nature human behaviour

Registered Report

https://doi.org/10.1038/s41562-024-01842-7

An experimental manipulation of the value of effort

Received: 8 July 2020

Accepted: 31 January 2024

Published online: 04 March 2024

Check for updates

Hause Lin [1,2], Andrew Westbrook [3], Frank Fan[4] & Michael Inzlicht[4,5]

People who take on challenges and persevere longer are more likely to succeed in life. But individuals often avoid exerting effort, and there is limited experimental research investigating whether we can learn to value effort. We developed a paradigm to test the hypothesis that people can learn to value effort and will seek effortful challenges if directly incentivized to do so. We also dissociate the effects of rewarding people for choosing effortful challenges and performing well. The results provide limited evidence that rewarding effort increased people's willingness to choose harder tasks when rewards were no longer offered (near transfer). There was also mixed evidence that rewarding effort increased willingness to choose harder tasks in another unrelated and unrewarded task (far transfer). These heterogeneous results highlight the need for further research to understand when this paradigm may be the most effective for increasing and generalizing the value of effort.

Many people value effort. By exerting effort and persevering over time, we learn to approach challenges, become resilient and develop self-regulatory skills. These qualities correlate with not only academic attainment[1–3] but also outcomes as disparate as health, wealth and criminal offending[4–6]. But despite growing awareness of the importance of these qualities for individuals and societies[7–11], it remains unclear whether and how individuals can learn to value effort and approach challenges.

Although effort is costly[12–17], much evidence suggests that humans and other animals can (learn to) value effort[18–21]: We sometimes willingly challenge ourselves with tasks that are difficult to perform (for example, endurance sports) precisely because these tasks require effort and perseverance, contrary to the notion that we avoid tasks with high perceived effort or low success likelihood[14,22,23]. This so-called effort paradox[19] raises questions about how people (and other animals) come to value effort for its own sake.

Our goal is to develop a paradigm to investigate whether people can learn to seek out effortful tasks in an experimental context. Our approach contrasts with and addresses problems with existing interventions that aim to enhance cognitive skills and performance[24–26]. These interventions typically improve performance on trained tasks (that is, near-transfer effects) but not untrained tasks, even when the tasks are closely related[27–32]. Crucially, it is unclear whether these interventions benefit everyday behaviour and cognition[33] (that is, far transfer), nor is it clear whether emphasizing performance helps or hinders the learning of the value of effort and willingness to embrace effortful challenges.

Interventions that focus directly on improving these qualities have also provided mixed results. For example, training self-control regularly (by repeatedly practising overriding specific habitual responses for extended periods) does not reliably improve self-control performance in the laboratory or willingness to exert effort in everyday life[33–35]. Similarly, some work suggests that instilling grit or growth mindsets in students improves academic outcomes[36–39]. However, these effects might be less robust than previously thought[40–43]. Moreover, these interventions often target many skills concurrently (for example, grit, goal setting and emotion regulation), making it difficult to evaluate how and why they have (not) worked and, thus, provide little insight into whether people can be trained to value effortful challenges.

Decades of research on learning suggests that rewards[44–46]—if they appropriately shape environmental contingencies—can change humans' and non-human animals' preferences for challenge and exerting effort. For example, when infants observe adults persevere longer to achieve their goals, they subsequently work harder on a novel task[47], suggesting that conducive environments can help individuals value effort[48,49].

[1]Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA, USA. [2]Department of Psychology, Cornell University, Ithaca, NY, USA. [3]Center for Advanced Human Brain Imaging Research, Rutgers University, Piscataway, NJ, USA. [4]Department of Psychology, University of Toronto, Toronto, Ontario, Canada. [5]Rotman School of Management, University of Toronto, Toronto, Ontario, Canada. e-mail: hauselin@gmail.com

 https://doi.org/10.1038/s41562-024-01842-7

So why do people tend to avoid exerting effort even when they know that it is associated with positive, rewarding outcomes? We suggest that people have acquired associations that inadvertently encourage effort-avoidant behaviour. Rewards are often used to motivate good performance rather than instil the value of taking on effortful challenges. Many studies show that rewards improve immediate task performance[50–56], but they often fail to consider how rewards could also reduce the intrinsic value of effort[57–59], especially when rewards are later removed. Moreover, performance-based rewards could make us avoid taking on challenges (and prefer easier tasks) because it is more difficult to perform well on harder tasks (and easier to perform well on easier tasks). Thus, performance-based rewards might make individuals avoid beneficial but effortful tasks (for example, doing homework) that are challenging or invoke feelings of frustration[60,61].

A more promising approach is to directly reinforce behaviours that emphasize the value of engaging in effortful challenges[62,63]. For example, when rewarded for performing harder tasks, children and rats persisted longer or worked harder on subsequent unrelated effortful tasks[64–66]. These studies might not have teased apart or controlled for the value of performing well, but they nevertheless provided initial evidence for the idea that rewards can be used to reinforce behaviours that emphasize the value of taking on effortful challenges.

Moreover, when the act of engaging in effortful challenges is repeatedly associated with reward, effort itself can become a secondary reinforcer, making it rewarding in its own right[19,45]. Effort becoming a secondary reinforcer might help explain why we often willingly engage in effortful tasks such as endurance sports or crossword puzzles. This approach can not only potentially increase effort's value but also provide insights into how to foster qualities such as perseverance and conscientiousness[67], which might be particularly important for at-risk populations (for example, low-income students[68,69]).

What is emerging is a simple idea: People should become more willing to choose to perform effortful tasks if directly incentivized to do so. Crucially, this idea can be tested experimentally. To do so, we designed a three-section experimental paradigm (pre-training, training and post-training; Fig. 1) that rewarded participants for either choosing to engage in hard tasks (effort condition) or performing well (performance condition) during training. We also included a neutral condition where participants received the same amount of rewards regardless of their choice and performance. This design allowed us to dissociate the effects of these three types of rewards. While the results of this experiment cannot offer strong conclusions regarding how to intervene in real-world settings[70], they could nevertheless establish if it is possible to reinforce choosing to engage in demanding tasks and provide insights into how to design future intervention studies.

We predict that rewarding willingness to choose and engage in hard tasks during training will instil in participants the value of effort, which will manifest in increased willingness to engage in hard tasks—even ones that have not been associated with rewards and in the absence of rewards. We test eight related hypotheses that provide different levels of evidence for our core idea: Participants who have been rewarded for engaging in effortful tasks (effort condition) will choose harder tasks over easier tasks more frequently than participants in the performance and neutral conditions. The outcome measures for our hypotheses are related measures of effort preferences (that is, the percentage of choices whereby participants choose the hard task).

Importantly, we also assess participants' baseline effort preferences during a no-reward pre-training section so we can include pre-training effort preferences as a covariate in our models. That is, we statistically control for idiosyncratic preferences that might influence choice (for example, participants might choose to perform a task because they like certain low-level stimulus properties associated with the task cue). This approach not only allows us to accurately estimate the direct effects of our experimental manipulations but also further increases statistical power and mitigates against false negatives. Specifically, after controlling for baseline effort preferences in the pre-training section, we expect the following (see Table 1 for details):

1. On rewarded trials during the training section (that is, rewarded trials are interleaved with probe trials), participants in the effort condition will choose the hard task more frequently than those in the performance and neutral conditions (Table 1, rows 1 and 2). This analysis mainly reveals whether participants have learned the task and reward contingencies.
2. On probe (that is, unrewarded) trials during the training section (that is, rewarded trials are interleaved with probe trials), participants in the effort condition will choose the hard task more frequently than those in the performance and neutral conditions (Table 1, rows 3 and 4), reflecting within-block near-transfer effects. Differences in effort preferences across conditions will probably reflect carryover effects (owing to the interleaved reward and probe trials within a block) and actual changes in effort preferences (near transfer).
3. On trials during the post-training section (that is, completely unrewarded section), participants in the effort condition will choose the hard task more frequently than those in the performance and neutral condition (Table 1, rows 5 and 6). This analysis examines across-block near-transfer effects: Differences in effort preferences across conditions on these trials are less likely to be driven by carryover effects, since rewards are completely absent in this section.
4. In a separate block during the post-training section (that is, completely unrewarded section), we will examine across-block far-transfer effects by presenting participants with a different cognitive task that had not been paired with rewards during the training section. We expect participants in the effort condition to choose the hard task more frequently than those in the performance and neutral conditions (Table 1, rows 7 and 8). Since this cognitive task has not been associated with rewards in the training section, this analysis provides evidence for whether the manipulations can lead to domain-general changes in effort preferences that transfer to unrelated tasks.

Two of the eight hypotheses in Table 1 (rows 6 and 8) most directly test our claims because they predict that, in the absence of rewards (that is, post-training section), effort preferences will be higher in the effort than neutral condition on the inhibition task (hypothesis 6) and updating task (hypothesis 8). If the Bayesian analyses provide positive evidence for only one of the two hypotheses, the results will not refute our overarching claims—instead, they will provide support for only across-block near-transfer (hypothesis 5) or across-block far-transfer effects (hypothesis 7). Nevertheless, any pattern of positive results for hypotheses 3–8 should also provide some evidence for transfer effects and will be theoretically informative, given that transfer effects are rarely reported in the literature. However, four broad patterns of results will be the most informative.

First, if we find positive results for hypotheses 3 and 4 but not 5–8, the results will provide evidence for only within-block near-transfer effects. That is, changes in effort preferences are only observed during the training block but not during the post-training block. Although the lack of far-transfer effects (hypotheses 5–8) is inconsistent with the core hypotheses, the within-block near-transfer effects (hypothesis 3 and 4) suggest that longer or more intense interventions (for example, longitudinal studies) might be better suited to testing our hypotheses and examining the possibility of across-block near- and far-transfer effects.

Second, in addition to hypotheses 3 and 4, we might also find support for hypotheses 5 and 6. Positive results for these four hypotheses provide evidence for within-block (hypotheses 3 and 4) and across-block (hypotheses 5 and 6) effects that are limited to only the task that has been associated with rewards during training (that is, only near but not far

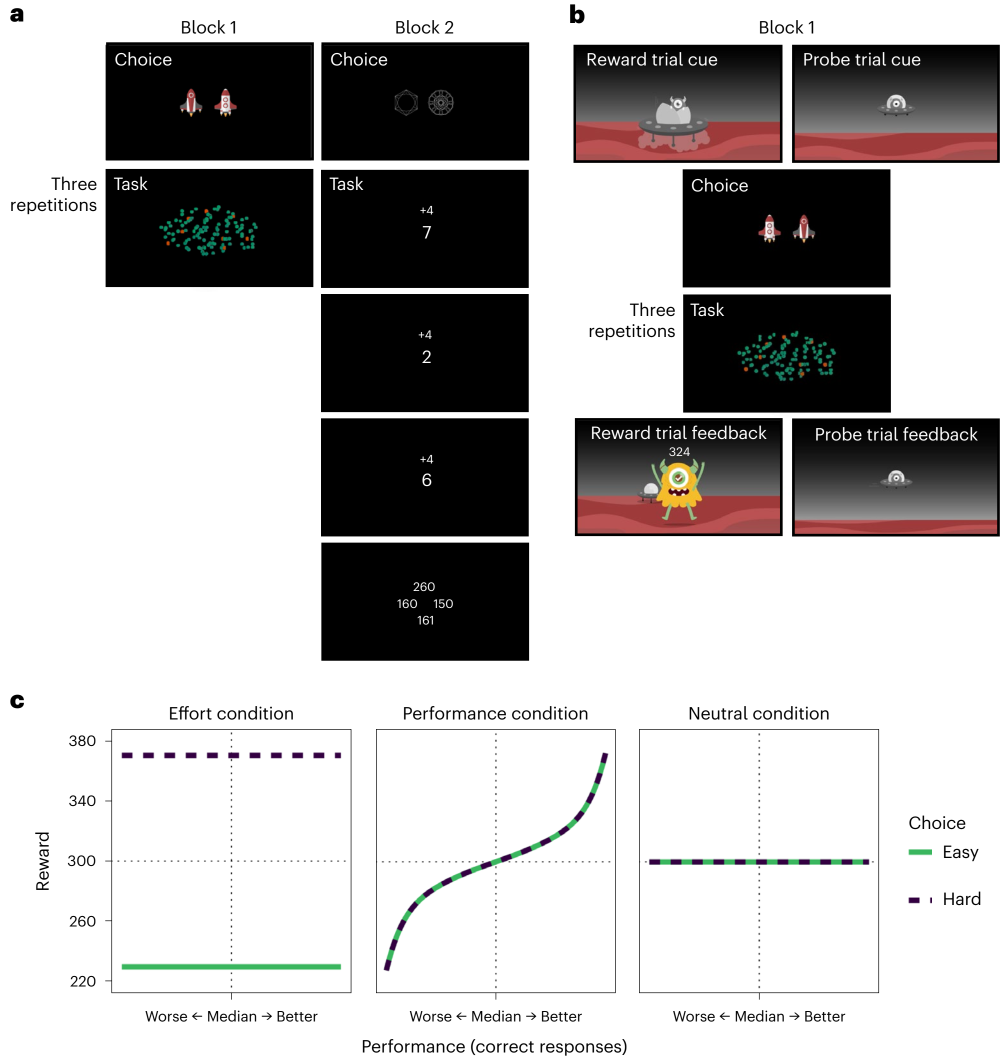

Fig. 1 | Design. The experiment has pre-training, training and post-training sections. Rewards are delivered only during the training section, but not the pre-training or post-training sections. **a**, The pre- and post-training sections consist of two demand selection task blocks each; these two blocks assess effort preferences on two unrelated cognitive tasks (to aid visualization, fewer and larger dots are shown than in the actual dot-motion inhibition task). On each trial, participants choose the version of the task (easy or hard) they prefer and then perform their chosen task. Effort preference is defined as the proportion of choices whereby participants choose the hard task. **b**, The training section has only one block that consists of two distinct and pseudo-randomly interleaved trial types: rewarded trial ($n = 40$) and probe trial ($n = 20$). At the beginning of each trial, participants will see one of two cues (reward trial cue or probe trial cue) explicitly signalling the presence or absence of rewards on that trial. Next, they will be presented with two options (easy or hard); they will choose the version they prefer and then perform their chosen task. If it is a rewarded trial (signalled by the reward-trial cue), participants can earn rewards or points; if it is a probe (that is, unrewarded) trial (signalled by the probe-trial cue), participants will be fully aware that they will not receive any points (see also probe trial feedback panel), no matter their choice and performance. **c**, Value functions showing how rewards on rewarded trials in the training section differ across the three experimental conditions. Vertical dotted lines reflect median performance (i.e. reaction time) for a given participant; horizontal dotted lines reflect the mean reward (number of points). Choice (easy versus hard) determines rewards in the effort condition. Performance (correct reaction times) determines rewards in the performance condition. Neither choice nor performance determines rewards in the neutral condition. Critically, the expected reward is identical across conditions. No points are given for incorrect or missed responses (reaction time deadlines are participant specific). All conditions have identical task instructions, structure, sequence and cues. The only difference is whether rewards are assigned based on the reward-effort value function (effort condition), reward-performance value function (performance condition) or a uniform distribution (neutral condition).

transfer). Thus, changes in effort preferences on one task carry over to the same task in a different block but do not generalize to a different task.

Third, we might obtain positive results only during the post-training block (hypotheses 5–8) and not during the training block (hypotheses 3 and 4), which provide evidence for across-block but not within-block transfer effects. If so, we speculate that consolidation processes that occur between the training and post-training sections might be necessary for changing effort preferences, and further analyses and studies will be necessary to examine this possibility.

Finally, we might find positive results only when contrasting the effort and performance conditions (hypotheses 1, 3, 5 and 7) but not the effort and neutral conditions (hypotheses 2, 4, 6 and 8). These results could suggest the presence of different processes: The reward-effort manipulation increases effort preferences, but the reward-performance manipulation decreases effort preferences.

Results

Rewarded and probe trials (hypotheses 1–4)

We found strong evidence for experimental hypotheses 1 and 2 (Fig. 2), suggesting that participants successfully learned the task and reward contingencies. On rewarded trials in the training section, participants in the effort condition chose the hard inhibition task more frequently

Table 1 | Design table

Question	Hypothesis	Sampling plan	Analysis	Interpretation given to different outcomes
1. Do effort preferences on rewarded trials in the training section differ between the effort and performance conditions?	Effort preferences will be higher in the effort condition than performance condition.	We will stop sampling or recruiting participants when the BF for hypotheses 6 and 8 (that is, rows 6 and 8 of this table) exceed a threshold (10 or 0.1) or when a total of 750 participants have been recruited.	The outcome measure is effort preferences (% hard choices) on rewarded trials in the training section. We will fit a Bayesian linear regression model with condition as the main regressor and effort preferences on the pre-training inhibition task as the covariate.	The results are consistent with the hypothesis if the BF ≥3. The results are consistent with the null hypothesis if the BF ≤0.3. Specifically, BFs larger than 1, 3 or 10 (or less than 1, 0.3 or 0.1) provide anecdotal, moderate and strong evidence for the experimental hypothesis (or null hypothesis).
2. Do effort preferences on rewarded trials in the training section differ between the effort and neutral conditions?	Effort preferences will be higher in the effort condition than neutral condition.	Same as above.	Same as the row above.	Same as above.
3. Do effort preferences on probe (unrewarded) trials in the training section differ between the effort and performance conditions?	Effort preferences will be higher in the effort condition than performance condition.	Same as above.	The outcome measure is effort preferences (% hard choices) on probe trials in the training section. We will fit a Bayesian linear regression model with condition as the main regressor and effort preferences on the pre-training inhibition task as the covariate.	Same as above.
4. Do effort preferences on probe (unrewarded) trials in the training section differ between the effort and neutral conditions?	Effort preferences will be higher in the effort condition than neutral condition.	Same as above.	Same as the row above.	Same as above.
5. Do effort preferences on the inhibition task in the post-training section differ between the effort and performance conditions?	Effort preferences will be higher in the effort condition than performance condition.	Same as above.	The outcome measure is effort preferences (% hard choices) on the inhibition task in the post-training section. We will fit a Bayesian linear regression model with condition as the main regressor and effort preferences on the pre-training inhibition task as the covariate.	Same as above.
6. Do effort preferences on the inhibition task in the post-training section differ between the effort and neutral conditions?	Effort preferences will be higher in the effort condition than neutral condition.	Same as above.	Same as the row above.	Same as above.
7. Do effort preferences on the updating task in the post-training section differ between the effort and performance conditions?	Effort preferences will be higher in the effort condition than performance condition.	Same as above.	The outcome is effort preferences (% hard choices) on the updating task in the post-training section. We will fit a Bayesian linear regression model with condition as the main regressor and effort preferences on the pre-training updating task as the covariate.	Same as above.
8. Do effort preferences on the updating task in the post-training section differ between the effort and neutral conditions?	Effort preferences will be higher in the effort condition than neutral condition.	Same as above.	Same as the row above.	Same as above.

than those in the performance condition (H1) (beta estimate, $b = 0.34$ [0.29, 0.38], Cohen's $d = 1.22$, Bayes factor (BF) > 100; $P < 0.001$). Effort preferences were also higher in the effort than neutral condition (H2) ($b = 0.26$ [0.22, 0.31], $d = 0.95$, BF > 100; $P < 0.001$). These effects were primarily driven by participants in the effort and performance conditions increasing and decreasing their effort preferences, respectively, relative to their preferences in the pre-training section (Fig. 3).

Hypotheses 3 and 4, where we examined the effects of the experimental manipulation on the interleaved probe (that is, unrewarded) trials in the training section, were not confirmed. The BFs provide anecdotal support for the null hypotheses: Effort preferences were not higher in the effort than performance condition (H3) ($b = 0.04$ [0.00, 0.09], $d = 0.17$, BF = 0.52; $P = 0.066$) and were also not higher relative to the neutral condition (H4) ($b = 0.05$ [0.01, 0.09], $d = 0.19$, BF = 0.83; $P = 0.036$). These findings suggest no carryover or near-transfer effects (that is, from reward to probe trials) within the training section, but they do not preclude across-block transfer effects (hypotheses 5–8), which also more directly test our hypotheses.

Post-training near transfer (hypotheses 5 and 6)

We then investigated the effects of the experimental manipulation on effort preferences in the post-training section whereby no rewards were provided for the inhibition task (that is, across-block near transfer). Participants completed a demand selection task that was identical to the one they completed in the pre-training section (Fig. 1). Participants in the effort condition chose the hard inhibition task more frequently

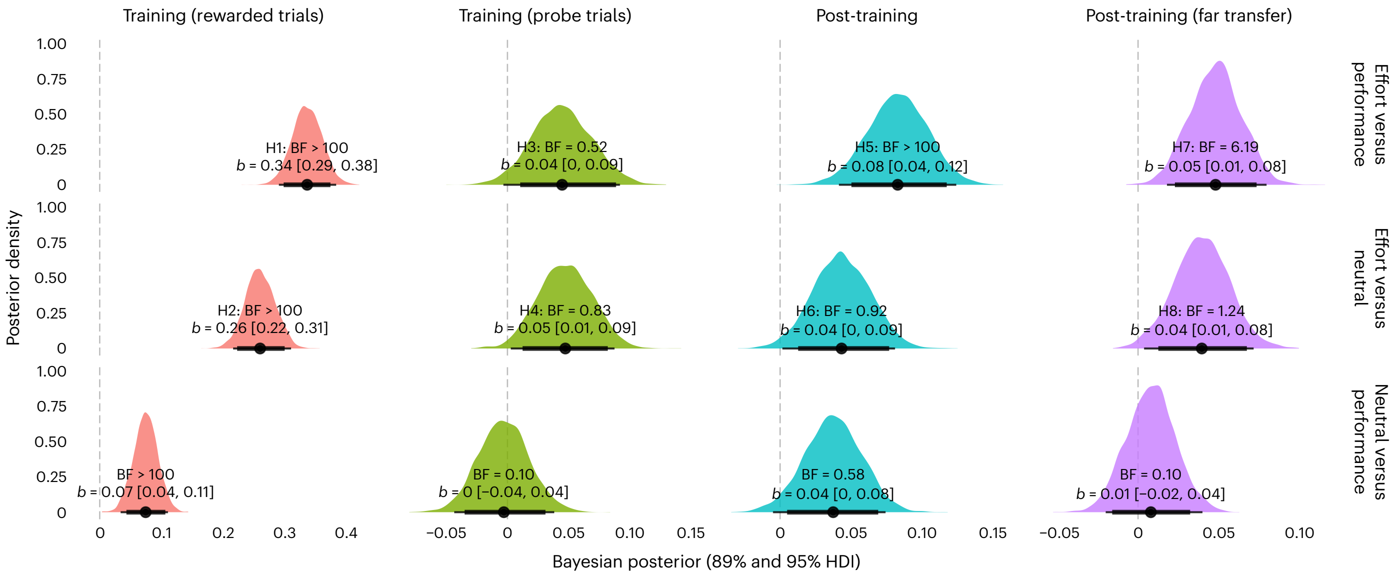

Fig. 2 | Bayesian posterior densities for the effect of condition on effort preferences. For the eight registered hypotheses (top and middle rows), the mass of the posterior distributions was mostly above zero, and five hypotheses had BFs greater than 1 (at least anecdotal evidence for the experimental hypothesis). Bottom rows analyses were not registered and are included here only to facilitate the interpretation of condition contrasts in the top and middle rows. Positive Bayesian posterior estimates are hypothesis-consistent results (that is, higher effort preference in the effort than the other condition). No result is consistent with the null hypothesis because no BF is smaller than the registered criteria of BF ≤ 0.3. Posterior means and HPD intervals (95% and 89%) are shown. Dashed vertical lines indicate the null value where effort preference is the same in both conditions (i.e. beta estimate = 0).

than those in the performance condition (H5; Fig. 2) (b = 0.08 [0.04, 0.12], d = 0.35, BF > 100; P < 0.001), confirming hypothesis 5. However, hypothesis 6 was not confirmed: There was anecdotal evidence favouring the null hypothesis when contrasting the effort and neutral conditions (H6; Fig. 2) (b = 0.04 [0.00, 0.09], d = 0.19, BF = 0.92; P = 0.032).

Post-training far transfer (hypotheses 7 and 8)

In the post-training section, participants also completed a separate block whereby no rewards were provided for performing a demand selection that required them to choose to perform either the easy or hard updating task (which they had also completed in one of the pre-training blocks, but not in the training block; Fig. 1). Effort preferences in this block allowed us to evaluate whether there was far transfer from the inhibition to updating task. The results were consistent with hypothesis 7 but provided relatively weaker support for hypothesis 8 (Fig. 2). Participants in the effort condition chose the hard updating task more frequently than those in the performance condition (H7) (b = 0.05 [0.01, 0.08], d = 0.26, BF = 6.19; P = 0.004), but there was anecdotal evidence suggesting that they chose the hard task more frequently than those in the neutral condition (H8) (b = 0.04 [0.01, 0.08], d = 0.21, BF = 1.24; P = 0.022).

Exploratory analyses

Magnitude of rewards received influenced preference. We examined whether the magnitude of rewards received influenced effort preference (model specification: effort preference ~ condition * (pre-training baseline preference + mean rewards received)). Note that we only contrast the effort and performance conditions in this analysis because there was no variation in rewards received in the neutral condition (300 points were delivered for all correct responses, regardless of choice and performance).

On rewarded trials, participants who received more rewards in the effort condition chose the hard task much more (b = 0.66 [0.62, 0.70], d = 4.32, BF > 100; P < 0.001), and there was a strong interaction between condition and rewards received (b = −0.73 [−0.77, −0.69], d = −4.81, BF > 100; P < 0.001), such that the positive relationship was attenuated (and became negative) for those in the performance condition. We found similar results on probe trials (positive relationship between rewards received and effort preference in the effort condition: b = 0.21 [0.14, 0.28], d = 0.79, BF > 100; P < 0.001; interaction between condition and points received: b = −0.23 [−0.30, −0.15], d = −0.85, BF > 100; P < 0.001), post-training near-transfer trials (positive relationship in effort condition: b = 0.14 [0.08, 0.20], d = 0.60, BF > 100; P < 0.001; interaction between condition and points received: b = −0.16 [−0.22, −0.10], d = −0.71, BF > 100; P < 0.001) and post-training far-transfer trials (positive relationship in effort condition: b = 0.10 [0.05, 0.14], d = 0.55, BF = 29.78; P < 0.001; interaction between condition and points effect: b = −0.11 [−0.16, −0.06], d = −0.59, BF > 100; P < 0.001).

The consistent patterns of results above indicate opposing effects: In the effort condition, participants who, on average, received more rewards also chose the hard task more frequently across all sections of the task and trial types. However, in the performance condition, participants who received more rewards chose the hard task less frequently.

Task performance. The posterior distributions were mostly centred above zero, and five hypotheses had BFs greater than 1 (Fig. 2). The BFs also provided evidence for two transfer effects (post-training near-transfer (H5) and far-transfer (H7) effects), providing limited evidence that rewards may be used to reinforce behaviours that emphasize the value of taking on effortful challenges.

However, an important question is whether performance also differed across conditions (note that controlling for pre-training task accuracy led to the same conclusions regarding the efficacy of our manipulations; Supplementary Fig. 2). Ideally, the manipulation should increase effort preferences without reducing accuracy, such that task accuracy should not differ across conditions. Alternatively, participants in the effort condition might perform worse because the manipulation could have caused them to seek rewards while sacrificing accuracy, and that they might continue to perform worse even when rewards were no longer available.

We fitted eight models (corresponding to the eight registered hypotheses; Fig. 4) to explore whether task accuracy differed between

 https://doi.org/10.1038/s41562-024-01842-7

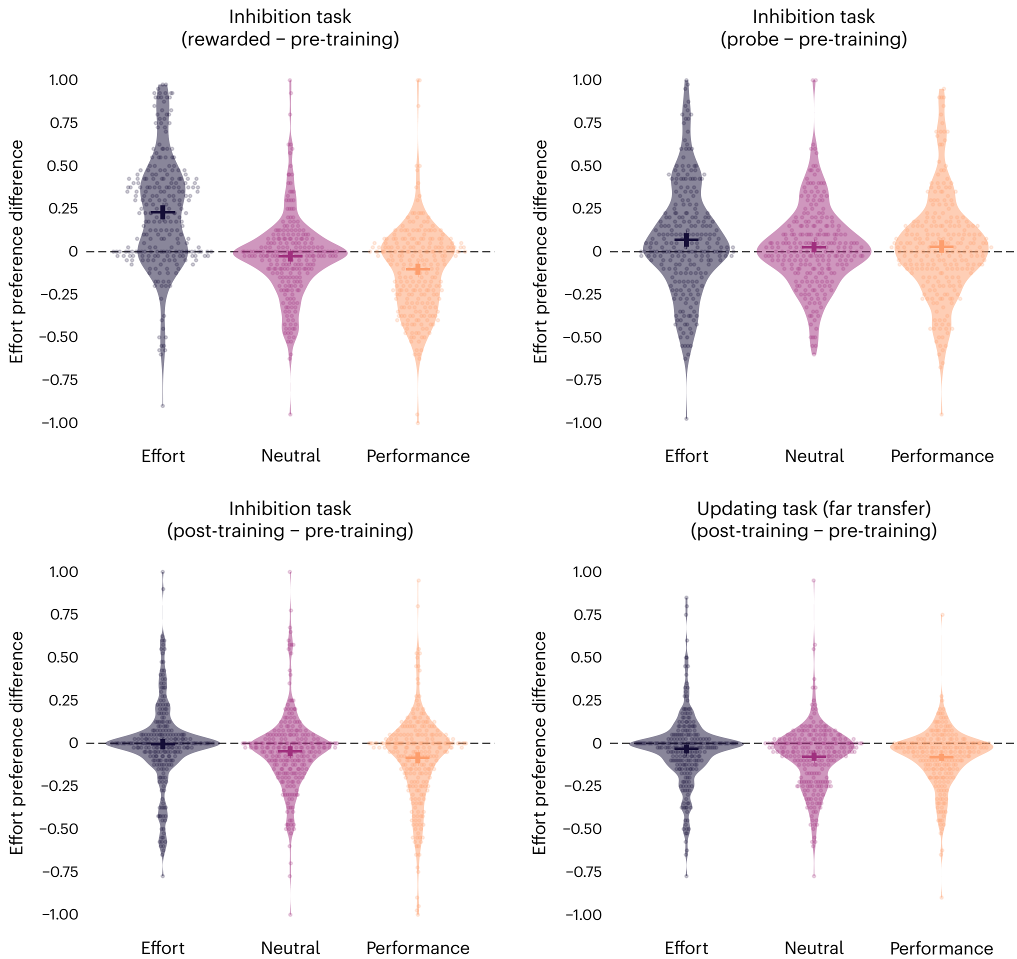

Fig. 3 | Effort preference relative to baseline preference in the pre-training section. The effort preference difference computed by subtracting the pre-training baseline effort preference from the effort preference for rewarded trials (training section), probe trials (training section), post-training inhibition task and post-training updating task. Zero indicates no change relative to pre-training baseline. Each dot is one participant. Mean and 95% credible intervals are shown.

conditions (model specification: accuracy ~ condition + pre-training baseline accuracy + objective task difficulty; see Supplementary Fig. 3 for task reaction time results and Supplementary Fig. 4 for accuracies for reaction time quartiles). For rewarded trials, accuracy was similar between the effort and performance conditions (anecdotal evidence supporting the null hypothesis) (b = 0.02 [0.00, 0.04], d = 0.15, BF = 0.73; P = 0.015) and also similar between the effort and neutral conditions (b = −0.01 [−0.03, 0.00], d = −0.09, BF = 0.18; P = 0.151). These results suggest that, even though participants in the effort condition chose the hard task more frequently, there is limited evidence that their performance differed from participants in the other two conditions. On probe trials, there is limited evidence that performance differed between the effort and performance conditions (b = 0.03 [0.01, 0.05], d = 0.16, BF = 0.90; P = 0.010). There was strong evidence that performance was similar between the effort and control conditions (b = −0.00 [−0.02, 0.02], d = −0.01, BF = 0.07; P = 0.817). These results for the probe trials were driven largely by a decrease in task accuracy in the performance condition (Fig. 4). For the remaining four models (trials in the post-training near- and far-transfer blocks; Fig. 4), task accuracy did not differ between conditions (b values <0.2, d values <0.11, BFs <0.30, P values >0.073).

Individual differences. In all three conditions, participants' effort preferences for the inhibition task in the pre-training, training (reward and probe trials) and post-training sections correlated positively (Pearson correlation coefficient r values >0.35, P values <0.001; Fig. 5). For example, those who chose the hard task more frequently in the pre-training block also chose the hard task more frequently in training and post-training sections. We fitted eight models (corresponding to the eight registered hypotheses; Supplementary Fig. 5) to explore whether condition and pre-training baseline effort preference interacted to predict effort preference and found strong evidence for interaction for hypothesis 5. The effect of condition (effort versus performance) was stronger for participants with higher effort preferences at baseline on the post-training inhibition trials (b = 0.29 [0.13, 0.44], d = 1.23, BF = 42.65; P < 0.001). That is, participants in the effort (versus performance) condition who chose the hard task more frequently during pre-training also chose the hard task more frequently during the post-training inhibition block.

To further identify heterogeneous treatment effects, we fitted eight causal forests[71,72] (one for each hypothesis; Supplementary Fig. 6). This non-parametric algorithm allows for data-driven covariate selection and flexible modelling of interactions in high dimensions, which guard against spurious heterogeneity or interactions common in classical approaches[73]. Following previous work[74], we first trained (separately for each hypothesis) a pilot forest on seven covariates (age, pre-training effort preference, Need for Cognition[75], Distress Tolerance[76], Conscientiousness[77], Grit[78] and Implicit Theories of Intelligence[79]). We then trained a second forest on the top three covariates that saw a reasonable number of splits in the first forest, which allows the second forest to make more splits on the most important covariates. For each hypothesis, we performed a conservative omnibus test for the presence of

 https://doi.org/10.1038/s41562-024-01842-7

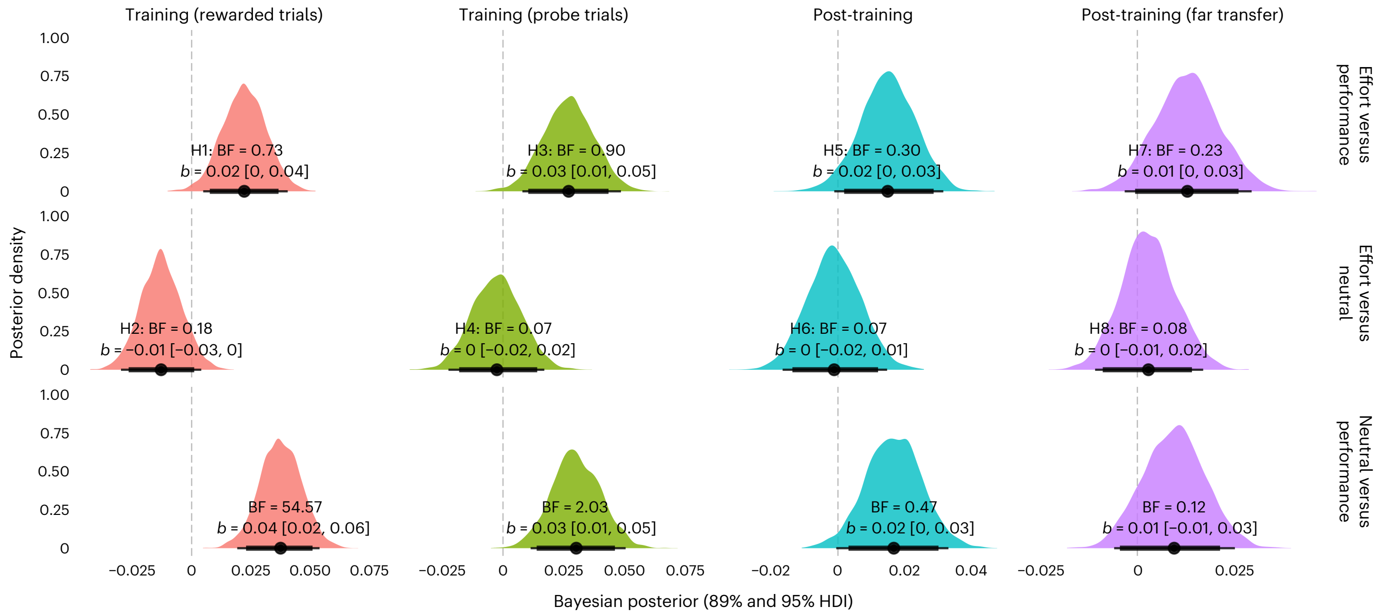

Fig. 4 | Exploratory analyses and Bayesian posterior densities for the effect of condition on task accuracy. Positive Bayesian posterior estimates indicate higher accuracy in the effort than the other condition. Model specification: accuracy ~ condition + pre-training baseline accuracy + objective task difficulty. Posterior means and HPD intervals (95% and 89%) are shown.

heterogeneity (coefficients significantly greater than zero indicate the presence of heterogeneity in the fitted forest).

For the reward trials (effort versus performance (H1) and effort versus neutral (H2)), the top three covariates were pre-training baseline, Need For Cognition and age, but there were heterogeneous effects only for hypothesis 2 (b = 1.13 [0.56], Student's t = 2.02, p = 0.022) but not hypothesis 1 (b = 0.62 [0.51], t = 1.21, P = 0.114). For the probe trials (H3 and H4), Need for Cognition, Grit and age were the top three covariates that moderated treatment effects and there were heterogeneous effects only for hypothesis 3 (b = 0.90 [0.40], t = 2.29, P = 0.011) but not hypothesis 4 (b = −0.57 [0.67], t = −0.85, P = 0.802). We also found heterogeneous effects for hypothesis 5 (b = 0.94 [0.41], t = 12.29, P = 0.011), such that higher pre-training baseline was associated with stronger treatment effects on post-training inhibition trials (Supplementary Fig. 6). However, we did not find significant heterogeneous effects for hypothesis 6 (b = −0.60 [0.81], t = −0.74, P = 0.769). Finally, for the post-training updating (far-transfer) trials (H7 and H8), there were significant heterogeneous effects for hypothesis 7 (top three covariates: grit, conscientiousness and distress tolerance; b = 0.75 [0.45], t = 1.64, P = 0.050) but not hypothesis 8 (b = −0.25 [0.75], t = −0.33, P = 0.631). Overall, there is some evidence suggesting that the treatment might be more effective for those who with lower levels of grit (for example, H3 and H7; see Supplementary Fig. 6 for details). These results suggest that the manipulation may be more effective for certain people and that future work should focus on the covariates highlighted in Supplementary Fig. 6.

Discussion

Together, the eight registered analyses provide limited evidence for the idea that people can learn to value effort[19,80,81]. Importantly, as preregistered, transfer effects are rarely reported in the literature, so the pattern of positive and null results observed here provides theoretically and practically important preliminary evidence for transfer effects.

The strong results for hypotheses 1 and 2 (BFs >100) indicate that participants successfully learned the task and reward contingencies in our paradigm. When rewarded for choosing hard over easy tasks, participants actively sought out the hard task and, importantly, did not sacrifice task performance.

Although the mass of the posterior distributions for hypotheses 3 and 4 was mostly positive, the Bayesian analyses found anecdotal evidence favouring the null hypothesis for hypotheses 3 (BF of 0.52) and 4 (BF of 0.83). That is, we did not find evidence for within-block near-transfer effects: Increases in effort preferences on rewarded trials did not reliably carry over or transfer to the probe trials. This finding is surprising, since we had expected the carryover effects to be stronger considering that the probe trials were randomly interleaved with rewarded trials. One possible explanation for this lack of positive effects is that participants in the effort condition—who had chosen to perform the hard task more frequently—might be strategically 'easing off' on the probe trials (for example, labour–leisure trade-off[82–89]). The exploratory finding that the treatment effects on probe trials were weaker for participants with higher Need for Cognition and Grit scores is consistent with the idea of strategic 'easing off'.

Relative to hypotheses 3 and 4, the evidence for hypothesis 5 (BF > 100) was very strong. However, for hypothesis 6 (BF = 0.92) the evidence anecdotally favoured the null, even though the posterior distribution was mostly positive (P = 0.032). This result suggests that rewarding effort increased effort preference, relative to rewarding performance (H5), in the post-training section. However, as there was no evidence that rewarding effort led to increased effort preference relative to the neutral condition (H6), we cannot definitely conclude that rewarding effort per se was sufficient to increase effort preference. As shown in Fig. 3, the mean effort preference difference between the post- and pre-training sections was close to 0 for the effort condition but negative in the neutral and performance conditions. That is, participants in the performance condition were 'depleted' or avoiding cognitive demand[14,90,91], whereas those in the effort condition may have been buffered against this negative effect.

Similarly, we also found moderate evidence supporting hypothesis 7 (BF = 6.19) but anecdotal evidence for hypothesis 8 (BF = 1.24). These two hypotheses focused on effort preferences on the updating task. Because participants did not train on this task, which also required a different set of cognitive processes than the trained inhibition task, this

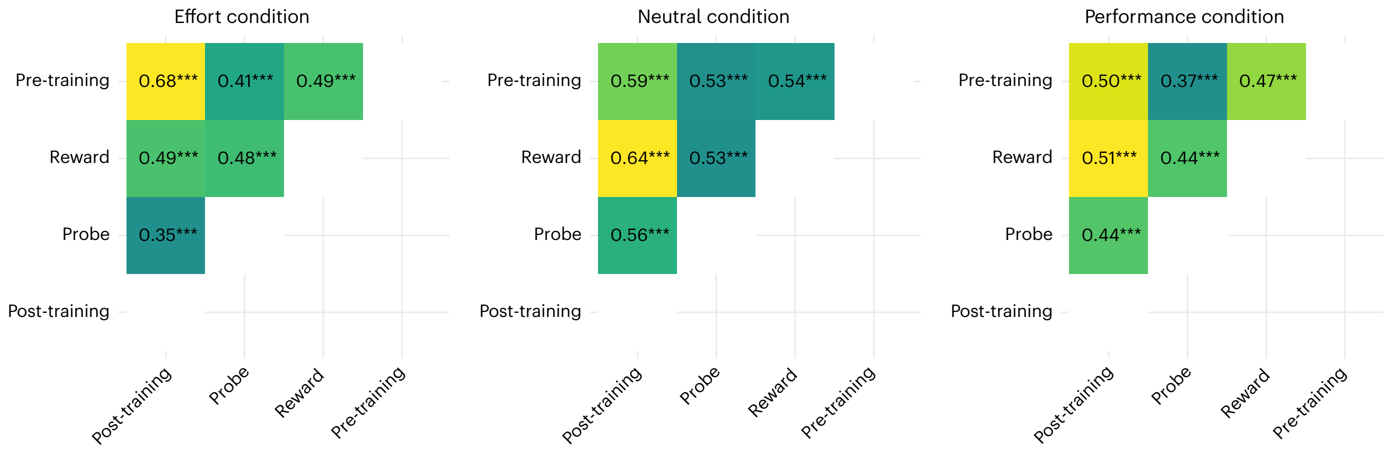

Fig. 5 | Effort preference correlations. Effort preferences for different trial types in different study sections correlated strongly with one another. ***$P < 0.001$.

task served as a far-transfer measure[92]. The results show that rewarding participants for choosing the hard inhibition task in the training section made them more willing to engage with a hard task that requires a different set of cognitive processes (see also refs. 93,94) compared with the performance condition (H7) but with weaker and more limited evidence for an effect compared with the neutral condition (H8). The pattern of results in Fig. 3 also suggests that participants in effort condition may have been buffered against a decline in willingness to choose and engage in the hard updating task.

As preregistered, hypotheses 6 (BF = 0.92) and 8 (BF = 1.24) test our claims most stringently (effort versus neutral condition contrast in the post-training section for the inhibition and updating tasks, respectively), and the evidence provides limited support for these two hypotheses. Nevertheless, as stated in our registered analytic plan, the pattern of results across all eight hypotheses—if we consider not just the BFs but also the posterior distributions, effect sizes and credible intervals, and frequentist P values—provide a mixed picture of evidence for possible near-transfer or far-transfer effects. We also found that, in the post-training sections (both inhibition and updating tasks), participants in the performance condition chose the hard task less frequently relative to their own pre-training choices and participants in the effort and neutral conditions.

The condition differences in effort preference across study sections suggest that participants in the performance and neutral conditions may have become less willing (relative to their own pre-training baseline) to choose the hard task (inhibition and updating tasks) over time[14]. Performance-contingent rewards may have reduced the intrinsic value of effort[57–59], especially when rewards were removed in the post-training sections, resulting in choice patterns resembling 'depletion'[90]. That is, even though performance-contingent rewards reliably affect task performance in many other studies[95–98], such rewards might also encourage effort-avoidant behaviour. Most importantly, participants in the effort condition mostly maintained their pre-training effort preferences in the post-training sections, suggesting that rewarding effort may be preferable if the goal is to reduce effort avoidance.

There is some evidence for heterogeneity in treatment effects. Results from exploratory linear interaction models and non-parametric causal forests[71,72] suggest that rewarding effort may be most effective for those who were already predisposed to choosing harder tasks before the experiment; note, however, that the treatment effects were still positive for those who were less predisposed. Future work is needed to identify factors that moderate treatment effects, which can help design tailored short- and long-term interventions in practical settings such as schools.

The main limitations of our study were the resource limitations (total registered $N = 750$) and the relatively brief manipulation (that is, the 40 rewarded trials in the training section lasted only about 7 min). Extending the present study to longitudinal and developmental studies in both quasi-experimental or field settings with stronger and repeated manipulations is an important direction for future work.

Methods

Participants

A total of 846 participants completed the study ($n_{\mathrm{effort}} = 282$, $n_{\mathrm{performance}} = 281$ and $n_{\mathrm{neutral}} = 283$). Following our registered sampling plan (BF > 10 for hypotheses 6 and 8, or a total of 750 participants), we excluded 85 participants (see 'Sampling plan' section for exclusion details). The final sample size was 761 ($n_{\mathrm{effort}} = 254$, $n_{\mathrm{performance}} = 255$ and $n_{\mathrm{neutral}} = 252$), 377 and 384 participants were female and male, respectively. The mean age was 35.97 years.

Ethics information

The research complied with the University of Toronto Research Ethics in Human Research Unit's regulations. All participants provided informed consent in accordance with the regulations of the ethics unit. Undergraduate participants at the university received course credits for completing the study, and participants recruited via online platforms (for example, Amazon Mechanical Turk, Prolific) were paid US$12.50 (only participants from Prolific were ultimately recruited). All participants also received a cash bonus (between US$1 and US$5), which was determined by the number of points earned.

Pilot data

Results from a pilot study using the dot-motion inhibition task demonstrated the feasibility of the proposed paradigm because the value functions (Fig. 1c) led to different effort preferences that are consistent with hypothesis 1. Participants in the effort condition ($n = 59$) had higher effort preferences than those in the performance condition ($n = 60$) ($b = 13.89$, 95% highest posterior density (HPD) of [5.62, 22.06], $d = 0.63$, BF = 29.79; $P < 0.001$). That is, participants who experienced the reward-effort value function (Fig. 1c, left) chose the hard version of the dot-motion inhibition task more frequently than those who experienced the reward-performance value function (Fig. 1c, right) (Supplementary Fig. 1).

Design

This experimental paradigm used a mixed (between–within subjects) design. All participants completed three sections in this order (within-subject): pre-training, training and post-training. The pre-training section presented participants with two unrelated cognitive tasks in two separate blocks to evaluate participants' baseline effort preferences

 https://doi.org/10.1038/s41562-024-01842-7

on these two tasks (Fig. 1a). Next, the training section consisted of a single block and was the only block in the experiment that delivered rewards (Fig. 1b). Participants were randomly assigned (between-subject assignment) to the effort, performance or neutral condition. Finally, the post-training section had two blocks that assessed whether the reward manipulation (Fig. 1c) during the previous training section affected effort preferences on the two tasks presented during pre-training (Fig. 1a). Data collection was performed blind to the conditions of the experiment. Data analysis, however, was not performed blind.

Pre-training section and post-training section. Both the pre-training the post-training sections each consisted of two blocks of demand selection tasks[14] (Fig. 1a). In each section, two cognitive tasks requiring primarily inhibition and updating abilities were presented in two separate blocks (order counterbalanced across participants). Each pre-training block had 40 trials. Behavioural indices (for example, choices, accuracy and reaction time) obtained from the two pre-training blocks provided baseline measures of effort preferences and performance on the two cognitive tasks. The post-training section (presented after the training section) was similar to the pre-training section, but only had 20 trials in each block. This block allowed us to measure how behaviour changes as a function of our experimental manipulation in the training section. On each trial in each block, participants pressed either the left or right key to select the cue shown on the left or right of the display (3,000 ms response deadline), which represented either the easy or hard version of the cognitive task. Participants then performed the task they selected, and no performance feedback was provided at the end of each trial (Fig. 1a). The mappings between cues and task difficulty were counterbalanced across participants, and the locations (left or right) of the cues were randomly determined on every trial. For each participant and each task, two cues were randomly chosen from a set of six cues to represent the easy and hard tasks (that is, different sets of six cues for the inhibition and updating tasks). The cues representing the inhibition and updating tasks were also visually distinct (Fig. 1a), ensuring that stimulus-driven carryover effects were minimal.

The inhibition task in one of the demand selection blocks was a Simon-like dot-motion conflict task[99,100]. After choosing a cue (reflecting either the easy or hard version of the inhibition task), participants performed three repetitions of the selected task (Fig. 1a). On each repetition, participants saw an array of coloured moving dots and had up to 1,500 ms to respond. Depending on their choice (easy or hard task), they had to press either the left or right key to indicate the dot motion direction (leftwards or rightwards; 300 dots with 75–100% motion coherence, sampled from a uniform distribution) or the colour of the dots (presented in one of four colours, with two colours mapped to each key). Another 20–50 (sampled from a uniform distribution) distractor dots in a different colour moved in a direction that was consistent with or opposite to the majority of the dots. The easy version required little controlled attention because participants simply had to indicate the motion direction of the majority of the dots while ignoring the colour of the dots. However, the hard version required controlled attention because participants had to indicate the colour of the dots while overriding their automatic tendency to indicate motion direction. Critically, the key for the correct colour response could be congruent with the dot motion direction or it could be incongruent with the dot motion direction (with a 65% chance that it was incongruent on each repetition). Colour response mappings were counterbalanced across participants.

The updating task in the other demand selection block was a working memory and attention control task[101]. After choosing one of the cues (reflecting either the easy or hard version of the updating task), participants performed the selected task (Fig. 1a). On each trial, participants added a digit to three serially presented digits. The easy version required participants to add 0 to each digit (for example, 7, 8 and 6) and to use the left, right, up or down key to choose the correct response (that is, 786) out of four similar responses (3,000 ms response deadline). The hard version required participants to add 3 or 4 to each digit (for example, add 4 to each digit, so 7 becomes 1, 8 becomes 2, 6 becomes 0, etc.). Two digits were used for the hard task to minimize practice effects, and the digit was selected randomly at the start of each trial.

Training section. In the training section, participants completed one block (60 trials) of the demand selection task with the inhibition task (but not the updating task; see explanation below). At the beginning of each trial, participants saw one of two cues—reward trial cue or probe trial cue (Fig. 1b)—explicitly signalling to participants the presence or absence of rewards on that trial, respectively. Next, participants chose to perform either the easy or hard version of the inhibition task (represented by the same cues seen in the pre-training section) and performed their chosen task. Finally, they saw either the reward or probe trial feedback, depending on whether it was a reward or probe trial, respectively. Rewarded ($n = 40$) and probe ($n = 20$) trials were pseudo-randomly interleaved in this block, such that the first trial was always a reward trial and probe trials occurred after every one to three reward trials.

To test our hypotheses, we randomly assigned participants to the effort, performance or neutral condition. The three conditions had identical task instructions, structure, sequence and cues (reward, probe and feedback). The only difference was how the points were delivered on rewarded trials (Fig. 1b, reward trial cue). On these trials, participants in the effort condition experienced the reward-effort value function (Fig. 1c, left). Those in the performance condition experienced the reward-performance function (Fig. 1c, middle). Those in the neutral condition received the same amount of rewards for choosing the easy or hard task and regardless of their performance (Fig. 1c, right).

Specifically, on rewarded trials (Fig. 1b, reward trial cue), participants in the effort condition received more points for choosing the hard relative to the easy task (for example, 370 versus 230 points, respectively, plus jitter drawn from a normal distribution, $N(\mu = 0, \sigma = 5)$), regardless of their reaction times (RTs). However, participants in the performance condition received points that scaled with their own RTs (plus jitter drawn from the same normal distribution as above), regardless of their choice (easy or hard version of the task). Participants in the neutral condition received the same points (plus jitter) regardless of their choice or performance. Critically, the expected reward across conditions was identical. Because participants performed three repetitions of the inhibition task after each choice (Fig. 1b), the feedback (Fig. 1b, reward trial feedback) received at the end of each rewarded trial was the mean of the points received on three repetitions of the inhibition task. On rewarded trials, 0 points were given for incorrect or missed responses (RT deadlines were participant specific; see next paragraph).

Participant-specific RT criteria were used to allocate points on rewarded trials, ensuring participants' RTs were evaluated against their own RT benchmarks on the inhibition task. These benchmarks were based on each participant's RTs across the easy and hard versions of the task. In addition, if participants responded incorrectly or too slowly, they received 0 points (that is, no 'empty praises'). This approach ensured incentive compatibility and that participants had to perform relatively well, even when RT performance was uncorrelated with points earned (Fig. 1c, left and right). That is, choosing the hard task and then slacking off was a bad strategy that probably resulted in no rewards.

The distribution of RTs used for benchmarking included RTs on only correct responses (that is, correct RTs) of the inhibition task in the pre-training section demand selection task (Fig. 1a). Very fast and slow correct RTs (±1 times the median absolute deviation) were also excluded[102]. RT deadlines were determined separately for each participant, and the easy and hard versions of the inhibition task had the same deadline (that is, maximum RT (after applying exclusion criteria above) plus 150 ms).

 https://doi.org/10.1038/s41562-024-01842-7

This RT benchmarking procedure helped to ensure each participant had neither too much nor too little time to respond on each trial. For example, two participants—one fast and one slow (with median RTs of 500 and 700 ms, respectively)—would have different RT deadlines (for example, 750 and 950 ms, respectively), but both would receive the same number of points on a trial if they responded correctly at their own median RTs (500 and 700 ms, respectively). Thus, task difficulty and rewards were tailored to each participant's ability, such that participants received, on average, the same number of points, regardless of individual differences in average RTs.

We chose to train participants on only the inhibition task but not the updating task, for two reasons. First, our paradigm required participants to feel efficacious on reward-performance trials (Fig. 1c, middle)—exerting more effort should lead to more accurate and faster responses, and the inhibition task allowed for a much tighter coupling between performance and rewards obtained than the updating task. Second, we are interested in whether the effects of training on a highly controlled inhibition task generalize to the updating task, which is more ecologically valid because solving mathematics problems resemble real-life problems more than indicating the motion or colour of dots.

Procedure. Participants were recruited to take part in a study titled 'What are your cognitive preferences and abilities?' and saw the following study description: 'We are examining how you make decisions and perform different cognitive tasks. You'll be doing cognitive tasks on the computer and answering a few questions about yourself after completing the tasks.'

At the beginning of the experiment, participants were told that they were adventurous space explorers who had to complete various missions in space. Before the pre-training section (Fig. 1a), participants practised and learned the four different cues associated with the easy and hard versions of the inhibition and updating tasks. Each practice block (maximum 80 trials per block) terminated when participants either completed all trials or performed well above chance levels in the preceding 20 trials (≥80% and 70% correct in the last 20 trials on the inhibition and updating tasks, respectively; see exclusion criteria in 'Sampling plan' section). During practice, participants were encouraged to be adventurous by trying different tasks instead of always choosing one task. Participants then completed the two actual blocks in the pre-training section (Fig. 1a).

Thereafter, they began the training section, where they first learned the cues associated with the reward and probe trials (Fig. 1b): 'If you see a landed spaceship, an alien will be delivering rewards. If the spaceship hasn't landed, you WON'T be receiving rewards.' They were told that in the upcoming mission, their 'goal is to earn as many points as possible' and that the earned points would be converted to a cash bonus at the end of the study. They also read the following instructions: 'How many points you could earn from the aliens depends on some combination of WHICH ROCKET YOU CHOOSE or/and HOW WELL YOU PERFORM (accuracy and reaction time). If you respond inaccurately or too slowly, you will receive 0 points for that response. So if the alien gave you very few points, it's likely because you made too many mistakes or/and were slow. Therefore, to maximize your earnings, try to use the feedback/points you receive from the aliens to improve HOW WELL YOU PERFORM and inform WHICH ROCKET TO CHOOSE in the future'. After completing ten practice trials (five rewarded trials, five probe trials) where they were encouraged to explore different strategies, they completed the actual training section (Fig. 1b).

Before beginning the post-training section, participants were briefly reminded of the four cues (from the pre-training section) associated with the easy and hard versions of the inhibition and updating tasks. To inform participants that no rewards would be delivered in this section, they were explicitly told that 'the aliens have retreated so they won't be around to deliver points or rewards' before they completed the two blocks in this section.

Finally, they completed demographic, personality and debriefing questionnaires: Need for Cognition[75], Distress Tolerance[76], Conscientiousness[77], Grit[78] and Implicit Theories of Intelligence[79]. They were then compensated for their participation and points earned during the training section.

Sampling plan

We used the BF design analysis approach[103] to determine the number of participants required to provide compelling evidence for hypotheses 6 and 8 (Table 1), which predicted that, relative to participants in the neutral condition, those in the effort condition would choose the hard over the easy version of the inhibition task (hypothesis 6) and updating task (hypothesis 8) more frequently in the post-training section. This criterion reflected our belief that these two hypotheses tested our ideas most directly.

We used the sequential design with maximum participant approach to recruit additional participants until either (1) the BF provided strong evidence for the null hypothesis (BF < 0.1) or alternative hypothesis (BF > 10) or (2) a total of 750 participants had been reached. That is, the BF for hypotheses 6 and 8 must each exceed one of the thresholds. If not, we stopped only after we had recruited 750 participants in total. We computed BFs using the BayesFactor package[104] for the R Environment for Statistical Computing[105] and used the default Jeffreys–Zellner–Siow prior[106]. We compared the full linear regression model (outcome ~ condition + baseline) against the null model (outcome ~ baseline) to determine whether the BF had exceeded either threshold.

Participants who failed to meet the criteria below were excluded from the sequential sampling procedure and all analyses. They were excluded if they performed poorly during pre-training practice blocks. Specifically, those who did not respond with at least 80% and 70% correct in the last 20 practice trials on the inhibition and updating tasks, respectively, were excluded from all analyses ($n = 26$). We also excluded participants who performed better (that is, faster median correct RT) on the hard relative to the easy version of either of the two cognitive tasks in the last 20 practice trials ($n = 61$). This criterion was necessary because our paradigm hinged on the fact that the hard tasks should require more effort (that is, as reflected in slower median correct RT) than the easy tasks, and thus, more rewards would be necessary to offset the costs associated with performing the hard tasks.

To further ensure that the data quality was high, we asked participants to indicate honestly the extent to which they had tried to follow all task instructions on a continuous slider scale with three equally spaced anchors (never, mostly and always). Participants who responded below the midpoint of the scale (mostly) were excluded from all analyses ($n = 1$). As a final check, participants were asked to recall the four tasks (easy and hard versions of the inhibition and updating tasks) and describe what the corresponding cues looked like. Participants who failed to describe the cues correctly were also excluded from all analyses ($n = 4$).

Analysis plan

We excluded trials with no responses (that is, because participants failed to respond in time) before performing any aggregation. To compute effort preferences on the different trial types in the different blocks, we calculated for each participant the proportion of choices whereby they chose the hard task. To test each hypothesis in Table 1, we fitted a Bayesian linear regression to test the effect of condition (effort versus performance or neutral condition) on effort preferences. To obtain the direct effect[107] of condition, we included effort preferences in the pre-training demand selection blocks as a covariate (that is, linear regression with one covariate and the condition regressor, that is, analysis of covariance). Specifically, effort preferences on the pre-training inhibition task was included as the covariate in hypotheses 1–6, whereas effort preferences on the pre-training updating task was the covariate in hypotheses 7 and 8.

We fitted the models with the R package brms[108] and report the following statistics that were calculated from the posterior samples (4,000 samples in total, drawn from 4 Markov chain Monte Carlo chains, 2000 iterations each and 1000 warm-up samples; all R-hat convergence diagnostic statistics were 1.00, indicating convergence): beta estimate (posterior mean), its Bayesian 95% HPD interval and Cohen's d effect size. We used default brms priors (that is, improper flat priors over the reals) because, with sufficient data, priors are unlikely to influence parameter estimates (but they always influence the BF). For each effect, we also report the BF, which was computed using the BayesFactor package and the default Jeffreys–Zellner–Siow prior[104,106]. BF = 1 indicates that the data do not favour either the experimental or null hypothesis. BFs between 3 and 10 provide moderate evidence for the experimental hypothesis, whereas BFs between 0.3 and 0.1 provide moderate evidence for the null hypothesis. BFs greater than 10 or smaller than 0.1 provide at least strong evidence for the experimental and null hypothesis, respectively[103]. We also report frequentist probability values.

Protocol registration

The Stage 1 protocol for this Registered Report was accepted in principle on 8 February 2021. The protocol, as accepted by the journal, can be found at https://doi.org/10.6084/m9.figshare.14230283.v1

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Data and materials are provided at the repository https://osf.io/9unj5 Source data are provided with this paper.

Code availability

Code is available at the repository https://osf.io/9unj5

References

1. Disabato, D. J., Goodman, F. R. & Kashdan, T. B. Is grit relevant to well-being and strengths? Evidence across the globe for separating perseverance of effort and consistency of interests. *J. Personal.* **87**, 194–211 (2019).
2. Shoda, Y., Mischel, W. & Peake, P. K. Predicting adolescent cognitive and self-regulatory competencies from preschool delay of gratification: identifying diagnostic conditions. *Dev. Psychol.* **26**, 978–986 (1990).
3. Watts, T. W., Duncan, G. J. & Quan, H. Revisiting the marshmallow test: a conceptual replication investigating links between early delay of gratification and later outcomes. *Psychol. Sci.* **29**, 1–19 (2018).
4. Credé, M., Tynan, M. C. & Harms, P. D. Much ado about grit: a meta-analytic synthesis of the grit literature. *J. Personal. Soc. Psychol.* **113**, 492–511 (2017).
5. de Ridder, D. T., Lensvelt-Mulders, G., Finkenauer, C., Stok, F. M. & Baumeister, R. F. Taking stock of self-control: a meta-analysis of how trait self-control relates to a wide range of behaviors. *Personal. Soc. Psychol. Rev.* **16**, 76–99 (2012).
6. Moffitt, T. E. et al. A gradient of childhood self-control predicts health, wealth, and public safety. *Proc. Natl Acad. Sci. USA* **108**, 2693–2698 (2011).
7. Inzlicht, M., Werner, K. M., Briskin, J. L. & Roberts, B. W. Integrating models of self-regulation. *Annu. Rev. Psychol.* **72**, 319–345 (2021).
8. Jackson, C. K. What do test scores miss? The importance of teacher effects on non–test score outcomes. *J. Political Econ.* **126**, 2072–2107 (2018).
9. Kwon, H. W. The sociology of grit: exploring grit as a sociological variable and its potential role in social stratification. *Sociol. Compass* **11**, e12544 (2017).
10. Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A. & Goldberg, L. R. The power of personality: the comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspect. Psychol. Sci.* **2**, 313–345 (2007).
11. Tough, P. How kids learn resilience. *Atlantic* **317**, 56–66 (2016).
12. Hull, C. L. *Principles of Behavior: an Introduction to Behavior Theory* (Appleton-Century-Crofts, 1943).
13. Kool, W. & Botvinick, M. Mental labour. *Nat. Hum. Behav.* **2**, 899–908 (2018).
14. Kool, W., McGuire, J. T., Rosen, Z. B. & Botvinick, M. M. Decision making and the avoidance of cognitive demand. *J. Exp. Psychol. Gen.* **139**, 665–682 (2010).
15. Kurzban, R. The sense of effort. *Curr. Opin. Psychol.* **7**, 67–70 (2016).
16. Shenhav, A. et al. Toward a rational and mechanistic account of mental effort. *Annu. Rev. Neurosci.* **40**, 99–124 (2017).
17. Westbrook, A., Kester, D. & Braver, T. S. What is the subjective cost of cognitive effort? Load, trait, and aging effects revealed by economic preference. *PLoS ONE* **8**, e68210 (2013).
18. Friedrich, A. M. & Zentall, T. R. Pigeons shift their preference toward locations of food that take more effort to obtain. *Behav. Process.* **67**, 405–415 (2004).
19. Inzlicht, M., Shenhav, A. & Olivola, C. Y. The effort paradox: effort is both costly and valued. *Trends Cogn. Sci.* **22**, 337–349 (2018).
20. Lydall, E. S., Gilmour, G. & Dwyer, D. M. Rats place greater value on rewards produced by high effort: an animal analogue of the 'effort justification' effect. *J. Exp. Soc. Psychol.* **46**, 1134–1137 (2010).
21. Norton, M. I., Mochon, D. & Ariely, D. The IKEA effect: when labor leads to love. *J. Consum. Psychol.* **22**, 453–460 (2012).
22. Dunn, T. L., Gaspar, C. & Risko, E. F. Cue awareness in avoiding effortful control. *Neuropsychologia* **123**, 77–91 (2019).
23. Dunn, T. L., Inzlicht, M. & Risko, E. F. Anticipating cognitive effort: roles of perceived error-likelihood and time demands. *Psychol. Res.* **83**, 1033–1056 (2017).
24. Constantinidis, C. & Klingberg, T. The neuroscience of working memory capacity and training. *Nat. Rev. Neurosci.* **17**, 1–12 (2016).
25. Enriquez-Geppert, S., Huster, R. J. & Herrmann, C. S. Boosting brain functions: improving executive functions with behavioral training, neurostimulation, and neurofeedback. *Int. J. Psychophysiol.* **88**, 1–16 (2013).
26. Jaeggi, S. M., Buschkuehl, M., Jonides, J. & Shah, P. Short- and long-term benefits of cognitive training. *Proc. Natl Acad. Sci. USA* **108**, 10081–10086 (2011).
27. Melby-Lervåg, M. & Hulme, C. Is working memory training effective? A meta-analytic review. *Dev. Psychol.* **49**, 270–291 (2013).
28. Melby-Lervag, M., Redick, T. S. & Hulme, C. Working memory training does not improve performance on measures of intelligence or other measures of 'far transfer': evidence from a meta-analytic review. *Perspect. Psychol. Sci.* **11**, 512–534 (2016).
29. Owen, A. M. et al. Putting brain training to the test. *Nature* **465**, 775–778 (2010).
30. Shipstead, Z., Redick, T. S. & Engle, R. W. Is working memory training effective? *Psychol. Bull.* **138**, 628–654 (2012).
31. Simons, D. J. et al. Do 'brain-training' programs work? *Psychol. Sci. Public Interest* **17**, 103–186 (2016).
32. von Bastian, C. C., Guye, S. & De Simoni, C. in *Cognitive and Working Memory Training: Perspectives from Psychology, Neuroscience, and Human Development* (eds Novick J. M. et al.) pp 1–23 (Oxford Univ. Press, 2019).
33. Miles, E. et al. Does self-control improve with practice? Evidence from a six-week training program. *J. Exp. Psychol. Gen.* **145**, 1075–1091 (2016).

34. Friese, M., Frankenbach, J., Job, V. & Loschelder, D. D. Does self-control training improve self-control? A meta-analysis. *Perspect. Psychol. Sci.* **12**, 1077–1099 (2017).
35. Lee, B. M. & Kemmelmeier, M. How reliable are the effects of self-control training? A re-examination using self-report and physical measures. *PLoS ONE* **12**, e0178814 (2017).
36. Alan, S., Boneva, T. & Ertac, S. Ever failed, try again, succeed better: results from a randomized educational intervention on grit. *Q. J. Econ.* **134**, 1–130 (2019).
37. Lazowski, R. A. & Hulleman, C. S. Motivation interventions in education: a meta-analytic review. *Rev. Educ. Res.* **86**, 602–640 (2016).
38. Yeager, D. S. et al. Using design thinking to improve psychological interventions: the case of the growth mindset during the transition to high school. *J. Educ. Psychol.* **108**, 374–391 (2016).
39. Yeager, D. S. et al. Teaching a lay theory before college narrows achievement gaps at scale. *Proc. Natl Acad. Sci. USA* **113**, E3341–E3348 (2016).
40. Foliano, F., Rolfe, H., Buzzeo, J., Runge, J. & Wilkinson, D. *Changing Mindsets: Effectiveness Trial* (National Institute of Economic and Social Research, 2019).
41. Li, Y. & Bates, T. C. You can't change your basic ability, but you work at things, and that's how we get hard things done: Testing the role of growth mindset on response to setbacks, educational attainment, and cognitive ability. *J. Exp. Psychol. Gen.* **148**, 1640–1655 (2019).
42. Sisk, V. F., Burgoyne, A. P., Sun, J., Butler, J. L. & Macnamara, B. N. To what extent and under which circumstances are growth mind-sets important to academic achievement? Two meta-analyses. *Psychol. Sci.* **29**, 549–571 (2018).
43. Yeager, D. S. et al. A national experiment reveals where a growth mindset improves achievement. *Nature* **573**, 364–369 (2019).
44. Crespi, L. P. Quantitative variation of incentive and performance in the white rat. *Am. J. Psychol.* **55**, 467–517 (1942).
45. Eisenberger, R. Learned industriousness. *Psychol. Rev.* **99**, 248–267 (1992).
46. O'Doherty, J. P., Cockburn, J. & Pauli, W. M. Learning, reward, and decision making. *Annu. Rev. Psychol.* **68**, 73–100 (2017).
47. Leonard, J. A., Lee, Y. & Schulz, L. E. Infants make more attempts to achieve a goal when they see adults persist. *Science* **357**, 1290–1294 (2017).
48. Gunderson, E. A. et al. Parent praise to 1- to 3-year-olds predicts children's motivational frameworks 5 years later. *Child Dev.* **84**, 1526–1541 (2013).
49. Lucca, K., Horton, R. & Sommerville, J. A. Infants rationally decide when and how to deploy effort. *Nat. Hum. Behav.* **4**, 372–379 (2020).
50. Boksem, M. A. & Tops, M. Mental fatigue: costs and benefits. *Brain Res. Rev.* **59**, 125–139 (2008).
51. Botvinick, M. & Braver, T. Motivation and cognitive control: from behavior to neural mechanism. *Annu. Rev. Psychol.* **66**, 83–113 (2015).
52. Kurniawan, I. T., Guitart-Masip, M., Dayan, P. & Dolan, R. J. Effort and valuation in the brain: the effects of anticipation and execution. *J. Neurosci.* **33**, 6160–6169 (2013).
53. Massar, S. A., Lim, J., Sasmita, K. S. & Chee, M. W. Rewards boost sustained attention through higher effort: a value-based decision making approach. *Biol. Psychol.* **120**, 21–27 (2016).
54. Padmala, S. & Pessoa, L. Reward reduces conflict by enhancing attentional control and biasing visual cortical processing. *J. Cogn. Neurosci.* **23**, 3419–3432 (2011).
55. Umemoto, A. & Holroyd, C. B. Task-specific effects of reward on task switching. *Psychol. Res.* **79**, 698–707 (2015).
56. Varazzani, C., San-Galli, A., Gilardeau, S. & Bouret, S. Noradrenaline and dopamine neurons in the reward/effort trade-off: a direct electrophysiological comparison in behaving monkeys. *J. Neurosci.* **35**, 7866–7877 (2015).
57. Deci, E. L., Koestner, R. & Ryan, R. M. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychol. Bull.* **125**, 627–668 (1999).
58. Kohn, A. *Punished by Rewards: the Trouble with Gold Stars, Incentive Plans, A's, Praise, and Other Bribes* (Houghton Mifflin Harcourt, 1999).
59. Tang, S. & Hall, V. C. The overjustification effect: a meta-analysis. *Appl. Cogn. Psychol.* **9**, 365–404 (1995).
60. Henderlong Corpus, J. & Lepper, M. R. The effects of person versus performance praise on children's motivation: gender and age as moderating factors. *Educ. Psychol.* **27**, 487–508 (2007).
61. Mueller, C. M. & Dweck, C. S. Praise for intelligence can undermine children's motivation and performance. *J. Personal. Soc. Psychol.* **75**, 33–52 (1998).
62. Amsel, A. Frustrative nonreward in partial reinforcement and discrimination learning: some recent history and a theoretical extension. *Psychol. Rev.* **69**, 306–328 (1962).
63. McCuller, T., Wong, P. T. P. & Amsel, A. Transfer of persistence from fixed-ratio barpress training to runway extinction. *Anim. Learn. Behav.* **4**, 53–57 (1976).
64. Eisenberger, R., Carlson, J. & Frank, M. Transfer of persistence to the acquisition of a new behaviour. *Q. J. Exp. Psychol.* **31**, 691–700 (1979).
65. Eisenberger, R., Carlson, J., Guile, M. & Shapiro, N. Transfer of effort across behaviors. *Learn. Motiv.* **10**, 178–197 (1979).
66. Eisenberger, R., Mitchell, M. & Masterson, F. A. Effort training increases generalized self-control. *J. Personal. Soc. Psychol.* **49**, 1294–1301 (1985).
67. Göllner, R. et al. Is doing your homework associated with becoming more conscientious. *J. Res. Personal.* **71**, 1–12 (2017).
68. Claro, S., Paunesku, D. & Dweck, C. S. Growth mindset tempers the effects of poverty on academic achievement. *Proc. Natl Acad. Sci. USA* **113**, 8664–8668 (2016).
69. Tough, P. *How Children Succeed: Grit, Curiosity, and the Hidden Power of Character* (Houghton Mifflin Harcourt, 2012).
70. Lin, H., Werner, K. M. & Inzlicht, M. Promises and perils of experimentation: the mutual-internal-validity problem. *Perspect. Psychol. Sci.* **16**, 854–863 (2021).
71. Athey, S., Tibshirani, J. & Wager, S. Generalized random forests. *Ann. Stat.* **47**, 1148–1178 (2019).
72. Athey, S. & Wager, S. Estimating treatment effects with causal forests: an application. *Obs. Stud.* **5**, 37–51 (2019).
73. Wager, S. & Athey, S. Estimation and inference of heterogeneous treatment effects using random forests. *J. Am. Stat. Assoc.* **113**, 1228–1242 (2018).
74. Basu, S., Kumbier, K., Brown, J. B. & Yu, B. Iterative random forests to discover predictive and stable high-order interactions. *Proc. Natl Acad. Sci. USA* **115**, 1943–1948 (2018).
75. Cacioppo, J. T. & Petty, R. E. The need for cognition. *J. Personal. Soc. Psychol.* **42**, 116 (1982).
76. Simons, J. S. & Gaher, R. M. The Distress Tolerance Scale: development and validation of a self-report measure. *Motiv. Emot.* **29**, 83–102 (2005).
77. John, O. P. & Srivastava, S. in *Handbook of Personality: Theory and Research* (eds Pervin, L. A. & John, O. P.) pp 102–138 (Guildford, 1999).
78. Duckworth, A. L. & Quinn, P. D. Development and validation of the Short Grit Scale (GRIT–S). *J. Personal. Assess.* **91**, 166–174 (2009).
79. De Castella, K. & Byrne, D. My intelligence may be more malleable than yours: the revised implicit theories of intelligence (self-theory) scale is a better predictor of achievement, motivation, and student disengagement. *Eur. J. Psychol. Educ.* **30**, 245–267 (2015).

80. Campbell, A. V., Chung, J. M. H. & Inzlicht, M. Meaningfulness of effort: deriving purpose from really trying. Preprint at *PsyArxiv* https://doi.org/10.31234/osf.io/sg3aw (2022).
81. Clay, G., Mlynski, C., Korb, F. M., Goschke, T. & Job, V. Rewarding cognitive effort increases the intrinsic value of mental labor. *Proc. Natl Acad. Sci. USA* **119**, e2111785119 (2022).
82. Dora, J., van Hooff, M. L. M., Geurts, S. A. E., Kompier, M. A. J. & Bijleveld, E. The effect of opportunity costs on mental fatigue in labor/leisure trade-offs. *J. Exp. Psychol. Gen.* **151**, 695–710 (2022).
83. Kurzban, R., Duckworth, A., Kable, J. W. & Myers, J. An opportunity cost model of subjective effort and task performance. *Behav. Brain Sci.* **36**, 661–679 (2013).
84. Ritz, H. & Shenhav, A. Humans reconfigure target and distractor processing to address distinct task demands. *Psychol. Rev.* https://doi.org/10.1037/rev0000442 (2023).
85. Ritz, H., Leng, X. & Shenhav, A. Cognitive control as a multivariate optimization problem. *J. Cogn. Neurosci.* **34**, 569–591 (2022).
86. Shenhav, A., Prater Fahey, M. & Grahek, I. Decomposing the motivation to exert mental effort. *Curr. Dir. Psychol. Sci.* **30**, 307–314 (2021).
87. Westbrook, A. et al. Dopamine promotes cognitive effort by biasing the benefits versus costs of cognitive work. *Science* **367**, 1362–1366 (2020).
88. Westbrook, A., Frank, M. J. & Cools, R. A mosaic of cost-benefit control over cortico-striatal circuitry. *Trends Cogn. Sci.* **25**, 710–721 (2021).
89. Westbrook, A. et al. Economic choice and heart rate fractal scaling indicate that cognitive effort is reduced by depression and boosted by sad mood. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **8**, 687–694 (2022).
90. Lin, H., Saunders, B., Friese, M., Evans, N. J. & Inzlicht, M. Strong effort manipulations reduce response caution: a preregistered reinvention of the ego-depletion paradigm. *Psychol. Sci.* **31**, 531–547 (2020).
91. Wu, R., Ferguson, A. M. & Inzlicht, M. Do humans prefer cognitive effort over doing nothing? *J. Exp. Psychol. Gen*. **152**, 1069–1079 (2023).
92. Miyake, A. & Friedman, N. P. The nature and organization of individual differences in executive functions: four general conclusions. *Curr. Dir. Psychol. Sci.* **21**, 8–14 (2012).
93. Judd, N. & Klingberg, T. Training spatial cognition enhances mathematical learning in a randomized study of 17,000 children. *Nat. Hum. Behav.* **5**, 1548–1554 (2021).
94. Scherer, R., Siddiq, F. & Sánchez Viveros, B. The cognitive benefits of learning computer programming: a meta-analysis of transfer effects. *J. Educ. Psychol.* **111**, 764–792 (2019).
95. Frömer, R., Lin, H., Wolf, C. K. D., Inzlicht, M. & Shenhav, A. Expectations of reward and efficacy guide cognitive control allocation. *Nat. Commun.* **12**1030 (2021).
96. Lin, H., Ristic, J., Inzlicht, M. & Otto, A. R. The average reward rate modulates behavioral and neural indices of effortful control allocation. *J. Cogn. Neurosci.* **34**, 2113–2126 (2022).
97. Inzlicht, M. & Campbell, A. V. Effort feels meaningful. *J. Exp. Psychol. Gen.* **26**, 1035–1037 (2022).
98. Dweck, C. S. & Yeager, D. S. Mindsets: a view from two eras. *Perspect. Psychol. Sci.* **14**, 481–496 (2019).
99. Danielmeier, C., Eichele, T., Forstmann, B. U., Tittgemeyer, M. & Ullsperger, M. Posterior medial frontal cortex activity predicts post-error adaptations in task-related visual and motor areas. *J. Neurosci.* **31**, 1780–1789 (2011).
100. Ritz, H. & Shenhav, A. Parametric control of distractor-oriented attention. In *Proc. 41st Annual Meeting of the Cognitive Science Society* (eds Dale, R. & Bender, A.) 967–973 (Wiley-Blackwell, 2019).
101. Kahneman, D., Tursky, B., Shapiro, D. & Crider, A. Pupillary, heart rate, and skin resistance changes during a mental task. *J. Exp. Psychol.* **79**, 164–167 (1969).
102. Leys, C., Ley, C., Klein, O., Bernard, P. & Licata, L. Detecting outliers: do not use standard deviation around the mean, use absolute deviation around the median. *J. Exp. Soc. Psychol.* **49**, 764–766 (2013).
103. Schönbrodt, F. D. & Wagenmakers, E. J. Bayes factor design analysis: planning for compelling evidence. *Psychon. Bull. Rev.* **25**, 128–142 (2018).
104. Morey, R. D. & Rouder, J. N. BayesFactor: computation of Bayes factors for common designs. R package version 0.9.12-4.2 (R Foundation for Statistical Computing, 2018).
105. R Core Team. R: a language and environment for statistical computing (R Foundation for Statistical Computing, 2019).
106. Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D. & Iverson, G. Bayesian *t* tests for accepting and rejecting the null hypothesis. *Psychon. Bull. Rev.* **16**, 225–237 (2009).
107. Pearl, J. Lord's paradox revisited – (Oh Lord! Kumbaya!). *J. Causal Inference* **4**, 1–13 (2016).
108. Bürkner, P. C. brms: an R package for Bayesian multilevel models using Stan. *J. Stat. Softw.* **80**, 1–28 (2017).

Acknowledgements

We acknowledge H. Ritz for providing feedback on the task design and N. Lin for designing the task stimuli. This work was supported by a grant from the Natural Sciences and Engineering Research Council of Canada (RGPIN-2019-05280) awarded to M.I. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

H.L., A.W. and M.I. contributed to the conception and design of the work. A.W. and M.I. provided critical oversight of and feedback on the work. H.L. and F.F. programmed the experiment and collected the data. H.L. wrote the manuscript. A.W. and M.I. provided critical feedback.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41562-024-01842-7.

Correspondence and requests for materials should be addressed to Hause Lin.

Peer review information *Nature Human Behaviour* thanks Masud Husain, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

nature portfolio

Corresponding author(s): Hause Lin

Last updated by author(s): May 30, 2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☒	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	We programmed the task using the JavaScript library jsPsych 6.2.0 and collected data using the web platform cognition.run.
Data analysis	We analyzed the data using R 4.2.2 (2022-10-31).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All data and code can be found on our Open Science Framework repository: https://osf.io/9unj5

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	Sex and gender were not considered in the design because we did not have a priori or registered analyses related to sex and gender. However, gender information was collected/determined based on self-reporting (377 female, 384 male). This information is in the publicly available data and participants provided consent for sharing individual-level data.
Reporting on race, ethnicity, or other socially relevant groupings	We did not use socially constructed or socially relevant categorization variables in our manuscript. Our study was an fully randomized experimental study, and we conducted exploratory causal forests analyses using self-reported covariates/confounding variables (e.g., conscientiousness, grit, distress tolerance, age, need for cognition) to examine heterogeneous treatment effects.
Population characteristics	See above.
Recruitment	Participants were recruited through Prolific.
Ethics oversight	University of Toronto

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	The study is an experiment that collected quantitative data.
Research sample	Online recruitment platform Prolific. Prolific participants tend to be more attentive than participants from other recruitment platforms, and attentiveness is important for our experiment. The mean age was 36 (SD = 13); 377 were female and 384 were male. The sample was not a nationally representative sample.
Sampling strategy	Our sampling strategy was pre-established and registered: We indicated we would stop data collection when the Bayes factor was > 10 or < .01 or when the sample size reached 750.
Data collection	Participants were recruited online via Prolific and they completed the experiment wherever they were, on their computers. The researcher was blind to experimental condition and the study hypothesis during data collection.
Timing	Data collection happened between April 2022 and July 2022.
Data exclusions	We excluded 85 participants based on the pre-established/registered exclusion criteria.
Non-participation	No participant dropped out/declined participation.
Randomization	Participants were randomly assigned to one of three experimental groups.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study	
☒	☐	Antibodies
☒	☐	Eukaryotic cell lines
☒	☐	Palaeontology and archaeology
☒	☐	Animals and other organisms
☒	☐	Clinical data
☒	☐	Dual use research of concern
☒	☐	Plants

Methods

n/a	Involved in the study	
☒	☐	ChIP-seq
☒	☐	Flow cytometry
☒	☐	MRI-based neuroimaging