AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: General

ISSN: 0096-3445

2025, Vol. 154, No. 6, 1628–1642
https://doi.org/10.1037/xge0001738

Effort Can Have Positive, Negative, and Nonmonotonic Impacts on Outcome Value in Economic Choice

Przemysław Marcowski[1, 2], Wojciech Białaszek[2], and Piotr Winkielman[2, 3]
[1] Swartz Center for Computational Neuroscience, Institute for Neural Computation, University of California, San Diego
[2] Institute of Psychology, SWPS University
[3] Department of Psychology, University of California, San Diego

Every action demands some effort, and its level influences decision making. Existing data suggest that in some decision contexts, effort devalues outcomes, but in other contexts, effort enhances outcome valuation. Here, we describe an empirical study and propose a model that incorporates negative, positive, and mixed impacts of effort on outcomes in different decision contexts and different participants. Participants chose between money and an item associated with varying levels of stair-climbing effort. Some participants had previous direct experience with a real physical effort and made decisions about a physically present reward. For other participants, the effort and the associated reward were always purely hypothetical. Furthermore, the decisions were framed as prospective or retrospective—before or after effort exertion. The key behavioral finding was that in the "real" condition, greater effort decreased outcome value when considered prospectively, but increased outcome value when considered retrospectively. Interestingly, even within the same decision context, individuals showed diverse relationships between effort and outcome value. These relationships ranged from those where greater effort increased value and decreased value to nonlinear patterns, where small effort initially increased outcome value but higher effort decreased it, or the other way around (initial decrease followed by a decrease). When our model was applied to participants' individual choices, it was able to capture the monotonic and nonmonotonic relationships and outperformed previous solutions.

Public Significance Statement

This study employs behavioral and modeling methods to demonstrate that effort can both decrease and increase the perceived value of effort-contingent outcomes. We show that these effects depend on the timing of effort and the tangibility of effort and rewards. Specifically, we found that outcomes are perceived as less valuable when greater effort is anticipated and more valuable when effort is already spent. These opposite effects of effort on valuation occur particularly in contexts involving tangible real effort and real rewards. Importantly, our novel computational model explains individual decisions better compared to previous models. One reason is that we model heterogeneity in individual responses to effort. For some people, there is a monotonic positive relationship between effort and value; for others, the relation is inverse, and some display shifting preferences. These findings emphasize the varied ways effort influences our choices, holding implications for fields where effort is a key variable influencing choice.

Keywords: effort, value, decision making, effort–reward trade-off

Supplemental materials: https://doi.org/10.1037/xge0001738.supp

This article was published Online First March 31, 2025.
Ross Otto served as action editor.
Przemysław Marcowski https://orcid.org/0000-0001-7342-6561
Wojciech Białaszek https://orcid.org/0000-0002-4672-4376
Piotr Winkielman https://orcid.org/0000-0003-2330-1802
The data that support the findings of this study and the code used for analysis are available on the Open Science Framework repository at https://osf.io/e3baz/. This study was not preregistered. This study was performed as part of the doctoral dissertation research of the corresponding author, Przemysław Marcowski. The authors have no known conflicts of interest to disclose.

This work was supported by the National Science Centre, Poland (https://www.ncn.gov.pl/), Grant 2018/28/T/HS6/00210 awarded to Przemysław Marcowski, by the SWPS University (https://english.swps.pl/) Grant ISDW/2017/DP/4 awarded to Przemysław Marcowski, and by the University of California, San Diego, Academic Senate Grant A1005308 awarded to Piotr Winkielman. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the article. The authors thank the members of the DecisionLab (SWPS University) and Winkielman Lab (University of California, San Diego), as well as Shlomi Sher (Pomona College) and Tim Brady (University of California, San Diego), for valuable discussions. This research was funded in whole, or in part, by the National Science Centre, Poland (NCN; Grant 2018/28/T/HS6/00210). For the purpose of open access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission.

Przemysław Marcowski: conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, writing–original draft. Wojciech Białaszek: conceptualization, methodology, supervision, writing–review and editing. Piotr Winkielman: conceptualization, writing–review and editing, interpretation, supervision.

continued

Choosing one action over another reflects the respective values of their outcomes. Because virtually every action involves some effort, the role of effort in choice is gaining increased attention from researchers in psychology, economics, biology, and decision neuroscience (Białaszek et al., 2017; Inzlicht et al., 2018; Massar et al., 2015; Otto et al., 2022; Westbrook & Braver, 2013, 2015). Effort is understood as motivation-driven intensification of activity (cognitive or physical) in service of, or associated with, obtaining an outcome (Białaszek et al., 2022; Eisenberger, 1992; Shenhav et al., 2017).

Previous research has shown that effort can have diverse impacts on outcome value in economic choice, as evidenced by various shapes of the function linking effort and value (e.g., linear, parabolic, hyperbolic) depending on the decision context and individual differences (Kool & Botvinick, 2018). Here we present new empirical data from a behavioral paradigm designed to elicit a variety of individual effort–value relationships in different decision scenarios. Central to the article is a new model aimed at capturing the diversity of individual effort profiles. Our model proposes that the evaluation of outcomes is a mixture of effort costs (negative effects) and effort benefits (positive effects), with the resulting net effect on choice reflecting situational and individual differences in the relative weight of those specific components. Before introducing the behavioral and modeling components of this research, it is worth considering reasons for why we may observe a diversity of effort–value relationships across different situations and different individuals.

Diverse Effects of Effort on Value

The classic "law of less work" (Hull, 1943) states that, when given a choice between two equivalent options differing only in the amount of effort they require, individuals opt for the less effortful option because the more effortful option is devalued by effort cost. Accordingly, if avoidable, greater effort will always reduce the choice of the option. Much evidence highlights that effort can be aversive or costly (Botvinick et al., 2009; Cavanagh & Frank, 2014; Dreisbach & Fischer, 2012; Elkins-Brown et al., 2016; Hartmann et al., 2013; Holroyd & Yeung, 2012; Inzlicht et al., 2015; Massar et al., 2015; McGuire & Botvinick, 2010; Nishiyama, 2014, 2016; Saunders et al., 2015; Sayalı & Badre, 2019; Shenhav et al., 2013, 2017; Winkielman et al., 2003).

However, in some circumstances, individuals prefer outcomes associated with higher effort (Alessandri et al., 2008; Schrift et al., 2016). For example, individuals value consumer products that they effortfully assembled themselves over their preassembled counterparts—aptly named the IKEA effect (Norton et al., 2012). One mechanism behind this effect reflects the sense that the exerted effort added value (subjective or objective) to the product. However, even without effort adding any objective value, the effort justification hypothesis, derived from cognitive dissonance theory, states that people value outcomes more when they were obtained as a result of exerting effort (Inzlicht et al., 2018). Similar arguments have been put forward to explain the sunk cost effect, where people are reluctant to abandon a course of action because of the previous effort that went into it (Arkes & Blumer, 1985). Note that in all these cases, value is assigned to outcomes retrospectively (after effort exertion). But there are also cases where positive value is assigned to more effortful outcomes prospectively (before effort exertion). For example, in the "martyrdom effect," the prospect of greater discomfort or greater effort increases donations for charity (Olivola & Shafir, 2018).

The idea that effort can be both costly and valued has been called the "effort paradox," suggesting the need for researchers to better understand the diverse effects of effort on outcome evaluation (Inzlicht et al., 2018). Indeed, as we discuss shortly, the diverse negative and positive effects of effort on value have been observed across different experimental paradigms and even within the same paradigm. Next, we consider in more detail how specific situational features change the effort–value relation and then discuss reasons for differences in individual effort profiles.

As already indicated, one situational feature that influences the effort–value relation is temporal orientation—prospective versus retrospective, that is, whether effort is anticipated or it is already spent (Inzlicht et al., 2018). Note that in the prospective situation, future effort can usually be avoided by choice. Accordingly, one should expect a generally negative effect of greater effort level on value of associated outcomes (law-of-less-work hypothesis). However, in the retrospective situation, effort is not avoidable because it is in the past. This invites potential inferences about effort adding to outcome value (e.g., "my exertion made this product better") or about effort justification (e.g., "I worked hard for this product so I must really like it"). After exerting effort, individuals may also be more likely to wonder about environmental correlations between different degrees of effort and value (e.g., "I guess products requiring more effort are typically better"). All these inferences would lead to a generally positive relation of effort level in the retrospective condition (effort-as-value hypothesis).

Another important feature shaping the relation between effort and value is whether a person has direct, personal experience with the type of effort required in the situation, as well as the outcome (reward) to which the effort is directed. Anecdotally, when people have little experiential knowledge of potential effort and potential reward, they sometimes eagerly commit to, or unwisely avoid, "effortful" actions. Empirically, there are demonstrations of some similarities between physiological responses to hypothetically considering (merely imagining) versus actually exerting tangible effort (Decety et al., 1991; Pellicano et al., 2021). However, it is also known that participants considering hypothetical versus

Przemysław Marcowski played a lead role in conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, and writing–original draft and an equal role in writing–review and editing. Wojciech Białaszek played a supporting role in conceptualization, methodology, and supervision and an equal role in writing–original draft and writing–review and editing. Piotr Winkielman played a supporting role in conceptualization and supervision and an equal role in writing–original draft and writing–review and editing.

Correspondence concerning this article should be addressed to Przemysław Marcowski, Swartz Center for Computational Neuroscience, Institute for Neural Computation, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, United States or Institute of Psychology, SWPS University, Chodakowska 19/31, 03-815 Warsaw, Poland. Email: p.marcowski@gmail.com

tangible effort focus on different task aspects (e.g., physical costs, mental costs, time costs; see Decety et al., 1989). There is also a large body of literature on hypothetical versus real rewards (see, e.g., Camerer & Mobbs, 2017; Dixon et al., 2013; Johnson & Bickel, 2002; Kang et al., 2011; Kühberger et al., 2002). Though the patterns reported in these works are complex (presumably reflecting multiple inferences and reactions), they suggest that the effort–value relation may vary depending on whether individuals have an actual experience (personal knowledge) of effort and associated rewards.

Crucially, a comprehensive model of effort-based choice should reflect the fact that the effort–value relation may not only depend on temporal orientation (prospective vs. retrospective) or the tangibility (real vs. hypothetical) of effort and outcome but also on an individual's preference and aptitude for engaging in various levels of effort (Gheza et al., 2023; Inzlicht et al., 2018; Westbrook & Braver, 2013, 2015). This can be illustrated by considering examples of some effort scenarios. For example, for one individual, greater effort might reduce the value of an item and make them choose a no-effort alternative. In contrast, for a different individual, greater effort may actually increase the value of the item. As such, the same situation can create opposite effort effects in different individuals. Even within the same individual, the impact on value may be a nonlinear function of effort intensity (Białaszek et al., 2022; Klein-Flügge et al., 2015). For example, for some individuals, low levels of effort may initially add to outcome value, but if the effort exceeds some threshold, it could become aversive and then subtract from outcome value. Such a scenario would result in a nonmonotonic value function, which is at first increasing and then shifts downward. The opposite is also possible. For example, for some individuals, low effort may be initially aversive (e.g., for a proclimber, an easy route may subtract value from the outcome of reaching the peak), but high effort may add value (e.g., for a proclimber, a difficult route adds value). This can be quantitatively captured by a function that is first decreasing, and changes to an increasing one. Of course, the two initially described simpler scenarios are also possible. That is, a prediction in line with the law of less work (monotonically decreasing function) and a prediction in line with the effort-as-value hypothesis (monotonically increasing function). Therefore, we propose that a more flexible functional model might better explain effort-based choice by differentiating among various effort preference profiles.

Before we introduce the details of the model, note that in a broader sense, our model reflects a classic insight that value functions can be nonmonotonic. There are many familiar examples in economics where marginal utility varies in sign (positive vs. negative) depending on the consumption level. That is, the utility (e.g., pleasure resulting from consumption) of one additional unit of a good (e.g., slice of cake) can be positive, zero, or negative, depending on the current level (e.g., one more slice yields positive utility when hungry but negative utility when satiated).

Note also that, in some cases, a nonmonotonic utility function may result from the integration of separate monotonic cost and benefit functions. For example, an extra hour of work may have positive but diminishing marginal utility, while opportunity costs (such as time spent away from family) may be a roughly linear function of time spent working. The net result of these monotonic cost and benefit functions may be a nonmonotonic utility function with an optimal allocation of time to work at an intermediate level. This example motivates our formal model of effort-based choice, as described next.

Dual-System Power Model of Effort-Based Choice

Based on previous empirical findings and the considerations presented above, here we propose a novel model of effort-based choice. Our model reflects previous research showing that the value function for effortful outcomes might not necessarily solely decay or rise but can also be nonmonotonic, which includes direction shifts (Białaszek et al., 2022; Klein-Flügge et al., 2015).

We built our model to account for situations where preference toward more or less effortful outcomes may shift as a function of situational and individual differences. A key feature of our model is that it incorporates dual (positive and negative) evaluations of effortful outcomes. That is, in our model, the integrated value of an effortful outcome results from evaluations by two subsystems, with each evaluation following a weighted power function. These subsystems allow for simultaneous evaluation of negative and positive aspects of effort.

In assuming the two subsystems, we are following a rich theoretical and empirical tradition in psychology. There are many well-known examples of dual models in the general domain of reasoning and social cognition, though the exact notion of a "system," "process," or "source" often varies (Evans, 2008; Kahneman, 2003; Smith & DeCoster, 2000). In closely related research on temporal (delay) discounting, previous work proposed two systems, one corresponding to faster and one to slow mechanisms (McClure et al., 2007). We also follow general research on evaluative space that shows the need to assume separate positive and negative affect systems when explaining the organism's choice to engage in approach or avoidance behavior (Cacioppo & Berntson, 1999). This assumption is often made in decision neuroscience, given evidence that appetitive and aversive values are initially represented in distinct brain regions and then integrated into a final value representation (Talmi et al., 2009).

We will return to this issue later, but for now, note that the dual can be used to understand the variation in effort effects on value within the same participant, and also the variability in effort effects across different participants in the same situation, and finally across different participants in different situations. As we show later, assuming a dual evaluation model is also empirically better able to account for the distribution of all empirical profiles (i.e., all profiles cannot be explained with a reduced model).

Our DPOWER (Equation 1) is defined as follows:

$$sv(x) = x \times [1 + \omega\delta_1 E^{\gamma_1} - (1 - \omega)\delta_2 E^{\gamma_2}], \tag{1}$$

where $sv(x)$ signifies the subjective value assigned to an effortful outcome of nominal value x (e.g., monetary value of the action outcome). Parameters δ_1 and δ_2 control the steepness of each system's contribution to the value function, while γ_1 and γ_2 determine the curvature, that is, how that contribution changes with increasing effort levels. This means that the constituents $\delta_1 E^{\gamma_1}$ and $\delta_2 E^{\gamma_2}$ stand for two subsystems that evaluate the effect (positive or negative, respectively) of effort E on the subjective value of outcome x. Meanwhile, ω determines the weight assigned to each subsystem, with its range lying between 0 and 1 ($0 < \omega < 1$). If ω equals 1 or 0, the model simplifies to a singular, flexible power function component. These parameters reflect different aspects of how effort influences decision making. The weight parameter ω represents that attention and importance can shift between positive and negative aspects of effort across individuals and situations. The steepness parameters δ_1 and δ_2 capture individual differences in sensitivity to effort's effects. The curvature parameters γ_1 and γ_2 reflect that effort effects scale nonlinearly across intensity levels—similar to many nonlinearities observed broadly in human judgment and decision making (Kahneman & Tversky, 1979; Killeen, 2009; see also Klein-Flügge et al., 2015).

In Figure 1, we illustrate the different shapes of the value function, also referred to as effort preference profiles, as predicted by the DPOWER model. The value assigned to an effortful outcome is subject to change based on the impact each subsystem has on the overall evaluation. This subjective value could either monotonically decrease ("decreasing" profile), monotonically increase ("increasing" profile), or take a nonmonotonic, mixed shape (decreasing–increasing or increasing–decreasing profile), as a function of the effort needed to obtain a particular outcome. To illustrate, when the first subsystem's contribution, represented as $\delta_1 E^{\gamma_1}$, dominates (with ω nearing 1), the function will increase, mirroring the increasing profile displayed in Figure 1. On the other hand, if the second subsystem's contribution, denoted as $\delta_2 E^{\gamma_2}$, is the dominant one, the value function will monotonically decrease, reflecting the decreasing profile. In both scenarios, the magnitude of the net value for each subsystem determines the steepness of the subjective value's increase or decrease. Notably, our model also incorporates nonmonotonic preferences. Specifically, a mixture of contributions from each subsystem, modulated by parameter ω, can produce a value function that initially increases then decreases past a certain change point (increasing–decreasing profile) or, inversely, first decreases then increases (decreasing–increasing profile). Therefore, parameter ω governs the proportion between the positive and negative components in the resulting effort evaluation.

We provide an interactive visualization tool at https://przemyslawmarcowski.com/projects/dual-power that demonstrates how different parameter combinations affect the system's value functions (see Supplemental Materials). The application includes preset configurations corresponding to the profiles shown in Figure 1 while also enabling exploration of additional value function shapes through dynamic parameter adjustment.

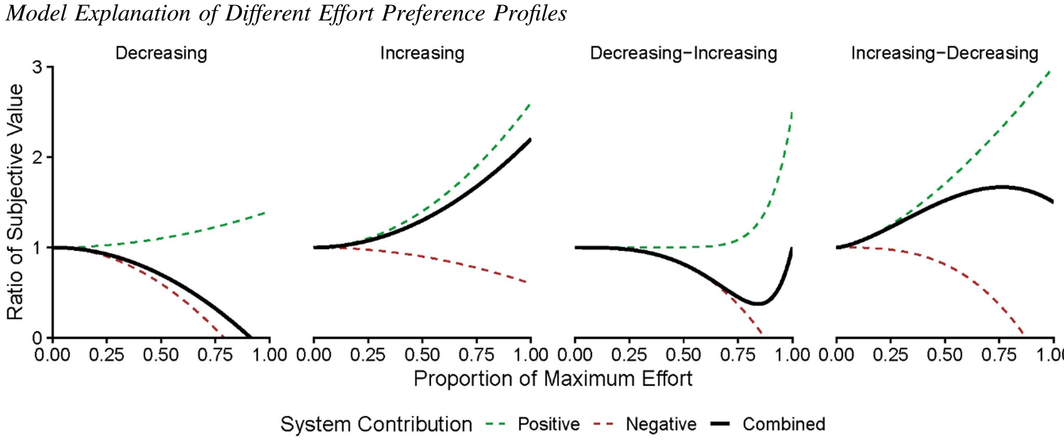
Let us make the model assumptions more concrete by using an example of different individuals walking to a café to get a cup of coffee. Note that in our example even the same individual may experience both positive and negative effects of effort. For example, an individual exerting effort to obtain their coveted coffee in the morning may not like the brisk walk in itself (negative effect) but appreciate the benefits of the exercise associated with it (positive effect). So, the composite effect is neutral. For another individual, they also may not like the brisk walk itself, but they also do not see any perceived benefits of effort (e.g., no positive value from exercise), so the composite outcome in the same situation is negative. These aversive and appetitive components are represented by the two weighted subsystems (positive and negative) in our model so that the composite evaluation is a mixture of these two components. As we demonstrate later, the interaction between positive and negative subsystems can produce all empirically observed patterns: Monotonic decreases when costs dominate, monotonic increases when benefits dominate, and nonmonotonic relationships when their relative influence shifts across effort levels.

Note that for the current model, it is not essential to distinguish whether effort is enjoyed or disliked per se (e.g., some people like and others do not like exertion) or whether effort is enjoyed or disliked because it brings about some associated effects (e.g., health benefits, or costs of being sweaty). Our model is not supposed to distinguish between these subtle differences, but simply help capture differences in effort-based choice, that is different outcome evaluation as a function of effort levels.

The primary goal of current research was to evaluate the performance of our DPOWER model, as explained next. However, to do so, we needed data that show how different degrees of effort influence value in different individuals in different situations. So, we collected data in choice scenarios that vary in temporal orientation (prospective

Figure 1

Model Explanation of Different Effort Preference Profiles



Note. Example value function shapes that illustrate the different preference profiles accounted for under the dual-system power model. Examples were constructed with parameters (δ_1, γ_1, δ_2, γ_2, ω): decreasing (2, 2, 2, 2, 0.2), increasing (2, 2, 2, 2, 0.8), decreasing–increasing (3, 10, 3, 3, 0.5), and increasing–decreasing (4, 1.5, 3, 3, 0.5). See the online article for the color version of this figure.

or retrospective) and tangibility of effort (experienced or hypothetical). In the task, explained in detail in the Method section, we had individuals make repeated choices between options that were either effortless or varied in the amount of effort. These choices involved monetary outcomes and a common item with different effort requirements to obtain it. The payoffs and efforts were being considered in a prospective or retrospective fashion (temporal orientation) and were either purely hypothetical or actually experienced in the experiment (task type). These data allow us to identify and compare individual effort profiles across the prospective and retrospective conditions, as well as within each condition. We expected to see a variety of empirical profiles even within a specific prospective or retrospective condition.

To determine if DPOWER explains effort-based choice while accounting for different effort–reward trade-off preference profiles better than prior solutions, we compared its performance against a selection of other prominent models of choice that do not account for the compound (positive and negative) evaluation of effort. Our selection of candidate models was not exhaustive (Doyle, 2012), but it represents solutions previously shown to perform well in prior works (Białaszek et al., 2017; Cavagnaro et al., 2016; Klein-Flügge et al., 2015; Wulff & van den Bos, 2018). This includes single models, including a heuristic solution (HRST, a heuristic choice model based on the intertemporal choice heuristic; Ericson et al., 2015), convex discount functions (HYPER, generalized hyperbolic model; SENS, constant-sensitivity model; Ebert & Prelec, 2007; Green & Myerson, 2004), concave discount functions (POWER, flexible power function; SIGM, sigmoidal model; Białaszek et al., 2017; Klein-Flügge et al., 2015). We also include dual-system models, including our novel DPOWER and the existing double-exponential model (McClure et al., 2007). As a control, we have also included a baseline intercept model which assumes the choice to depend on a simple utility difference between the two available options. Model equations and details are available in Supplemental Material.

The Present Study

Based on prior studies on effort-based choice we predict that effort may add value, subtract it, or have mixed effects. These effects should occur in different decision scenarios, reflecting the presence of multiple individual effort profiles. In order to vary the decision scenarios, we elicited decisions in retrospective as well as prospective conditions. We also had some participants experience real effort and real rewards, whereas other participants made decisions about effort and rewards that were purely hypothetical. Our predictions were straightforward and consistent with previous literature. In general, greater effort should impact value negatively in the prospective condition and positively in the retrospective condition. Furthermore, effort effects should vary in the "real" versus hypothetical conditions (previous literature does not allow for more precise predictions).

Most importantly, we predict that effort may not always impact outcome evaluation in a linear fashion, that is, it may add value at lower or moderate levels and subtract it as it becomes more demanding (or vice versa). Therefore, we expected that individuals in similar choice situations can exhibit a nonmonotonic preference for effortful outcomes, further differentiated by whether the effort–outcome relationship is prospective or retrospective and whether the effort is purely hypothetical or real. Although we compare our model to those tested previously, we do not claim it to be the only possible model. Our primary and most important goal here is to demonstrate that the relationship between effort and value may be more complex than previously anticipated. We thus aim to offer a framework able to capture the seemingly contradictory pattern of results from different paradigms, which suggests that effort may sometimes add or subtract value from outcomes.

Method

Participants

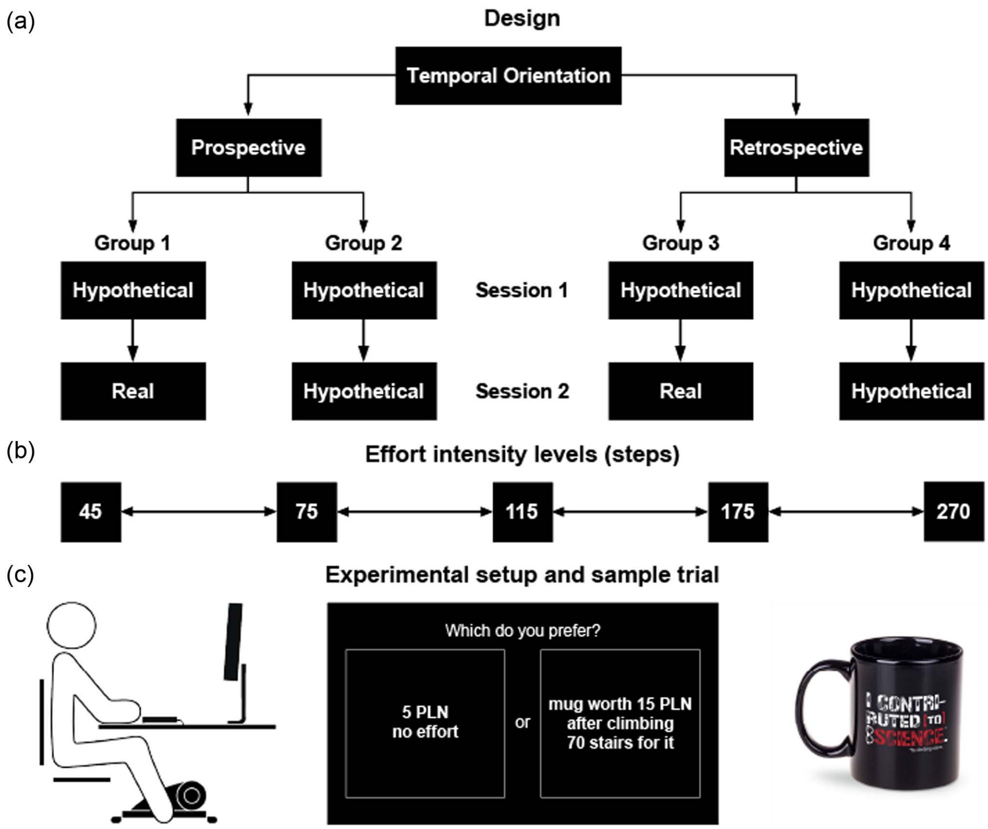
A total of 91 individuals participated in the experiment (55 females and 36 males; 28.0 ± 7.74, $M_{\text{age}} \pm SD$). Participants were recruited from the SWPS University pool, and all signed an informed consent form beforehand. The experimental procedures were approved by the ethics committee of the Faculty of Psychology, SWPS University, Warsaw, Poland, and performed in line with the relevant guidelines and regulations. All procedures were computer-assisted and performed individually at the university behavioral laboratory. The sample size was predetermined and was based on previous research with a similar paradigm (Klein-Flügge et al., 2015).

Procedure

The experiment used a mixed design with repeated measures (see Figure 2): 2 (temporal orientation: prospective or retrospective; between subjects) × 2 (task type: hypothetical or real; within subjects) × 5 (effort intensity: [45, 75, 115, 175, 270] stairs; within subjects). The experimental procedure was programmed in the PsychoPy (J. W. Peirce, 2007, 2009; J. Peirce et al., 2019) software library for Python. The design employed a modified version of the psychophysical procedure called Parametric Estimation by Sequential Testing (Klein-Flügge et al., 2015; Taylor & Creelman, 1967). This procedure allows for the collection of data on choices made by participants with parametrically changing properties of each decision, which in this case included the nominal value and effort associated with the choice alternatives.

Participants made a series of choices between an effortless monetary payoff and an item associated with a given level of effort. The effort was presented as climbing a given number of stairs (Figure 2b). There were five blocks that varied in the amount of effort (45, 75, 115, 175, 270 steps), with the order randomized for each participant. The nominal value of the effortful option (custom mug designed for this study) was kept constant (PLN 15) throughout the experiment, while the value of the effortless option was adjusted trial-by-trial using the parametric estimation by sequential testing algorithm. Depending on if the effortful option was accepted or rejected, the value of the effortless option was increased or decreased, respectively, by an amount accordant to the parametric estimation by sequential testing algorithm (initial value of the adjusting option was set to PLN 7.5; 1 USD was equal to approximately PLN 3.84 at the time). By adjusting the value of the effortless option, the procedure aims to elicit the point of subjective indifference between the alternatives, that is, assigning equal value to both options. As such, it yields a wide array of choices, with distal and proximal alternatives as values converge onto the point of subjective indifference. This procedure was repeated for each of five levels of effort. As a manipulation check,

Figure 2
Experimental Design



Note. (a) Schema of the experimental design with group assignment. (b) Effort intensity levels included in the study (presented randomly). (c) Setup (elliptical stepper device, used to produce effort in the real experimental condition); sample experiment trial with a choice between the effortless and effortful option (retrospective temporal orientation); custom mug used as the effortful option representing a monetary value of 15. See the online article for the color version of this figure.

participants were asked to evaluate their final choice in each effort block, that is, the choice that resulted in the point of subjective indifference between the effortless option (money) and the effortful option (item), in terms of how much they liked each option and how confident they were in their final choice (for details, see Supplemental Material). These evaluations were performed on a Likert scale ranging from 1 to 7. Validating our assumption that choices align with preferences, participants liked what they chose more than the unchosen option (so they liked an item more when they chose the item). After completion of the main experiment, participants were administered questionnaires assessing general psychological functioning, measures of personality, and affective dispositions. These data were collected but were not analyzed and are not further discussed in this study.

Each participant was tested in two sessions with different experimental conditions depending on a pseudorandom group assignment (see Figure 2a). Participants were assigned to one of four groups. Those assigned to Groups 1 and 2 performed their choices in the prospective condition, while those assigned to Groups 3 and 4 performed the choices in the retrospective condition. In addition, participants assigned to Groups 1 and 3 were first tested in the hypothetical and then real condition, while those assigned to Groups 2 and 4 were tested in the hypothetical condition in both sessions. Our design had two sequential hypothetical conditions (Groups 2 and 4) to prevent participants in the second session from being influenced by their choices in the real condition of the first session. There was a 14- to 21-day interval between sessions.

In the prospective condition, participants were instructed that in each choice, they could receive the mug after performing a specific effort or instead receive a sum of money without any effort. The instruction in the retrospective condition differed from the prospective condition in that participants were informed that they had obtained the presented mug after performing a specific effort. Participants were asked to decide whether to keep their mug or give it up for a certain amount of money. Specifically, the instructions for the prospective/retrospective conditions were as follows:

> Choices in the study involve the mug presented below, worth 15 PLN.
>
> You [can obtain/own] such a mug, but to obtain it, you [must/had to] climb a certain number of stairs.
>
> You can [obtain/keep] this mug or receive a certain amount of money instead without any effort and immediately.

In addition, in the hypothetical (prospective and retrospective) conditions, the participants were instructed to imagine both the mug and the effort as well as the possible amounts of money by providing the following instruction:

> The money, mug, and effort presented in the study are purely hypothetical, i.e., imaginary.

Before the decision phase of the experiment, in the real (prospective and retrospective) conditions, participants were first asked to perform the full range of possible effort, that is, 270 steps, using a stair simulator. Because this effort was performed before the main experimental procedure, separately from the later decisions, at that time of effort performance participants were unaware of its later role in the study. This arrangement ensured that all "real" condition participants had direct, experiential knowledge of what the effort felt like at all levels (all performed 270 steps). Following this, participants in the real condition were shown the actual mug and given instructions corresponding to their assigned temporal orientation condition (prospective or retrospective). In the prospective condition, participants were informed they could obtain the mug by performing the effort corresponding to the current experiment block. In the retrospective condition, participants were told they had already obtained the mug by performing the given level of effort in the block. Additionally, each participant in the real condition was informed that after the procedure, they would receive the result of one choice (real money or real mug), which would be selected at random from choices made by them during the experiment. Sample choice and the materials used in the experiment are presented in Figure 2c. This design allowed for a clear differentiation between conditions based on direct personal experience with the effort and outcome (mug) while maintaining similar overall situations for all participants. We address the limitations of the design later.

We designed our experiment to minimize the potentially confounding effect of real effort performance time on choice. However, to ensure that increasing the number of steps effectively enhances perceived effort demand, we conducted a supplementary experiment (see Supplemental Material). Results showed that effort demand ratings significantly increased with the number of steps performed, even when controlling for temporal demand. This validation confirms that our task manipulation influences perceived effort as intended, independent of performance time.

Behavioral Modeling

All inferential procedures were performed in the R computational environment (R Core Team, 2024) and were adapted from the approach by Wulff and van den Bos (2018). Models were evaluated using Monte Carlo Cross-Validation by repeatedly fitting the models to a subsampled set of choices and predicting the remaining set while obtaining the prediction error and proportion of choices correctly predicted (Wulff & van den Bos, 2018). Each model was first fitted to a sampled 70% of the choice data (training set) using maximum likelihood estimation. Model fitting used wide parameter bounds: [1e−7, 1 − 1e−7] for the system weight parameter ω and [1e−7, 1e+2] for both the steepness (δ1, δ2) and curvature (γ_1, γ_2) parameters. These bounds allow substantial flexibility in capturing different value function shapes while maintaining numerical stability. The log-likelihood function (Equation 2) for each model was defined as follows:

$$\ell(\theta,y) = \sum_{i=1}^{n}[y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)], \tag{2}$$

where y_i and $\hat{y}_i$ are the observed choice i and predicted probability of choosing the effortful option, respectively. The softmax decision rule was used to transform option values into choice probabilities (Equation 3), defined as follows:

$$\hat{y} = \frac{1}{1 + e^{-\tau}[sv_{E2} - sv_{E1}]}, \tag{3}$$

where sv_{E2} and sv_{E1} correspond to the subjective value of the effortful and effortless options, respectively, predicted by a given model; τ is the softmax temperature that determines the steepness of the softmax function, that is, how sensitive is the choice probability to the differences in the subjective value of both options. To reduce overfitting and account for the randomness in the decision-making process, random noise was introduced during modeling by blending the predicted choice probability with the chance level (at ε = 0.001), thus balancing prediction and randomness while promoting better generalization to unseen data. Best-fitting model parameters $\hat{\theta}(l(\theta,y))$ were obtained by fitting each model 10 times with different starting values. The remaining 30% of the choices (test set) were then predicted using $\hat{\theta}$. This procedure was performed with 100 repetitions. Models were evaluated for each individual per condition separately. Different parameter starting values were drawn from a set obtained by fitting each model in each participant a single time. Two evaluation criteria were computed based on the choices in the test set and predicted choice probabilities: logistic loss (log-loss, $J(y_i,\hat{y}_i)$; Equation 4), corresponding to prediction error and matched to the estimator used to fit the models, defined as follows:

$$J(y_i,\hat{y}_i) = -\frac{1}{n}\sum_{i=1}^{n}\begin{cases} -\log(\hat{y}), & y \sim \hat{y} \\ -\log(1-\hat{y}), & y \nsim \hat{y} \end{cases}, \tag{4}$$

and π_{correct}, the proportion of choices correctly predicted (complementary of the zero-one loss; Equation 5), defined as follows:

$$\pi_{\text{correct}} = -\frac{1}{n}\sum_{i=1}^{n} I(y_i,\hat{y}_i), I(y_i,\hat{y}_i) = \begin{cases} 1, & y \sim \hat{y} \\ 0, & y \nsim \hat{y} \end{cases}. \tag{5}$$

In addition to cross-validation, we employed the corrected Akaike information criterion (AICc; Akaike, 1974; Hurvich & Tsai, 1989; see also Burnham & Anderson, 2002) to explicitly account for model complexity in our comparisons. The AICc extends the standard Akaike information criterion by including a correction for finite sample sizes, penalizing models for additional parameters while accounting for the number of observations. Models were fitted to individual choices and compared using ΔAICc values, calculated as the difference between each model's AICc and the minimum AICc value in the set of candidate models. These ΔAICc values were then averaged across participants for the model comparison (see Supplemental Material).

Finally, to further validate the explanatory and predictive power of the proposed model, we conducted a model recovery analysis using an agent-based simulation. This involved simulating choices using individual DPOWER parameters estimated from participant choices within the same experimental procedure. For parameter recovery, we refitted the model to the simulated choices and compared the recovered parameters with the original empirical parameters using permutation tests.

Transparency and Openness

This study adheres to the transparency and openness guidelines. Our data, analysis code, and experimental materials are publicly available on the Open Science Framework repository at https://osf.io/e3baz/. While not preregistered, we disclose all details of our experimental design and analysis methods.

Results

Behavioral Effort Preferences

Before evaluating our model, we present behavioral results. To remind, participants performed a series of choices between an effortless monetary payoff and an item (mug) that involved different amounts of effort. Participants made these choices in a prospective or retrospective fashion (temporal orientation factor). Furthermore, some choices scenarios involved purely hypothetical effort and rewards or involved an initial experience of real effort and dealt with real rewards (task type factor). A logit model was fitted to participants' choices (with subject-level random intercepts) to determine how the value of the effortless option and the effort associated with the effortful option influenced choice probabilities in prospective or retrospective conditions and in hypothetical or real conditions (also controlling for the experimental session order). All inferences described in this section are based on the logit model described here. Confidence intervals (CI) are 95%. Uppercase P refers to choice probability, and lowercase p stands for statistical significance.

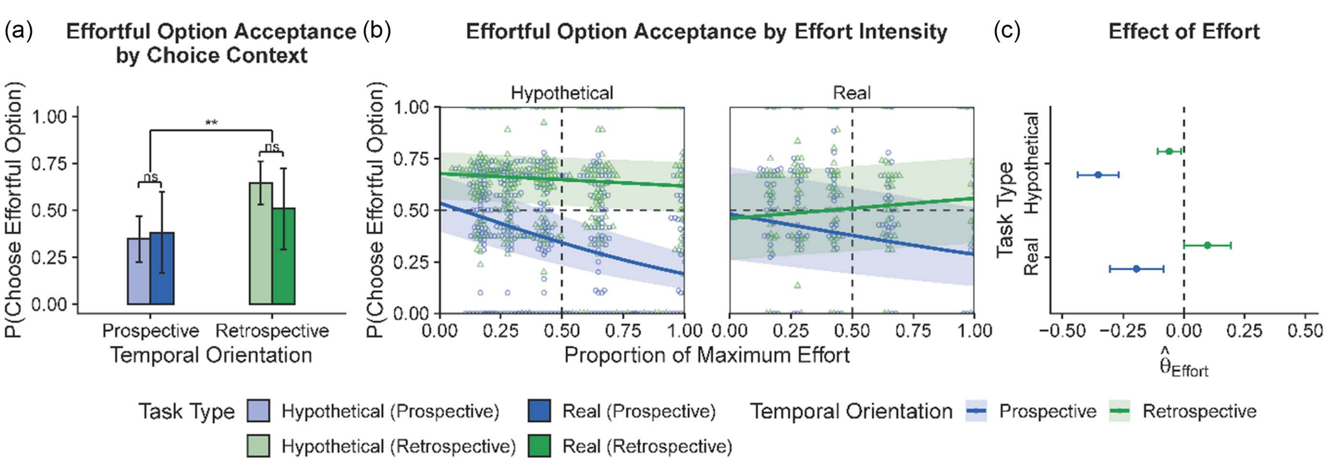
Not surprisingly, we found that the probability of selecting the effortful option was significantly predicted by the value of the effortless option ($\chi^2 = 600.03$, $p < .001$) and, crucially, the effort intensity ($\chi^2 = 57.38$, $p < .001$; $\hat{\theta}_E = -0.13$, 95% CI [−0.17, −0.08]).

In addition, we observed an effect of the temporal orientation ($\chi^2 = 9.84$, $p = .002$), but not an effect of the task type ($\chi^2 = 0.35$, $p = .555$), session order ($\chi^2 = 1.33$, $p = .248$), or an interaction of temporal orientation and task type ($\chi^2 = 1.12$, $p = .290$). Specifically, marginal probabilities are presented in Figure 3a. Note that in general, the participants were overall more likely to choose the effortful option in the retrospective condition ($P = .58$, 95% CI [0.46, 0.69]) than in the prospective condition ($P = .36$, 95% CI [0.25, 0.48]; $c = -0.21$, CI [−0.37, −0.05], $t = -2.62$, $p = .009$). This overall difference could represent the endowment effect (higher value of a mug that is already owned) and is interesting to study in future research, though not essential to our interest in the shape of effort–value function here.

Note also that overall choice probabilities did not differ between the hypothetical and real conditions in either prospective ($P = .35$, 95% CI [0.22, 0.47] and $P = .38$, 95% CI [0.16, 0.60], respectively; $c = -0.03$, 95% CI [−0.29, 0.23]; $t = -0.26$, $p = .796$) or retrospective temporal orientation ($P = .65$, 95% CI [0.53, 0.76] and $P = .51$, 95% CI [0.29, 0.72], respectively; $c = 0.14$, 95% CI [−0.11, 0.39]; $t = 1.08$, $p = .282$).

Critically, we discovered an interaction of effort intensity and temporal orientation ($\chi^2 = 69.67$, $p < .001$). We also found an interaction between effort intensity and task type ($\chi^2 = 16.31$, $p < .001$). To unpack these interaction effects further, we inspected the linear trends of the influence of effort on choices. As illustrated in Figure 3b and 3c, effort intensity influenced choices differently across conditions. In the prospective condition, higher effort decreased the probability of selecting the effortful option. This decrease was steeper in the hypothetical condition ($\hat{\theta}_{EPH} = -0.35$, 95% CI [−0.44, −0.27]) than in the real condition ($\hat{\theta}_{EPR} = -0.19$,

Figure 3
Effortful Option Acceptance



Note. (a) Marginal probabilities of accepting the effortful option estimated across conditions. (b) Probabilities of accepting the effortful option are estimated as a function of effort intensity by temporal orientation (prospective or retrospective) and task type (hypothetical or real). Bee swarms represent mean empirical acceptance probabilities of the effortful option for each participant at given effort intensity. (c) Linear trends of the estimated effect of effort across conditions. Error bars and bands represent 95% confidence intervals. Effort intensity values are represented as the proportion of the maximum effort intensity used in the study. P = probability. See the online article for the color version of this figure.
** $p < .01$. ns $p > .05$.

95% CI [−0.31, −0.08]; $c = -0.16$, 95% CI [−0.30, −0.02]; $t = -2.22$, $p = .026$). Interestingly, as depicted in Figure 3b, in the retrospective condition, greater effort also negatively influenced the probability of selecting the effortful option, but only when the type of task was hypothetical ($\hat{\theta}_{ERH} = -0.06$, 95% CI [−0.11, −0.01]). Conversely, greater effort positively influenced the choice probability when the task type was real ($\hat{\theta}_{ERR} = 0.10$, 95% CI [0.01, 0.19]; $c = -0.16$, 95% CI [−0.26, −0.05], $t = -2.87$, $p = .004$).

Moreover, as shown in Figure 3c, the effect of effort on choices in the purely hypothetical task was more negative ($\hat{\theta}_{EPH} = -0.35$, 95% CI [−0.44, −0.27]) in the prospective condition than in the retrospective condition ($\hat{\theta}_{ERH} = -0.06$, 95% CI [−0.11, −0.01]), with the contrast for the difference being significant ($c = -0.29$, 95% CI [−0.39, −0.20]; $t = -6.01$, $p < .001$). However, our most crucial finding (Figure 3b and 3c) is that when the task was real, the effect of effort reversed in the prospective versus retrospective condition, that is, it was negative in the prospective condition ($\hat{\theta}_{EPR} = -0.20$, 95% CI [−0.31, −0.08]) but positive ($\hat{\theta}_{ERR} = 0.10$, 95% CI [0.01, 0.19]) in the retrospective condition, with the difference being significant ($c = -0.29$, 95% CI [−0.44, −0.15]; $t = -3.91$, $p < .001$). This behavioral result in the real condition (tangible effort and tangible reward) demonstrates that effort has the opposite effect on value depending on whether the decision involves consideration of future effort or consideration of past effort.

As just described, the behavioral data highlight that average trends depend on specific decision settings. Specifically, they demonstrate the opposite effects of effort on value as a function of temporal orientation (in the real condition). They also show the role of tangibility of effort and rewards. Importantly, these average trends hide the variability in the preference profiles across different choice contexts, as described next.

Modeling Effort-Based Choice

To further explore choice dynamics, candidate models were first evaluated using our participants' behavioral data. Model evaluation shows that DPOWER consistently outperforms other candidate models, regardless of whether the choices were made in prospective, retrospective, hypothetical, or real conditions. This can be seen in multiple indicators. DPOWER yielded lower prediction error when compared to other models (Figure 4a). DPOWER was followed by POWER and then by double-exponential models in the prospective condition and by HRST and POWER in the retrospective condition, yielding second and third lowest error, respectively. For formal model selection, we performed pairwise comparisons based on bootstrapped differences in mean prediction error across all models (with 1,000 resamples; Tibshirani & Efron, 1993) with the false discovery rate multiplicity correction (Benjamini & Hochberg, 1995). In all conditions, DPOWER performed significantly better compared to other models—indicated by the highest error-based rank, averaged across Monte Carlo Cross-Validation repetitions (Figure 4b). Finally, we computed relative model frequencies to inspect the distribution of cases where a given model performed best (and second and third best; Figure 4c). Again, DPOWER yielded the highest proportion of cases where it performed best, achieving Rank 1 in 33.5% and 35.5% of cases in the prospective and retrospective conditions, respectively. This was followed by the HRST and POWER models, achieving Rank 1 in 26.4% and 25.6% and 19.8% and 11.6% of the cases in the prospective and retrospective conditions, respectively.

To illustrate model performance more intuitively, the proportion of choices correctly predicted ($\pi_{correct}$) was computed and averaged over each condition (Figure 4d). Consistent with prior results, DPOWER correctly predicted the highest percentage of choices in both conditions (72.8% and 71.1% in the prospective and retrospective conditions, respectively). As for the other models, in the prospective condition, DPOWER was followed by POWER (67.4%) and double-exponential model (66.5%), and by HRST (67.1%) and POWER (62.5%) in the retrospective.

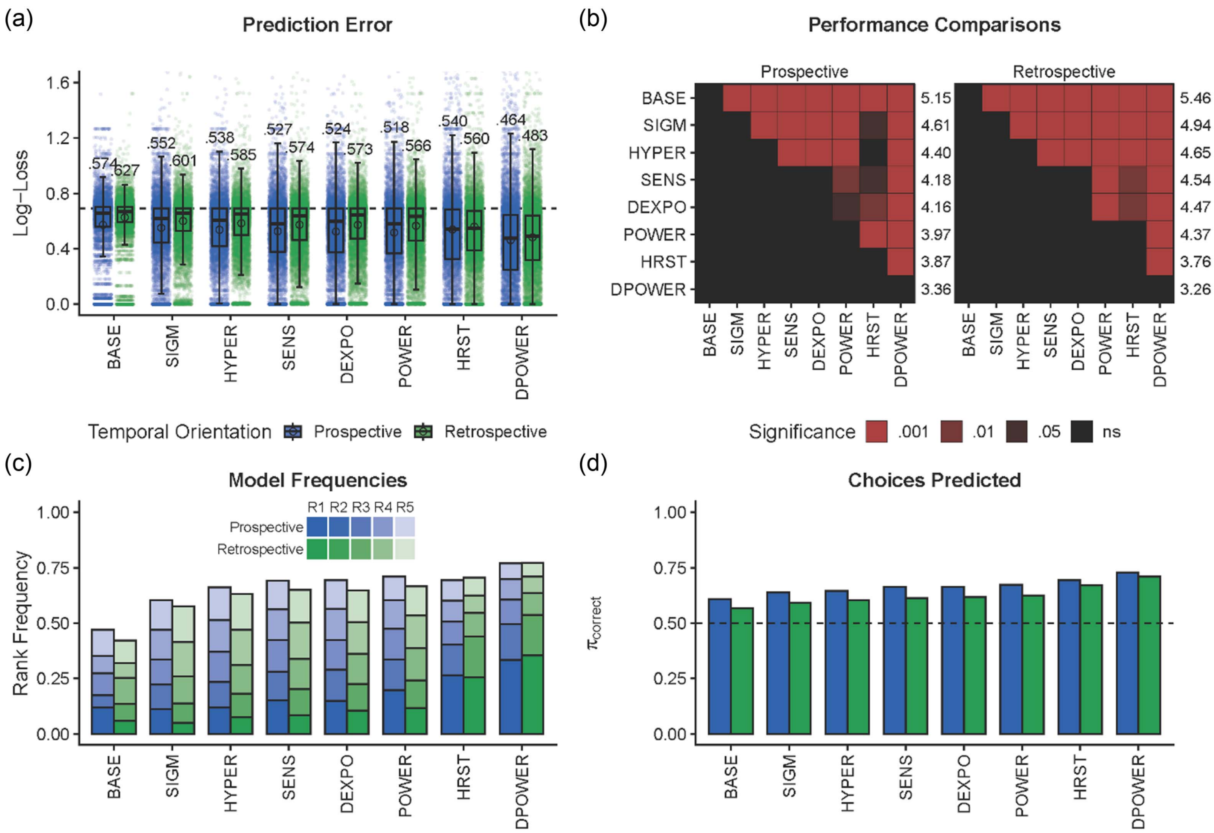
Next, we fitted the DPOWER model to individual choices. The analysis using AICc further confirmed DPOWER's superior performance relative to other candidate models, even when penalizing for additional parameters (mean ΔAICc = 5.56, compared to the next best models HRST = 17.24 and POWER = 13.21; see Supplemental Material).

Figure 5a shows the estimated parameter values derived for both prospective and retrospective conditions. We evaluated these estimates by comparing bootstrapped differences in means between the prospective and retrospective conditions. This analysis revealed an important discovery: the values of the subsystem weight parameter, ω, varied between the conditions ($c = -0.11$, 95% CI [−0.21, −0.01]; $t = -66.76$, $p < .001$). In essence, this suggests that the positive subsystem's influence was more pronounced in the retrospective condition, while the negative subsystem's impact was more substantial in the prospective condition.

Estimated parameters were then used to identify four distinct preference profiles. The profiles were established by inspecting the shape of the value functions derived for each participant. This means that for each individual, the monotonicity of their value function was assessed point-by-point, and any existing function change points were detected. In accordance with the assumptions of our model, we partitioned our participants into profiles based on the shape of their value function: decreasing, monotonic decrease; increasing, monotonic increase; decreasing–increasing, value function initially downward with an upward shift after reaching a change point; increasing–decreasing, value function initially upward with a downward shift after reaching a change point. It is important to note that these categories represent examples of the variation in effort preferences (successfully captured by our model), rather than discrete, mutually exclusive groups. For examples of observed profiles, see Figure 5b. Note that examples illustrating nonmonotonic effort–value relationships represent cases where for the same individual the effort–value relation flips from being positive to negative (increasing–decreasing) or from being negative to positive (decreasing–increasing). This is interesting because it shows that changing levels of effort can lead to a form of preference reversal (e.g., an individual prefers an effortful option at low-to-high effort, but effortless option at very high effort).

The profile proportions, as illustrated in Figure 5c, were entered into a Fisher's exact test, which revealed differences in the incidence of increasing and decreasing profiles between prospective and retrospective conditions ($p = .044$). The decreasing profile occurred more frequently in the prospective condition, as compared to the retrospective condition ($P = .41$ vs. $P = .25$). Conversely, the increasing profile occurred more often in the retrospective condition rather than the prospective one ($P = .16$ vs. $P = .08$).

Figure 4
Model Comparison



Note. (a) Boxplots and points in color represent the distributions of prediction error (log-loss) across a total of 18,200 Monte Carlo Cross-Validation repetitions per model. Hollow points and labels represent mean prediction error. Dashed line references a chance-level predictive performance, that is, log-loss equal to −log(.5), or 0.693. (b) Tiles represent pairwise comparisons that test if model performance differed across models. Right-hand side values correspond to the mean rank obtained by each model based on the produced prediction error. (c) Relative rank frequencies based on the prediction error produced by each model. Stacked bars represent the proportion of cases in which a given model achieved Rank 1 (R1) through Rank 5 (R5), that is, performed best, second best, and third best. Increasing transparency represents increasing rank (i.e., decreasing model performance). (d) In this panel, π_{correct}, represents the proportion of choices correctly predicted by each candidate model averaged across all Monte Carlo Cross-Validation repetitions. BASE = baseline intercept model; SIGM = sigmoidal model; HYPER = hyperboloid model; SENS = constant-sensitivity model; DEXPO = double-exponential model; POWER = flexible power function; HRST = heuristic choice model; DPOWER = dual-system power model; R = rank. See the online article for the color version of this figure.

It is important to note that different effort profiles occurred within each temporal orientation condition (Figure 5c). This means that even under the same circumstances (temporal orientation), some participants show a positive relation between effort and value, some show negative, and yet others exhibit nonmonotonic preference profiles. Further analysis of profile distributions across temporal orientation and task type conditions (real vs. hypothetical) shows that each type of profile is present throughout all experimental conditions (see Supplemental Figure S5).

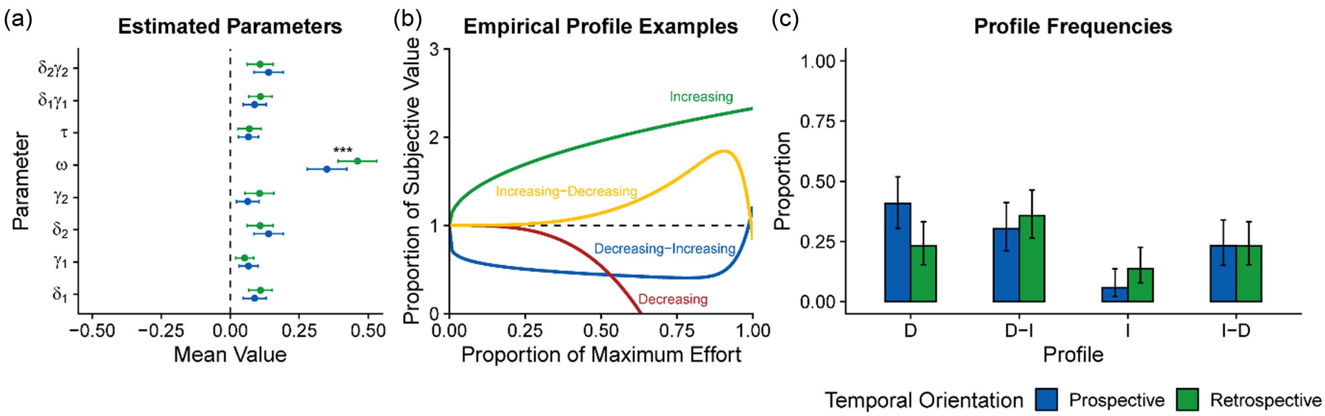
Model Simulation

Comparison of the distributions of empirical and simulated choices revealed no significant differences, $\chi^2(1) = 0.85$, $p = .357$. This indicates a high degree of similarity between the choice patterns observed in the empirical data and those generated in the simulation. Model parameters were then recovered from the simulated choice data. Using a permutation test of differences, we compared the empirical and recovered values for each parameter. An offset was applied to parameter values before analysis to avoid negative values, followed by a log transformation to stabilize the variance. The results, summarized in Table 1, showed no significant differences for any of the parameters.

The agreement between empirical and recovered parameters is visualized with a Bland–Altman analysis in Figure 6a. The analysis indicated that most differences fell within the 95% limits of agreement, further demonstrating model robustness.

Finally, to determine if key behavioral effects observed in empirical data can be replicated in an agent-based simulation, we fitted the same logit model to the simulated data. Consistent with our empirical findings, the value of the effortless option ($\chi^2 = 948.26$, $p < .001$) and effort intensity ($\chi^2 = 104.56$, $p < .001$) significantly predicted the probability of selecting the effortful option. Additionally, we replicated the effect of temporal orientation

Figure 5
Dual-System Power Model Predictions



Note. (a) Estimated model parameter value in the prospective and retrospective condition. Presented are mean normalized values of the parameters in each subsystem, as well as the subsystem weight (ω), choice sensitivity (softmax temperature, τ), and the magnitudes of each of the two subsystem terms: $\delta_1 E^{\gamma_1}$ and $\delta_2 E^{\gamma_2}$ at $E = 1$. (b) Example value functions based on individual parameter estimates of the dual-system power model. Shown are value functions that decrease (D = decreasing) or increase (I = increasing) monotonically or initially decrease or increase and then reverse in evaluation after given effort is reached (D-I = decreasing-increasing or I-D = increasing-decreasing profile). (c) Proportion of each preference profile in the prospective or retrospective condition. Effort intensity values are represented as the proportion of the maximum effort intensity used in the study. Error bars represent 95% confidence intervals. See the online article for the color version of this figure.
$^{***}p < .001$; Except for a significant difference in ω values between conditions, all remaining parameter differences did not reach significance.

($\chi^2 = 8.95$, $p = .003$), but not the effect of task type ($\chi^2 = 0.10$, $p = .749$) or session order ($\chi^2 = 0.73$, $p = .394$). Importantly, we also observed the interaction between effort intensity and temporal orientation ($\chi^2 = 31.24$, $p < .001$), as well as between effort intensity and task type ($\chi^2 = 18.47$, $p < .001$). However, the interaction between temporal orientation and task type ($\chi^2 = 1.60$, $p = .206$) was not significant, mirroring the empirical results. These results indicate that our model successfully captures behavioral dynamics observed in the empirical data. Figure 6b illustrates the probabilities of accepting the effortful option estimated from simulated data as a function of effort intensity.

Discussion

Here we reported new behavioral findings and proposed a novel computational approach that allows for the effort associated with the outcome to have positive, negative, or nonmonotonic effects on choice. We designed our experiment and built our model to address the diversity of relations between effort and outcome values in different situations and different individuals.

Our results and previous empirical findings show that the value of an outcome can be decreased but also enhanced by increased effort. The exact relation can depend on the specific temporal orientation of the choice, that is, prospective or retrospective evaluation. Presumably, this reflects the fact that in future choices effort is avoidable, so if the effort itself is largely experienced as unpleasant, greater effort level reduces the choice of the associated item. However, in retrospective evaluation, the effort has already been spent, so the evaluations presumably are more likely to incorporate inferences about effort adding subjective or objective value, being correlated with higher outcome value, or with inferences aimed at effort justification. All these inferences may result in greater effort in increasing the choice of the associated item. This is supported by our observation indicating that the DPOWER model's positive subsystem had a more pronounced contribution to the overall evaluation in the retrospective condition, compared to the prospective one. Conversely, the impact of the negative subsystem was greater in the prospective rather than the retrospective condition.

Critically, the effort–value relationship also varies across individuals, presumably reflecting their idiosyncratic hedonic value for mere effort itself and associated benefits (e.g., health benefits of exercise) or costs (e.g., sweating). Accordingly, some individuals devalue effortful outcomes, some prefer effortful options, and others show mixed profiles. One challenging question for future work on effort-based choice will be to distinguish the relative role of changes in the value of "mere" effort and the value of effort-associated benefits (Shenhav et al., 2017). For example, a nonmonotonic profile could reflect changes in mere hedonics of effort (a short run feels good, a long run feels bad) or in its associated cost (a short run breaks no sweat, a long one does).

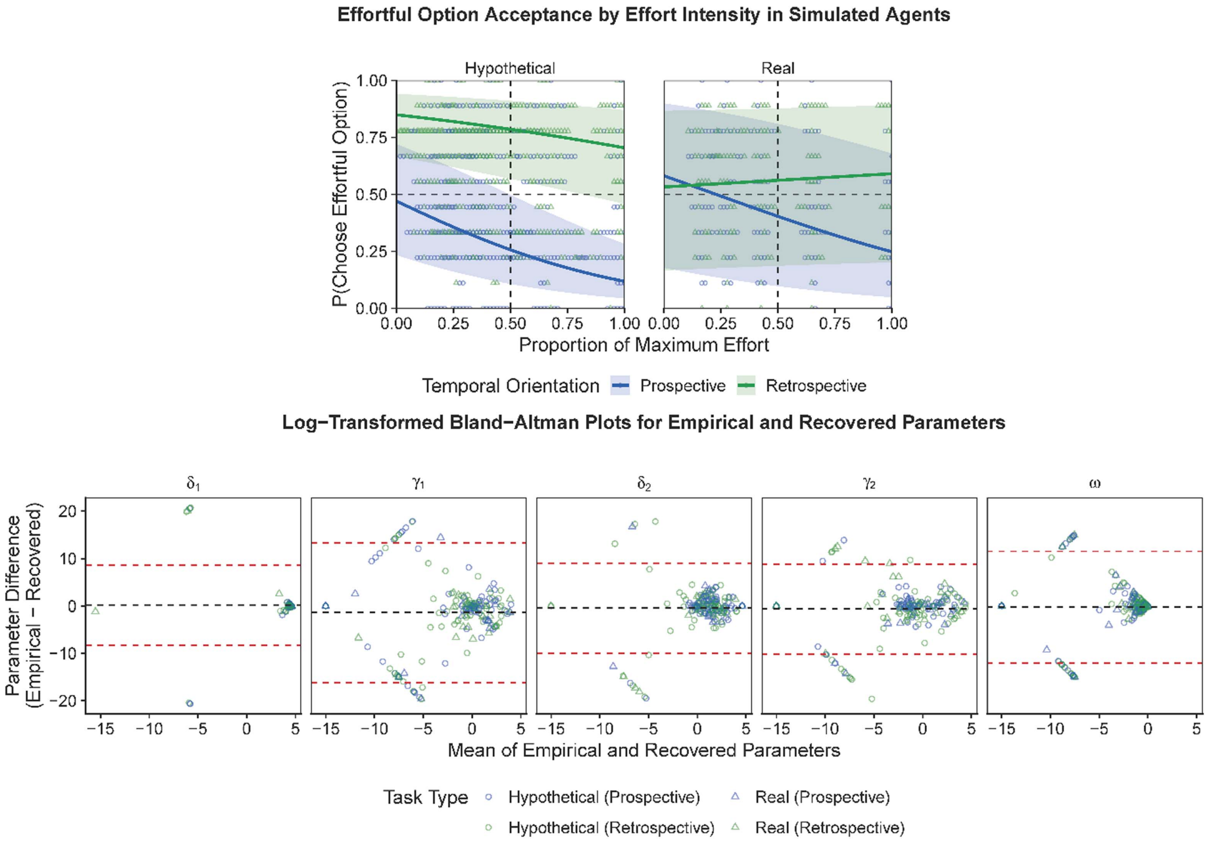
Table 1
Results of the Permutation Tests Comparing Empirical and Recovered Parameters

Parameter	Δ	*p*
δ_1	1.02	.853
γ_1	0.45	.208
δ_2	0.68	.339
γ_2	0.65	.480
ω	1.08	.893

Note. Δ = observed mean difference between the empirical and recovered values for each parameter; presented values are back-transformed to the original scale; p = p value from the permutation test.

Figure 6

Model Recovery and Results of the Agent-Based Simulation



Note. (a) Estimated probabilities of accepting the effortful option as a function of effort intensity. Choices were simulated using the dual-system power model in the same experimental procedure, with each agent created using individual parameter values obtained from corresponding participants. Bee swarms indicate the mean acceptance probabilities of the effortful option for each agent at given effort intensities. (b) Bland–Altman plots showing the agreement between empirical and recovered parameters across experimental conditions. The *x*-axis represents the mean of the empirical and recovered values, while the *y*-axis shows their difference. The dashed black line is the mean difference, and the dashed red lines indicate the 95% limits of agreement ($M \pm 1.96$ *SD*). Points within these limits indicate good agreement. P = probability. See the online article for the color version of this figure.

The behavioral data presented here also show the role of task type—whether it was purely hypothetical or involved some personal experience with effort and reward. Note that the overall difference between the hypothetical or real choice conditions was not significant. However, task type interacted with temporal orientation. When the temporal orientation was prospective, effort had an overall negative effect that was larger in hypothetical than real tasks. In contrast, when the orientation was retrospective, effort had a negative effect only in the hypothetical condition, but a positive effect when the task effort and rewards were real. This reversal of the effort–value relation as a function of temporal orientation in the real condition is important because it suggests that the dissociations between the negative and positive effect of effort on value are more likely when the components of choice (i.e., decision costs and benefits) are tangible and presumably more salient to the decision maker than in the hypothetical condition. Future investigations could explore this issue further.

It is also important to acknowledge the limitations of our task paradigm. As mentioned, in the "real" condition, participants had a direct experience only with one level of effort (270 steps), exerted in the initial stage of the experiment, before they made any choices. While our methodological decision ensured that in the "real" condition participants understood the feeling of effort and had a subjective anchor for its maximal value, it also means that their actual choices were still hypothetical and were not tied to the performed effort. These features could have reduced any differences between our "hypothetical" and "real" conditions. Furthermore, because participants experienced only one "real" level of effort, their decisions could be based primarily on that level and thus less sensitive to effort level variations. Future studies could address this issue by having participants experience all effort levels for real, directly preceding or following the choices. In addition, in our studies, only the "real" condition had a tangible reward. This could make choices in the condition with both hypothetical rewards and hypothetical effort less consequential. Future studies could independently manipulate the hypothetical versus real nature of both effort and reward.

Importantly, despite these limitations, we still found robust effects of experimental conditions on average choice proportions.

Furthermore, we found that these condition averages contain a diversity of effort profiles, with each choice scenario including a sizable proportion of each type of effort profile (see Supplemental Materials for more detailed information on profile distributions in each condition). As such, our data are consistent with recent calls to go beyond inferences from averaged data that may hide the heterogeneity of underlying processes (Haaf & Rouder, 2019). Reflecting this, our model is designed to capture the situational and individual differences in effort profiles.

We compared our model against other models previously used to explain individual differences in intertemporal, risky, or effortful choices—but did not allow for a simultaneous positive and negative evaluation of outcome contingencies. In all experimental conditions, regardless of whether decision scenarios were purely hypothetical or real, DPOWER outperformed all other candidate models by yielding the lowest prediction error. The superior explanatory power of DPOWER stems from its ability to accommodate diverse value function shapes, which allows it to capture the complexity of individual effort preferences. By emerging from the integration of two model subsystems, DPOWER allows for a more flexible representation of effort preference dynamics. Furthermore, the model recovery results demonstrate that DPOWER not only reproduces original empirical responses but also accurately regenerates individual parameter estimates, further suggesting the model's robustness and applicability to variation in empirical contexts (prospective, retrospective, hypothetical, and real). Having said this, it is important that future research explores the applicability and generalizability of the model to other empirical data sets and different decision scenarios.

Our finding showing the variability of effort profiles, including the discovery of positive relations between effort and choice, aligns with previous work. For example, Kool et al. (2010) found that while most participants consistently chose options associated with lower cognitive demand, others showed no such preference. Ma et al. (2014) showed that exerting more effort on a task led to larger feedback-related negativity and P300 responses during subsequent reward feedback, indicating increased subjective valuation of rewards following greater effort expenditure. Zerna et al. (2023) observed a preference for more difficult tasks in a subset of their sample, coupled with individuals with higher need for cognition scores reporting lower subjective task load and less aversion to more effortful tasks, further suggesting heterogeneity in effort preferences. Finally, recent neuroimaging work by Bogdanov et al. (2022) has also shown that cognitive effort exertion can increase the subjective value of subsequent outcomes, as reflected in enhanced neural responses to both gains and losses following high-effort trials. More broadly, recent reviews of research on the role of mental effort in social evaluations point out the diversity of relationships, with processing difficulty sometimes decreasing but sometimes increasing value judgments (Schwarz et al., 2021; Yoo et al., 2024).

In our study, we proposed a novel formal model capturing the complex relationship between effort and outcome valuation. The innovation comes from multiple features. Unlike prior models that focused on the devaluation of effortful outcomes, our approach allows for the possibility that effort enhances value. This feature allows our model to capture a range of individual effort profiles: some individuals may reject more effortful outcomes (effort discounting), while others may opt for them (effort as value). Another important feature of our model is its ability to capture changes and even reversals in outcome valuation after specific effort thresholds. These nonmonotonic relationships could reflect situations where, for a given individual, effort finally reaches a level at which it starts enhancing value (e.g., when effort during a walk starts offering exercise benefits) or reaches a level at which it starts decreasing value (e.g., when effort starts to exceed one's physical capacity). Such intraindividual reversals in the effort–value relationships offer an exciting avenue for future work.

Our model's ability to capture complex relationships between effort and value comes from the assumption that net effort effects reflect a combination of two evaluation systems, which on their own operate using simpler principles. These assumptions are consistent with psychological theories, suggesting relative independence of positive and negative evaluations (e.g., Cacioppo & Berntson, 1999) and work in decision neuroscience on the separability of different evaluative systems (e.g., McClure et al., 2007; Talmi et al., 2009), though we leave the validation of underlying mechanisms for future work. Mechanistic claims about dual models (whether they are about two systems, two processes, or two sources) require additional evidence of separability (e.g., manipulations that selectively impact only one system, process, or source).

In conclusion, our work fills a theoretical and empirical gap by providing direct evidence that in reasonably similar situations, effort can be both costly (the law of less work) and prized (effort as value). Using a between-subject design, our study shows that the situational differences in temporal perspective (prospective or retrospective) change the effort–value relationship. Future research on temporal perspective should use within-subject designs to explore whether these flips in effort–value relation can occur intraindividually, that is, whether the same individual can exhibit different profiles purely as a function of considering the effort in the past or the future. We expect this to be the case, especially given our discovery of individuals with nonmonotonic effort profiles. Critically, we showed that differences in effort profiles across situations but also within individuals can be captured by a relatively parsimonious formal model. Overall, our findings provide a more nuanced understanding of how individuals evaluate effortful outcomes across various decision contexts and effort levels.

Constraints on Generality

Our study investigates fundamental behavioral processes related to the role of effort in decision making, which are applicable across diverse populations. We did not target any specific population, aiming instead to balance male and female participants in our sample and ensure random sampling from our participant pool. While our findings are expected to generalize across various demographics, future research could verify this by replicating the study with more diverse samples, including different age groups, cultural backgrounds, and socioeconomic statuses.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705

Alessandri, J., Darcheville, J.-C., Delevoye-Turrell, Y., & Zentall, T. R. (2008). Preference for rewards that follow greater effort and greater delay. *Learning & Behavior*, *36*(4), 352–358. https://doi.org/10.3758/LB.36.4.352

Arkes, H. R., & Blumer, C. (1985). The psychology of sunk cost. *Organizational Behavior and Human Decision Processes*, *35*(1), 124–140. https://doi.org/10.1016/0749-5978(85)90049-4

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B, Statistical Methodology*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Białaszek, W., Marcowski, P., & Mizak, S. (2022). Everything comes at a price: Considerations in modeling effort-based choice. *Behavioural Processes*, *200*, Article 104692. https://doi.org/10.1016/j.beproc.2022.104692

Białaszek, W., Marcowski, P., & Ostaszewski, P. (2017). Physical and cognitive effort discounting across different reward magnitudes: Tests of discounting models. *PLOS ONE*, *12*(7), Article e0182353. https://doi.org/10.1371/journal.pone.0182353

Bogdanov, M., Renault, H., LoParco, S., Weinberg, A., & Otto, A. R. (2022). Cognitive effort exertion enhances electrophysiological responses to rewarding outcomes. *Cerebral Cortex*, *32*(19), 4255–4270. https://doi.org/10.1093/cercor/bhab480

Botvinick, M. M., Huffstetler, S., & McGuire, J. T. (2009). Effort discounting in human nucleus accumbens. *Cognitive, Affective & Behavioral Neuroscience*, *9*(1), 16–27. https://doi.org/10.3758/CABN.9.1.16

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach* (2nd ed.). Springer-Verlag. https://doi.org/10.1016/j.ecolmodel.2003.11.004

Cacioppo, J. T., & Berntson, G. G. (1999). The affect system: Architecture and operating characteristics. *Current Directions in Psychological Science*, *8*(5), 133–137. https://doi.org/10.1111/1467-8721.00031

Camerer, C., & Mobbs, D. (2017). Differences in behavior and brain activity during hypothetical and real choices. *Trends in Cognitive Sciences*, *21*(1), 46–56. https://doi.org/10.1016/j.tics.2016.11.001

Cavagnaro, D. R., Aranovich, G. J., McClure, S. M., Pitt, M. A., & Myung, J. I. (2016). On the functional form of temporal discounting: An optimized adaptive test. *Journal of Risk and Uncertainty*, *52*(3), 233–254. https://doi.org/10.1007/s11166-016-9242-y

Cavanagh, J. F., & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends in Cognitive Sciences*, *18*(8), 414–421. https://doi.org/10.1016/j.tics.2014.04.012

Decety, J., Jeannerod, M., Germain, M., & Pastene, J. (1991). Vegetative response during imagined movement is proportional to mental effort. *Behavioural Brain Research*, *42*(1), 1–5. https://doi.org/10.1016/S0166-4328(05)80033-6

Decety, J., Jeannerod, M., & Prablanc, C. (1989). The timing of mentally represented actions. *Behavioural Brain Research*, *34*(1–2), 35–42. https://doi.org/10.1016/S0166-4328(89)80088-9

Dixon, M. R., Lik, N. M. K., Green, L., & Myerson, J. (2013). Delay discounting of hypothetical and real money: The effect of holding reinforcement rate constant. *Journal of Applied Behavior Analysis*, *46*(2), 512–517. https://doi.org/10.1002/jaba.42

Doyle, J. R. (2012). Survey of time preference, delay discounting models. *SSRN Electronic Journal*, *8*(2), 116–135. https://doi.org/10.2139/ssrn.1685861

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

Ebert, J. E. J., & Prelec, D. (2007). The fragility of time: Time-insensitivity and valuation of the near and far future. *Management Science*, *53*(9), 1423–1438. https://doi.org/10.1287/mnsc.1060.0671

Eisenberger, R. (1992). Learned industriousness. *Psychological Review*, *99*(2), 248–267. https://doi.org/10.1037/0033-295X.99.2.248

Elkins-Brown, N., Saunders, B., & Inzlicht, M. (2016). Error-related electromyographic activity over the corrugator supercilii is associated with neural performance monitoring. *Psychophysiology*, *53*(2), 159–170. https://doi.org/10.1111/psyp.12556

Ericson, K. M., White, J. M., Laibson, D., & Cohen, J. D. (2015). Money earlier or later? Simple heuristics explain intertemporal choices better than delay discounting does. *Psychological Science*, *26*(6), 826–833. https://doi.org/10.1177/0956797615572232

Evans, J. S. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, *59*(1), 255–278. https://doi.org/10.1146/annurev.psych.59.103006.093629

Gheza, D., Kool, W., & Pourtois, G. (2023). Need for cognition moderates the relief of avoiding cognitive effort. *PLOS ONE*, *18*(11), Article e0287954. https://doi.org/10.1371/journal.pone.0287954

Green, L., & Myerson, J. (2004). A discounting framework for choice with delayed and probabilistic rewards. *Psychological Bulletin*, *130*(5), 769–792. https://doi.org/10.1037/0033-2909.130.5.769

Haaf, J. M., & Rouder, J. N. (2019). Some do and some don't? Accounting for variability of individual difference structures. *Psychonomic Bulletin & Review*, *26*(3), 772–789. https://doi.org/10.3758/s13423-018-1522-x

Hartmann, M. N., Hager, O. M., Tobler, P. N., & Kaiser, S. (2013). Parabolic discounting of monetary rewards by physical effort. *Behavioural Processes*, *100*, 192–196. https://doi.org/10.1016/j.beproc.2013.09.014

Holroyd, C. B., & Yeung, N. (2012). Motivation of extended behaviors by anterior cingulate cortex. *Trends in Cognitive Sciences*, *16*(2), 122–128. https://doi.org/10.1016/j.tics.2011.12.008

Hull, C. L. (1943). *Principles of behavior: An introduction to behavior theory*. Appleton-Century. https://psycnet.apa.org/record/1944-00022-000

Hurvich, C. M., & Tsai, C.-L. (1989). Regression and time series model selection in small samples. *Biometrika*, *76*(2), 297–307. https://doi.org/10.1093/biomet/76.2.297

Inzlicht, M., Bartholow, B. D., & Hirsh, J. B. (2015). Emotional foundations of cognitive control. *Trends in Cognitive Sciences*, *19*(3), 126–132. https://doi.org/10.1016/j.tics.2015.01.004

Inzlicht, M., Shenhav, A., & Olivola, C. Y. (2018). The effort paradox: Effort is both costly and valued. *Trends in Cognitive Sciences*, *22*(4), 337–349. https://doi.org/10.1016/j.tics.2018.01.007

Johnson, M. W., & Bickel, W. K. (2002). Within-subject comparison of real and hypothetical money rewards in delay discounting. *Journal of the Experimental Analysis of Behavior*, *77*(2), 129–146. https://doi.org/10.1901/jeab.2002.77-129

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, *58*(9), 697–720. https://doi.org/10.1037/0003-066X.58.9.697

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–292. https://doi.org/10.2307/1914185

Kang, M. J., Rangel, A., Camus, M., & Camerer, C. F. (2011). Hypothetical and real choice differentially activate common valuation areas. *The Journal of Neuroscience*, *31*(2), 461–468. https://doi.org/10.1523/JNEUROSCI.1583-10.2011

Killeen, P. R. (2009). An additive-utility model of delay discounting. *Psychological Review*, *116*(3), 602–619. https://doi.org/10.1037/a0016414

Klein-Flügge, M. C., Kennerley, S. W., Saraiva, A. C., Penny, W. D., & Bestmann, S. (2015). Behavioral modeling of human choices reveals dissociable effects of physical effort and temporal delay on reward devaluation. *PLOS Computational Biology*, *11*(3), Article e1004116. https://doi.org/10.1371/journal.pcbi.1004116

Kool, W., & Botvinick, M. (2018). Mental labour. *Nature Human Behaviour*, *2*(12), 899–908. https://doi.org/10.1038/s41562-018-0401-9

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

Kühberger, A., Schulte-Mecklenbeck, M., & Perner, J. (2002). Framing decisions: Hypothetical and real. *Organizational Behavior and Human Decision Processes*, *89*(2), 1162–1175. https://doi.org/10.1016/S0749-5978(02)00021-3

Ma, Q., Meng, L., Wang, L., & Shen, Q. (2014). I endeavor to make it: Effort increases valuation of subsequent monetary reward. *Behavioural Brain Research*, *261*, 1–7. https://doi.org/10.1016/j.bbr.2013.11.045

Massar, S. A. A., Libedinsky, C., Weiyan, C., Huettel, S. A., & Chee, M. W. L. (2015). Separate and overlapping brain areas encode subjective value during delay and effort discounting. *NeuroImage*, *120*, 104–113. https://doi.org/10.1016/j.neuroimage.2015.06.080

McClure, S. M., Ericson, K. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2007). Time discounting for primary rewards. *The Journal of Neuroscience*, *27*(21), 5796–5804. https://doi.org/10.1523/JNEUROSCI.4246-06.2007

McGuire, J. T., & Botvinick, M. M. (2010). Prefrontal cortex, cognitive control, and the registration of decision costs. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(17), 7922–7926. https://doi.org/10.1073/pnas.0910662107

Nishiyama, R. (2014). Response effort discounts the subjective value of rewards. *Behavioural Processes*, *107*, 175–177. https://doi.org/10.1016/j.beproc.2014.08.002

Nishiyama, R. (2016). Physical, emotional, and cognitive effort discounting in gain and loss situations. *Behavioural Processes*, *125*, 72–75. https://doi.org/10.1016/j.beproc.2016.02.004

Norton, M. I., Mochon, D., & Ariely, D. (2012). The IKEA effect: When labor leads to love. *Journal of Consumer Psychology*, *22*(3), 453–460. https://doi.org/10.1016/j.jcps.2011.08.002

Olivola, C. Y., & Shafir, E. (2018). Blood, sweat, and cheers: The martyrdom effect increases willingness to sponsor others' painful and effortful prosocial acts. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3101447

Otto, A. R., Braem, S., Silvetti, M., & Vassena, E. (2022). Is the juice worth the squeeze? Learning the marginal value of mental effort over time. *Journal of Experimental Psychology: General*, *151*(10), 2324–2341. https://doi.org/10.1037/xge0001208

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1–2), 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, *2*, Article 10. https://doi.org/10.3389/neuro.11.010.2008

Pellicano, A., Mingoia, G., Ritter, C., Buccino, G., & Binkofski, F. (2021). Respiratory function modulated during execution, observation, and imagination of walking via SII. *Scientific Reports*, *11*(1), Article 23752. https://doi.org/10.1038/s41598-021-03147-5

R Core Team. (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.r-project.org/

Saunders, B., Milyavskaya, M., & Inzlicht, M. (2015). What does cognitive control feel like? Effective and ineffective cognitive control is associated with divergent phenomenology. *Psychophysiology*, *52*(9), 1205–1217. https://doi.org/10.1111/psyp.12454

Sayalı, C., & Badre, D. (2019). Neural systems of cognitive demand avoidance. *Neuropsychologia*, *123*, 41–54. https://doi.org/10.1016/j.neuropsychologia.2018.06.016

Schrift, R. Y., Kivetz, R., & Netzer, O. (2016). Complicating decisions: The work ethic heuristic and the construction of effortful decisions. *Journal of Experimental Psychology: General*, *145*(7), 807–829. https://doi.org/10.1037/xge0000171

Schwarz, N., Jalbert, M., Noah, T., & Zhang, L. (2021), Metacognitive experiences as information: Processing fluency in consumer judgment and decision making. *Consumer Psychology Review*, *4,* 4–25. https://doi.org/10.1002/arcp.1067

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

Shenhav, A., Musslick, S., Lieder, F., Kool, W., Griffiths, T. L., Cohen, J. D., & Botvinick, M. M. (2017). Toward a rational and mechanistic account of mental effort. *Annual Review of Neuroscience*, *40*(1), 99–124. https://doi.org/10.1146/annurev-neuro-072116-031526

Smith, E. R., & DeCoster, J. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review*, *4*(2), 108–131. https://doi.org/10.1207/S15327957PSPR0402_01

Talmi, D., Dayan, P., Kiebel, S. J., Frith, C. D., & Dolan, R. J. (2009). How humans integrate the prospects of pain and reward during choice. *The Journal of Neuroscience*, *29*(46), 14617–14626. https://doi.org/10.1523/JNEUROSCI.2026-09.2009

Taylor, M. M., & Creelman, C. D. (1967). PEST: Efficient estimates on probability functions. *The Journal of the Acoustical Society of America*, *41*(4A), 782–787. https://doi.org/10.1121/1.1910407

Tibshirani, R. J., & Efron, B. (1993). *An introduction to the bootstrap* (1st ed.). CRC Press.

Westbrook, J. A., & Braver, T. S. (2015). Cognitive effort: A neuroeconomic approach. *Cognitive, Affective & Behavioral Neuroscience*, *15*(2), 395–415. https://doi.org/10.3758/s13415-015-0334-y

Westbrook, J. A., & Braver, T. S. (2013). The economics of cognitive effort. *Behavioral and Brain Sciences*, *36*(6), 704–705. https://doi.org/10.1017/S0140525X13001179

Winkielman, P., Schwarz, N., Fazendeiro, T. A., & Reber, R. (2003). The hedonic marking of processing fluency: Implications for evaluative judgment. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 189–217). Lawrence Erlbaum.

Wulff, D. U., & van den Bos, W. (2018). Modeling choices in delay discounting. *Psychological Science*, *29*(11), 1890–1894. https://doi.org/10.1177/0956797616664342

Yoo, J., Jasko, K., & Winkielman, P. (2024). Fluency, prediction and motivation: How processing dynamics, expectations and epistemic goals shape aesthetic judgements. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *379*(1895), Article 20230326. https://doi.org/10.1098/rstb.2023.0326

Zerna, J., Scheffel, C., Kührt, C., & Strobel, A. (2023). Need for cognition is associated with a preference for higher task load in effort discounting. *Scientific Reports*, *13*(1), Article 19501. https://doi.org/10.1038/s41598-023-44349-3

Received November 20, 2023
Revision received December 25, 2024
Accepted January 10, 2025 ■