed at cognitive tasks, even under difficult circumstances. Prior research on self-control forecasts has found that people overestimate their ability to exert self-control in the presence of temptations (Nordgren et al., 2009). It could also be the case that people overestimate their ability to exert self-control even immediately after difficult or fatiguing prior activities. Optimism biases could thus minimize the extent to which people forecast negative after-effects.

1.4.2 | Impact bias

Forecasts are also affected by *impact bias*—people often overestimate the impact (both positive and negative) that experiences will have on our mood and well-being (Wilson & Gilbert, 2003). Impact biases are common when forecasting hedonic events, such as making a purchase or missing a train (Gilbert et al., 2004)—we overestimate how happy we will be after the former and how upset we will be after the latter. However, impact biases do not occur in response to every forecasted situation; occasionally people instead underestimate how much an event will affect their state, especially when the events are mild or low-intensity and short in duration (Buechel et al., 2017).

To our knowledge, potential impact biases have not been explored in the context of forecasting after-effects on cognition (either positive or negative). The ubiquity of impact biases would suggest that people may overestimate the extent to which prior activities will affect their subsequent performance, either positively or negatively, depending on their beliefs. People who expect a task to improve their subjective energy and subsequent task performance may overestimate the beneficial impact, and people who expect a task to hinder their performance may likewise overestimate the detrimental impact. However, negative after-effects might be overestimated to a greater extent than positive after-effects, due to a negative valence effect (Christophe & Hansenne, 2021). Alternatively, because many tasks proposed to cause after-effects are relatively trivial events with modest durations, impact biases may instead play a minimal role in people's forecasts of after-effects.

1.4.3 | Accuracy

Even biased forecasts can have a degree of accuracy. A group of three students may all underestimate how long their homework will take (optimism bias), yet one student may accurately predict that her history homework will take longer than chemistry while another accurately predicts that chemistry will be slower. Indeed, while people overestimate the emotional impact of events, there is substantial convergence between people's expected and actual emotional responses (Christophe & Hansenne, 2021; Robinson & Clore, 2001; Wilson &

Gilbert, 2003). Even if after-effect forecasts are biased, people's individual forecasts may still predict their personal after-effects. We are thus able to simultaneously examine whether people's forecasts are biased and examine whether people's forecasts may predict their own personal experiences in a sequential task paradigm.

Aim 3: *Describe the bias and accuracy in people's forecasts of after-effects. Are after-effect forecasts affected by optimism bias, impact bias, or both? Do people's forecasts accurately predict their personal experiences? (Study 2)*

1.5 | Consequences of forecasts

Forecasts and beliefs about one's self-regulation are important because they are used for decision-making and planning for the future (Scholer et al., 2018). For example, when people overestimate their ability to resist a cigarette, they put themselves in tempting situations and are more likely to smoke (Nordgren et al., 2009). When people expect that exercise will be unenjoyable, they report weaker intentions to exercise in the future (Loehr & Baldwin, 2014). When people imagine feeling depleted and fatigued, they select less complex and challenging movies to watch (Eden et al., 2018). Forecasts of after-effects could themselves be used when people are making decisions about their futures. Unfortunately, biases or inaccuracies in forecasts can sometimes result in non-optimal decision-making (Armor & Taylor, 1998; Shepperd et al., 2017).

Aim 4: *Examine whether after-effect forecasts influence decision-making, and whether bias in one's forecasts may result in suboptimal decisions. (Study 3)*

1.6 | Approach of current research

1.6.1 | Task selection

The current studies explore people's particular expectations or forecasts of how they would be affected by a variety of tasks. Across studies, we prioritized tasks that are familiar to participants and are ecologically valid, resembling activities beyond the lab. We intentionally included a larger number of tasks (six to fourteen) than are typically studied in depletion or exercise after-effect literatures, and statistically model 'task' as a random effect where possible. This method allows us to generalize our claims somewhat beyond the particular tasks included (Judd et al., 2012; Yarkoni, 2022).

While our approach and task-selection were inspired by the depletion and reverse-depletion literatures, we did not restrict our selection of preceding tasks to 'self-control' tasks. We are interested broadly in people's forecasts of after-effects on subjective energy and cognition, beyond the scope of the existing depletion and exercise cognition literatures, and are agnostic of the mechanisms underlying any actual after-effects (beneficial or detrimental) that may or may not occur.

1.6.2 | Within-subject designs

By using within-subject designs throughout, we can examine the correspondence between people's forecasts and their actual experiences. Such designs also allow us to separate bias and accuracy in people's forecasts and have improved statistical power. Furthermore, by examining both forecasts and actual states and performance as within-subject processes, these studies can speak directly to within-subject processes, which is typically what scholars care most about (Fisher et al., 2018).

2 | STUDIES 1A AND 1B

These studies describe people's forecasts of how various specific tasks would help or hinder their subsequent performance on a speeded maths-test (Study 1a) and how these tasks would change their likelihood of subsequently behaving inconsistently with a personal goal (Study 1b). We examined the variability within-person and within-task, what task characteristics are associated with forecasts of after-effects (on task performance), and how forecasts are related to one's willpower theory. In particular, we examined whether people expect tasks to have detrimental effects on their subsequent task performance when they expected the task to be more demanding (Dang et al., 2017; Lin et al., 2020), less rewarding (Francis & Inzlicht, 2016; Kurzban et al., 2013), and unsupportive of self-efficacy (Chow et al., 2015).

We also investigated whether limited willpower theorists (or, instead, people with low trait self-control) were more likely to forecast detrimental after-effects overall, or whether they were more likely to forecast detrimental after-effects for activities that they perceived as demanding.

2.1 | Methods

We report all manipulations, measures, and exclusions in all studies. Data, materials, and analysis code for all studies are available at https://osf.io/63u4r/. Studies 1a and 1b were not preregistered.

2.1.1 | Participants

Sensitivity analysis suggests each study had 80% power to detect within-subject effects of r = 0.08 or larger (repeated-measures $r_{Study1a}$ = .31 and $r_{Study1b}$ = .26), and between-subject effects of r = .24 (Study 1a) and r = .22 (Study 1b).

Study 1a. Study 1a was completed online by 136 participants; 120 from Amazon's Mechanical Turk,[1] and 16 from Reddit.com/r/sample size. Eligibility criteria used for the study on Mechanical Turk included

[1] Note that Studies 1a and 1b, which used samples recruited from Mechanical Turk, were conducted in February 2016 and November 2017, prior to the increase in non-disclosed non-native English speakers and 'bots' that decreased data quality in the summer of 2018 (Chmielewski & Kucker, 2020).

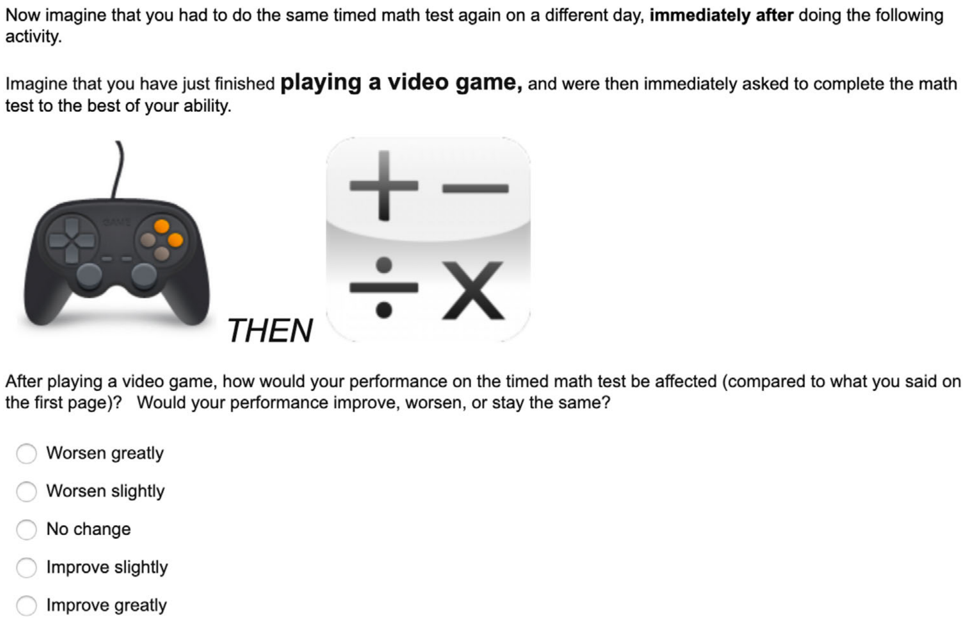

FIGURE 1 Main dependent variable for Study 1a.

being in the United States, having a HIT approval rate of 98% or higher, and having had at least 100 approved HITs. Participants were, on average, 35.2 years old (*SD* = 13.26), 56% were male, and 86% reported having at least some college (including 40% undergraduate degree and 13% with graduate degree). MTurk participants received a small payment for their time; Reddit participants were not compensated.

Study 1b. Online participants were recruited through Amazon's Mechanical Turk, using the same eligibility criteria as Study 1a. Of 180 participants who completed the survey, 152 participants successfully answered an attention check question and were included in our analyses. The analysed sample consisted of 81 men, 70 women, and one other-identified, and were an average of 36.6 years old (*SD* = 10.38, 19 to 66).

2.1.2 | Study 1a procedure

After providing informed consent, participants were asked to imagine that they had 10 minutes to complete as many maths questions as they could to the best of their ability. The questions would start with multiplying two single-digit numbers together, and then gradually progress in difficulty. The participants were told that this task would require them to focus their attention. They then indicated their expected baseline performance on this task (see Supplemental Materials S1).

Next, participants were asked to imagine that they had to do an equivalent maths test again on a different day, immediately after finishing a different activity (Figure 1). Participants indicated whether they thought that their maths performance would improve, worsen, or stay the same (5-point scale). In a random order, participants saw 14 real-world activities, including babysitting, playing video games, writing an exam, and playing sports (full list in Appendix).

Participants later indicated how much they enjoyed each activity (7-point scale), how demanding each is (5-point scale from 'not at all taxing' to 'extremely taxing'), and how frequently they have experienced each (from 0 'never' to 3, 'often/regularly'; see Supplemental Materials S2).

Participants then completed the 6-item implicit theory of willpower scale on strenuous mental activity (Job et al., 2010). Items were measured on scales from 1 to 6 and summed ($\alpha = .93$, $M = 23.24$, $SD = 6.90$). This scale includes items such as 'After a strenuous mental activity, your energy is depleted and you must rest to get it refuelled again' and, 'After a strenuous mental activity, you feel energized for further challenging activities' (reverse-scored). Participants lastly indicated their age, gender, education level, and how much they like maths (on a scale from 1 to 7; $M = 4.17$, $SD = 1.87$).

2.1.3 | Study 1b procedure

Study 1b followed the same procedure, except that the participants forecasted how their behaviour regarding a personally relevant goal would be affected by various preceding activities, rather than forecasting how their performance on a speeded maths test would be affected. They also completed additional measures about the characteristics of the preceding activities, and two additional trait measures.

Participants chose a goal that was most self-relatable from six options: improve physical activity/exercise (N = 40), increase study/work (N = 35), be more patient (N = 20), reduce spending (N = 24), improve eating (N = 26), or reduce substance use (N = 7). Participants indicated to what extent they were actively working on that goal. Most participants were actively (N = 77) or very actively (N = 42) working on it. They then indicated the likelihood, in general, that they would engage in a corresponding undesired behaviour (e.g., procrastinating, spending money unnecessarily).

In the next portion of the study, participants were asked how each of thirteen activities (presented in a random order) would affect their likelihood of behaving in a goal-inconsistent manner immediately afterwards, relative to their own baseline (similar to Figure 1), on a 5-point scale (from 'much more likely', to 'much less likely').

Participants indicated how much they liked or disliked each preceding activity (1–7), how taxing they thought each activity was (1–5), how frustrating (1–5), and how accomplished/successful they would feel afterwards (1–5). Finally, participants again completed the 6-item willpower theories scale of strenuous mental activity (Job et al., 2010; $\alpha = .94$; $M = 23.08$, $SD = 7.31$) and their demographics. Unlike Study 1a, they also completed the 13-item brief trait self-control scale (Tangney et al., 2012) with items measured on a seven-point scale ($\alpha = .92$, $M = 4.51$, $SD = 1.25$) and an 8-item trait goal self-efficacy measure (Chen et al., 2001) with responses provided on a five-point scale ($\alpha = .96$; M = 3.86, $SD = 0.76$).

2.1.4 | Analysis

We conducted multi-level models with random intercepts for each participant. Because we wanted our conclusions to speak to activities in general, we also included random intercepts and random slopes for each activity. Within-subject predictors (e.g., task demand) were person-centred, while between-subject predictors (e.g., willpower theory) were grand-mean centred.

2.2 | Results

2.2.1 | Descriptive statistics of forecasts

Across studies, both helpful and hindering after-effects were often forecasted (Figure 2); it was also common for participants to expect no after-effect. The full range of forecasted after-effects (from 'greatly improve' to 'greatly worsen') was observed for each task. Only 5.4% (Study 1a) and 1.8% (Study 1b) of variability was explained by the task itself, suggesting low consensus across participants in how a particular task would affect one's subsequent performance on a maths task or at resisting temptation. Slightly more variability was explained by participants (9.7% Study 1a; 7.9% Study 1b), suggesting some people tended to rate most activities as hindering subsequent performance, while others had the reverse tendency.

2.2.2 | Task characteristics predicting forecasted after-effects

In Study 1a, both higher subjective task demand and lower subjective enjoyment of the preceding task were related to forecasting more impairment on the subsequent maths task (demand $b = -0.08$, $SE = 0.03$, $t(11) = -3.08$, $p = .011$, $r = -0.68$; enjoyment $b = 0.09$, $SE = 0.02$, $t(31) = 3.75$, $p < .001$, $r = 0.56$). These relationships were also true on the between-person level (see Supplemental Materials S3). Subjective demand and enjoyment were significantly negatively correlated, making it difficult to disentangle the relative effects of each characteristic (person-centred variables, $r = -.67$).

In Study 1b, subjective task demand, enjoyment, and frustration (frustration was not measured in Study 1a) were all highly correlated with one another (*r*s over .70); we thus created a composite variable from these three variables (reverse-coding enjoyment; $\alpha = .86$). This composite variable strongly predicted forecasted after-effects (Figure 3; $b = -0.20$, $SE = 0.03$, $t(18) = -5.84$, $p < .001$, $r = -0.81$), such that more demanding/unenjoyable tasks were expected to hinder subsequent goal-consistent behaviours. Task self-efficacy (measured only in Study 1b) was not significantly related to forecasted effects on subsequent behaviour ($p = .88$).

2.2.3 | Moderation by willpower theory and trait self-control

We hypothesized that participants with more limited willpower theories would be (1) more likely to expect tasks to have detrimental after-effects, and (2) especially more likely to expect *demanding* tasks to have detrimental after-effects—in other words, we predicted an interaction between willpower theory and the perceived demand of the preceding task on forecasted after-effects. We also explored whether willpower theories were related to expectations of how demanding and enjoyable tasks would be.

In Study 1a, participants with a more limited theory of willpower were more likely to expect their maths performance to be negatively affected by prior activities ($b = -0.02$, $SE = 0.004$, $t(133) = -4.48$, $p < .001$, $r = 0.36$). However, in Study 1b, there was no overall relationship between willpower theory and forecasted after-effects on goal behaviour ($b = -0.005$, $SE = 0.005$, $t(150) = -0.92$, $p = .357$, $r = -0.08$). Instead, trait self-control predicted after-effect forecasts ($b = 0.12$, $SE = 0.03$, $t(150) = 3.86$, $p < .001$, $r = 0.30$), such that those with lower self-control expected that prior tasks, in general, would increase their likelihood of behaving in a goal-incongruent way. Trait goal self-efficacy, which was strongly correlated with trait self-control ($r = .59$), showed the same relationship ($p = .004$, $r = 0.24$).

Unexpectedly, willpower theory did not significantly moderate the relationship between perceptions of a preceding activity as demanding and expectations of that activity hindering subsequent self-control. The hypothesized interaction did not reach significance in either Study 1a ($b = -0.004$, $SE = 0.002$, $t(1722) = -1.92$, $p = .055$, $r = 0.05$) or Study 1b ($b = -0.003$, $SE = 0.003$, $t(1812) = -1.10$, $p = .273$, $r = -0.03$).

Instead, willpower theory predicted how demanding and enjoyable people expected activities to be. Those with a more limited willpower theory saw tasks as more effortful (Study 1a, $b = 0.02$, $SE = 0.006$, $t(133) = 3.46$, $p < .001$, $r = 0.29$) and less enjoyable (Study 1a $= -0.02$, $SE = 0.008$, $t(133) = -2.44$, $p = .016$, $r = -0.21$). These relationships replicated in Study 1b (willpower theory predicting 3-item composite:

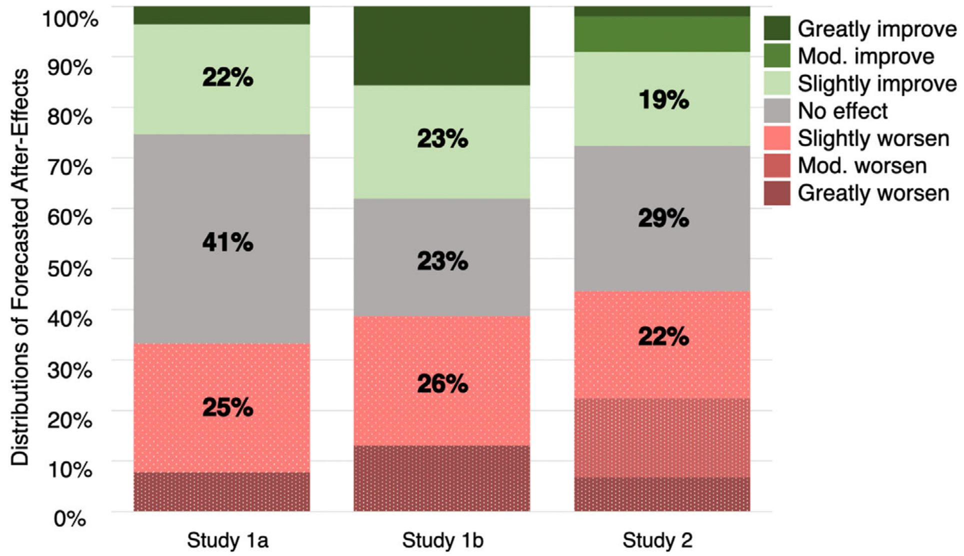

FIGURE 2 Percentages of forecasted after-effects (beneficial or detrimental) across tasks. *Note*. Study 1a and 1b used five response-options, while Study 2 used seven response-options (addition of *moderately worsen* and *moderately improve* options).

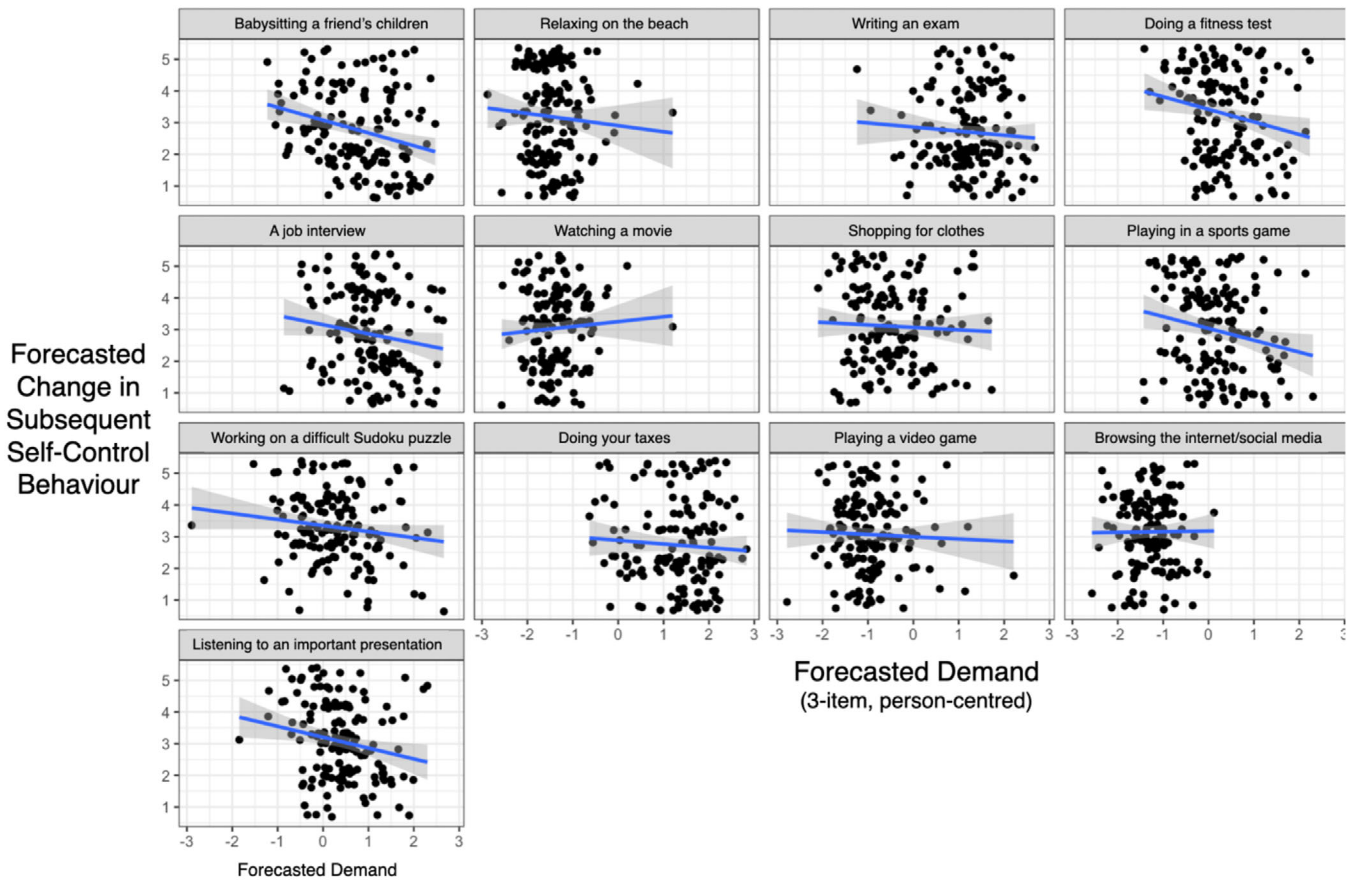

FIGURE 3 Relationships between perceived demand and forecasted after-effects for each task. *Note*. Forecasted demand was measured as a 3-item aggregate of the preceding activity's forecasted demand, forecasted enjoyment (reverse-scored), and forecasted frustration. Forecasted change was measured from 1 = worsen greatly to 5 = improve greatly.

$b = 0.01$, $SE = 0.004$, $t(150) = 2.60$, $p = .010$, $r = 0.21$). Willpower theory was not significantly related to perceived task self-efficacy ($b = -0.01$, $SE = 0.007$, $t(150) = -1.78$, $p = .077$, $r = -0.14$).

2.3 | Discussion

Overall, people had variable expectations about which tasks would cause depleting or beneficial after-effects; there were substantial differences in how people thought of these tasks and how they believed these tasks would affect them. Some of the variation in beliefs about after-effects was explained by variation in perceptions of the tasks themselves—people who perceived an activity (like completing Sudoku puzzles) to be more difficult and less enjoyable tended to predict that this activity would hinder their subsequent performance.

While task enjoyment was a superior predictor of forecasted after-effect than task demand in Study 1a (when enjoyment and demand were simultaneous predictors), these task characteristics were highly correlated in both studies. Research on self-control and motivation tends to focus on the role of task demand and difficulty (Dang et al., 2017; Lin et al., 2020), but the intrinsic rewards and enjoyment of a task may be similarly or even more important to how people perceive and experience activities in daily life (Francis & Inzlicht, 2016; Werle et al., 2015; Werner et al., 2016).

Interestingly, nearly all participants expected that subjectively difficult and unenjoyable activities would hinder their subsequent self-control performance, regardless of their willpower theory. In neither Study 1a nor Study 1b was the relationship between perceived task demand and forecasted after-effect moderated by willpower theory. Instead, these studies provided some insight into *why* limited theorists may more often expect to experience depletion—they expect activities themselves to be more demanding and less enjoyable, while non-limited theorists see the same activities as less demanding and more enjoyable.

A limitation to these studies was that detailed descriptions of the tasks were not provided. Participants may have had insufficient information about the tasks to be able to judge how the task would affect their subsequent behaviour or performance—for instance, the length of the tasks was not stated and could vary widely for some activities (e.g., one could babysit for 20 minutes or for 8 hours). This lack of detail may have resulted in participants making various assumptions about the tasks and may have caused uncertainty in participants' forecasts. We address this limitation by providing comprehensive descriptions of the tasks for participants in Study 2 and 3, including specifying the task length.

These first studies support the idea that people can have specific forecasts about what activities they *believe* will help or hurt their subsequent self-control performance and that these forecasts correspond with expectations of the activity's characteristics. But to what extent are these forecasts accurate? In the next study, we test the accuracy and bias of peoples' predictions.

3 | STUDY 2

In this study, participants again made specific forecasts of whether their performance on a maths test would be impaired or improved by six different preceding activities. We selected activities that undergraduate students should be familiar with, while also drawing on tasks that have been used in the depletion and exercise cognition literatures. After creating forecasts for all six tasks, the same participants came into the lab on three later occasions to experience three of the six situations. In this way, using a within-subject design, we could analyse and compare participants' *predictions* of after-effects with their *actual* in-lab performance and subjective states.

3.1 | Methods

3.1.1 | Participants

Introduction to psychology students at the University of Toronto Scarborough received course credit for their participation. Participants were primarily female (76%) with an average age of 18.8 years ($SD = 2.47$). Participants ($N = 167$) completed an online portion of the study, and $N = 124$ of them then attended at least one subsequent in-lab session ($M = 2.4$ sessions). An additional six students completed lab sessions without having completed the online portion of the study, resulting in $N = 311$ in-lab sessions, 300 of which have corresponding (within-subject) predictions. Each analysis uses all available data.

A minimum sample size of $N = 160$ was planned based on an initial power analysis (the assumptions of which were unfortunately not saved), and data collection continued until the end of the year. Analyses were only conducted after all data collection was complete. Sensitivity analyses conducted in G*Power (Faul et al., 2009) suggest that the sample of 124 in-lab participants has 80% power to detect within-subject effects of $r = 0.09$ (average of 2.4 measurements, repeated measures $r = .70$) and 80% power to detect a between-subject correlation of $r = .25$.

3.1.2 | Tasks

For this study, we looked at potential after-effects for six different tasks, spanning three different domains: writing, emotion, and physical. Within each domain, one task was relatively more demanding and one was less demanding. Full descriptions of these tasks, as provided to the participants, are available in the Appendix.

The two writing tasks—restricted and unrestricted story-writing—are classic manipulations in the depletion literature. The more demanding writing task consisted of writing a story for 6 minutes without using the letters 'A' or 'N', while the easier writing task consisted writing a story for 6 minutes with no restriction (Carter et al., 2015; Schmeichel, 2007).

The next two tasks—emotion suppression and experiencing emotion—have also been commonly used in depletion studies (Carter et al., 2015). Both tasks consisted of watching 'a 2-½minute long movie-clip, portraying one of the saddest scenes in movie history. The movie shows a young boy, full of emotion, crying next to his father who has just died'. For the emotion suppression task, participants were to suppress all internal reactions to the movie and suppress all external signs of their feelings. For the sad video task, participants were to watch the movie clip without any additional instructions. When participants came into the lab, they then watched a scene from *The Champ* (Lovell & Zeffirelli, 1979), which has previously used as the stimuli for an emotion suppression task (e.g., Wang et al., 2014).

The final tasks involved physical activity. For the more demanding task, participants completed a 7-minute fitness routine, to the best of their ability, which included 60 seconds each of jumping jacks, push-ups, a wall-sit, a handgrip task, and burpees, with 20-second breaks between each. This physical routine primarily involves aerobic exercises, which are most commonly used in the study of the beneficial after-effects of acute exercise on cognition (Pontifex et al., 2019). For the other task, participants walked through the university halls along a set indoor route, which took approximately seven minutes to complete.

3.2 | Procedure

3.2.1 | Online forecasting survey

At least two days before coming into the lab, participants completed the online forecasting portion of the study.

Participants first indicated how they currently felt, on six 100-point slider scales measuring: (i) subjective energy (*tired* to *energized*), (ii) subjective affect (*unpleasant* to *pleasant*), (iii) subjective control, (iv) accomplishment, (v) frustration, and (vi) stress (*calm* to *stressed*).

They then had 8 minutes to complete a baseline 15-item multiple-choice SAT maths test, similar to standardized maths tests successfully used as the dependent variable in depletion studies (Carter et al., 2015; Dang, 2018). This baseline task provided participants with the experience to judge how their performance on a similar set of questions might be affected by preceding tasks. Maths tests can vary substantially in their complexity and difficulty from one another, so we chose to have participants gain experience in this particular format of maths test to increase their ability to accurately forecast their performance on future iterations. Participants were immediately provided with their score and were asked to refer to their baseline score when forecasting.

Participants were then shown descriptions of six tasks, one at a time (see Appendix). These tasks were the same tasks that participants would later complete in the lab. They consisted of: a restricted writing task, an unrestricted story-writing task, walking around campus, a 7-minute exercise routine, watching a sad movie clip, and an emotion-suppression task. For each, they first forecasted characteristics of the task (expected difficulty, interest, fun, unpleasant, challenge, requiring self-control, self-efficacy) on 6-point scales. They next forecasted how they would feel after completing the given task, using the same six scales described above, to measure forecasted energy, subjective affect, sense of control, accomplishment, frustration, and stress.

Next, participants were asked to imagine that, immediately after the task in question, they completed a 15-question maths test, equivalent to the one done earlier. They were asked how many questions they believed they would get right, from 0 to 15. On this page, we provided their own baseline score as an anchor (e.g., 'If you did equally well to how you performed today, you might expect to answer around 13–15 questions correctly'). We also asked, subjectively, whether they thought that the task would improve, worsen, or have no effect on their ability to do the maths test (see Figure 2).

Finally, participants completed the willpower theory questionnaires (Bernecker & Job, 2017; Job et al., 2010; strenuous mental activity $\alpha = .85$, $M = 4.26$, SD = 1.00 on 1–6 scale), the brief self-control scale (Tangney et al., 2004; $\alpha = .83$, $M = 3.94$, $SD = 0.93$ on 1–7 scale), an eight-item general self-efficacy scale (Chen et al., 2001; $\alpha = .91$, $M = 3.69$, $SD = 0.70$ on 1–5 scale), and indicated their age and gender.

3.2.2 | In-lab sessions

Participants each came to three in-lab sessions, spaced one week apart. Each participant was assigned to one easier condition and two demanding conditions, and only experienced one task for each of the three domain categories (writing, emotion, physical; see Appendix).

After arriving at the lab, participants responded to six questions measuring their subjective state (energy, subjective affect, sense of control, accomplishment, frustration, and stress), using the identical scales used in the online forecasting portion of the study. They then experienced one of six activities. After this activity, they completed the subjective state questions again. They then had 8 minutes to complete the 15 SAT maths questions. Participants saw a different maths test (each version consisting of 15 comparable SAT questions) during each in-lab session, and the version of the maths test was counter-balanced and fully crossed with task and task-order. Finally, they were asked about the characteristics of the focal task.[2]

3.2.3 | Analysis

Analyses were conducted in R using multi-level models (Kuznetsova et al., 2017), with random intercepts for participant and random intercepts for each task. While we included random slopes within each activity for Study 1a and 1b, Study 2 included half as many activities (six rather than 13 or 14) and, as such, including random slopes within each activity significantly reduced the available degrees of freedom and statistical power of the analyses. We thus did not model 'task' as a random variable.

[2] The task characteristics in the in-lab portion of the study could unfortunately not be analysed, because an unknown number of participants misunderstood the instructions and described how they felt the SAT maths task was difficult, frustrating, fun, and so on, rather than describing the task characteristics of the preceding target task (emotion suppression task, writing task, etc.).

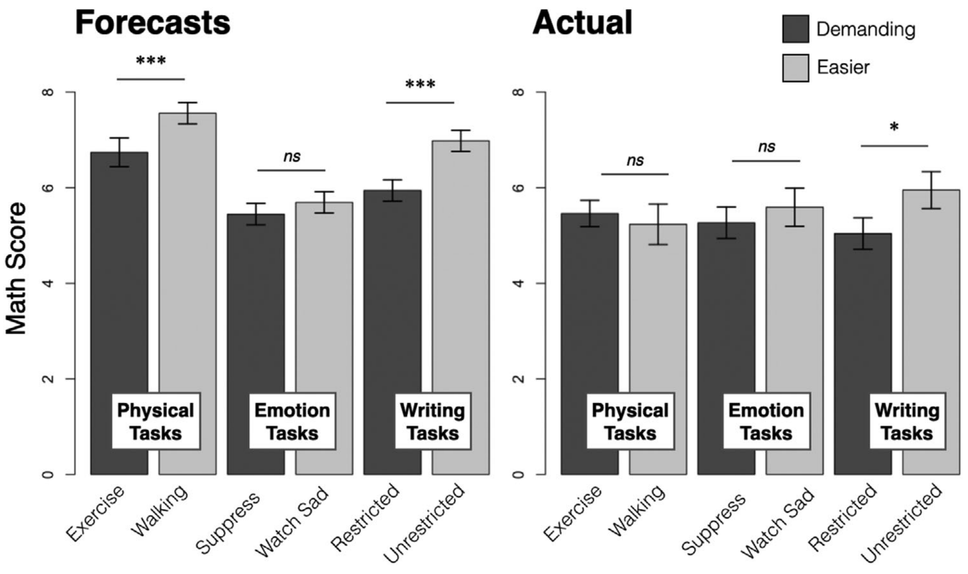

FIGURE 4 Forecasted and actual performance on a timed 15-question SAT maths test. *Note*. Performance on the maths task, always completed immediately after six different tasks. Error bars show the standard errors for the effect of task. Contrasts between the more effortful task and its less effortful counterpart are shown: *** = $p < .001$; * = $.01 < p < .05$.

Within-subject variables, including task characteristics and subjective state measures, were participant-centred and individual difference measures were grand-mean centred. To interpret interactions, we report the simple slopes at +1*SD* and −1*SD* or at specific levels of a categorical variable (Preacher et al., 2004).

3.3 | Results

Replicating Study 1, activities that were expected to be more demanding and less enjoyable were also expected to worsen one's subsequent maths performance; these perceived relationships were again not moderated by willpower theory (see Supplemental Materials S4).

3.3.1 | Forecasted after-effects of tasks on maths performance

First, we examined the correspondence between participants' forecasted maths performance and their actual maths performance. In general, participants overestimated how well they would do on the maths test ($b = 0.77$, $SE = 0.14$, $t(885) = -5.37$, $p < .001$, $r = 0.18$) relative to how they actually performed. In other words, there was an overall optimism or self-enhancement bias in participants' predictions of their maths ability.

However, participants' forecasted maths scores were only significantly over-optimistic for three of the six task conditions (Figure 4): walking ($b = 2.04$, $SE = 0.38$, $p < .001$), story-writing ($b = 1.21$, $SE = 0.37$, $p = .001$), and exercise ($b = 1.03$, $SE = 0.30$, $p < .001$). The remaining three tasks—which were typically perceived as fatiguing and difficult—all lacked significant self-enhancement bias in the maths score forecasts (restricted writing $b = 0.52$, $SE = 0.30$, $p = .082$; emotion suppression $b = -0.19$, $SE = 0.30$, $p = .51$; sad video $b = -0.19$, $SE = 0.39$, $p = .63$). These less biased predictions of one's maths performance were likely because people's unrealistic optimism was counteracted by their overestimated expectation of a depleting after-effect (i.e., an impact bias).

We can also examine comparisons between particular tasks. Participants expected to perform relatively better on a maths test after either type of physical exertion relative to after writing or emotional tasks (Figure 4, $p < .001$), and to perform especially well after a short walk. These forecasts are consistent with research finding beneficial after-effects of exercise on cognition. Actual benefits of exercise on subsequent maths performance were not observed in-lab. On average, participants typically expected to perform more poorly after a restricted writing task than after a normal story-writing task ($b = 1.04$, $SE = 0.22$, $t(580) = 4.65$, $p < .001$, $r = 0.19$), reflecting expectations of a depletion effect—and, in fact, a small depletion effect on maths performance was significant when contrasting these two in-lab conditions ($b = 0.89$, $SE = 0.39$, $t(261) = 2.27$, $p = .024$, $r = 0.14$).[3] While participants still overestimated the magnitude of this detrimental after-effect, as a group, participants' forecasts were relatively consistent with reality. Finally, participants expected both the emotion suppression task and the unrestricted viewing of the sad movie to comparably impair their subsequent maths performance. In actuality, not only was there no observed difference in maths performance between the two sad movie conditions ($b = 0.31$, $SE = 0.41$, $t(250) = 0.77$, $p = .441$, $r = 0.05$),[4] but these two emotion tasks also did not impair

[3] This depletion effect (the contrast between restricted and non-restricted writing conditions) on actual subsequent maths performance was not significantly moderated by willpower lay theory, $b = 0.63$, $SE = 0.44$, $t(244) = 1.43$, $p = .154$, $r = 0.09$.

[4] The contrast between the emotion suppression video and the freely watched video conditions on actual subsequent maths performance was not significantly moderated by willpower lay theory, $b = -0.24$, $SE = 0.47$, $t(237) = -0.51$, $p = .614$, $r = 0.03$.

subsequent maths performance relative to the two exercise conditions ($b = -0.08$, $SE = 0.23$, $t(192) = -0.33$, $p = .739$).

3.3.2 | Impact biases on subjective state

On average, participants expected to feel less pleasant after the lab tasks than they actually were ($b = -4.66$, $SE = 1.68$, $t(1045) = 2.77$, $p = .006$, $r = 0.09$), and also overestimated how tired ($b = -5.31$, $SE = 1.64$, $t(1040) = 3.23$, $p = .001$, $r = 0.10$) and how 'out of control' they would feel ($b = -6.40$, $SE = 1.75$, $t(1028) = 3.65$, $p < .001$, $r = 0.11$). However, these impact biases only significant occurred for some tasks.

Participants significantly overestimated their fatigue after restricted writing ($b_{energy} = -12.46$, $SE = 3.40$, $p < .001$), emotion suppression ($b = -10.35$, SE = 3.41, $p = .002$), and the sad video ($b = -11.30$, $SE = 4.61$, $p = .01$). For the other three tasks, there were no significant differences between participants forecasted and actual reported fatigue (exercise $b = -4.08$, $SE = 3.65$, $p = .265$; walking $b = 6.52$, $SE = 4.39$, $p = .138$; story-writing $b = -0.55$, $SE = 4.35$, $p = .90$).

3.3.3 | Relationships between subjective fatigue and maths performance: Forecasted and actual

Subjective fatigue and maths performance were more highly related in participants' forecasts, compared to among their actual in-lab experiences (interaction, $b = 0.03$, $SE = 0.005$, $t(837) = 6.40$, $p < .001$, $r = 0.22$). When participants forecasted that an activity would be especially fatiguing, they expected to perform more poorly on the subsequent maths task ($b = 0.03$, $SE = 0.003$, $t(840) = 9.78$, $p < .001$, $r = 0.32$). In the lab, however, self-reported energy did not significantly predict performance on the subsequent maths test ($b = -0.002$, $SE = 0.004$, $t(860) = -0.42$, $p = .672$, $r = 0.01$). While forecasted subjective fatigue was the best predictor of forecasted maths performance, other forecasted subjective states also predicted forecasted maths performance (Table 1). No experienced subjective states significantly predicted in-lab maths performance (Table 1).

3.3.4 | Predictive validity (accuracy) of forecasts

Even if their estimates were biased, participants might be able to correctly identify which particular task would hurt or hinder their subsequent maths performance, or make them feel energized or fatigued. We here examine the correspondence between forecasted and actual states and performance.

Participants who expected to generally do better on the maths task did, in fact, tend to score higher on the task ($b = 0.51$, $SE = 0.05$, $t(96) = 10.33$, $p < .001$, $r = 0.72$). In other words, people with stronger maths skills correctly predicted that they would generally perform well.

However, participants' specific forecasts—their forecasts for how they would do on a maths test in a specific situation, after a particular preceding task—did *not* predict how well they would actually perform under those circumstances ($p = .650$, Table 2 column 1). Even though participants had specific beliefs about how prior activities would help or hinder their maths performance, these beliefs did not predict reality.

Participants were better at predicting how a given task would make them feel subjectively. Their expectations of feeling pleasant, energetic, accomplished, and in-control all related to these actual emotional states reported after the activity (Table 2 column 1). For example, if participants expected an activity to make them feel one scale-point more pleasant, that activity actually made them feel 0.33 scale-points more pleasant.

3.3.5 | Moderation of bias by individual differences

The difference between participants' forecasted fatigue and actual reported fatigue—the degree of bias—was not significantly moderated by either willpower theory ($b = 2.70$, $SE = 1.63$, $t(1057) = 1.66$, $p = .097$, $r = 0.05$) or trait self-control ($b = 0.22$, $SE = 1.73$, $t(1034) = 0.13$, $p = .90$, $r = 0.004$). While those with more limited willpower theories expected to have less energy after various activities ($b = -3.69$, $SE = 1.27$, $t(190) = -2.91$, $p = .004$, $r = 0.21$), they also tended to report less energy in the lab ($b = -6.51$, $SE = 1.77$, $t(549) = -3.68$, $p < .001$, $r = 0.16$), resulting in comparable level of bias in their forecasts of fatigue.

The difference between participants' forecasted maths scores and their actual maths scores was significantly moderated by willpower theory ($b = 0.30$, $SE = 0.07$, $t(891) = 4.27$, $p < .001$, $r = 0.14$), such that the self-enhancement bias (or optimism bias) was larger for those with non-limited willpower theories ($b = -0.65$, $SE = 0.10$, $t(895) = -6.48$, $p < .001$, $r = 0.21$) and this bias did not reach statistical significant for those with more limited willpower theories ($b = -0.05$, $SE = 0.10$, $t(889) = -0.56$, $p = .577$, $r = 0.02$). Because those with more limited willpower theories expected to do more poorly than non-limited theorists, they were ultimately more accurate (with less optimistic self-inflation). Similarly, the difference between forecasted and actual maths scores was significantly moderated by trait self-control ($b = -0.23$, $SE = 0.07$, $t(881) = -3.08$, $p = .002$, $r = 0.10$), such that the self-enhancement bias was larger for those with higher trait self-control. Participants with higher trait self-control expected to do much better on the maths test than they actually did ($b = 1.12$, $SE = 0.20$, $t(884) = 5.65$, $p < .001$, $r = 0.19$), while participants with lower trait self-control had more modest expectations of themselves and thus a smaller difference between their forecasted and actual maths scores ($b = 0.28$, $SE = 0.19$, $t(883) = 1.44$, $p = .150$, $r = 0.05$). Both moderations by willpower theory and trait self-control remained statistically significant when acting as simultaneous moderators.

3.4 | Discussion

Participants again forecast that engaging in activities that they personally believed to be unenjoyable and difficult would hinder their

TABLE 1 Subjective states predicting maths performance (forecasted and actual)

Fixed predictor (Person-centred)	Model 1. Forecasted maths scores B (SE)	Model 2. Actual maths scores B (SE)	Model 3. Interaction (Forecasted vs. Actual) B (SE)
Energy	0.018 (0.004)***	−0.004 (0.007)	$t(684) = 2.955$*
Mood	0.000 (0.004)	0.001 (0.007)	$t(680) = 0.077$
Frustration	−0.009 (0.004) *	−0.001 (0.007)	$t(700) = -1.261$
Accomplished	0.011 (0.004) **	0.001 (0.007)	$t(565) = 1.498$
Stress	−0.013 (0.004)**	−0.003 (0.007)	$t(688) = -1.539$
In Control	0.003 (0.003)	−0.003 (0.007)	$t(700) = 0.740$
Random effects	**ICC**	**ICC**	**ICC**
Person	0.793	0.488	0.673
Task	0.005	0.000	0.000

Note. Model 1 shows how the six forecasted states predict forecasted maths score, when the six subjective states are entered as simultaneous predictors. Model 2 shows how the six actual experienced states (do not) predict forecasted maths score, when entered as simultaneous predictors. Model 3 shows the significance of the interaction term of the subjective state and 'forecast versus actual' (i.e., Model 3 shows whether Model 1 and Model 2 are significantly different from one another). *** = $p < .001$, ** = $p < .01$, * = $p < .05$.

TABLE 2 Correspondence between forecasts and actual subjective states and maths performance

DV	Forecast for specific task (Person-centred) b (SE)	Average forecast (Person-mean) B (SE)	Variability accounted for by task ICC	Variability accounted for by person ICC	Variability by maths test version ICC
Pleasant	0.332 (0.057)***	0.561 (0.099)***	0.011	0.176	–
Energy	0.167 (0.061)**	0.639 (0.104)***	0.005	0.282	–
Accomplished	0.135 (0.065)*	0.486 (0.092)***	0.074	0.182	–
In control	0.129 (0.058)*	0.593(0.089)***	0.003	0.232	–
Frustrated	0.058 (0.066)	0.552 (0.115)***	0.065	0.223	–
Stressed	0.063 (0.063)	0.487 (0.120)***	0.016	0.352	–
Maths Score	−0.032 (0.069)	.0513 (0.050)***	0.003	0.223	0.108

Note. Each row shows the results of a model with the actual experienced state or performance as the dependent variable. The outcome is then predicted by the person-centred forecast for that particular task and the person-average across all six tasks. *** = $p < .001$, ** = $p < .01$, * = $p < .05$.

performance on a subsequent maths task and negatively affect their subjective state. Consistent with affective forecasting studies (e.g., Christophe & Hansenne, 2021; Robinson & Clore, 2001), participants had some insight into how different tasks would affect their subjective state afterwards—participants' forecasts of how fatigued, pleasant, accomplished, and in-control they would feel after a given task did predict how they actually reported feeling after that same task. Also consistent with prior research, participants significantly overestimated the extent to which their subjective states would be affected by the prior tasks, evidence of an impact bias (Wilson & Gilbert, 2003).

Impact biases did not only affect people's forecasts of their affective states. Participants also expected their maths performance to fluctuate substantially across different tasks (Figure 4) and expected that they would perform more poorly when feeling more subjective fatigue (Table 2; Figure 5). Neither of these patterns actually consistently occurred in the lab. When forecasting, participants significantly overestimated the degree to which various tasks would negatively affect their subjective state and hinder their subsequent performance, and overestimated the association between subjective fatigue and subsequent maths performance.

Furthermore, participants' predictions of their maths performance after specific tasks had no relationship to how their actual maths performance changed across tasks. For example, participants' actual maths scores did not differ when comparing their performance after tasks that participants thought would *worsen* their performance (M = 5.41, SD = 2.48) to their performance after tasks that participants thought would *improve* their performance (M = 5.34, SD = 2.19). This study demonstrates that people do not only overestimate impacts on how they will *feel*, but also overestimate how their *objective behaviour* will be affected. Given that empirical research has not convincingly established that these detrimental carry-over

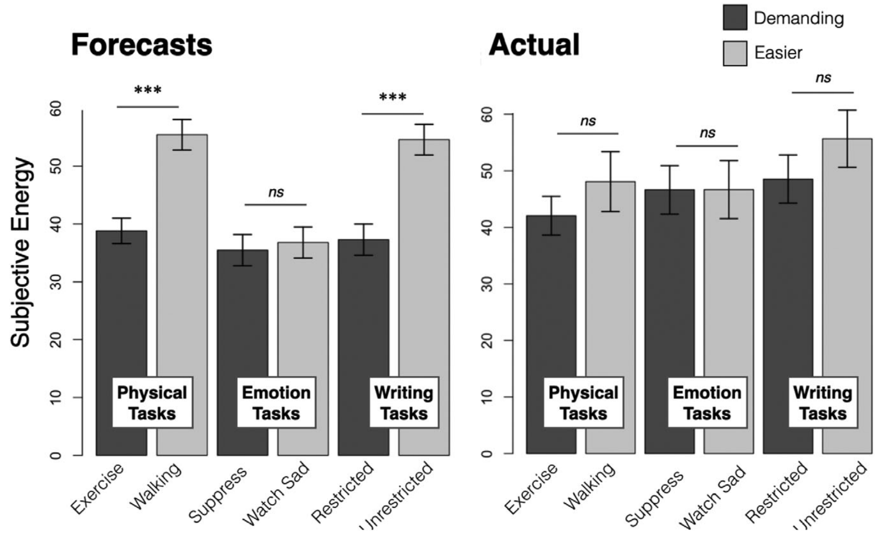

FIGURE 5 Forecasted and actual subjective mental energy. *Note*. Subjective energy was measured on a 1–100 slider scale from 'tired' to 'energized' after completing each task. Error bars show the standard errors for the effect of task. Contrasts between the more effortful self-control task and its less effortful counterpart are shown: *** = $p < .001$.

effects on cognition occur (e.g., Vohs et al., 2021), it is noteworthy that participants forecasted that these after-effects would occur.

While participants overestimated the *degree* to which their maths performance would be affected, participants' forecasts for the relative *direction* of these after-effects often aligned with empirical research. Participants generally expected both physical tasks to be more beneficial to their subsequent maths performance and expected to perform best on the maths test after mild physical activity—these expectations resemble empirical research on how physical activity can benefit cognition (Pontifex et al., 2019). Participants accurately expected restricted writing to result in depletion, compared to unrestricted writing; this effect has also been predicted and sometimes demonstrated in the depletion literature (Carter et al., 2015; Schmeichel, 2007). However, unlike prior depletion research, participants expected viewing a sad film to hinder their subsequent maths performance, regardless of whether or not they were asked to suppress their emotions. This might partially be because people plan to suppress their sadness even without instructions to do so.

Demand effects may exaggerate participants' expected differences between tasks, due to contrast effects (Baird & Lucas, 2011); participants may more readily forecast after-effects when they are directly asked about them. Study 3 partially addresses this limitation by not directly asking participants about whether prior tasks affect their subsequent maths performance.

4 | STUDY 3

In this study, we examined whether people use their (biased) after-effect forecasts when making decisions for the future. Participants made economic decisions about whether or not to accept a higher rate of pay to complete a performance-rewarded maths test *after* another activity, rather than completing the performance-rewarded maths test with lower rate of pay *before* the other activity. If participants expect their performance on the maths test to be negatively affected by activities that they perceive as difficult, as in Study 2, then they may choose to forego additional bonus pay.

4.1 | Methods

Study details, including the sample size, research questions, analysis code, exclusion criteria, and stopping rules, were preregistered and are available at https://osf.io/fzj8b.

4.2 | Participants

This study was conducted online with 215 students enrolled in Introduction to Psychology, who participated in exchange for course credit. Students who had participated in Study 2 were not eligible. The sample consisted of 141 women and 72 men, with an average age of 18.77 years old ($SD = 1.21$). This sample size has 80% power to detect within-subject effects at $r = 0.07$ (repeated-measures $r = .41$, observed) and between-subject effects of at least $r = .19$.

4.3 | Procedure

Participants were told that we were interested in what they would prefer to do during a potential in-lab study, which may be scheduled for later in the semester. They were informed that their study responses may be used for this future study, and they created a unique ID code to link their responses with their future in-lab session. Participants could thus expect that the choices they made may be implemented in the future (although the future in-lab study did not occur).

As in Study 2, participants completed a baseline SAT maths test (five questions in 3 minutes) to give them experience with the test and were given their score. They then gave their baseline prediction for how many questions they generally would expect to answer on a future fifteen-question test.

Participants next read that during a future study, they would be doing both the maths task and another task. They would be paid for each maths test question that they answered correctly, but would not be paid based on their performance on the other task. We wanted to know whether they would rather do the maths task *before* or *after* the other task, and whether we could pay them a higher payrate to switch task-order.

Participants were then presented with descriptions of each of the tasks used in Study 2, in a random order (see Appendix), and briefly wrote what they thought the task would be like. On the same screen, they indicated which task order they preferred: the maths test done *before* the other activity, or the maths test done *after*. To determine participants' 'price to switch', they were asked whether they would rather (i) be paid \$0.50 per correct maths test question and do the maths test in the order that they had just chosen, or (ii) be paid more (\$0.51, \$0.55, \$0.60, \$0.70, or \$0.80) for each correct maths test question, but perform the tasks in their non-preferred order. Participants were asked this question for each of these five increased pay rates, in a random order. We used these five questions to calculate their 'point of indifference' (see Supplemental Materials S5 for details)—this variable had twelve possible values: −\$0.31 (strong preference to complete the maths task *afterwards*), −\$0.21, −\$0.11, −\$0.06, −\$0.02, −\$0.01, \$0.01, \$0.02, \$0.06, \$0.11, \$0.21, \$0.31 (strong preference for to complete maths test *first*).

We preregistered that we would exclude trials where these five decisions were inconsistent—for example, if a participant agreed to switch for \$0.60, but then refused to switch for \$0.70. Out of 1290 observations, 118 were excluded. Additionally, as preregistered, we removed all data from participants who had two or more inconsistent trials out of six ($n = 24$).

After making all choices for all six tasks, participants indicated how difficult they thought the tasks would be, on scales from 1 (not at all) to 6 (extremely). They then used a slider bar to indicate how they believe they would feel after each activity, from 0 (tired) to 100 (energized).

Lastly, participants completed the willpower theory questionnaires used in Study 2 (strenuous mental activity $\alpha = .83$), as well as the 13-item trait self-control scale ($\alpha = .85$), the goal self-efficacy scale ($\alpha = .88$), and questions on their age and gender. They were then debriefed and received credit for the study.

4.4 | Results

4.4.1 | Predicting binary choices

For the first preregistered analyses, the outcome variable was participants' binary choice to complete the scored maths task first (1) or maths second (0). On average, participants preferred completing the maths test first (intercept $b = 0.70$, $SE = 0.25$, $z = 2.85$, $p = .004$, $r = 0.19$). However, participants' likelihood of choosing maths first differed across the six tasks ($\chi^2(2, 7) = 78.51$, $p < .001$; Figure 6).

Modelling task as a random effect, participants were significantly more likely to want to complete the maths test first when they expected the other activity to be more difficult ($b = 0.42$, $SE = 0.09$, $z = 4.90$, $p < .001$). Furthermore, participants were more likely to choose to complete the maths test first if they expected to feel less energetic (more fatigued) after completing the other task ($b = -0.02$, $SE = 0.004$, $z = -5.27$, $p < .001$).

4.4.2 | Predicting foregone costs

For the next preregistered analysis, we examined how much money participants were willing to give up in order to complete the maths test *first*, rather than after the other activity (or their switch-cost to maintain the reverse order, resulting in a negative score).

On average, participants were willing to forego 4 cents per question to complete the maths test first ($b = 0.04$, $SE = 0.02$, $t(6) = 2.55$, $p = .043$). Like the binomial outcome, participants' switch costs varied across tasks ($\chi^2(3, 8) = 72.03$, $p < .001$). Participants were not willing to incur any financial cost to specify the order of the maths test when paired with unrestricted story-writing or walking around campus ($ps > .24$; Figure 6). However, participants chose to forego money to complete the maths test first for the other four task conditions ($ps < .01$). For example, when the maths task would be done along with watching a short sad video, participants chose to complete the maths task first and be paid \$0.50 per question rather than completing the maths task afterwards and receiving \$0.58 per question, reducing their payrate by 14%.

Modelling task as a random effect, participants were willing to forego extra payment to complete the maths test first if the other task was seen as more difficult (Figure 7; $b = 0.03$, $SE = 0.004$, $t(4) = 7.02$, $p = .002$). They were also willing to forego a higher pay rate if they expected the other task to make them feel more tired ($b = -0.002$, $SE = 0.001$, $t(4) = -6.00$, $p = .003$).

4.4.3 | Moderation by willpower theory

As exploratory analyses (not preregistered), we examined whether trait self-control or willpower theory was involved in participants' estimates of how difficult these activities would be and their subsequent decisions.

As in earlier studies, participants with more limited willpower theories expected activities, in general, to be more difficult ($p < .001$, $r = 0.27$). Limited willpower theorists also held stronger preferences to do these demanding activities after the maths test (Figure 8). The same pattern of moderation was true when predicting foregone cost (Supplemental Materials S6).

4.5 | Discussion

While Study 2 found that engaging in prior effortful tasks did not actually impede students' performance on a subsequent maths test, Study 3 found that participants seemed to base their economic

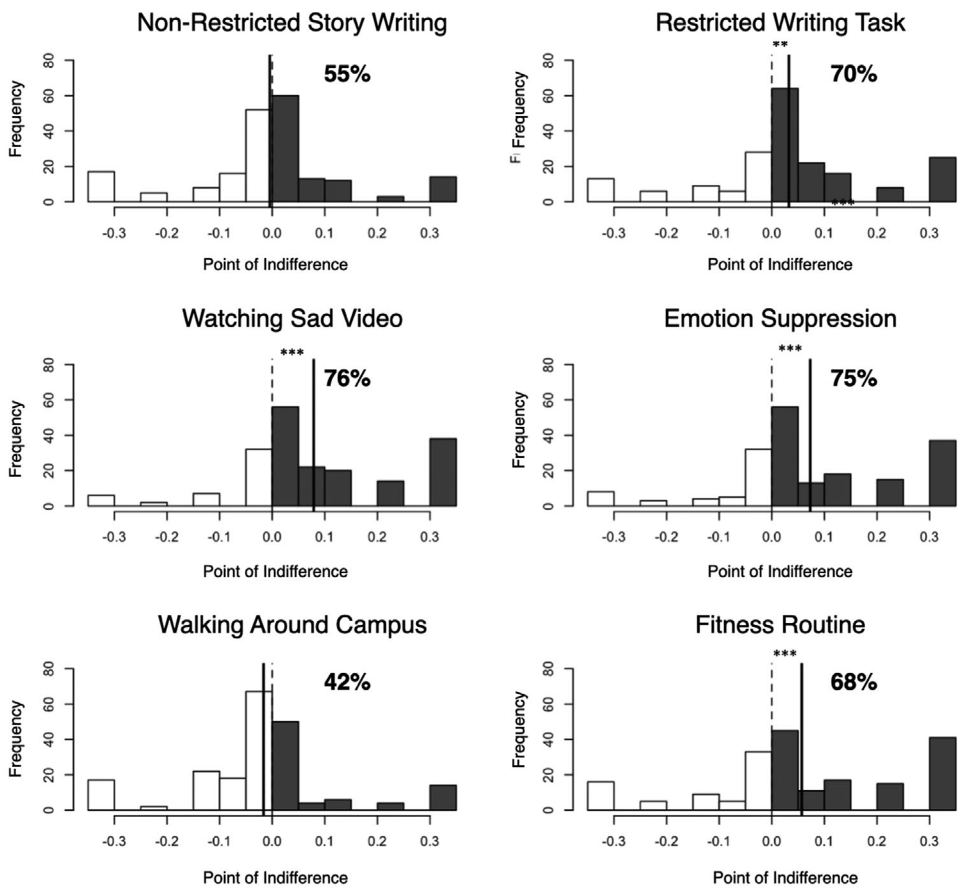

FIGURE 6 Histograms of points of indifference for each task. *Note*. Higher values (dark bars) reflect a preference for doing the maths test first and lower values (white bars) reflect a preference for doing the maths test *after* the other activity (allowing the test to be impacted by beneficial or hindering after-effects). The dashed line shows the average preference for that task. Percentages of 'maths first' choices are displayed. Asterisks show whether the average preference is significantly different from zero (solid line), at $p < .001$ (***) or at $p < .01$ (**).

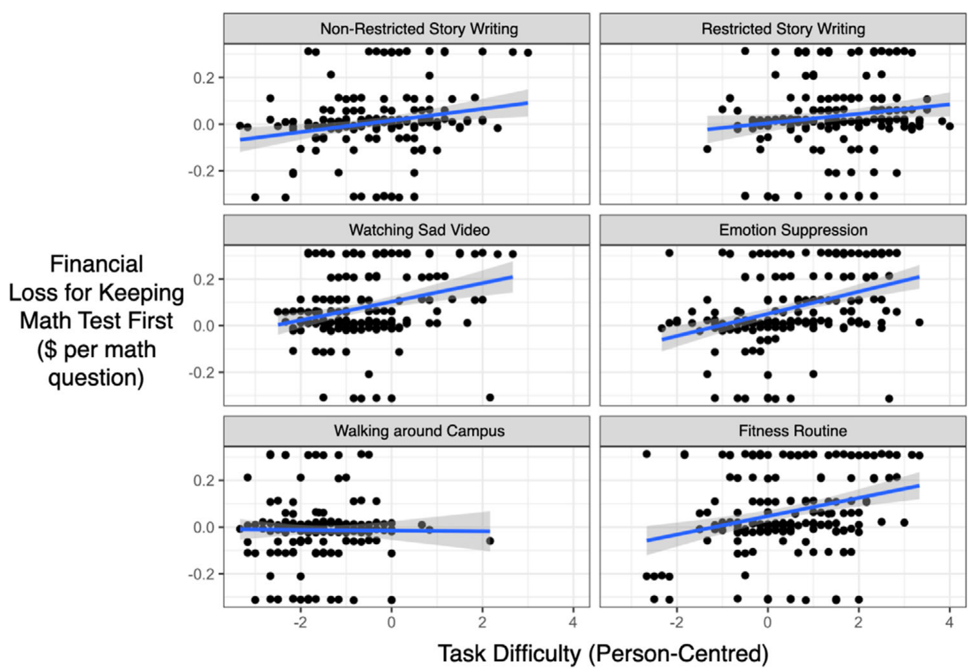

FIGURE 7 Relationship between perceived task difficulty and points of indifference.

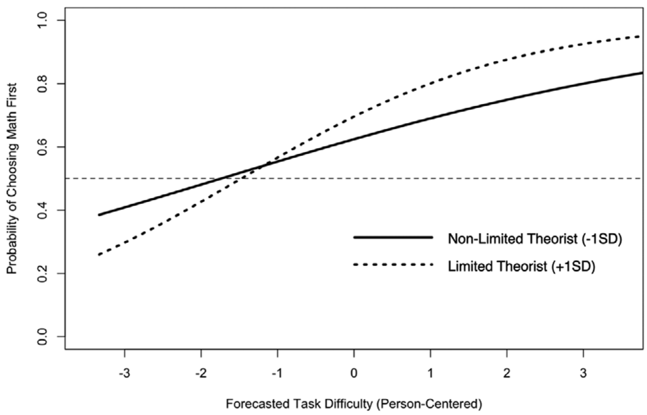

FIGURE 8 How forecasted task difficulty affects probability of task-order choice. *Note*. Forecasted difficulty of the other task significantly predicts the probability that a participant chooses to do a performance-rewarded maths test prior to the other task (both simple slopes $p < .001$). This was especially true for participants with more limited willpower theories (interaction $p < .01$).

decisions on similar (incorrect and exaggerated) expectations that difficult preceding tasks would result in mental fatigue and would hurt their maths performance. When given a choice to complete a performance-rewarded maths task and another (non-rewarded) task in either order, participants did not always show a general preference for one order or the other—they were largely indifferent for easy tasks. However, when the paired task was expected to be difficult, participants preferred to complete the performance-rewarded maths task beforehand. Furthermore, participants were willing to forego additional bonus pay to maintain their preferred task-order. Prior reports of null depletion effects (Hagger et al., 2016; Vohs et al., 2021) and the results from Study 2 suggest that this is likely to be a poor economic choice. Inaccuracies and biases in people's forecasts may, in some situations, have unintended negative consequences.

Additionally, while limited willpower theorists were especially likely to avoid effortful exertion prior to a rewarded maths-test, non-limited willpower theorists showed the same pattern of decisions. Even people who hold non-limited willpower theories seem to incorporate forecasts of mental depletion in their decision-making.

5 | GENERAL DISCUSSION

Across studies, participants often forecasted that engaging in everyday types of tasks would affect their subsequent ability to perform on a maths task or behave in a goal-consistent manner. When someone perceived an activity as unenjoyable and demanding, they typically expected that activity to hinder their subsequent performance; when someone perceived an activity as easy and enjoyable, they expected that activity to help their subsequent performance. However, compared to people's actual experiences, forecasts significantly overestimated the magnitude of after-effects. While people's subjective energy levels did vary somewhat depending on the preceding task, actual self-reported energy varied much less across conditions than what people had forecast. Furthermore, people's actual maths performance rarely varied depending on the preceding task, yet people forecast that their maths performance would be substantially altered depending on the preceding task. These patterns suggest that people's forecasts of after-effects are subject to impact biases (Morewedge et al., 2010; Wilson & Gilbert, 2003). While people may generally overestimate their self-control capacity (Nordgren et al., 2009), this self-enhancement bias does not extend to people believing that they are immune to depletion effects—instead, people expect to be much more affected by depletion than they actually are. While we did find some evidence of self-enhancement (in that, overall, people predicted higher scores than they received), this optimism bias was fully suppressed by the overestimation of how much a difficult prior activity would hurt subsequent performance.

These biased forecasts were also apparent in an economic decision-making task, with participants frequently making the suboptimal choice to reduce their rate-of-pay in order to avoid having to perform under potentially fatigued conditions.

5.1 | Implications

These studies have both theoretical and practical implications. First, by measuring both specific forecasts and the willpower theory measure, these studies provide a more detailed understanding of lay theories and expectations of after-effects. Unlike what we hypothesized, across studies, limited willpower theorists were not the only ones to expect demanding cognitive activities to impair subsequent performance; willpower theory did not consistently moderate the associations between perceptions of the prior task's difficulty and forecasted after-effects (no moderation in Studies 1a, 1b, or 2). While non-limited theorists were slightly less likely than limited theorists to make decisions based on anticipated after-effects (Study 3), non-limited theorists still typically expected effortful activities to reduce their subsequent performance and made decisions accordingly. Instead of moderating forecasts of after-effects, willpower theory was associated with perceptions of the prior activities as more demanding and less enjoyable. These findings shift our understanding of what the willpower theory scale captures—perhaps it does not primarily capture participants' lay conceptions of ego depletion, as previously assumed (Francis & Job, 2018), but instead captures the extent to which participants experience daily tasks as difficult and unrewarding.

Study 2 demonstrated that both unrealistic optimism biases and impact biases affect people's forecasts of their cognitive performance after prior tasks. Applying research on affective forecasts to a new domain of after-effects on cognition, we found that (i) people tend to be overly optimistic about their future performance (Weinstein, 1980), (ii) simultaneously, people overestimate the degree to which contextual factors will (negatively) affect their performance, indicative of an impact bias (Morewedge et al., 2010; Wilson & Gilbert, 2003), and (iii) despite these biases, participants' forecasts of their future subjective states still correspond—in relative terms—to their actual experienced subjective states (Christophe & Hansenne, 2021). While biases in affective forecasting are thought to have motivational benefits (Miloyan & Suddendorf, 2015), it is less clear whether similar biases in forecasting of *objective* outcomes are equally desirable.

Inaccuracies in forecasts of objective outcomes often come with concrete costs (Kruger & Evans, 2004; Roger et al., 1994; Shepperd et al., 2017), including suboptimal planning and decision-making. In daily life, overestimating detrimental after-effects (e.g., depletion effects) could have both negative and positive impacts. Study 3 demonstrates one potential negative consequence; participants were willing to incur a financial cost to avoid detrimental after-effects that would not have actually occurred (or, at least, would not have occurred to the extent that people expected). Similar consequences may occur in the real world if people refuse or postpone good opportunities due to the incorrect belief that they would be too fatigued to perform well or take advantage of those opportunities (e.g., Delose et al., 2015; Eden et al., 2018). On the other hand, the overestimation of negative after-effects could help people plan effectively, by preventing people from overextending themselves or overestimating their abilities—as Study 2 demonstrated, anticipation of depletion can counteract optimism or self-enhancement biases (Nordgren et al., 2009). Believing that one will experience depletion may lead people to instead proactively engage in situation selection or modification strategies to help their goal pursuit, for example (Jia et al., 2019). Future research should further explore practical implications of forecasted after-effects, and should further examine whether conscious after-effect forecasts mediate decisions to avoid or enter effortful but beneficial situations (Gieseler et al., 2020).

Next, heterogeneity in prior studies of depletion after-effects has led to the hypothesis that perhaps *only some* people experience negative after-effects on cognition in particular situations, and that researchers must do a better job of identifying and customizing experiments to be sufficiently depleting (Dang et al., 2017; Lin et al., 2020). In line with this prediction, we did find substantial variability in people's expectations about whether a given task would make them depleted or not. And, promisingly, participants were relatively accurate in predicting how pleasant and fatigued they would personally feel after conducting a given task; the heterogeneity in actual experienced subjective fatigue could be partially explained by participants' self-knowledge and expectations. However, across the set of tasks used in Study 2, participants were not able to accurately predict whether or how a task would affect their actual subsequent performance on a maths test. In other words, if it *is* the case that a specific task causes some people to subsequently perform more poorly and others to benefit, participants were not able to prospectively state which group they would be in. This suggests that if null effects in ego depletion studies are truly due to 'hidden moderators', then the hidden moderators are very well hidden indeed (Wenzel et al., 2019). People themselves could not accurately predict whether a particular task would improve or hinder their subsequent performance.

5.2 | Future directions: Processes underlying after-effect forecasts

In line with experimental findings on the relationship between demand and depletion (Dang et al., 2017; Lin et al., 2020), participants typically expected more demanding and unenjoyable tasks to hurt their subsequent performance, and easier or enjoyable tasks to improve their subsequent performance. These associations do not necessarily mean that participants themselves hold a meta-cognitive belief of a relationship between a task's demands (or rewards) and resulting after-effects—people may or may not be consciously aware of *why* they believe some tasks to help or hurt their subsequent performance. The associations could occur without conscious awareness if someone has positive or negative attitudes towards an activity, and their like or dislike of that activity generalizes to how they evaluate other characteristics of that activity (Bargh et al., 1992; Forgas, 1995)—'I don't like this activity, so it is probably also difficult and tiring and bad and boring, and will make me worse at everything'. Alternatively, because we focused on activities that were relatively common, participants could have created their forecasts by remembering (accurately or inaccurately) characteristics of past experiences and the results of those experiences (Meyvis et al., 2010; Mitchell et al., 1997). For example, someone might remember feeling fatigued after watching a sad movie, and may also recall that they do not enjoy watching sad movies. Or, some participants may have the explicit belief that a lack of enjoyment or difficulty is what leads to hindering after-effects. Future experiments could investigate the characteristics that participants use to create after-effect forecasts in more controlled ways, by using activities that participants do not have personal experience with and systematically varying characteristics of these tasks.

We also found differences in predictions of *affective* state verses predictions of *objective* maths performance, with participants being more accurate in their prediction of subjective states than their prediction of maths performance. Affective forecasting may be easier because people have more information from past experiences—we nearly always have feedback on how we are feeling, while only sometimes do we receive feedback on our objective performance. Additionally, salient emotional information is more likely to be remembered than memories about exactly how one performed on an exam or other cognitive task (Kensinger, 2004; Talarico et al., 2004). Someone may thus remember an experience of feeling mentally fatigued and depleted, and the memory of their subjective state may be recalled more readily than information about what tasks they successfully completed while in that fatigued state. Future research could investigate how people recall and remember past experiences of after-effects; do people tend to remember being more fatigued and cognitively impaired than they actually were? Such exaggerations in memory (retrospective impact biases; Mitchell et al., 1997; Wilson et al., 2003) may contribute to the impact biases seen in after-effect forecasts.

Finally, these studies did not investigate the mechanisms by which participants think their performance is affected. People may expect their performance on a maths test to be affected through changes in mood, motivation, physiology, or other mechanisms. While Study 2 showed that forecasted subjective energy was the best predictor of forecasted maths performance (compared to stress, frustration, mood, etc.), we cannot know whether participants believe that lower energy results in a reduced *capacity* to perform or a reduced *motivation* to perform (Inzlicht et al., 2014).

5.3 | Limitations

The results of these studies have limited generalizability. While we included a larger number of tasks than most studies on after-effects, and modelled task as a random effect in three of four studies, we still cannot speak to all types of preceding activities and tasks (Yarkoni, 2022). All tasks were relatively short (for Studies 2 and 3, between 3 and 8 minutes long) and forecasters expected them to occur in the upcoming few weeks. Forecasting tasks that are significantly longer, more impactful, less likely, or further in the future may affect the magnitude of observed impact bias (Buechel et al., 2017).

These studies also did not focus on the task that people expect to be affected *by* the preceding activities. We used only three dependent tasks: a speeded multiplication task, a standardized maths test, and resisting a goal-incongruent behaviour. While many results were consistent across these three tasks (e.g., consistent relationships between expected prior-task difficulty and forecasted negative after-effects), people may have varying beliefs about what types of tasks are readily influenced by preceding tasks. Different tasks can result in different after-effects depending on the aspect of cognition tested (for example, exercise does not affect all aspects of cognition equally; Pontifex et al., 2019), and the similarity between the preceding and subsequent tasks may also play a role (Primoceri et al., 2021). Future research might explore people's beliefs about what types of tasks are susceptible to after-effects.

The choice of dependent task may also have affected our ability to measure actual after-effects on performance. In Study 2, we measured performance using an SAT maths test because (i) prior research suggested that standardized tests are well suited to measure depletion (Carter et al., 2015; Dang, 2018), (ii) participants would be familiar with the task, and (iii) multiple standardized versions of the test allowed for repeated measures. There was an observed depletion effect when comparing the restricted-writing versus story-writing tasks, suggesting that this measure was sufficiently granular to demonstrate small group differences. However, as with all null effects, it is difficult to determine whether other null results are due to true null effects or due to insufficient sensitivity of the measure.

Finally, these studies only sampled from American workers on Mechanical Turk and undergraduate students at a Canadian university. While our findings were consistent across these two different samples, these results may not generalize to all groups of people; beliefs about mental energy are known to vary culturally (Savani & Job, 2017; Sun et al., 2019). Furthermore, people from different cultures are unequally susceptible to impact biases (Lam et al., 2005).

6 | CONCLUSIONS

People do not all agree that babysitting, writing an exam, or suppressing their emotions will hurt their subsequent cognitive performance—in part because people have varying expectations about how effortful and enjoyable these activities will be. However, across four North American samples, people consistently expected demanding and unenjoyable activities to hinder their subsequent performance, and easier, more enjoyable activities to help their subsequent performance, regardless of their willpower theory. People's personal forecasts of how various tasks would affect their subjective feelings of mood, energy, and accomplishment were relatively accurate to their actual experiences, reflecting good self-knowledge. However, people significantly overestimated the detrimental effects that demanding activities would have on both their subjective state and their subsequent cognitive performance, reflecting a significant bias that can negatively affect decision-making and can even result in financial costs. People forecast detrimental after-effects much more strongly than they experience them.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Data, materials, and analysis code for all studies are available at https://osf.io/63u4r/.

ETHICS STATEMENT

This research was approved by the University of Toronto's institutional review board and adheres to APA guidelines for ethical research.

ORCID

Zoë Francis https://orcid.org/0000-0002-9488-9132

REFERENCES

Armor, D. A., & Taylor, S. E. (1998). Situated optimism: Specific outcome expectancies and self-regulation. *Advances in Experimental Social Psychology*, *30*(C), 309–379. https://doi.org/10.1016/S0065-2601(08)60386-X

Baird, B. M., & Lucas, R. E. (2011). … And how about now?': Effects of item redundancy on contextualized self-reports of personality. *Journal of Personality*, *79*(5), 1081–1112. https://doi.org/10.1111/j.1467-6494.2011.00716.x

Bargh, J. A., Chaiken, S., Govender, R., & Pratto, F. (1992). The generality of the automatic attitude activation effect. *Journal of Personality and Social Psychology*, *62*(6), 893–912. https://doi.org/10.1037/0022-3514.62.6.893

Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, *74*(5), 1252–1265. https://doi.org/10.1037//0022-3514.74.5.1252

Bernecker, K., & Job, V. (2017). Implicit theories about willpower in resisting temptations and emotion control. *Journal of Personality*, *225*, 157–166. https://doi.org/10.1027/2151-2604/a000292

Brown, I. D. (1986). Self-enhancement biases in social judgments. *Social Cognition*, *4*(4), 353–376.

Buechel, E. C., Zhang, J., & Morewedge, C. K. (2017). Impact bias or underestimation? Outcome specifications predict the direction of affective forecasting errors. *Journal of Experimental Psychology: General*, *146*(5), 746–761. https://doi.org/10.1037/xge0000306

Carruth, N. P., Ramos, J. A., & Miyake, A. (2018). Does willpower mindset really moderate the ego-depletion effect?: A preregistered direct replication of job, dweck, and walton's (2010). Study 1.

Carter, E. C., Kofler, L. M., Forster, D. E., & McCullough, M. E. (2015). A series of meta-analytic tests of the depletion effect: Self-control does not seem to rely on a limited resource. *Journal of Experimental Psychology: General*, *144*(4), 796–815. https://doi.org/10.1037/xge0000083

Chen, G., Gully, S. M., & Eden, D. (2001). Validation of a new general self-efficacy scale. *Organisational Research Methods*, *4*(1), 62–83. https://doi.org/10.1177/109442810141004

Chmielewski, M., & Kucker, S. C. (2020). An MTurk crisis? Shifts in data quality and the impact on study results. *Social Psychological and Personality Science*, *11*(4), 464–473. https://doi.org/10.1177/1948550619875149

Chow, J. T., Hui, C. M., & Lau, S. (2015). A depleted mind feels inefficacious: Ego-depletion reduces self-efficacy to exert further self-control. *European Journal of Social Psychology*, *45*(6), 754–768. https://doi.org/10.1002/ejsp.2120

Christophe, V., & Hansenne, M. (2021). Negative valence effect in affective forecasting: The unique impact of the valence among dispositional and contextual factors for certain life events. *Europe's Journal of Psychology*, *17*(2), 117–130. https://doi.org/10.5964/ejop.1945

Clarke, V. A., Lovegrove, H., Williams, A., & Machperson, M. (2000). Unrealistic optimism and the health belief model. *Journal of Behavioral Medicine*, *23*(4), 367–376. https://doi.org/10.1023/A:1005500917875

Converse, P. D., & Deshon, R. P. (2009). A tale of two tasks: Reversing the self-regulatory resource depletion effect. *The Journal of Applied Psychology*, *94*(5), 1318–1324. https://doi.org/10.1037/a0014604

Dang, J. (2016). Commentary: A multilab preregistered replication of the ego-depletion effect. *Frontiers in Psychology*, *7*(AUG), 1–2. https://doi.org/10.3389/fpsyg.2016.01155

Dang, J. (2018). An updated meta-analysis of the ego depletion effect. *Psychological Research*, *82*(4), 645–651. https://doi.org/10.1007/s00426-017-0862-x

Dang, J., Liu, Y., Liu, X., & Mao, L. (2017). The ego could be depleted, providing initial exertion is depleting. *Social Psychology*, *48*(4), 242–245. https://doi.org/10.1027/1864-9335/a000308

Dang, J., Zerhouni, O., Imhoff, R., Jia, L., Giacomantonio, M., Sevincer, A., Buchholz, N., Lange, F., Kubiak, T., Wenzel, M., Berkman, E., Ludwig, R., Livingston, J., Buczny, J., Gong, R., Shi, J., Barker, P., Rassi, N., Cristofaro, V. D. E., … H, B. (2021). A multilab replication of the ego depletion effect. *Social Psychological and Personality Science*, *12*(1), 14–24. https://doi.org/10.1177/1948550619887702

Delose, J. E., VanDellen, M. R., & Hoyle, R. (2015). First on the list: Effectiveness at self-regulation and prioritizing difficult exercise goal pursuit. *Self and Identity*, *14*(3), 271–289. https://doi.org/10.1080/15298868.2014.983442

Eden, A., Johnson, B. K., & Hartmann, T. (2018). Entertainment as a creature comfort: Self-control and selection of challenging media. *Media Psychology*, *21*(3), 352–376. https://doi.org/10.1080/15213269.2017.1345640

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

Fayard, J. V., Roberts, B. W., Robins, R. W., & Watson, D. (2012). Uncovering the affective core of conscientiousness: The role of self-conscious emotions. *Journal of Personality*, *80*(1), 1–32. https://doi.org/10.1111/j.1467-6494.2011.00720.x

Fisher, A. J., Medaglia, J. D., & Jeronimus, B. F. (2018). *Lack of group-to-individual generalizability is a threat to hu*man subjects research. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(27), E6106–E6115. https://doi.org/10.1073/pnas.1711978115

Forgas, J. P. (1995). Mood and judgment: The affect infusion model (AIM). *Psychological Bulletin*, *117*(1), 39–66. https://doi.org/10.1037/0033-2909.117.1.39

Francis, Z., & Inzlicht, M. (2016). Proximate and ultimate causes of ego depletion. In E. R. Hirt, J. J. Clarkson, & L. Jia, (Eds.), *Self-regulation and ego control* (pp. 373–398). Elsevier. https://doi.org/10.1016/B978-0-12-801850-7.00018-4

Francis, Z., & Job, V. (2018). Lay theories of willpower. *Social and Personality Psychology Compass*, *12*(4), e12381. https://doi.org/10.1111/spc3.12381

Franz, S. I. (1897). Six reviews of articles on mental fatigue and performance. *Psychological Review*, *4*(5), 558–561.

Gieseler, K., Inzlicht, M., & Friese, M. (2020). Do people avoid mental effort after facing a highly demanding task? *Journal of Experimental Social Psychology*, *90*(2019). 104008. https://doi.org/10.1016/j.jesp.2020.104008

Gilbert, D. T., Morewedge, C. K., Risen, J. L., & Wilson, T. D. (2004). Looking forward to looking backward: The misprediction of regret. *Psychological Science*, *15*(5), 346–350. https://doi.org/10.1111/j.0956-7976.2004.00681.x

Hagger, M. S., Chatzisarantis, N. L. D., Alberts, H., Anggono, C. O., Batailler, C., Birt, A., Brand, R., Brandt, M. J., Brewer, G. A., Bruyneel, S., Calvillo, D. P., Campbell, K., Cannon, P. R., Carlucci, M., Carruth, N., Cheung, T., Crowell, A., De Ridder, D. T. D., Dewitte, S., & Zweinenberg, M. (2016). A multi-lab pre-registered replication of the ego-depletion effect. *Perspectives on Psychological Science*, *11*, 546–573. https://doi.org/10.1177/0956797614526415.Data

Inzlicht, M., & Friese, M. (2019). The past, present, and future of ego depletion. *Social Psychology*, *50*(5–6), 370–378. https://doi.org/10.1027/1864-9335/a000398

Inzlicht, M., Schmeichel, B. J., & Macrae, C. N. (2014). Why self-control seems (but may not be) limited. *Trends in Cognitive Sciences*, *18*(3), 127–133. https://doi.org/10.1016/j.tics.2013.12.009

Jia, L., Hirt, E. R., & Fishbach, A. (2019). Protecting an important goal: When prior self-control increases motivation for active goal pursuit. *Journal of Experimental Social Psychology*, *85*, 103875. https://doi.org/10.1016/j.jesp.2019.103875

Job, V., Dweck, C. S., & Walton, G. M. (2010). Ego depletion—Is it all in your head? Implicit theories about willpower affect self-regulation. *Psychological Science*, *21*(11), 1686–1693. https://doi.org/10.1177/0956797610384745

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

Kensinger, E. A. (2004). Remembering emotional experiences: The contribution of valence and arousal. *Reviews in the Neurosciences*, *15*(4), 241–251. https://doi.org/10.1515/REVNEURO.2004.15.4.241

Kotabe, H. P., Righetti, F., & Hofmann, W. (2019). How anticipated emotions guide self-control judgments. *Frontiers in Psychology*, *10*, 1614. https://doi.org/10.3389/fpsyg.2019.01614

Kruger, J., & Evans, M. (2004). If you don't want to be late, enumerate: Unpacking reduces the planning fallacy. *Journal of Experimental Social Psychology*, *40*(5), 586–598. https://doi.org/10.1016/j.jesp.2003.11.001

Kurzban, R., Duckworth, A. L., Kable, J. W., & Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *The Behavioral and Brain Sciences*, *36*(6), 661–679. https://doi.org/10.1017/S0140525X12003196

Kuznetsova, A., Brockhoff, P., & Christensen, R. H. B. (2017). lmerTest Package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26.

Lam, K. G. H., Buehler, R., McFarland, C., Ross, M., & Cheung, I. (2005). Cultural differences in affective forecasting: The role of focalism. *Personality and Social Psychology Bulletin*, *31*(9), 1296–1309. https://doi.org/10.1177/0146167205274691

Lin, H., Saunders, B., Friese, M., Evans, N. J., & Inzlicht, M. (2020). Strong effort manipulations reduce response caution: A preregistered reinvention of the ego-depletion paradigm. *Psychological Science*, *31*(5), 531–547. https://doi.org/10.1177/0956797620904990

Loehr, V. G., & Baldwin, A. S. (2014). Affective forecasting error in exercise: Differences between physically active and inactive individuals. *Sport, Exercise, and Performance Psychology*, *3*(3), 177–183. https://doi.org/10.1037/spy0000006

Lovell, D., (Producer), & Zeffirelli, F. (Director), (1979). The champ [Motion picture]. United States: Metro-Goldwyn-Mayer Studios Inc.

Meyvis, T., Ratner, R. K., & Levav, J. (2010). Why don't we learn to accurately forecast feelings? How misremembering our predictions blinds us to past forecasting errors. *Journal of Experimental Psychology: General*, *139*(4), 579–589. https://doi.org/10.1037/a0020285

Miloyan, B., & Suddendorf, T. (2015). Feelings of the future. *Trends in Cognitive Sciences*, *19*(4), 196–200. https://doi.org/10.1016/j.tics.2015.01.008

Mitchell, T. R., Thompson, L., Peterson, E., & Cronk, R. (1997). Temporal adjustments in the evaluation of events: The 'rosy view. *Journal of Experimental Social Psychology*, *33*(4), 421–448. https://doi.org/10.1006/jesp.1997.1333

Morewedge, C. K., Gilbert, D. T., Myrseth, K. O. R., Kassam, K. S., & Wilson, T. D. (2010). Consuming experience: Why affective forecasters overestimate comparative value. *Journal of Experimental Social Psychology*, *46*(6), 986–992. https://doi.org/10.1016/j.jesp.2010.07.010

Mukhopadhyay, A., & Johar, G. V. (2005). Where there is a will, is there a way? Effects of lay theories of self-control on setting and keeping resolutions. *Journal of Consumer Research*, *31*(4), 779–786. https://doi.org/10.1086/426611

Nordgren, L. F., van Harreveld, F., & van der Pligt, J. (2009). The restraint bias. *Psychological Science*, *20*(12), 1523–1528. https://doi.org/10.1111/j.1467-9280.2009.02468.x

Pontifex, M. B., McGowan, A. L., Chandler, M. C., Gwizdala, K. L., Parks, A. C., Fenn, K., & Kamijo, K. (2019). A primer on investigating the after effects of acute bouts of physical activity on cognition. *Psychology of Sport and Exercise*, *40*, 1–22. https://doi.org/10.1016/j.psychsport.2018.08.015

Preacher, K. J., Curran, P. J., & Bauer, D. J. (2004). *Simple intercepts, simple slopes, and regions of significance in MLR 3-way interactions*, Unpublished, 1–5.

Primoceri, P., Ramer, N., Ullrich, J., & Job, V. (2021). The role of task similarity for ego depletion: A registered report. *Journal of Experimental Social Psychology*, *95*, 104133. https://doi.org/10.1016/j.jesp.2021.104133

Regan, P. C., Snyder, M., & Kassin, S. M. (1995). Unrealistic optimism: Self-enhancement or person positivity? *Personality and Social Psychology Bulletin*, *21*(10), 1073–1082. https://doi.org/10.1177/01461672952110008

Robinson, M. D., & Clore, G. L. (2001). Simulation, scenarios, and emotional appraisal: Testing the convergence of real and imagined reactions to emotional stimuli. *Personality and Social Psychology Bulletin*, *27*(11), 1520–1532. https://doi.org/10.1177/01461672012711012

Roger, B., Dale, G., & Michael, R. (1994). Exploring the 'planning fallacy': Why people underestimate their task completion times. *Journal of Personality and Social Psychology*, *67*(3), 366–381.

Savani, K., & Job, V. (2017). Reverse ego-depletion: Acts of self-control can improve subsequent performance in Indian cultural contexts. *Journal of Personality and Social Psychology*, *113*, 589–607. https://doi.org/10.1037/pspi0000099

Schmeichel, B. J. (2007). Attention control, memory updating, and emotion regulation temporarily reduce the capacity for executive control. *Journal of Experimental Psychology. General*, *136*(2), 241–255. https://doi.org/10.1037/0096-3445.136.2.241

Scholer, A. A., Miele, D. B., Murayama, K., & Fujita, K. (2018). New directions in self-regulation: The role of metamotivational beliefs. *Current Directions in Psychological Science*, *27*(6), 437–442. https://doi.org/10.1177/0963721418790549

Shepperd, J. A., Pogge, G., & Howell, J. L. (2017). Assessing the consequences of unrealistic optimism: Challenges and recommendations. *Consciousness and Cognition*, *50*, 69–78. https://doi.org/10.1016/j.concog.2016.07.004

Sun, X., Cortina, K. S., Miller, K. F., & Ning, H. (2019). Willpower as cultural construct: Do chinese students believe less in its depletion? *Frontiers in Psychology*, *10*, 1–7. https://doi.org/10.3389/fpsyg.2019.00988

Talarico, J. M., Labar, K. S., & Rubin, D. C. (2004). Emotional intensity predicts autobiographical memory experience. *Memory and Cognition*, *32*(7), 1118–1132. https://doi.org/10.3758/BF03196886

Tangney, J., Baumeister, R. F., & Boone, A. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality*, *72*(2), 271–322. https://doi.org/10.1111/j.0022-3506.2004.00263.x

Vohs, K. D., Schmeichel, B. J., Lohmann, S., Gronau, Q. F., Finley, A., Wagenmakers, E. J., Ainsworth, S. E., Alquist, J., L, M., D, B., A, B., A, B., Grant, J., Campbell, C., Capaldi, J., Cau, C., Chambers, H., Chatzisarantis, N., Christensen, W., … Albarracind (2021). A preregistered paradigmatic test of the ego depletion effect. *Psychological Science*, *32*(10), 1566–1581. https://doi.org/10.1177/0956797621989733

Wang, Y., Yang, L., & Wang, Y. (2014). Suppression (but not reappraisal) impairs subsequent error detection: An ERP study of emotion regulation's resource-depleting effect. *Plos One*, *9*(4), e96339. https://doi.org/10.1371/journal.pone.0096339

Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology*, *39*(5), 806–820. https://doi.org/10.1037//0022-3514.39.5.806

Wenzel, M., Rowland, Z., Zahn, D., & Kubiak, T. (2019). Let there be variance: Individual differences in consecutive self-control in a laboratory setting and daily life. *European Journal of Personality*, *33*(4), 468–487. https://doi.org/10.1002/per.2208

Werle, C. O., Wansink, B., & Payne, C. R. (2015). Is it fun or exercise? The framing of physical activity biases subsequent snacking. *Marketing Letters*, *26*(4), 691–702. https://doi.org/10.1007/s11002-014-9301-6

Werner, K. M., Milyavskaya, M., Foxen-Craft, E., & Koestner, R. (2016). Some goals just feel easier: Self-concordance leads to goal progress through subjective ease, not effort. *Personality and Individual Differences*, *96*, 237–242. https://doi.org/10.1016/j.paid.2016.03.002

Wilson, T. D., & Gilbert, D. T. (2003). Affective forecasting. *Advances in Experimental Social Psychology*, *35*, 345–411. https://doi.org/10.1016/S0065-2601(03)01006-2

Wilson, T. D., Meyers, J., & Gilbert, D. T. (2003). How happy was i, anyway?' A retrospective impact bias. *Social Cognition*, *21*(6), 421–446. https://doi.org/10.1521/soco.21.6.421.28688

Yarkoni, T. (2022). The generalizability crisis. *Behavioral and Brain Sciences*, *45*, 1–16. https://doi.org/10.1017/S0140525X20001685

Zhang, C., Smolders, K. C. H. J., Lakens, D., & IJsselsteijn, W. A. (2018). Two experience sampling studies examining the variation of self-control capacity and its relationship with core affect in daily life. *Journal of Research in Personality*, *74*, 102–113. https://doi.org/10.1016/j.jrp.2018.03.001

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Francis, Z., & Inzlicht, M. (2023). Expecting tasks to help or hurt subsequent cognitive performance: Variability, accuracy, and bias in forecasted after-effects. *European Journal of Social Psychology*, *53*, 531–551. https://doi.org/10.1002/ejsp.2921

APPENDIX

Tasks used for Studies 1a and 1b

Writing an exam	A fitness test
Watching a movie	A job interview
Relaxing on the beach	Playing in a sports game
Doing your taxes	Listening to an important presentation at work
Babysitting a friend's children	Working on a difficult Sudoku puzzle
Playing a video game	Deciding what to order at a restaurant*
Shopping for clothes	
Browsing the internet/social media	*(in *Study 1a only*)

Task descriptions used for Studies 2 and 3

Imagine that you were presented with a 2-½minute long movie-clip, portraying one of the saddest scenes in movie history. The movie shows a young boy, full of emotion, crying next to his father who has just died. While watching this movie, you are instructed to suppress all internal reactions to the movie and suppress all external signs of your feelings. A camera is set up on the table recording your facial expressions, to confirm that you are not showing any visible emotions.	Imagine that you were presented with a 2-½minute long movie-clip, portraying one of the saddest scenes in movie history. The movie shows a young boy, full of emotion, crying next to his father who has just died. You watch this movie clip. There are no additional instructions, and you watch the movie freely.
Imagine that you completed a 7-minute fitness routine. The fitness routine involves: - doing jumping jacks for 60 seconds - doing as many pushups as possible within 60 seconds - doing a wall-sit for 60 seconds, holding a plank (e.g. a straight-body push-up position, without moving) for 60 seconds - holding a handgrip closed for 60 seconds (or as long as possible) - doing 'burpees' for 60 seconds, You would get 20 second breaks between each task, and the exercise routine would be done by yourself, on a yoga mat, with one researcher present.	Imagine that you spent 7 minutes walking through a hallway in the Science Wing at UTSC (e.g. from the far end of SY to the other end of SW, and back again). You can walk at whatever speed you prefer, and you are not carrying your backpack or any other bags/objects.
Imagine that you were asked to write a story about a recent trip you have taken. It may be a trip to the store, to Kingston, or to another country - wherever! Importantly, though, you can NOT use the letters 'A' or 'N' anywhere in your story. Instead, you must find another way to express your thoughts (e.g. write 'yellow fruit' instead of 'banana'). Imagine that you spend 6 minutes writing your story, without using the letters A or N. Imagine that you write this story on paper (with a pencil).	Imagine that you were asked to spend 6 minutes writing a story about a recent trip you have taken. It may be a trip to the store, to Kingston, or to another country—wherever! Imagine that you write this story on paper (with a pencil). There are no additional instructions, and you write your story freely.