Trading Mental Effort for Confidence in the Metacognitive Control of Value-Based Decision-Making

Douglas Lee[1,2] Jean Daunizeau[2,3]

[1] Sorbonne University, Paris, France
[2] Paris Brain Institute (ICM), Paris, France
[3] Translational Neuromodeling Unit (TNU), ETH, Zurich, Switzerland

Address for correspondence:
Jean Daunizeau
Motivation, Brain, and Behavior Group
Brain and Spine Institute
47, bvd de l'Hopital, 75013, Paris, France

Tel: +33 1 57 27 47 19
E-mail: jean.daunizeau@gmail.com

ABSTRACT

Why do we sometimes opt for actions or items that we do not value the most? Under current neurocomputational theories, such preference reversals are typically interpreted in terms of errors that arise from the unreliable signaling of value to brain decision systems. But, an alternative explanation is that people may change their mind because they are reassessing the value of alternative options while pondering the decision. So, why do we carefully ponder some decisions, but not others? In this work, we derive a computational model of the metacognitive control of decisions or MCD. In brief, we assume that fast and automatic processes first provide initial (and largely uncertain) representations of options' values, yielding prior estimates of decision difficulty. These uncertain value representations are then refined by deploying cognitive (e.g., attentional, mnesic) resources, the allocation of which is controlled by an effort-confidence tradeoff. Importantly, the anticipated benefit of allocating resources varies in a decision-by-decision manner according to the prior estimate of decision difficulty. The ensuing MCD model predicts response time, subjective feeling of effort, choice confidence, changes of mind, and choice-induced preference change and certainty gain. We test these predictions in a systematic manner, using a dedicated behavioral paradigm. Our results provide a quantitative link between mental effort, choice confidence, and preference reversals, which could inform interpretations of related neuroimaging findings.

1. INTRODUCTION

Why do we carefully ponder some decisions, but not others? Decisions permeate every aspect of our lives—what to eat, where to live, whom to date, etc.—but the amount of effort that we put into different decisions varies tremendously. Rather than processing all decision-relevant information, we often rely on fast habitual and/or intuitive decision policies, which can lead to irrational biases and errors (Kahneman et al., 1982). For example, snap judgments about others are prone to unconscious stereotyping, which often has enduring and detrimental consequences (Greenwald and Banaji, 1995). Yet we don't always follow the fast but negligent lead of habits or intuitions. So, what determines how much time and effort we invest when making decisions?

Biased and/or inaccurate decisions can be triggered by psychobiological determinants such as stress (Porcelli and Delgado, 2009; Porcelli et al., 2012), emotions (Harlé and Sanfey, 2007; Martino et al., 2006; Sokol-Hessner et al., 2013), or fatigue (Blain et al., 2016). But, in fact, they also arise in the absence of such contextual factors. That is why they are sometimes viewed as the outcome of inherent neurocognitive limitations on the brain's decision processes, e.g., bounded attentional and/or mnemonic capacity (Giguère and Love, 2013; Lim et al., 2011; Marois and Ivanoff, 2005), unreliable neural representations of decision-relevant information (Drugowitsch et al., 2016; Wang and Busemeyer, 2016; Wyart and Koechlin, 2016), or physiologically-constrained neural information transmission (Louie and Glimcher, 2012; Polanía et al., 2019). However, an alternative perspective is that the brain has a preference for efficiency over accuracy (Thorngate, 1980). For example, when making perceptual or motor decisions, people frequently trade accuracy for speed, even when time constraints are not tight (Heitz, 2014; Palmer et al., 2005). Related neural and behavioral data are best explained by "accumulation-to-bound" process models, in which a decision is emitted when the accumulated perceptual evidence reaches a bound (Gold and Shadlen, 2007; O'Connell et al., 2012; Ratcliff and McKoon, 2008; Ratcliff et al., 2016). Further computational work demonstrated that, if the bound is properly set, these models actually implement an optimal solution to speed-accuracy tradeoff problems (Ditterich, 2006; Drugowitsch et al., 2012). From a theoretical standpoint, this implies that accumulation-to-bound policies can be viewed as an evolutionary adaptation, in response to selective pressure that favors efficiency (Pirrone et al., 2014).

This line of reasoning, however, is not trivial to generalize to value-based decision making, for which objective accuracy remains an elusive notion (Dutilh and Rieskamp, 2016; Rangel et al., 2008). This is because, in contrast to evidence-based (e.g., perceptual) decisions, there are no right or wrong value-based decisions. Nevertheless, people still make choices that deviate from subjective reports of value, with a rate that decreases with value contrast. From the perspective of accumulation-to-bound models, these preference reversals count as errors and arise from the unreliable signaling of value to decision systems in the brain (Lim et al., 2013). That value-based variants of accumulation-to-bound models are able to capture the neural and behavioral effects of, e.g., overt attention (Krajbich et al., 2010; Lim et al., 2011), external time pressure (Milosavljevic et al., 2010), confidence (De Martino et al., 2013) or default preferences (Lopez-Persem et al., 2016) lends empirical support to this type of interpretation. Further credit also comes from theoretical studies showing that these process models, under some simplifying assumptions, optimally solve the problem of efficient value comparison (Tajima et al., 2016, 2019). However, they do not solve the issue of adjusting the optimal amount of effort to invest in reassessing an uncertain prior preference with yet-unprocessed value-relevant information. Here, we propose an alternative computational model of value-based decision-making that focuses on how value representations are modified

– as opposed to compared – while pondering decisions (Slovic, 1995; Tversky and Thaler, 1990; Warren et al., 2011).

We start from the premise that the brain generates representations of options' value in a quick and automatic manner, even before attention is engaged for making a decision (Lebreton et al., 2009). The brain also encodes the certainty of such value estimates (Lebreton et al., 2015), from which *a priori* feelings of choice difficulty and confidence could, in principle, be derived. Importantly, people are reluctant to make a choice that they are not confident about (De Martino et al., 2013). Thus, when faced with a difficult decision, people should reassess option values until they reach a satisfactory level of confidence about their preference. This effortful mental deliberation would engage neurocognitive resources, such as attention and memory, in order to process value-relevant information. In line with recent proposals regarding the strategic deployment of cognitive control (Musslick et al., 2015; Shenhav et al., 2013), we assume that the amount of allocated resources optimizes a tradeoff between expected effort cost and confidence gain. The main issue here is that the impact of yet-unprocessed information on value representations is *a priori* unknown. Critically, we show how the system can anticipate the expected benefit of allocating resources before having processed value-relevant information. The ensuing *metacognitive control of decisions* or *MCD* thus adjusts mental effort on a decision-by-decision basis, according to prior decision difficulty and importance (cf. Figure 1 below).

===== Insert Figure 1 here. =====

As we will see, the MCD model makes clear quantitative predictions about several key decision variables (cf. Model section below). We test these predictions by asking participants to report their judgments about each item's subjective value and their subjective certainty about their value judgements, both before and after choosing between pairs of the items. Note that we also measure choice confidence, response time, and subjective effort for each decision.

The objective of this work is to show how most non-trivial properties of value-based decision making can be explained with a minimal (and mutually consistent) set of assumptions. The MCD model predicts response time, subjective effort, choice confidence, probability of changing one's mind, and choice-induced preference change and certainty gain, out of two properties of pre-choice value representations, namely: value ratings and value certainty ratings. Relevant details regarding the model derivations, as well as the decision-making paradigm we designed to evaluate those predictions, can be found in the Model and Methods sections below. In the subsequent section, we present our main dual computational/behavioral results. Finally, we discuss our results in light of the existing literature on value-based decision making.

2. THE MCD MODEL

In what follows, we derive a computational model of the metacognitive control of decisions or MCD. In brief, we assume that the amount of cognitive resources that is deployed during a decision is controlled by an effort-confidence tradeoff. Critically, this tradeoff relies on a prospective anticipation of how these resources will perturb the internal representations of subjective values. As we will see, the MCD model eventually predicts how cognitive effort expenditure depends upon prior estimates of decision difficulty, and what impact this will have on post-choice value representations.

2.1 Deriving the expected value of decision control

Let z be the amount of cognitive (e.g., executive, mnemonic, or attentional) resources that serve to process value-relevant information. Allocating these resources will be associated with both a benefit $B(z)$, and a cost $C(z)$. As we will see, both are increasing functions of z: $B(z)$ derives from the refinement of internal representations of subjective values of alternative options or actions that compose the choice set, and $C(z)$ quantifies how aversive engaging cognitive resources is (mental effort). In line with the framework of expected value of control or EVC (Musslick et al., 2015; Shenhav et al., 2013), we assume that the brain chooses to allocate the amount of resources $\hat{z}$ that optimizes the following cost-benefit trade-off:

$$\hat{z} = \arg\max_z E\left[B(z) - C(z) \right] \tag{1}$$

where the expectation accounts for predictable stochastic influences that ensue from allocating resources (this will be clarified below). Note that the benefit term $B(z)$ is the (weighted) choice confidence $P_c(z)$:

$$B(z) = R \times P_c(z) \tag{2}$$

where the weight R is analogous to a reward and quantifies the importance of making a confident decision (see below). As we will see, $P_c(z)$ plays a pivotal role in the model, in that it captures the efficacy of allocating resources for processing value-relevant information. So, how do we define choice confidence?

We assume that decision makers may be unsure about how much they like/want the alternative options that compose the choice set. In other words, the internal representations of values V_i of alternative options are probabilistic. Such a probabilistic representation of value can be understood in terms of, for example, an uncertain prediction regarding the to-be-experienced value of a given option. Without loss of generality, the probabilistic representation of option value takes the form of Gaussian probability density functions, as follows:

$$p(V_i) = N(\mu_i, \sigma_i) \quad (3)$$

where μ_i and σ_i are the mode and the variance of the probabilistic value representation, respectively (and i indexes alternative options in the choice set).

This allows us to define choice confidence P_c as the probability that the (predicted) experienced value of the (to be) chosen item is higher than that of the (to be) unchosen item:

$$\begin{aligned} P_c &= \begin{cases} P(V_1 > V_2) & \text{if item \#1 is chosen} \\ P(V_2 > V_1) & \text{if item \#2 is chosen} \end{cases} \\ &= \begin{cases} P(V_1 > V_2) & \text{if } \Delta\mu > 0 \\ P(V_2 > V_1) & \text{if } \Delta\mu < 0 \end{cases} \\ &\approx s\left(\frac{\pi|\Delta\mu|}{\sqrt{3(\sigma_1 + \sigma_2)}} \right) \end{aligned} \quad (4)$$

where $s(x) = 1/1 + e^{-x}$ is the standard sigmoid mapping. Here the second line derives from assuming that the choice follows the sign of the preference $\Delta\mu = \mu_1 - \mu_2$, and the last line derives from a moment-matching approximation to the Gaussian cumulative density function (Daunizeau, 2017).

As stated in the Introduction section, we assume that the brain valuation system automatically generates uncertain estimates of options' value (Lebreton et al., 2009, 2015), before cognitive effort is invested in decision making. In what follows, μ_i^0 and σ_i^0 are the mode and variance of the ensuing prior value representations (we treat them as inputs to the MCD model). We also assume that these prior representations neglect existing value-relevant information that would require cognitive effort to be retrieved and processed (Lopez-Persem et al., 2016).

Now, how does the system anticipate the benefit of allocating resources to the decision process? Recall that the purpose of allocating resources is to process (yet unavailable) value-relevant information. The critical issue is thus to predict how both the uncertainty σ_i and the modes μ_i of value representations will eventually change, before having actually allocated the resources (i.e., without having processed the information). In brief, allocating resources essentially has two impacts: (i) it decreases the uncertainty σ_i, and (ii) it perturbs the modes μ_i in a stochastic manner.

The former impact derives from assuming that the amount of information that will be processed increases with the amount of allocated resources. Here, this implies that the variance of a given probabilistic value representation decreases in proportion to the amount of allocated effort, i.e.:

$$\sigma_i \cong \sigma_i(z) = \frac{1}{\frac{1}{\sigma_i^0} + \beta z} \quad (5)$$

where σ_i^0 is the prior variance of the representation (before any effort has been allocated), and β controls the efficacy with which resources increase the precision of the value representation. Formally speaking, Equation 5 has the form of a Bayesian update of the belief's precision in a Gaussian-likelihood model, where the precision of the likelihood term is βz. More precisely, β is the precision increase that follows from allocating a unitary amount of resources z. In what follows, we will refer to β as the "*type #1 effort efficacy*".

The latter impact follows from acknowledging the fact that the system cannot know how processing more value-relevant information will affect its preference before having allocated the corresponding resources. Let $\delta_i(z)$ be the change in the position of the mode of the i^{th} value representation, having allocated an amount z of resources. The direction of the mode's perturbation $\delta_i(z)$ cannot be predicted because it is tied to the information that would be processed. However, a tenable assumption is to consider that the magnitude of the perturbation increases with the amount of information that will be processed. This reduces to stating that the variance of $\delta_i(z)$ increases in proportion to z, i.e.:

$$\begin{aligned} &\mu_i(z) = \mu_i^0 + \delta_i \\ &\quad \delta_i \; : \; N(0, \gamma z) \end{aligned} \tag{6}$$

where μ_i^0 is the mode of the value representation before any effort has been allocated, and γ controls the relationship between the amount of allocated resources and the variance of the perturbation term δ. The higher γ, the greater the expected perturbation of the mode for a given amount of allocated resources. In what follows, we will refer to γ as the "*type #2 effort efficacy*". Note that Equation 6 treats the impact of future information processing as a non-specific random perturbation on the mode of the prior value representation. Our justification for this assumption is twofold: (i) it is simple, and (ii) and it captures the idea that the MCD controller does not know how the allocated resources will be used (here, by the value-based decision system downstream). We will see that, in spite of this, the MCD controller can still make quantitative predictions regarding the expected benefit of allocating resources.

Taken together, Equations 5 and 6 imply that predicting the net effect of allocating resources onto choice confidence is not trivial. On the one hand, allocating effort will increase the precision of value representations (cf. Equation 5), which mechanically increases choice confidence, all other things being equal. On the other hand, allocating effort can either increase or decrease the absolute difference $|\Delta\mu(z)|$ between the modes. This, in fact, depends upon the sign of the perturbation terms δ, which are not known in advance. Having said this, it is possible to derive the *expected* absolute difference between the modes that would follow from allocating an amount z of resources:

$$E\left[\left|\Delta\mu\right|\middle|z\right]=2\sqrt{\frac{\gamma z}{\pi}}\exp\left(-\frac{\left|\Delta\mu^{0}\right|^{2}}{4\gamma z}\right)+\Delta\mu^{0}\left(2\times s\left(\frac{\pi\,\Delta\mu^{0}}{\sqrt{6\gamma z}}\right)-1\right) \tag{7}$$

where we have used the expression for the first-order moment of the so-called "folded normal distribution", and the second term in the right-hand side of Equation 7 derives from the same moment-matching approximation to the Gaussian cumulative density function as above. The expected absolute means' difference $E\left[\left|\Delta\mu\right|\middle|z\right]$ depends upon both the absolute prior mean difference $\left|\Delta\mu^{0}\right|$ and the amount of allocated resources z. This is depicted on Figure 2 below.

===== Insert Figure 2 here. =====

One can see that $E\left[\left|\Delta\mu\right|\middle|z\right]-\left|\Delta\mu^{0}\right|$ is always greater than 0 and increases with z (and if $z=0$, then $E\left[\left|\Delta\mu\right|\middle|z\right]=\left|\Delta\mu^{0}\right|$). In other words, allocating resources is expected to increase the value difference, despite the fact that the impact of the perturbation term can go either way. In addition, the expected gain in value difference afforded by allocating resources decreases with the absolute prior means' difference.

Similarly, the variance $V\left[\left|\Delta\mu\right|\middle|z\right]$ of the absolute means' difference is derived from the expression of the second-order moment of the corresponding folded normal distribution:

$$V\left[\left|\Delta\mu\right|\middle|z\right]=2\gamma z+\left|\Delta\mu^{0}\right|^{2}-E\left[\left|\Delta\mu\right|\middle|z\right]^{2} \tag{8}$$

One can see on Figure 2 that $V\left[\left|\Delta\mu\right|\middle|z\right]$ increases with the amount z of allocated resources (but if $z=0$, then $V\left[\left|\Delta\mu\right|\middle|z\right]=0$).

Knowing the moments of the distribution of $|\Delta\mu|$ now enables us to derive the expected confidence level $\bar{P}_c(z)$ that would result from allocating the amount of resource z :

$$\begin{aligned}\bar{P}_c(z) &\triangleq E\left[P_c|z\right] \\ &= E\left[s\left(\frac{\pi|\Delta\mu|}{\sqrt{6\sigma(z)}}\right)\middle| z\right] \\ &\approx s\left(\frac{\pi E\left[|\Delta\mu|\middle| z\right]}{\sqrt{6\left(\sigma(z)+\frac{1}{2}V\left[|\Delta\mu|\middle| z\right]\right)}}\right)\end{aligned} \tag{9}$$

where we have assumed, for the sake of conciseness, that both prior value representations are similarly uncertain (i.e., $\sigma_1^0 \approx \sigma_2^0 \triangleq \sigma^0$). It turns out that the expected choice confidence $\bar{P}_c(z)$ always increase with z, irrespective of the efficacy parameters, as long as $\beta \neq 0$ or $\gamma \neq 0$. These, however, control the magnitude of the confidence gain that can be expected from allocating an amount z of resources. Equation 9 is important, because it quantifies the expected benefit of resource allocation, before having processed the ensuing value-relevant information. More details regarding the accuracy of Equation 9 can be found in section 1 of the Appendix. In addition, section 2 of the Appendix summarizes the dependence of MCD-optimal choice confidence on $|\Delta\mu^0|$ and σ^0.

To complete the cost-benefit model, we simply assume that the cost of allocating resources to the decision process linearly scales with the amount of resources, i.e.:

$$C(z) = \alpha z \tag{10}$$

where α determines the effort cost of allocating a unitary amount of resources z. In what follows, we will refer to α as the "effort unitary cost". We note that weak nonlinearities in the cost function (e.g., quadratic terms) would not qualitatively change the model predictions.

In brief, the MCD-optimal resource allocation $\hat{z} \triangleq \hat{z}(\alpha,\beta,\gamma)$ is simply given by:

$$\hat{z} = \arg\max_{z} \left[R \times \bar{P}_c(z) - \alpha z \right] \quad (11)$$

which does not have any closed-form analytic solution. Nevertheless, it can easily be identified numerically, having replaced Equations 7-9 into Equation 11. We refer the readers interested in the impact of model parameters $\{\alpha, \beta, \gamma\}$ on the MCD-optimal control to section 2 of the Appendix.

At this point, we do not specify how Equation 11 is solved by neural networks in the brain. Many alternatives are possible, from gradient ascent (Seung, 2003) to winner-take-all competition of candidate solutions (Mao and Massaquoi, 2007). We will also comment on the specific issue of prospective (offline) versus reactive (online) MCD processes in the Discussion section.

Note: implicit in the above model derivation is the assumption that the allocation of resources is similar for both alternative options in the choice set (i.e. $z_1 \approx z_2 @z$). This simplifying assumption is justified by eye-tracking data (cf. section 8 of the Appendix).

2.2 Corollary predictions of the MCD model

In the previous section, we derived the MCD-optimal resource allocation $\hat{z}$, which effectively best balances the expected choice confidence with the expected effort costs, given the predictable impact of stochastic perturbations that arise from processing value-relevant information. This quantitative prediction is effectively shown in Figures 5 and 6 of the Results section below, as a function of (empirical proxies for) the prior absolute difference between modes $\left|\Delta\mu^0\right|$ and the prior certainty $1/\sigma^0$ of value representations. But, this resource allocation mechanism has a few interesting corollary implications.

To begin with, note that knowing $\hat{z}$ enables us to predict what confidence level the system should eventually reach. In fact, one can define the MCD-optimal confidence level as the expected confidence evaluated at the MCD-optimal amount of allocated resources, i.e., $\bar{P}_c(\hat{z})$. This is important, because it implies that the model can predict both the effort the system will invest and its associated confidence, on a decision-by-decision basis. The impact of the efficacy parameters on this quantitative prediction is detailed in section 2 of the Appendix.

Additionally, $\hat{z}$ determines the expected improvement in the certainty of value representations (hereafter: the "certainty gain"), which trivially relates to type #2 efficacy, i.e.: $1/\sigma(\hat{z}) - 1/\sigma^0 = \beta\hat{z}$. This also means that, under the MCD model, no choice-induced value certainty gain can be expected when $\beta = 0$.

Similarly, one can predict the MCD-optimal probability of changing one's mind. Recall that the probability $Q(z)$ of changing one's mind depends on the amount of allocated resources z, i.e.:

$$
\begin{aligned}
Q(z) &\triangleq P\left(sign(\Delta\mu) \neq sign(\Delta\mu^0) \middle| z\right) \\
&= \begin{cases} P(\Delta\mu > 0 | z) & \text{if } \Delta\mu^0 < 0 \\ P(\Delta\mu < 0 | z) & \text{if } \Delta\mu^0 > 0 \end{cases} \\
&\approx s\left(-\frac{\pi|\Delta\mu^0|}{\sqrt{6\gamma z}}\right)
\end{aligned}
\tag{12}
$$

One can see that the MCD-optimal probability of changing one's mind $Q(\hat{z})$ is a simple monotonic function of the allocated effort $\hat{z}$. Importantly, $Q(z) = 0$ when $\gamma = 0$. This implies that MCD agents do not change their minds when effort cannot change the relative position of the modes of the options' value representations (irrespective of type #1 effort efficacy). In retrospect, this is critical because there should be no incentive at all to invest resources in deliberation, were one to have no possibility of changing one's pre-deliberation preference.

Lastly, we can predict the magnitude of choice-induced preference change, i.e., how value representations are supposed to spread apart during the decision. Such an effect is typically measured in terms of the so-called "spreading of alternatives" or SoA, which is defined as follows:

$$\begin{aligned} SOA &= \left(\mu_{chosen}^{(post-choice)} - \mu_{unchosen}^{(post-choice)}\right) - \left(\mu_{chosen}^{(pre-choice)} - \mu_{unchosen}^{(pre-choice)}\right) \\ &= \begin{cases} \Delta\mu(z) - \Delta\mu^0 & \text{if } \Delta\mu(z) > 0 \\ \Delta\mu^0 - \Delta\mu(z) & \text{if } \Delta\mu(z) < 0 \end{cases} \\ &= \begin{cases} \Delta\delta(z) & \text{if } \Delta\delta(z) > -\Delta\mu^0 \\ -\Delta\delta(z) & \text{if } \Delta\delta(z) < -\Delta\mu^0 \end{cases} \end{aligned} \tag{13}$$

where $\Delta\delta(z) : N(0, 2\gamma z)$ is the cumulative perturbation term of the modes' difference. Taking the expectation of the right-hand term of Equation 13 under the distribution of $\Delta\delta(z)$ and evaluating it at $z = \hat{z}$ now yields the MCD-optimal spreading of alternatives $\overline{SOA}(\hat{z})$:

$$\begin{aligned} \overline{SOA}(\hat{z}) &= E\left[SOA \middle| \hat{z}\right] \\ &= E\left[\Delta\delta(\hat{z}) \middle| \Delta\delta(\hat{z}) > -\Delta\mu^0\right] P\left(\Delta\delta(\hat{z}) > -\Delta\mu^0\right) \\ &\quad - E\left[\Delta\delta(\hat{z}) \middle| \Delta\delta(\hat{z}) < -\Delta\mu^0\right] P\left(\Delta\delta(\hat{z}) < -\Delta\mu^0\right) \\ &= 2\sqrt{\frac{\gamma\hat{z}}{\pi}} \exp\left(-\frac{\left|\Delta\mu^0\right|^2}{4\gamma\hat{z}}\right) \end{aligned} \tag{14}$$

where the last line derives from the expression of the first-order moment of the truncated Gaussian distribution. Note that the expected preference change also increases monotonically with the allocated effort $\hat{z}$. Here again, under the MCD model, no preference change can be expected when $\gamma = 0$.

We note that all of these corollary predictions essentially capture choice-induced modifications of value representations. This is why we will refer to choice confidence, value certainty gain, change of mind and spreading of alternatives as “decision-related” variables.

2.3 Correspondence between model variables and empirical measures

In summary, the MCD model predicts cognitive effort (or, more properly, the amount of allocated resources) and decision-related variables, given the prior absolute difference between modes $\left|\Delta\mu^0\right|$ and the prior certainty $1/\sigma^0$ of value representations. In other words, the inputs to the MCD model are the prior moments of value representations, whose trial-by-trial variations determine variations in model predictions. Here, we simply assume that pre-choice value and value certainty ratings provide us with an approximation of these prior moments. More precisely, we use ΔVR^0 and VCR^0 (cf. section 3.3 below) as empirical proxies for $\Delta\mu^0$ and $1/\sigma^0$, respectively. Accordingly, we consider post-choice value and value certainty ratings as empirical proxies for the posterior mean $\mu(\hat{z})$ and precision $1/\sigma(\hat{z})$ of value representations, at the time when the decision was triggered (i.e., after having invested the effort $\hat{z}$). Similarly, we match expected choice confidence $\bar{P}_c(z)$ (i.e., after having invested the effort $\hat{z}$) with empirical choice confidence.

Note that the MCD model does not specify *what* the allocated resources are. In principle, both mnesic and attentional resources may be engaged when processing value-relevant information. Nevertheless, what really matters is assessing the magnitude z of decision-related effort. We think of z as the cumulative engagement of neurocognitive resources, which varies both in terms of duration and intensity. Empirically, we relate $\hat{z}$ to two different "effort-related" empirical measures, namely: subjective feeling of effort and response time. The former relies on the subjective cost incurred when deploying neurocognitive resources, which would be signaled by experiencing mental effort. The latter makes sense if one thinks of response time in terms of effort duration. Although it is a more objective measurement than subjective rating of effort, response time only approximates $\hat{z}$ if effort intensity shows relatively small variations. We will comment on this in the Discussion section.

Finally, the MCD model is also agnostic about the definition of "decision importance", i.e. the weight R in Equation 2. In this work, we simply investigate the effect of decision importance by comparing subjective effort and response time in “neutral” versus “consequential” decisions (cf. section 3.4 below). We will also comment on this in the Discussion section.

3. METHODS

3.1 Participants

Participants for our study were recruited from the RISC (*Relais d’Information sur les Sciences de la Cognition*) subject pool through the ICM (*Institut du Cerveau et de la Moelle – Paris Brain Institute*). All participants were native French speakers, with no reported history of psychiatric or neurological illness. A total of 41 people (28 female; age: mean=28, stdev=5, min=20, max=40) participated in this study. The experiment lasted approximately 2 hours, and participants were paid a flat rate of 20€ as compensation for their time, plus a bonus, which was given to participants to compensate for potential financial losses in the “penalized” trials (see below). More precisely, in "penalized" trials, participants lost 0.20€ (out of a 5€ bonus) for each second that they took to make their choice. This resulted in an average 4€ bonus (across participants). One group of 11 participants was excluded from the cross-condition analysis only (see below), due to technical issues.

3.2 Materials

Written instructions provided detailed information about the sequence of tasks within the experiment, the mechanics of how participants would perform the tasks, and images illustrating what a typical screen within each task section would look like. The experiment was developed using Matlab and PsychToolbox, and was conducted entirely in French. The stimuli for this experiment were 148 digital images, each representing a distinct food item (50 fruits, 50 vegetables, 48 various snack items including nuts, meats, and cheeses). Food items were selected such that most items would be well known to most participants.

Eye gaze position and pupil size were continuously recorded throughout the duration of the experiment using The Eye Tribe eye-tracking devices. Participants' head positions were fixed using stationary chinrests. In case of incidental movements, we corrected the pupil size data for distance to screen, separately for each eye.

3.3 Task design

Prior to commencing the testing session of the experiment, participants underwent a brief training session. The training tasks were identical to the experimental tasks, although different stimuli were used (beverages). The experiment itself began with an initial section where all individual items were displayed in a random sequence for 1.5 seconds each, in order to familiarize the participants with the set of options they would later be considering and form an impression of the range of subjective value for the set. The main experiment was divided into three sections, following the classic Free-Choice Paradigm protocol (e.g., Izuma and Murayama, 2013): pre-choice item ratings, choice, and post-choice item ratings. There was no time limit for the overall experiment, nor for the different sections, nor for the individual trials. The item rating and choice sections are described below (see Figure 3).

===== Insert Figure 3 here. =====

Item rating (same for pre-choice and post-choice sessions): Participants were asked to rate the entire set of items in terms of how much they liked each item. The items were presented one at a time in a random sequence (pseudo-randomized across participants). At the onset of each trial, a fixation cross appeared at the center of the screen for 750ms. Next, a solitary image of a food item appeared at the center of the screen. Participants had to respond to the question, “How much do you like this item?” using a horizontal slider scale (from “I hate it!” to “I love it!”) to indicate their value rating for the item. The middle of the scale was the point of neutrality (“I don’t care about it.”). Hereafter, we refer to the reported value as the "pre-choice value rating". Participants then had to respond to the question, “What degree of certainty do you have?” (about the item’s value) by expanding a solid bar symmetrically around the cursor of the value slider scale to indicate the range of possible value ratings that would be compatible with their subjective feeling. We measured participants' certainty about value rating in terms of the percentage of the value scale that is not occupied by the reported range of compatible value ratings. We refer to this as the "pre-choice value certainty rating". At that time, the next trial began.

Note: In the Results section below, ΔVR^0 is the difference between pre-choice value ratings of items composing a choice set. Similarly, VCR^0 is the average pre-choice value certainty ratings across items composing a choice set. Both value and value certainty rating scales range from 0 to 1 (but participants were unaware of the quantitative units of the scales).

Choice: Participants were asked to choose between pairs of items in terms of which item they preferred. The entire set of items was presented one pair at a time in a random sequence. Each item appeared in only one pair. At the onset of each trial, a fixation cross appeared at the center of the screen for 750ms. Next, two images of snack items appeared on the screen: one towards the left and one towards the right. Participants had to respond to the question, "Which do you prefer?" using the left or right arrow key. We measured response time in terms of the delay between the stimulus onset and the response. Participants then had to respond to the question, "Are you sure about your choice?" using a vertical slider scale (from "Not at all!" to "Absolutely!"). We refer to this as the report of choice confidence. Finally, participants had to respond to the question, "To what extent did you think about this choice?" using a horizontal slider scale (from "Not at all!" to "Really a lot!"). We refer to this as the report of subjective effort. At that time, the next trial began.

3.4 Task conditions

We partitioned the task trials into three conditions, which were designed to test the following two predictions of the MCD model: all else equal, effort should increase with decision importance and decrease with related costs. We aimed to check the former prediction by asking participants to make some decisions where they knew that the choice would be real, i.e. they would actually have to eat the chosen food item at the end of the experiment. We refer to these trials as "consequential" decisions. To check the latter prediction, we imposed a financial penalty that increased with response time. More precisely, participants were instructed that they would lose 0.20€ (out of a 5€ bonus) for each second that they would take to make their choice. The choice section of the experiment was composed of 60 "neutral" trials, 7 "consequential" trials, and 7 "penalized" trials, which were randomly intermixed. Instructions for both "consequential" and "penalized" decisions were repeated at each relevant trial, immediately prior to the presentation of the choice items.

3.5 Probabilistic model fit

The MCD model predicts trial-by-trial variations in the probability of changing one's mind, choice confidence, spreading of alternatives, certainty gain, response time, and subjective effort ratings (MCD outputs) from trial-by-trial variations in value rating difference ΔVR^0 and mean value certainty rating VCR^0 (MCD inputs). Together, three unknown parameters control the quantitative relationship between MCD inputs and outputs: the *effort unitary cost* α, *type #1 effort efficacy* β, and *type #2 effort efficacy* γ. However, additional parameters are required to capture variations induced by experimental conditions. Recall that we expect "consequential" decisions to be more important than "neutral" ones, and "penalized" decisions effectively include an extraneous cost-of-time term. One can model the former condition effect by making R (cf. Equation 2) sensitive to whether the decision is consequential or not. We proxy the latter condition effect by making the effort unitary cost α a function of whether the decision is penalized (where effort induces both intrinsic and extrinsic costs) or not (intrinsic effort cost only). This means that condition effects require one additional parameter each.

In principle, all of these parameters may vary across people, thereby capturing idiosyncrasies in people's (meta-)cognitive apparatus. However, in addition to estimating these five parameters, fitting the MCD model to each participant's data also requires a rescaling of the model's output variables. This is because there is no reason to expect the empirical measure of these variables to match their theoretical scale. We thus inserted two additional nuisance parameters per output MCD variable, which operate a linear rescaling (affine transformation, with a positive constraint on slope parameters). Importantly, these nuisance parameters cannot change the relationship between MCD inputs and outputs. In other terms, the MCD model really has only five degrees of freedom.

For each subject, we fit all MCD dependent variables concurrently with a single set of MCD parameters. Within-subject probabilistic parameter estimation was performed using the variational Laplace approach (Daunizeau, 2018; Friston et al., 2007), which is made available from the VBA toolbox (Daunizeau et al., 2014). We refer the reader interested in the mathematical details of within-subject MCD parameter estimation to the section 3 of the Appendix (this also includes a parameter recovery analysis). In what follows, we compare empirical data to MCD-fitted dependent variables (when binned according to ΔVR^0 and VCR^0). We refer to the latter as "postdictions", in the sense that they derive from a posterior predictive density that is conditional on the corresponding data.

We also fit the MCD model on reduced subsets of dependent variables (e.g., only "effort-related" variables), and report proper out-of-sample predictions of data that were not used for parameter estimation (e.g., "decision-related" variables). We note that this is a strong test of the model, since it does not rely on any train/test partition of the predicted variable (see next section below).

4. RESULTS

Here, we test the predictions of the MCD model. We note that basic descriptive statistics of our data, including measures of test-retest reliability and replications of previously reported effects on confidence in value-based choices (De Martino et al., 2013), are appended in sections 5, 6 and 7 of the Appendix.

4.1 Within-subject model fit accuracy and out-of-sample predictions

To capture idiosyncrasies in participants' metacognitive control of decisions, the MCD model was fitted to subject-specific trial-by-trial data, where all MCD outputs (namely: change of mind, choice confidence, spreading of alternatives, value certainty gain, response time, and subjective effort ratings) were considered together. In the next section, we present summary statistics at the group level, which validate the predictions that can be derived from the MCD model, when fitted to all dependent variables. But can we provide even stronger evidence that the MCD model is capable of predicting all dependent variables at once? In particular, can the model make out-of-sample predictions regarding effort-related variables (i.e., RT and subjective effort ratings) given decision-related variables (i.e., choice confidence, change of mind, spreading of alternatives, and certainty gain), and *vice versa*?

To address this question, we performed two partial model fits: (i) with decision-related variables only, and (ii) with effort-related variables only. In both cases, out-of-sample predictions for the remaining dependent variables were obtained directly from within-subject parameter estimates. For each subject, we then estimated the cross-trial correlation between each pair of observed and predicted variables. Figure 4 below reports the ensuing group-average correlations, for each dependent variable and each model fit. In this context, the predictions derived when fitting the full dataset only serve as a reference point for evaluating the accuracy of out-of-sample predictions. For completeness, we also show chance-level prediction accuracy (i.e. the 95% percentile of group average correlations between observed and predicted variables under the null).

===== Insert Figure 4 here. =====

In what follows, we refer to model predictions on dependent variables that were actually fitted by the model as "postdictions" (full data fits: all dependent variables, partial model fits: variables included in the fit). As one would expect, the accuracy of postdictions is typically higher than that of out-of-sample predictions. Slightly more interesting, perhaps, is the fact that the accuracy of model predictions/postdictions depends upon which output variable is considered. For example, choice confidence is always better predicted/postdicted than spreading of alternatives. This is most likely because the latter data has lower reliability.

But the main result of this analysis is the fact that out-of-sample predictions of dependent variables perform systematically better than chance. In fact, all across-trial correlations between observed and predicted (out-of-sample) data where statistically significant at the group-level (all $p<10^{-3}$). In particular, this implies that the MCD model makes accurate out-of-sample predictions regarding effort-related variables given decision-related variables, and reciprocally.

4.2 Predicting effort-related variables

In what follows, we inspect the three-way relationships between pre-choice value and value certainty ratings and each effort-related variable: namely, RT and subjective effort rating. The former can be thought of as a proxy for the duration of resource allocation, whereas the latter is a metacognitive readout of resource allocation cost. Unless stated otherwise, we will focus on both the absolute difference between pre-choice value ratings (hereafter: $|\Delta VR^0|$) and the mean pre-choice value certainty rating across paired choice items (hereafter: VCR^0). Under the MCD model, increasing $|\Delta VR^0|$ and/or VCR^0 will decrease the demand for effort, which should result in smaller expected RT and subjective effort rating. We will now summarize the empirical data and highlight the corresponding quantitative MCD model postdictions and out-of-sample predictions (here: predictions are derived from model fits on decision-related variables only, i.e. all dependent variables except RT and subjective effort rating).

First, we checked how RT relates to pre-choice value and value certainty ratings. For each subject, we regressed (log-) RT data against $|\Delta VR^0|$ and VCR^0, and then performed a group-level random-effect analysis on regression weights. The results of this model-free analysis provide a qualitative summary of the impact of trial-by-trial variations in pre-choice value representations on RT. We also compare RT data with both MCD model postdictions (full data fit) and out-of-sample predictions. In addition to summarizing the results of the model-free analysis, Figure 5 below shows empirical, predicted, and postdicted RT data, when median-split (within subjects) according to both $|\Delta VR^0|$and VCR^0.

===== Insert Figure 5 here. =====

One can see that RT data behave as expected under the MCD model, i.e. RT decreases when $|\Delta VR^0|$and/or VCR^0 increases. The random effect analysis shows that both variables have a significant negative effect at the group level ($|\Delta VR^0|$: mean standardized regression weight=-0.16, s.e.m.=0.02, $p<10^{-3}$; CR^0: mean standardized regression weight=-0.08, s.e.m.=0.02, $p<10^{-3}$; one-sided t-tests). Moreover, MCD postdictions are remarkably accurate at capturing the effect of both $|\Delta VR^0|$and VCR^0 variables in a quantitative manner. Although MCD out-of-sample predictions are also very accurate, they tend to slightly underestimate the quantitative effect of $|\Delta VR^0|$. This is because this effect is typically less pronounced in decision-related variables than in effort-related variables (see below), which then yield MCD parameter estimates that eventually attenuate the impact of $|\Delta VR^0|$ on effort.

Second, we checked how subjective effort ratings relate to pre-choice value and value certainty ratings. We performed the same analysis as above, the results of which are summarized in Figure 6 below.

===== Insert Figure 6 here. =====

Here as well, subjective effort rating data behave as expected under the MCD model, i.e. subjective effort decreases when $|\Delta VR^0|$and/or VCR^0 increases. The random effect analysis shows that both variables have a significant negative effect at the group level ($|\Delta VR^0|$: mean standardized regression weight=-0.21, s.e.m.=0.03, $p<10^{-3}$; CR^0: mean regression weight=-0.05, s.e.m.=0.02, p=0.027; one-sided t-tests). One can see that MCD postdictions and out-of-sample predictions accurately capture the effect of both $|\Delta VR^0|$and VCR^0 variables. More quantitatively, we note that MCD postdictions slightly overestimate the effect VCR^0, whereas out-of-sample predictions also tend to underestimate the effect of $|\Delta VR^0|$.

At this point, we note that the MCD model makes two additional predictions regarding effort-related variables, which relate to our task conditions. In brief, all else equal, effort should increase in “consequential” trials, while it should decrease in “penalized” trials. To test these predictions, we modified the model-free regression analysis of RT and subjective effort ratings by including two additional subject-level regressors, encoding consequential and penalized trials, respectively. Figure 7 below shows the ensuing augmented set of standardized regression weights for both RT and subjective effort ratings.

===== Insert Figure 7 here. =====

First, we note that accounting for task conditions does not modify the statistical significance of the impact of $|\Delta VR^0|$ and VCR^0 on effort-related variables, except for the effect of VCR^0 on subjective effort ratings (p=0.09, one-sided t-test). Second, one can see that the impact of “consequential” and “penalized” conditions on effort-related variables globally conforms to MCD predictions. More precisely, both RT and subjective effort ratings were significantly higher for "consequential" decisions than for "neutral" decisions (log-RT: mean standardized regression weight=0.07, s.e.m.=0.03, p=0.036; effort ratings: mean standardized regression weight=0.12, s.e.m.=0.03, $p<10^{-3}$; one-sided t-tests). In addition, response times are significantly faster for "penalized" than for "neutral" decisions (mean standardized regression weight=-0.26, s.e.m.=0.03, $p<10^{-3}$; one-sided t-test). However, the difference in subjective effort ratings between "neutral" and "penalized" decisions does not reach statistical significance (mean effort difference=0.012, s.e.m.=0.024, p=0.66; two-sided t-test). We will comment on this in the Discussion section.

4.3 Predicting decision-related variables

Under the MCD model, “decision-related” dependent variables (i.e., choice confidence, change of mind, spreading of alternatives, and value certainty gain) are determined by the amount of resources allocated to the decision. However, their relationship to features of prior value representation is not trivial (see section 2 of the Appendix for the specific case of choice confidence). For this reason, we will recapitulate the qualitative MCD prediction that can be made about each of them, prior to summarizing the empirical data and its corresponding postdictions and out-of-sample predictions. Note that here, the latter are obtained from a model fit on effort-related variables only.

First, we checked how choice confidence relates to $|\Delta VR^0|$ and VCR^0. Under the MCD model, choice confidence reflects the discriminability of the options’ value representations after optimal resource allocation. Recall that more resources are allocated to the decision when either $|\Delta VR^0|$ or VCR^0 decreases. However, under moderate effort efficacies, this does not overcompensate decision difficulty, and thus choice confidence should decrease. As with effort-related variables, we regressed trial-by-trial confidence data against $|\Delta VR^0|$ and VCR^0, and then performed a group-level random-effect analysis on regression weights. The results of this analysis, as well as the comparison between empirical, predicted, and postdicted confidence data is shown in Figure 8 below.

===== Insert Figure 8 here. =====

The results of the group-level random effect analysis confirm our qualitative predictions. In brief, both $|\Delta VR^0|$ (mean standardized regression weight=0.25, s.e.m.=0.02, $p<10^{-3}$; one-sided t-test) and VCR^0 (mean standardized regression weight=0.16, s.e.m.=0.03, $p<10^{-3}$; one-sided t-test) have a significant positive impact on choice confidence. Here again, MCD postdictions and out-of-sample predictions are remarkably accurate at capturing the effect of both $|\Delta VR^0|$ and VCR^0 variables (though predictions slightly underestimate the effect of $|\Delta VR^0|$).

Second, we checked how change of mind relates to $|\Delta VR^0|$ and VCR^0. Note that we define a change of mind according to two criteria: (i) the choice is incongruent with the prior preference inferred from the pre-choice value ratings, and (ii) the choice is congruent with the posterior preference inferred from post-choice value ratings. The latter criterion distinguishes a change of mind from a mere "error", which may arise from attentional and/or motor lapses. Under the MCD model, we expect no change of mind unless type #2 efficacy $\gamma \neq 0$. In addition, the rate of change of mind should decrease when either $|\Delta VR^0|$ or VCR^0 increases. This is because increasing $|\Delta VR^0|$ and/or VCR^0 will decrease the demand for effort, which implies that the probability of reversing the prior preference will be smaller. Figure 9 below shows the corresponding model predictions/postdictions and summarizes the corresponding empirical data.

===== Insert Figure 9 here. =====

Note that, on average, the rate of change of mind reaches about 14.5% (s.e.m.=0.008, $p<10^{-3}$, one-sided t-test), which is significantly higher than the rate of "error" (mean rate difference=2.3%, s.e.m.=0.01, p=0.032; two-sided t-test). The results of the group-level random effect analysis confirm our qualitative MCD predictions. In brief, both $|\Delta VR^0|$ (mean standardized regression weight=-0.17, s.e.m.=0.02, $p<10^{-3}$; one-sided t-test) and VCR^0 (mean standardized regression weight=-0.08, s.e.m.=0.03, $p<10^{-3}$; one-sided t-test) have a significant negative impact on the rate of change of mind. Again, MCD postdictions and out-of-sample predictions are remarkably accurate at capturing the effect of both $|\Delta VR^0|$ and VCR^0 variables (though predictions slightly underestimate the effect of $|\Delta VR^0|$).

Third, we checked how spreading of alternatives relates to $|\Delta VR^0|$ and VCR^0. Recall that spreading of alternatives measures the magnitude of choice-induced preference change. Under the MCD model, the reported value of alternative options cannot spread apart unless type #2 efficacy $\gamma \neq 0$. In addition, and as with change of mind, spreading of alternatives should globally follow the optimal effort allocation, i.e. it should decrease when $|\Delta VR^0|$ and/or VCR^0 increase. Figure 10 below shows the corresponding model predictions/postdictions and summarizes the corresponding empirical data.

===== Insert Figure 10 here. =====

One can see that there is a significant positive spreading of alternatives (mean=0.04 A.U., s.e.m.=0.004, $p<10^{-3}$, one-sided t-test). This is reassuring, because it dismisses the possibility that $\gamma = 0$ (which would mean that effort does not perturb the mode of value representations). In addition, the results of the group-level random effect analysis confirm that both $|\Delta VR^0|$ (mean standardized regression weight=-0.09, s.e.m.=0.03, p=0.001; one-sided t-test) and VCR^0 (mean standardized regression weight=-0.04, s.e.m.=0.02, p=0.03; one-sided t-test) have a significant negative impact on spreading of alternatives. Note that this replicates previous findings on choice-induced preference change (Lee and Coricelli, 2020; Lee and Daunizeau, 2020). Finally, MCD postdictions and out-of-sample predictions accurately capture the effect of both $|\Delta VR^0|$and VCR^0 variables in a quantitative manner (though predictions slightly underestimate the effect of $|\Delta VR^0|$).

Fourth, we checked how $|\Delta VR^0|$ and VCR^0 impact value certainty gain. Under the MCD model, the certainty of value representations cannot improve unless type #1 efficacy $\beta \neq 0$. In addition, value certainty gain should globally follow the optimal effort allocation, i.e. it should decrease when $|\Delta VR^0|$ and/or VCR^0 increase. Figure 11 below shows the corresponding model predictions/postdictions and summarizes the corresponding empirical data.

===== Insert Figure 11 here. =====

Importantly, there is a small but significantly positive certainty gain (mean=0.11 A.U., s.e.m.=0.06, p=0.027, one-sided t-test). This is reassuring, because it dismisses the possibility that $\beta = 0$ (which would mean that effort does not increase the precision of value representation). This time, the results of the group-level random effect analysis only partially confirm our qualitative MCD predictions. In brief, although VCR^0 has a very strong negative impact on certainty gain (mean standardized regression weight=-0.61, s.e.m.=0.04, $p<10^{-3}$; one-sided t-test), the effect of $|\Delta VR^0|$ does not reach statistical significance (mean standardized regression weight=-0.009, s.e.m.=0.01, p=0.35; one-sided t-test). We note that a simple regression-to-the-mean artifact (Stigler, 1997) likely inflates the observed negative correlation between VCR^0 and certainty gain, beyond what would be predicted under the MCD model. Accordingly, both MCD postdictions and out-of-sample predictions clearly underestimate the effect of VCR^0 (and overestimate the effect of $|\Delta VR^0|$).

5. DISCUSSION

In this work, we have presented a novel computational model of decision-making that explains the intricate relationships between effort-related variables (response time, subjective effort) and decision-related variables (choice confidence, change of mind, spreading of alternatives, and choice-induced value certainty gain). This model assumes that deciding between alternative options whose values are uncertain induces a demand for allocating cognitive resources to value-relevant information processing. Cognitive resource allocation then optimally trades mental effort for confidence, given the prior discriminability of the value representations.

Such metacognitive control of decisions or MCD provides an alternative theoretical framework to accumulation-to-bound models of decision-making, e.g., drift-diffusion models or DDMs (Milosavljevic et al., 2010; Ratcliff et al., 2016; Tajima et al., 2016). Recall that DDMs assume that decisions are triggered once the noisy evidence in favor of a particular option has reached a predefined bound. Standard DDM variants make quantitative predictions regarding both response times and decision outcomes, but are agnostic about choice confidence, spreading of alternatives, value certainty gain, and/or subjective effort ratings. We note that simple DDMs are significantly less accurate than MCD at making out-of-sample predictions on dependent variables common to both models (e.g., change of mind). We refer the reader interested in the details of the MCD-DDM comparison to section 9 of the Appendix.

But how do MCD and accumulation-to-bound models really differ? For example, if the DDM can be understood as an optimal policy for value-based decision making (Tajima et al., 2016), then how can these two frameworks both be optimal? The answer lies in the distinct computational problems that they solve. The MCD solves the problem of finding the optimal amount of effort to invest under the possibility that yet-unprocessed value-relevant information might change the decision maker's mind. In fact, this resource allocation problem would be vacuous, would it not be possible to reassess preferences during the decision process. In contrast, the DDM provides an optimal solution to the problem of efficiently comparing option values, which may be unreliably signaled, but remain nonetheless stationary. Of course, the DDM decision variable (i.e., the "evidence" for a given choice option over the alternative) may fluctuate, e.g. it may first drift towards the upper bound, but then eventually reach the lower bound. This is the typical DDM's explanation for why people change their mind over the course of deliberation (Kiani et al., 2014; Resulaj et al., 2009). But, critically, these fluctuations are not caused by changes in the underlying value signal (i.e., the DDM's drift term). Rather, the fluctuations are driven by neural noise that corrupts the value signals (i.e., the DDM's diffusion term). This is why the DDM cannot predict choice-induced preference changes, or changes in options' values more generally. This distinction between MCD and DDM extends to other types of accumulation-to-bound models, including race models (De Martino et al, 2013; Tajima et al, 2019). We note that either of these models (DDM or race) could be equipped with pre-choice value priors (initial bias), and then driven with "true" values (drift term) derived from post-choice ratings. But then, simulating these models would require both pre-choice and post-choice ratings, which implies that choice-induced preference changes cannot be *predicted* from pre-choice ratings using a DDM. In contrast, the MCD model assumes that the value representations themselves are modified during the decision process, in proportion to the effort expenditure. Now the latter is maximal when prior value difference is minimal, at least when type #2 efficacy dominates (γ-effect, see section 2 of the Appendix). In turn, the MCD model predicts that the magnitude of (choice-induced) value spreading should decrease when the prior value difference increases (cf. Equation 14). Together with (choice-induced) value certainty gain, this quantitative prediction is unique to the MCD framework, and cannot be derived from existing variants of DDM.

As a side note, the cognitive essence of spreading of alternatives has been debated for decades. Its typical interpretation is that of "cognitive dissonance" reduction: if people feel uneasy about their choice, they later convince themselves that the chosen (rejected) item was actually better (worse) than they originally thought (Bem, 1967; Harmon-Jones et al., 2009; Izuma and Murayama, 2013). In contrast, the MCD framework would rather suggest that people tend to reassess value representations until they reach a satisfactory level of confidence prior to committing to their choice. Interestingly, recent neuroimaging studies have shown that spreading of alternatives can be predicted from brain activity measured during the decision

(Colosio et al, 2017; Jarcho, Berkman, & Lieberman, 2010; Kitayama et al, 2013; van Veen et al, 2009, Voigt et al, 2018). This is evidence against the idea that spreading of alternatives only occurs after the choice has been made. In addition, key regions of the brain's valuation and cognitive control systems are involved, including: the right inferior frontal gyrus, the ventral striatum, the anterior insula and the anterior cingulate cortex (ACC). This further corroborates the MCD interpretation, under the assumption that the ACC is involved in controlling the allocation of cognitive effort (Musslick et al., 2015; Shenhav et al., 2013). Having said this, both MCD and cognitive dissonance reduction mechanisms may contribute to spreading of alternatives, on top of its known statistical artifact component (Chen and Risen, 2010). The latter is a consequence of the fact that pre-choice value ratings may be unreliable, and is known to produce an apparent spreading of alternatives that decreases with pre-choice value difference (Izuma and Murayama, 2013). Although this pattern is compatible with our results, the underlying statistical confound is unlikely to drive our results. The reason is twofold. First, effort-related variables yield accurate within-subject out-of-sample predictions about spreading of alternatives (cf. Figure 10). Second, we have already shown that the effect of pre-choice value difference on spreading of alternatives is higher here than in a control condition where the choice is made after both rating sessions (Lee and Daunizeau, 2020).

A central tenet of the MCD model is that involving cognitive resources in value-related information processing is costly, which calls for an efficient resource allocation mechanism. A related notion is that information processing resources may be limited, in particular: value-encoding neurons may have a bounded firing range (Louie and Glimcher, 2012). In turn, "efficient coding" theory assumes that the brain has evolved adaptive neural codes that optimally account for such capacity limitations (Barlow, 1961; Laughlin, 1981). In our context, efficient coding implies that value-encoding neurons should optimally adapt their firing range to the prior history of experienced values (Polanía et al., 2019). When augmented with a Bayesian model of neural encoding/decoding (Wei and Stocker, 2015), this idea was successful in explaining the non-trivial relationship between choice consistency and the distribution of subjective value ratings. Both MCD and efficient coding frameworks assume that value representations are uncertain, which stresses the importance of metacognitive processes in decision-making control (Fleming and Daw, 2017). However, they differ in how they operationalize the notion of efficiency. In efficient coding, the system is "efficient" in the sense that it changes the physiological properties of value-encoding neurons to minimize the information loss that results from their limited firing range. In MCD, the system is "efficient" in the sense that it allocates the amount of resources that optimally trades effort cost against choice confidence. These two perspectives may not be easy to reconcile. A possibility is to consider, for example, energy-efficient population codes (Hiratani and Latham, 2020; Yu et al., 2016), which would tune the amount of neural resources involved in representing value to optimally trade information loss against energetic costs.

Now, let us highlight that the MCD model offers a plausible alternative interpretation for the two main reported neuroimaging findings regarding confidence in value-based choices (De Martino et al., 2013). First, the ventromedial prefrontal cortex or vmPFC was found to respond positively to both value difference (i.e., ΔVR^0) and choice confidence. Second, the right rostrolateral prefrontal cortex or rRLPFC was more active during low-confidence versus high-confidence choices. These findings were originally interpreted through a so-called "race model", in which a decision is triggered whenever the first of option-specific value accumulators reaches a bound. Under this model, choice confidence is defined as the final gap between the two value accumulators. We note that this scenario predicts the same three-way relationship between response time, choice outcome, and choice confidence as the MCD model (see section 7 of the Appendix). In brief, rRLPFC was thought to perform a readout of choice confidence (for the purpose of subjective metacognitive report) from the racing value accumulators hosted in the vmPFC. Under the MCD framework, the contribution of the vmPFC to value-based decision control might rather be to construct item values, and to anticipate and monitor the benefit of effort investment (i.e., confidence). This would be consistent with recent fMRI studies suggesting that vmPFC confidence computations signal the attainment of task goals (Hebscher and Gilboa, 2016; Lebreton et al., 2015). Now, recall that the MCD model predicts that confidence and effort should be anti-correlated. Thus, the puzzling negative correlation between choice confidence and rRLPFC activity could be simply explained under the assumption that rRLPFC provides the neurocognitive resources that are instrumental for processing value-relevant information during decisions (and/or to compare item values). This resonates with the known involvement of rRLPFC in reasoning (Desrochers et al., 2015; Dumontheil, 2014) or memory retrieval (Benoit et al., 2012; Westphal et al., 2019).

At this point, we note that the current MCD model clearly has limited predictive power. Arguably, this limitation is partly due to the imperfect reliability of the data, and to the fact that MCD does not model all decision-relevant processes. In addition, assigning variations in many effort- and/or decision-related variables to a unique mechanism with few degrees of freedom necessarily restricts the model's expected predictive power. Nevertheless, the MCD model may also not yield a sufficiently tight approximation to the mechanism that it focuses on. In turn, it may unavoidably distort the impact of prior value representations and other decision input variables. The fact that it can only explain 81% of the variability in dependent variables that can be captured using simple linear regressions against ΔVR0 and VCR0 (see section 11 of the Appendix) supports this notion. A likely explanation here is that the MCD model includes constraints that prevent it from matching the model-free postdiction accuracy level. In turn, one may want to extend the MCD model with the aim of relaxing these constraints. For example, one may allow for deviations from the optimal resource allocation framework, e.g., by considering candidate systematic biases whose magnitudes would be controlled by specific additional parameters. Having said this, some of these constraints may be necessary, in the sense that they derive from the modeling assumptions that enable the MCD model to provide a unified explanation for all dependent variables (and thus make out-of-sample predictions). What follows is a discussion of what we perceive as the main limitations of the current MCD model, and the directions of improvement they suggest.

First, we did not specify what determines decision "importance", which effectively acts as a weight for confidence against effort costs (cf. R in Equation 2 of the Model section). We know from the comparison of "consequential" and "neutral" choices that increasing decision importance eventually increases effort, as predicted by the MCD model. However, decision importance may have many determinants, such as, for example, the commitment duration of the decision (e.g., life partner choices), the breadth of its repercussions (e.g., political decisions), or its instrumentality with respect to the achievement of superordinate goals (e.g., moral decisions). How these determinants are combined and/or moderated by the decision context is virtually unknown (Locke and Latham, 2002, 2006). In addition, decision importance may also be influenced by the prior (intuitive/emotional/habitual) appraisal of choice options. For example, we found that, all else equal, people spent more time and effort deciding between two disliked items than between two liked items (results not shown). This reproduces recent results regarding the evaluation of choice sets (Shenhav and Karmarkar, 2019). One may also argue that people should care less about decisions between items that have similar values (Oud et al., 2016). In other terms, decision importance would be an increasing function of the absolute difference in pre-choice value ratings. However, this would predict that people invest fewer resources when deciding between items of similar pre-choice values, which directly contradicts our results (cf. Figures 5 and 6). Importantly, options with similar values may still be very different from each other, when decomposed on some value-relevant feature space. For example, although two food items may be similarly liked and/or wanted, they may be very different in terms of, e.g., tastiness and healthiness, which would induce some form of decision conflict (Hare et al., 2009). In such a context, making a decision effectively implies committing to a preference about feature dimensions. This may be deemed to be consequential, when contrasted with choices between items that are similar in all regards. In turn, decision importance may rather be a function of options' feature conflict. In principle, this alternative possibility is compatible with our results, under the assumption that options' feature conflict is approximately orthogonal to pre-choice value difference. Considering how decision importance varies with feature conflict may significantly improve the amount of explained trial-by-trial variability in the model's dependent variables. We note that the brain's quick/automatic assessment of option features may also be the main determinant of the prior value representations that eventually serve to compute the MCD-optimal resource allocation. Probing these computational assumptions will be the focus of forthcoming publications.

Second, our current version of the MCD model relies upon a simple variant of resource costs and efficacies. One may thus wonder how sensitive model predictions are to these assumptions. For example, one may expect that type #2 efficacy saturates, i.e. that the magnitude of the perturbation $\delta(z)$ to the modes $\mu(z)$ of the value representations eventually reaches a plateau instead of growing linearly with z (cf. Equation 6). We thus implemented and tested such a model variant. We report the results of this analysis in section 10 of the Appendix. In brief, a saturating type #2 efficacy brings no additional explanatory power for the model's dependent variables. Similarly, rendering the cost term nonlinear (e.g., quadratic) does not change the qualitative nature of the MCD predictions. More problematic, perhaps, is the fact that we did not consider distinct types of effort, which could, in principle, be associated with different costs and/or efficacies. For example, the efficacy of allocating attention may depend upon which option is considered. In turn, the brain may dynamically refocus its attention on maximally-uncertain options when prospective information gains exceed switch costs (Callaway et al., 2021; Jang et al., 2021). Such optimal adjustment of divided attention might eventually explain systematic decision biases and shortened response times for “default” choices (Lopez-Persem et al., 2016). Another possibility is that effort might be optimized along two canonical dimensions, namely: duration and intensity. The former dimension essentially justifies the fact that we used RT as a proxy for the amount of allocated resources. This is because, if effort intensity stays constant, then longer RT essentially signals greater resource expenditure. In fact, as is evident from the comparison between “penalized” and “neutral” choices, imposing an external penalty cost on RT reduces, as expected, the ensuing effort duration. More generally, however, the dual optimization of effort dimensions might render the relationship between effort and RT more complex. For example, beyond memory span or attentional load, effort intensity could be related to processing speed. This would explain why, although "penalized" choices are made much faster than "neutral" choices, the associated subjective feeling of effort is not as strongly impacted as RT (cf. Figure 7). In any case, the relationship between effort and RT might depend upon the relative costs and/or efficacies of effort duration and intensity, which might themselves be partially driven by external availability constraints (cf. time pressure or multitasking). We note that the essential nature of the cost of mental effort in cognitive tasks (e.g., neurophysiological cost, interferences cost, opportunity cost) is still a matter of intense debate (Kurzban et al., 2013;

Musslick et al., 2015; Ozcimder et al., 2017). Progress towards addressing this issue will be highly relevant for future extensions of the MCD model.

Third, we did not consider the issue of identifying plausible neuro-computational implementations of MCD. This issue is tightly linked to the previous one, in that distinct cost types would likely impose different constraints on candidate neural network architectures (Feng et al., 2014; Petri et al., 2017). For example, underlying brain circuits are likely to operate MCD in a more reactive manner, eventually adjusting resource allocation from the continuous monitoring of relevant decision variables (e.g., experienced costs and benefits). Such a reactive process contrasts with our current, prospective-only variant of MCD, which sets resource allocation based on anticipated costs and benefits. We already checked that simple reactive scenarios, where the decision is triggered whenever the online monitoring of effort or confidence reaches the optimal threshold, make predictions qualitatively similar to those we have presented here. We tend to think however, that such reactive processes should be based upon a dynamic programming perspective on MCD, as was already done for the problem of optimal efficient value comparison (Tajima et al., 2016, 2019). We will pursue this and related neuro-computational issues in subsequent publications.

DATA AVAILABILITY

The data that support the findings of this study are available for download at https://doi.org/10.5061/dryad.7h44j0zsg.

CODE AVAILABILITY

The computer code and algorithms that support the findings of this study will soon be made available from the open academic freeware VBA (http://mbb-team.github.io/VBA-toolbox/). Until then, they are available from the corresponding author upon reasonable request.

ETHICAL COMPLIANCE

This study complies with all relevant ethical regulations and received formal approval from the INSERM Ethics Committee (CEEI-IRB00003888, decision no 16-333). In particular, in accordance with the Helsinki declaration, all participants gave written informed consent prior to commencing the experiment, which included consent to disseminate the results of the study via publication.

FUNDING

DL was supported by a grant from the Laboratory of Excellence of Biology for Psychiatry (LabEx BIO-PSY, Paris, France).

REFERENCES

Barlow, H. (1961). Possible principles underlying the transformations of sensory messages. Sens. Commun. 217–234.

Bem, D.J. (1967). Self-perception: An alternative interpretation of cognitive dissonance phenomena. Psychol. Rev. *74*, 183–200.

Benoit, R.G., Gilbert, S.J., Frith, C.D., and Burgess, P.W. (2012). Rostral Prefrontal Cortex and the Focus of Attention in Prospective Memory. Cereb. Cortex *22*, 1876–1886.

Blain, B., Hollard, G., and Pessiglione, M. (2016). Neural mechanisms underlying the impact of daylong cognitive work on economic decisions. Proc. Natl. Acad. Sci. *113*, 6967–6972.

Callaway, F., Rangel, A., and Griffiths, T.L. (2021). Fixation patterns in simple choice reflect optimal information sampling. PLOS Comput. Biol. *17*, e1008863.

Chen, K.M., and Risen, J.L. (2010). How choice affects and reflects preferences: Revisiting the free-choice paradigm. J. Pers. Soc. Psychol. *99*, 573–594.

Daunizeau, J. (2017). Semi-analytical approximations to statistical moments of sigmoid and softmax mappings of normal variables. ArXiv170300091 Q-Bio Stat.

Daunizeau, J. (2018). The variational Laplace approach to approximate Bayesian inference. ArXiv170302089 Q-Bio Stat.

Daunizeau, J., Adam, V., and Rigoux, L. (2014). VBA: A Probabilistic Treatment of Nonlinear Models for Neurobiological and Behavioural Data. PLoS Comput Biol *10*, e1003441.

De Martino, B., Fleming, S.M., Garrett, N., and Dolan, R.J. (2013). Confidence in value-based choice. Nat. Neurosci. *16*, 105–110.

Desrochers, T.M., Chatham, C.H., and Badre, D. (2015). The necessity of rostrolateral prefrontal cortex for higher-level sequential behavior. Neuron *87*, 1357–1368.

Ditterich, J. (2006). Evidence for time-variant decision making. Eur. J. Neurosci. *24*, 3628–3641.

Drugowitsch, J., Moreno-Bote, R., Churchland, A.K., Shadlen, M.N., and Pouget, A. (2012). The Cost of Accumulating Evidence in Perceptual Decision Making. J. Neurosci. *32*, 3612–3628.

Drugowitsch, J., Wyart, V., Devauchelle, A.-D., and Koechlin, E. (2016). Computational Precision of Mental Inference as Critical Source of Human Choice Suboptimality. Neuron *92*, 1398–1411.

Dumontheil, I. (2014). Development of abstract thinking during childhood and adolescence: The role of rostrolateral prefrontal cortex. Dev. Cogn. Neurosci. *10*, 57–76.

Dutilh, G., and Rieskamp, J. (2016). Comparing perceptual and preferential decision making. Psychon. Bull. Rev. *23*, 723–737.

Feng, S.F., Schwemmer, M., Gershman, S.J., and Cohen, J.D. (2014). Multitasking versus multiplexing: Toward a normative account of limitations in the simultaneous execution of control-demanding behaviors. Cogn. Affect. Behav. Neurosci. *14*, 129–146.

Fleming, S.M., and Daw, N.D. (2017). Self-Evaluation of Decision-Making: A General Bayesian Framework for Metacognitive Computation. Psychol. Rev. *124*, 91–114.

Friston, K., Mattout, J., Trujillo-Barreto, N., Ashburner, J., and Penny, W. (2007). Variational free energy and the Laplace approximation. NeuroImage *34*, 220–234.

Giguère, G., and Love, B.C. (2013). Limits in decision making arise from limits in memory retrieval. Proc. Natl. Acad. Sci. *110*, 7613–7618.

Gold, J.I., and Shadlen, M.N. (2007). The neural basis of decision making. Annu. Rev. Neurosci. *30*, 535–574.

Greenwald, A.G., and Banaji, M.R. (1995). Implicit social cognition: attitudes, self-esteem, and stereotypes. Psychol. Rev. *102*, 4–27.

Hare, T.A., Camerer, C.F., and Rangel, A. (2009). Self-control in decision-making involves modulation of the vmPFC valuation system. Science *324*, 646–648.

Harlé, K.M., and Sanfey, A.G. (2007). Incidental sadness biases social economic decisions in the Ultimatum Game. Emot. Wash. DC *7*, 876–881.

Harmon-Jones, E., Amodio, D.M., and Harmon-Jones, C. (2009). Action-based model of dissonance: A review, integration, and expansion of conceptions of cognitive conflict. In Advances in Experimental Social Psychology, Vol 41, (San Diego, CA, US: Elsevier Academic Press), pp. 119–166.

Hebscher, M., and Gilboa, A. (2016). A boost of confidence: The role of the ventromedial prefrontal cortex in memory, decision-making, and schemas. Neuropsychologia *90*, 46–58.

Heitz, R.P. (2014). The speed-accuracy tradeoff: history, physiology, methodology, and behavior. Front. Neurosci. *8*.

Hiratani, N., and Latham, P.E. (2020). Developmental and evolutionary constraints on olfactory circuit selection. BioRxiv 2020.12.22.423799.

Izuma, K., and Murayama, K. (2013). Choice-Induced Preference Change in the Free-Choice Paradigm: A Critical Methodological Review. Front. Psychol. *4*.

Jang, A.I., Sharma, R., and Drugowitsch, J. (2021). Optimal policy for attention-modulated decisions explains human fixation behavior. ELife *10*, e63436.

Kahneman, D., Slovic, P., and Tversky, A. (1982). Judgment Under Uncertainty: Heuristics and Biases (Cambridge University Press).

Kiani, R., Cueva, C.J., Reppas, J.B., and Newsome, W.T. (2014). Dynamics of neural population responses in prefrontal cortex indicate changes of mind on single trials. Curr. Biol. CB *24*, 1542–1547.

Krajbich, I., Armel, C., and Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. Nat. Neurosci. *13*, 1292–1298.

Kurzban, R., Duckworth, A., Kable, J.W., and Myers, J. (2013). An opportunity cost model of subjective effort and task performance. Behav. Brain Sci. *36*, 661–679.

Laughlin, S. (1981). A simple coding procedure enhances a neuron's information capacity. Z. Naturforsch. [C] *36*, 910–912.

Lebreton, M., Jorge, S., Michel, V., Thirion, B., and Pessiglione, M. (2009). An Automatic Valuation System in the Human Brain: Evidence from Functional Neuroimaging. Neuron *64*, 431–439.

Lebreton, M., Abitbol, R., Daunizeau, J., and Pessiglione, M. (2015). Automatic integration of confidence in the brain valuation signal. Nat. Neurosci. *18*, 1159–1167.

Lee, D., and Coricelli, G. (2020). An Empirical Test of the Role of Value Certainty in Decision Making. Front. Psychol. *11*.

Lee, D., and Daunizeau, J. (2020). Choosing what we like vs liking what we choose: How choice-induced preference change might actually be instrumental to decision-making. PLOS ONE *15*, e0231081.

Lee, D., and Usher, M. (2020). Value Certainty in Drift-Diffusion Models of Preferential Choice. BioRxiv 2020.08.22.262725.

Lim, S.-L., O'Doherty, J.P., and Rangel, A. (2011). The Decision Value Computations in the vmPFC and Striatum Use a Relative Value Code That is Guided by Visual Attention. J. Neurosci. *31*, 13214–13223.

Lim, S.-L., O'Doherty, J.P., and Rangel, A. (2013). Stimulus Value Signals in Ventromedial PFC Reflect the Integration of Attribute Value Signals Computed in Fusiform Gyrus and Posterior Superior Temporal Gyrus. J. Neurosci. *33*, 8729–8741.

Locke, E.A., and Latham, G.P. (2002). Building a practically useful theory of goal setting and task motivation. A 35-year odyssey. Am. Psychol. *57*, 705–717.

Locke, E.A., and Latham, G.P. (2006). New Directions in Goal-Setting Theory. Curr. Dir. Psychol. Sci. *15*, 265–268.

Lopez-Persem, A., Domenech, P., and Pessiglione, M. (2016). How prior preferences determine decision-making frames and biases in the human brain. ELife *5*, e20317.

Louie, K., and Glimcher, P.W. (2012). Efficient coding and the neural representation of value. Ann. N. Y. Acad. Sci. *1251*, 13–32.

Mao, Z.-H., and Massaquoi, S.G. (2007). Dynamics of winner-take-all competition in recurrent neural networks with lateral inhibition. IEEE Trans. Neural Netw. *18*, 55–69.

Marois, R., and Ivanoff, J. (2005). Capacity limits of information processing in the brain. Trends Cogn. Sci. *9*, 296–305.

Martino, B.D., Kumaran, D., Seymour, B., and Dolan, R.J. (2006). Frames, Biases, and Rational Decision-Making in the Human Brain. Science *313*, 684–687.

Milosavljevic, M., Malmaud, J., Huth, A., Koch, C., and Rangel, A. (2010). The drift diffusion model can account for value-based choice response times under high and low time pressure. Judgm. Decis. Mak. *5*, 437–449.

Musslick, S., Shenhav, A., Botvinick, M., and D Cohen, J. (2015). A Computational Model of Control Allocation based on the Expected Value of Control. p.

O'Connell, R.G., Dockree, P.M., and Kelly, S.P. (2012). A supramodal accumulation-to-bound signal that determines perceptual decisions in humans. Nat. Neurosci. *15*, 1729–1735.

Oud, B., Krajbich, I., Miller, K., Cheong, J.H., Botvinick, M., and Fehr, E. (2016). Irrational time allocation in decision-making. Proc. R. Soc. B Biol. Sci. *283*, 20151439.

Ozcimder, K., Dey, B., Musslick, S., Petri, G., Ahmed, N.K., Willke, T.L., and Cohen, J.D. (2017). A Formal Approach to Modeling the Cost of Cognitive Control. ArXiv170600085 Q-Bio.

Palmer, J., Huk, A.C., and Shadlen, M.N. (2005). The effect of stimulus strength on the speed and accuracy of a perceptual decision. J. Vis. *5*, 376–404.

Petri, G., Musslick, S., Dey, B., Ozcimder, K., Ahmed, N.K., Willke, T., and Cohen, J.D. (2017). Universal limits to parallel processing capability of network architectures. ArXiv170803263 Q-Bio.

Pirrone, A., Stafford, T., and Marshall, J.A.R. (2014). When natural selection should optimize speed-accuracy trade-offs. Front. Neurosci. *8*.

Polanía, R., Woodford, M., and Ruff, C.C. (2019). Efficient coding of subjective value. Nat. Neurosci. *22*, 134–142.

Porcelli, A.J., and Delgado, M.R. (2009). Acute stress modulates risk taking in financial decision making. Psychol. Sci. *20*, 278–283.

Porcelli, A.J., Lewis, A.H., and Delgado, M.R. (2012). Acute Stress Influences Neural Circuits of Reward Processing. Front. Neurosci. *6*.

Rangel, A., Camerer, C., and Montague, P.R. (2008). A framework for studying the neurobiology of value-based decision making. Nat. Rev. Neurosci. *9*, 545–556.

Ratcliff, R., and McKoon, G. (2008). The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks. Neural Comput. *20*, 873–922.

Ratcliff, R., Smith, P.L., Brown, S.D., and McKoon, G. (2016). Diffusion Decision Model: Current Issues and History. Trends Cogn. Sci. *20*, 260–281.

Resulaj, A., Kiani, R., Wolpert, D.M., and Shadlen, M.N. (2009). Changes of mind in decision-making. Nature *461*, 263–266.

Seung, H.S. (2003). Learning in Spiking Neural Networks by Reinforcement of Stochastic Synaptic Transmission. Neuron *40*, 1063–1073.

Shenhav, A., and Karmarkar, U.R. (2019). Dissociable components of the reward circuit are involved in appraisal versus choice. Sci. Rep. *9*, 1958.

Shenhav, A., Botvinick, M.M., and Cohen, J.D. (2013). The Expected Value of Control: An Integrative Theory of Anterior Cingulate Cortex Function. Neuron *79*, 217–240.

Slovic, P. (1995). The construction of preference. Am. Psychol. *50*, 364–371.

Sokol-Hessner, P., Camerer, C.F., and Phelps, E.A. (2013). Emotion regulation reduces loss aversion and decreases amygdala responses to losses. Soc. Cogn. Affect. Neurosci. *8*, 341–350.

Stigler, S.M. (1997). Regression towards the mean, historically considered. Stat. Methods Med. Res. *6*, 103–114.

Tajima, S., Drugowitsch, J., and Pouget, A. (2016). Optimal policy for value-based decision-making. Nat. Commun. *7*, 12400.

Tajima, S., Drugowitsch, J., Patel, N., and Pouget, A. (2019). Optimal policy for multi-alternative decisions. Nat. Neurosci. *22*, 1503–1511.

Thorngate, W. (1980). Efficient decision heuristics. Behav. Sci. *25*, 219–225.

Tversky, A., and Thaler, R.H. (1990). Anomalies: Preference Reversals. J. Econ. Perspect. *4*, 201–211.

Wang, Z., and Busemeyer, J.R. (2016). Interference effects of categorization on decision making. Cognition *150*, 133–149.

Warren, C., McGraw, A.P., and Van Boven, L. (2011). Values and preferences: defining preference construction. Wiley Interdiscip. Rev. Cogn. Sci. *2*, 193–205.

Wei, X.-X., and Stocker, A.A. (2015). A Bayesian observer model constrained by efficient coding can explain "anti-Bayesian" percepts. Nat. Neurosci. *18*, 1509–1517.

Westphal, A.J., Chow, T.E., Ngoy, C., Zuo, X., Liao, V., Storozuk, L.A., Peters, M.A.K., Wu, A.D., and Rissman, J. (2019). Anodal Transcranial Direct Current Stimulation to the Left Rostrolateral Prefrontal Cortex Selectively Improves Source Memory Retrieval. J. Cogn. Neurosci. *31*, 1380–1391.

Wyart, V., and Koechlin, E. (2016). Choice variability and suboptimality in uncertain environments. Curr. Opin. Behav. Sci. *11*, 109–115.

Yu, L., Zhang, C., Liu, L., and Yu, Y. (2016). Energy-efficient population coding constrains network size of a neuronal array system. Sci. Rep. *6*, 19369.

Figure Captions

Figure 1. The Metacognitive Control of Decisions. First, automatic processes provide a "pre-effort" belief about option values. This belief is probabilistic, in the sense that it captures an uncertain prediction regarding the to-be-experienced value of a given option. This pre-effort belief serves to identify the anticipated impact of investing costly cognitive resources (i.e., effort) in the decision. In particular, investing effort is expected to increase decision confidence beyond its pre-effort level. But how much effort it should be worth investing depends upon the balance between expected confidence gain and effort costs. The system then allocates resources into value-relevant information processing up until the optimal effort investment is reached. At this point, a decision is triggered based upon the current post-effort belief about option values (in this example, the system has changed its mind, i.e. its preference has reversed). Note: we refer to the ensuing increase in the value difference between chosen and unchosen items as the “spreading of alternatives” (cf. Methods section).

Figure 2. The expected impact of allocated resources onto value representations. Left panel: the expected absolute mean difference $E\left[|\Delta\mu(z)|\middle|z\right]$ (y-axis) is plotted as a function of the absolute prior mean difference $|\Delta\mu^0|$ (x-axis) for different amounts z of allocated resources (color code), having set type #2 effort efficacy to unity (i.e. $\gamma = 1$). **Right panel**: Variance $V\left[|\Delta\mu(z)|\middle|z\right]$ of the absolute mean difference ; same format.

Figure 3. Experimental design. **Left**: pre-choice item rating session: participants are asked to rate how much they like each food item and how certain they are about it (value certainty rating). **Center**: choice session: participants are asked to choose between two food items, to rate how confident they are about their choice, and to report the feeling of effort associated with the decision. **Right**: post-choice item rating session (same as pre-choice item rating session).

Figure 4: Accuracy of model postdictions and out-of-sample predictions. The mean within-subject (across-trial) correlation between observed and predicted/postdicted data (y-axis) is plotted for each variable (x-axis, from left to right: choice confidence, spreading of alternatives, change of mind, certainty gain, RT and subjective effort ratings), and each fitting procedure (grey: full data fit, blue: decision-related variables only, and red: effort-related variables only). Errorbars depict standard error of the mean, and the horizontal dashed black line shows chance-level prediction accuracy.

Figure 5. Three-way relationship between RT, value, and value certainty. **Left panel:** Mean standardized regression weights for $|\Delta VR^0|$ and VCR^0 on log-RT (*cst* is the constant term); errorbars represent s.e.m. **Right panel:** Mean z-scored log-RT (y-axis) is shown as a function of $|\Delta VR^0|$ (x-axis) and VCR^0 (color code: blue=0-50% lower quantile, green= 50-100% upper quantile); solid lines indicate empirical data (errorbars represent s.e.m.), star-dotted lines show out-of-sample predictions and diamond-dashed lines represent model postdictions.

Figure 6. Three-way relationship between subjective effort rating, value, and value certainty. Same format as Figure 5.

Figure 7. Impact of consequential and penalized conditions on effort-related variables. **Left panel:** log-RT: mean standardized regression weights (same format as Figure 4 – left panel, *cons* = "consequential" condition, *pena* = "penalized" condition). **Right panel:** subjective effort ratings: same format as left panel.

Figure 8. Three-way relationship between choice confidence, value, and value certainty. Same format as Figure 5.

Figure 9. Three-way relationship between change of mind, value, and value certainty. Same format as Figure 5.

Figure 10. Three-way relationship between spreading of alternatives, value, and value certainty. Same format as Figure 5.

Figure 11. Three-way relationship between value certainty gain, value, and value certainty. Same format as Figure 5.

Appendix-Figure 1: Quality of the analytical approximation to $\bar{P}$. Upper left panel: the Monte-Carlo estimate of $\bar{P}$ (colour-coded) is shown as a function of both the mean $\mu \in [-4,4]$ (y-axis) and the variance $\sigma^2 \in [0,4]$ (x-axis) of the parent process $x \sim N(\mu, \sigma^2)$. **Upper right panel**: analytic approximation to $\bar{P}$ as given by Equation A3 (same format). **Lower left panel**: the error, i.e. the difference between the Monte-Carlo and the analytic approximation (same format). **Lower right panel**: the analytic approximation (y-axis) is plotted as a function of the Monte-Carlo estimate (x-axis) for each pair of moments $\{\mu, \sigma^2\}$ of the parent distribution.

Appendix-Figure 2. The β-effect: MCD-optimal effort and confidence when effort has no impact on the value difference. MCD-optimal effort (left) and confidence (right) are shown as a function of the absolute prior mean difference $|\Delta\mu^0|$ (x-axis) and prior variance σ^0 (y-axis).

Appendix-Figure 3. The γ-effect: MCD-optimal effort and confidence when effort has no impact on value precision. Same format as Appendix-Figure 2.

Appendix-Figure 4. MCD-optimal effort and confidence when both types of effort efficacy are operant. Same format as Appendix-Figure 2.

Appendix-Figure 5: Comparison of simulated and estimated MCD parameters. Left panel: estimated parameters (y-axis) are plotted against simulated parameters (x-axis). Each dot is a Monte-Carlo simulation and different colors indicate distinct parameters (blue: efficacy type #1, red: efficacy type #2, yellow: unknown weight of consequential choices on decision importance, violet: intrinsic cost of effort, green: unknown weight of penalized choices on effort cost). The black dotted line indicates the identity line (perfect estimation). **Right panel:** Parameter recovery matrix**:** each line shows the squared partial correlation coefficient between a given estimated parameter and each simulated parameter (across 1000 Monte-Carlo simulations). Diagonal elements of the recovery matrix measure "correct estimation variability", i.e. variations in the estimated parameters that are due to variations in the corresponding simulated parameter. In contrast, non-diagonal elements of the recovery matrix measure "incorrect estimation variability", i.e. variations in the estimated parameters that are due to variations in other parameters. Perfect recovery would thus exhibit a diagonal structure, where variations in each estimated parameter are only due to variations in the corresponding simulated parameter. In contrast, strong non-diagonal elements in recovery matrices signal pairwise non-identifiability issues.

Appendix-Figure 6. Relationship between choices, pre-choice value ratings and choice confidence. Left panel: the probability of choosing the item on the right (y-axis) is shown as a function of the pre-choice value difference (x-axis), for high- (blue) versus low- (red) confidence trials. The plain lines show the logistic prediction that would follow from group-averages of the corresponding slope estimates. **Right panel**: the corresponding logistic regression slope (y-axis) is shown for both high- (blue) and low- (red) confidence trials (group means +/- s.e.m.).

Appendix-Figure 7. Relationship between pre-choice value ratings, choice confidence and response times. Left panel: response times (y-axis) are plotted as a function of low- and high- $|\Delta VR^0|$ (x-axis) for both low- (red) and high- (blue) confidence trials. Errorbars represent s.e.m. **Right panel**: A heatmap of mean z-scored confidence is shown as a function of both response time (x-axis) and $|\Delta VR^0|$ (y-axis).

Appendix-Figure 8. Correlation between pupil size and subjective effort ratings during decision time. Left panel: Mean (+/- s.e.m.) correlation between pupil size and subjective effort (y-axis) is plotted as a function of peristimulus time (x-axis). Here, epochs are co-registered w.r.t. stimulus onset (the green line indicates stimulus onset and the red dotted line indicates the average choice response). **Right panel**: Same, but for epochs co-registered w.r.t. choice response (the green line indicates choice response and the red dotted line indicates the average stimulus onset).

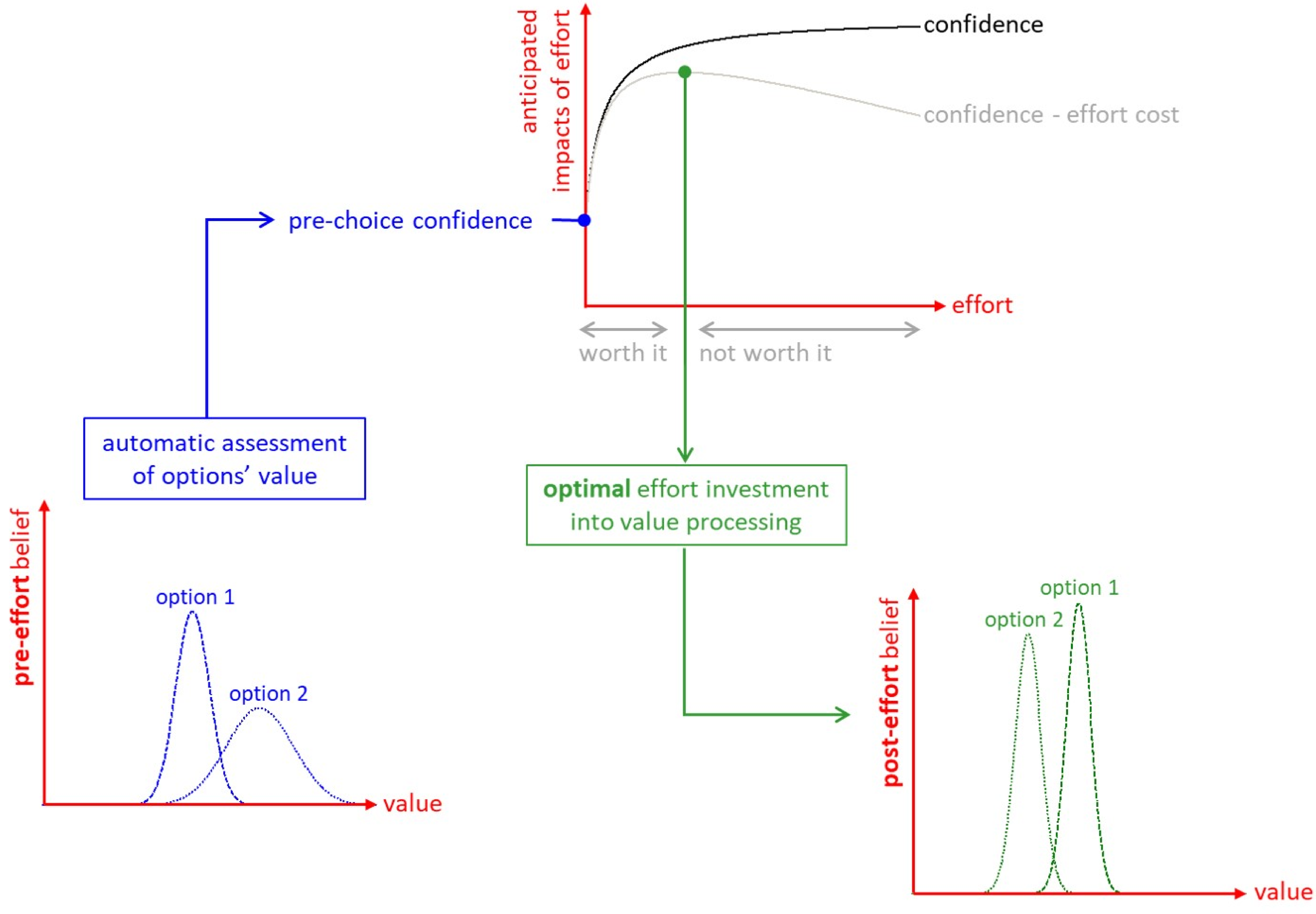
Appendix-Figure 9. Gaze bias for low and high effort trials. Mean (+/- s.e.m.) gaze bias is plotted for both low (left) and high (right) effort trials.

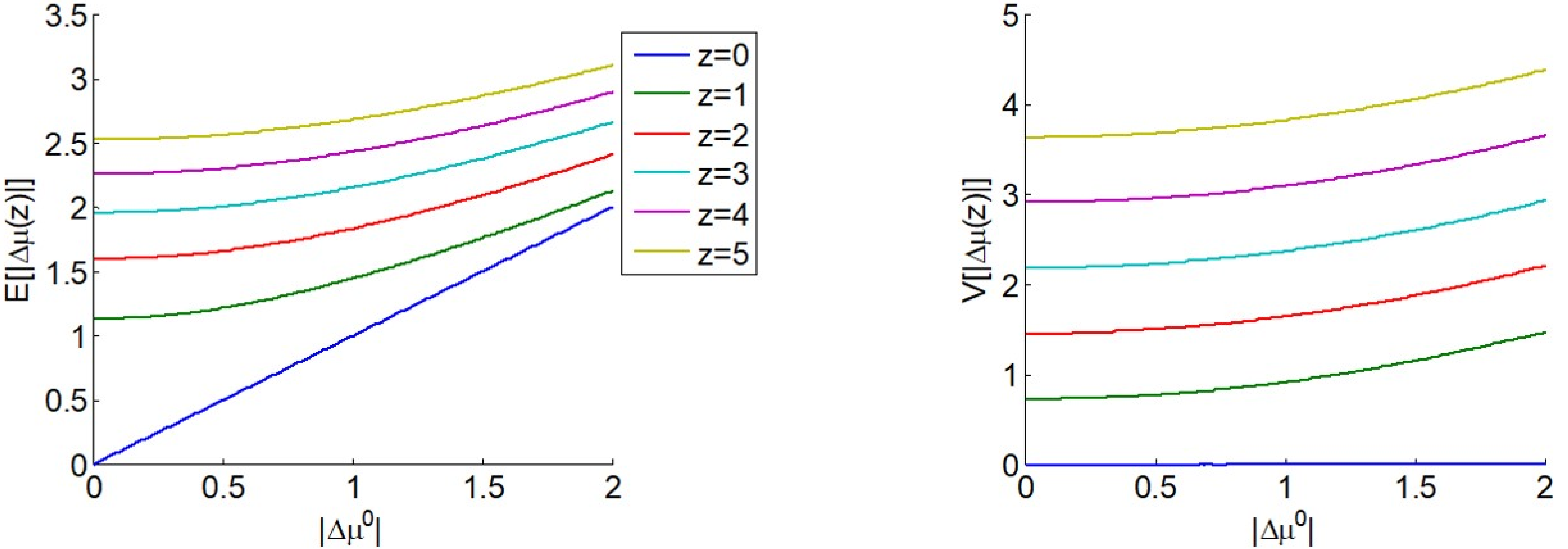
Appendix-Figure 10: Accuracy of RT postdictions. **Left panel:** The mean within-subject (across-trial) correlation between observed and postdicted RT data (y-axis) is plotted for each model (grey: MCD, blue: DDM1 and DDM2); errorbars depict s.e.m. **Right panel:** Mean z-scored log-RT (y-axis) is shown as a function of $|\Delta VR^0|$ (x-axis) and VCR^0 (color code: blue=0-50% lower quantile, green= 50-100% upper quantile); solid lines indicate empirical data (errorbars represent s.e.m.), diamond-dashed lines represent DDM1 postdictions and star-dotted lines show DDM2 postdictions.

Appendix-Figure 11: Accuracy of out-of-sample change of mind postdictions. Same format as Appendix-Figure 10.

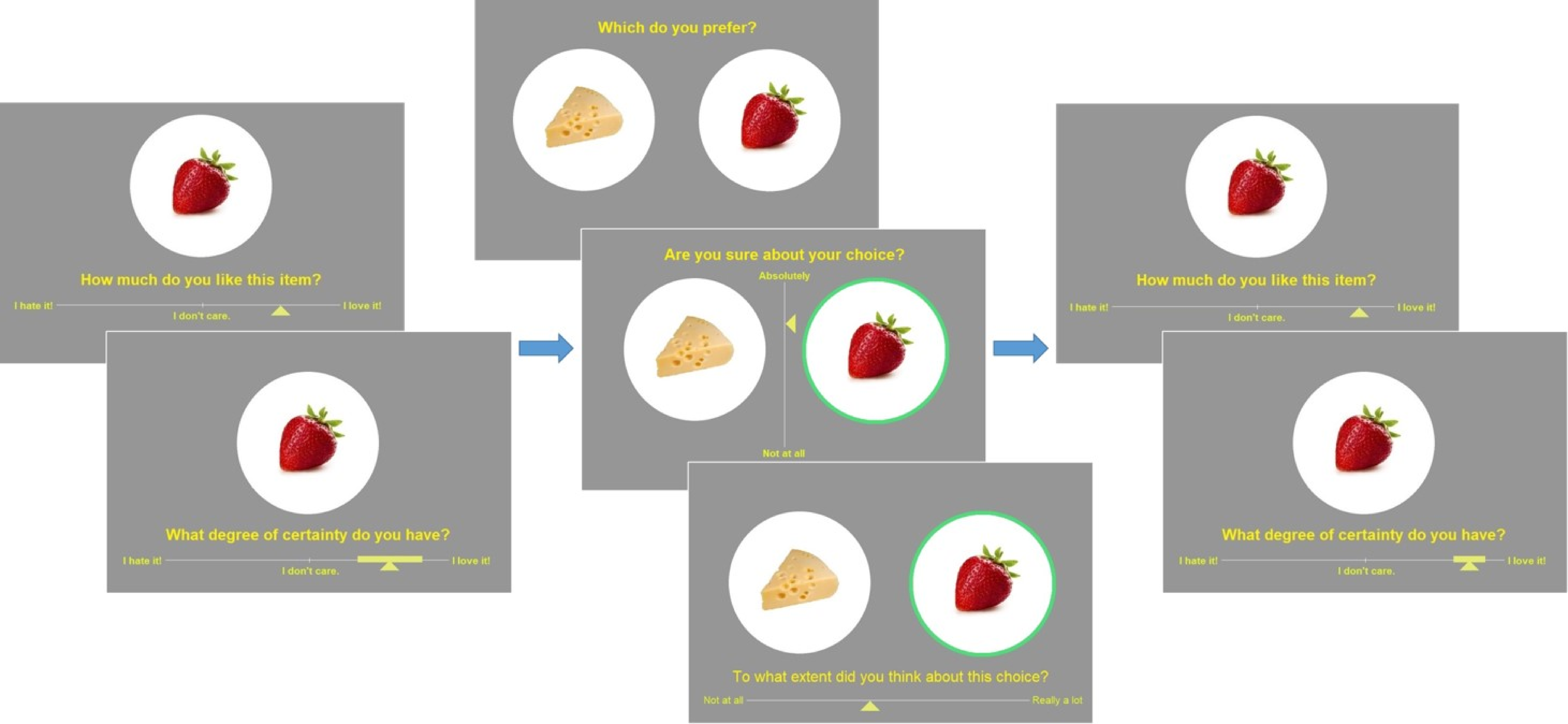
Appendix-Figure 12: Comparisons of MCD model with linear and saturating γ-effects. **Left panel:** The mean within-subject (across-trial) correlation between observed and postdicted data (y-axis) is plotted for dependent variable (x-axis, from left to right: choice confidence, spreading of alternatives, change of mind, certainty gain, RT and subjective effort ratings) and each model (grey: MCD with linear efficacy, blue: MCD with saturating efficacy); errorbars depict s.e.m. **Right panel:** Estimated model frequencies from the random-effect group-level Bayesian model comparison; errorbars depict posterior standard deviations.

Appendix-Figure 13: Comparisons of MCD and model-free postdiction accuracies. The mean within-subject (across-trial) correlation between observed and postdicted data (y-axis) is plotted for each variable (x-axis, from left to right: choice confidence, spreading of alternatives, change of mind, certainty gain, RT and subjective effort ratings), and each fitting procedure (grey: MCD full data fit, white: MCD 1-variable fit, and black: linear regression). Errorbars depict standard error of the mean.

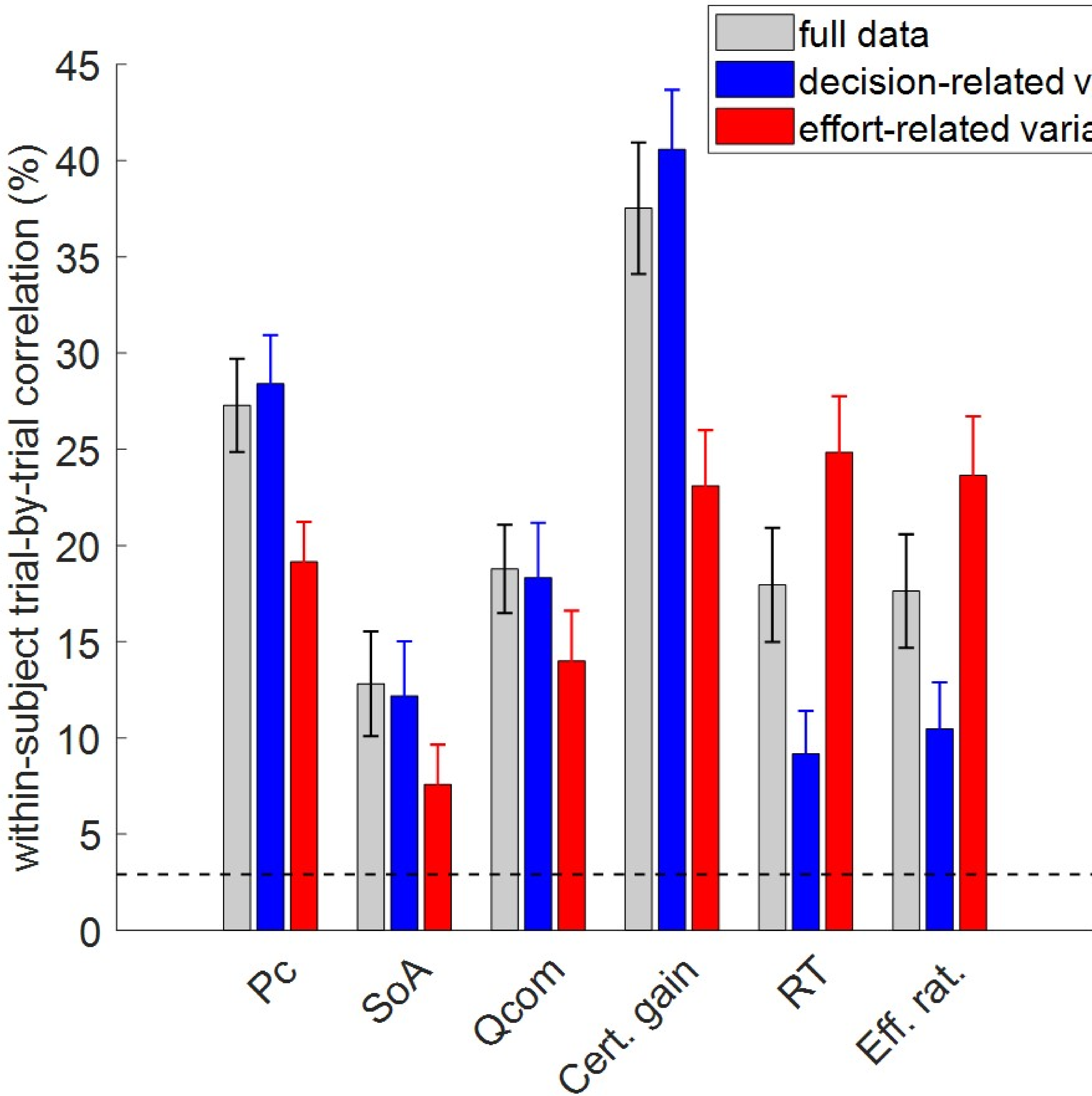

anticipated impacts of effort
confidence
confidence - effort cost
pre-choice confidence
effort
worth it
not worth it
automatic assessment of options' value
optimal effort investment into value processing
pre-effort belief
option 1
option 2
value
post-effort belief
option 1
option 2
value



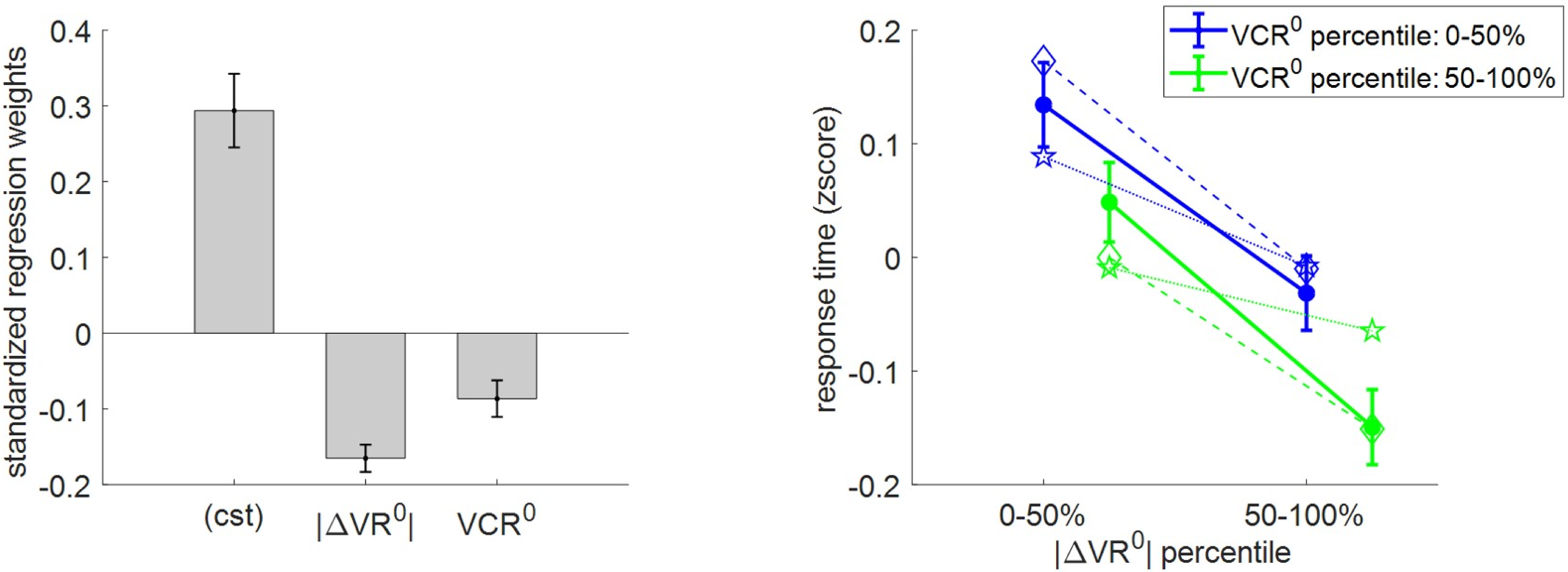
E[|Δμ(z)|]
|Δμ⁰|
z=0
z=1
z=2
z=3
z=4
z=5
V[|Δμ(z)|]
|Δμ⁰|



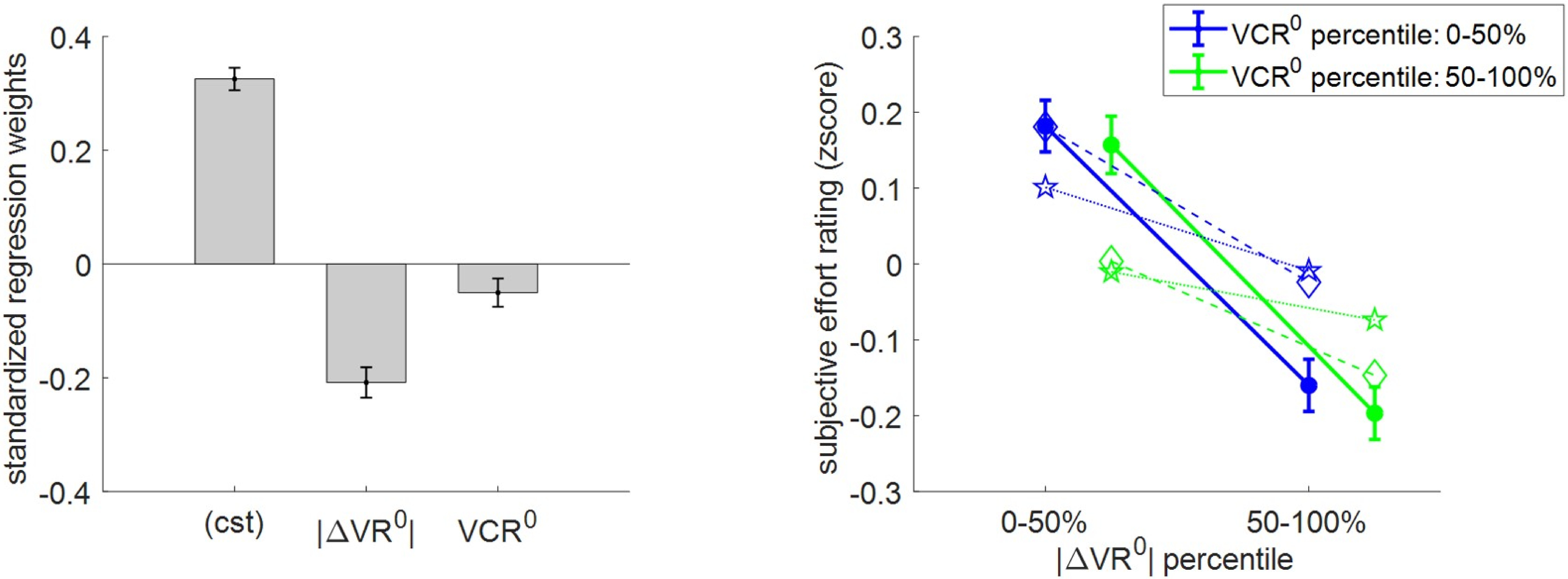
How much do you like this item?
I hate it!
I don't care.
I love it!
What degree of certainty do you have?
I hate it!
I don't care.
I love it!
Which do you prefer?
Are you sure about your choice?
Absolutely
Not at all
To what extent did you think about this choice?
Not at all
Really a lot
How much do you like this item?
I hate it!
I don't care.
I love it!
What degree of certainty do you have?
I hate it!
I don't care.
I love it!



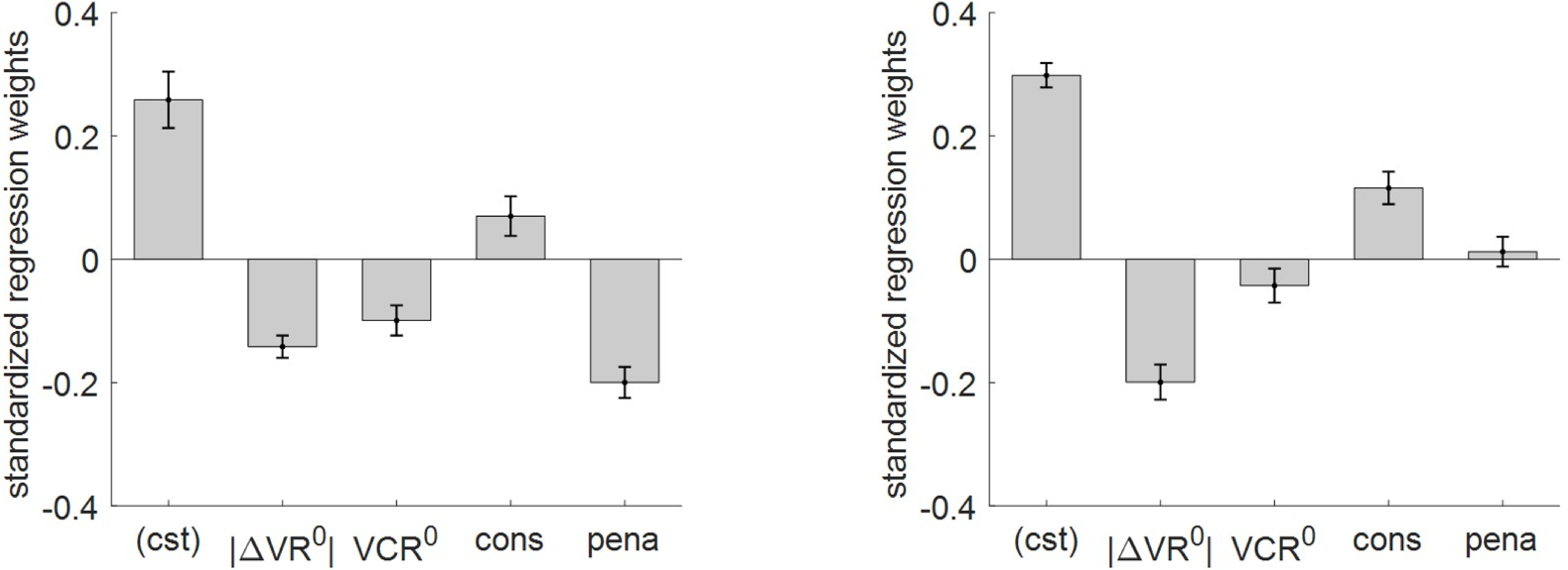
full data
decision-related variables only
effort-related variables only
within-subject trial-by-trial correlation (%)
45
40
35
30
25
20
15
10
5
0
Pc
SoA
Qcom
Cert. gain
RT
Eff. rat.



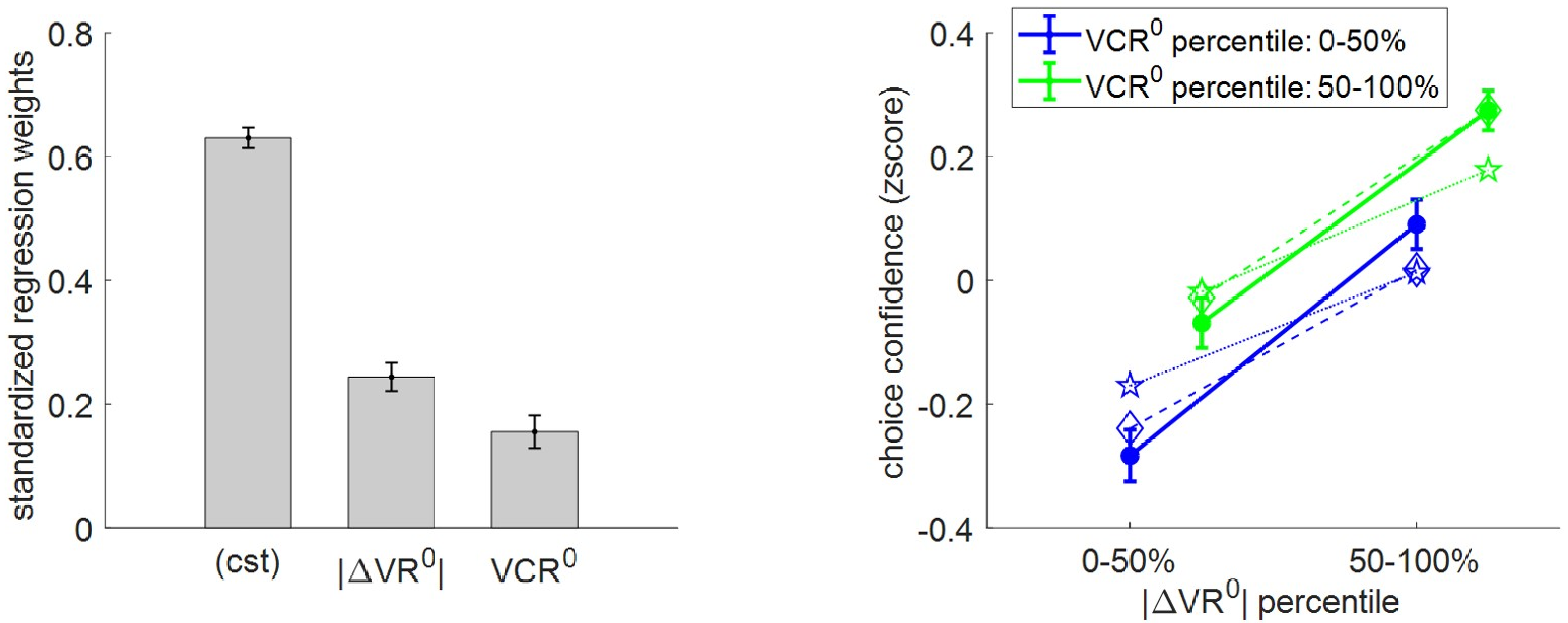
standardized regression weights
0.4
0.3
0.2
0.1
0
-0.1
-0.2
(cst)
|ΔVR⁰|
VCR⁰
response time (zscore)
0.2
0.1
0
-0.1
-0.2
0-50%
50-100%
|ΔVR⁰| percentile
VCR⁰ percentile: 0-50%
VCR⁰ percentile: 50-100%



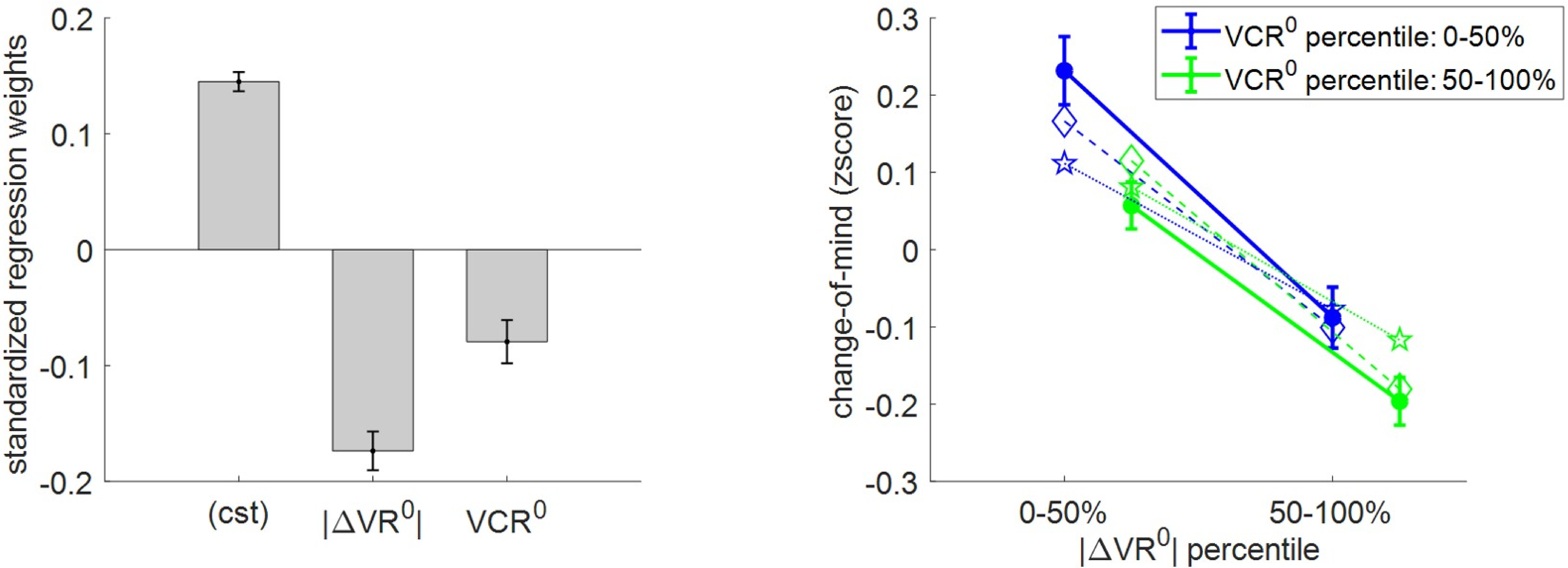
standardized regression weights
0.4
0.2
0
-0.2
-0.4
(cst)
|ΔVR0|
VCR0
VCR0 percentile: 0-50%
VCR0 percentile: 50-100%
subjective effort rating (zscore)
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
0-50%
50-100%
|ΔVR0| percentile



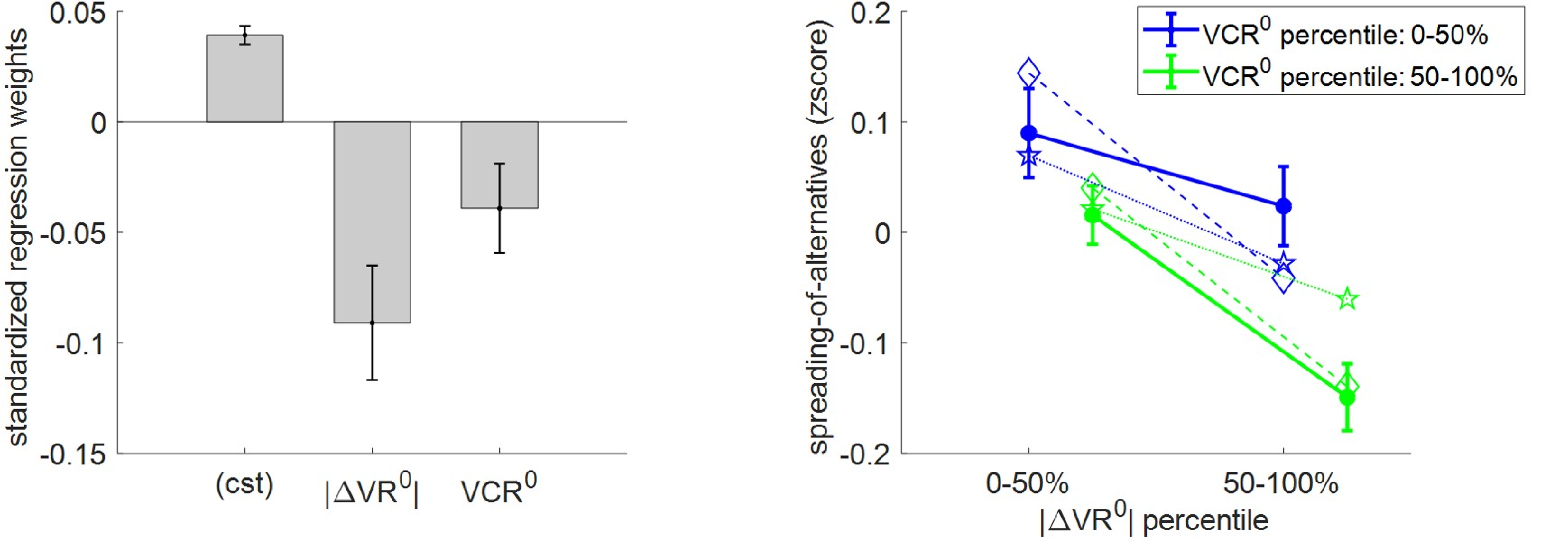
0.4
0.2
0
-0.2
-0.4
standardized regression weights
(cst)
|ΔVR⁰|
VCR⁰
cons
pena
0.4
0.2
0
-0.2
-0.4
standardized regression weights
(cst)
|ΔVR⁰|
VCR⁰
cons
pena



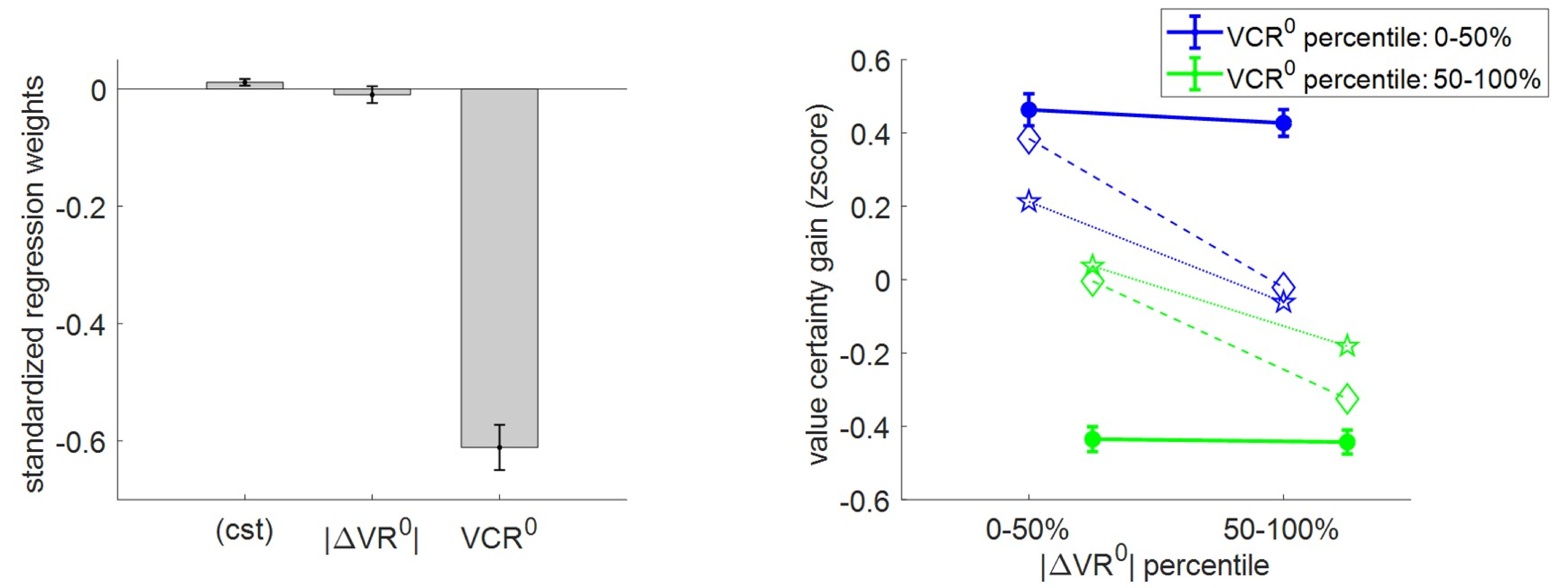
standardized regression weights
0.8
0.6
0.4
0.2
0
(cst)
|ΔVR⁰|
VCR⁰
VCR⁰ percentile: 0-50%
VCR⁰ percentile: 50-100%
choice confidence (zscore)
0.4
0.2
0
-0.2
-0.4
0-50%
50-100%
|ΔVR⁰| percentile



standardized regression weights
0.2
0.1
0
-0.1
-0.2
(cst)
|ΔVR0|
VCR0
change-of-mind (zscore)
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
VCR0 percentile: 0-50%
VCR0 percentile: 50-100%
0-50%
50-100%
|ΔVR0| percentile



standardized regression weights
0.05
0
-0.05
-0.1
-0.15
(cst)
|ΔVR⁰|
VCR⁰
spreading-of-alternatives (zscore)
0.2
0.1
0
-0.1
-0.2
VCR⁰ percentile: 0-50%
VCR⁰ percentile: 50-100%
0-50%
50-100%
|ΔVR⁰| percentile



standardized regression weights
0
-0.2
-0.4
-0.6
(cst)
|ΔVR⁰|
VCR⁰
VCR⁰ percentile: 0-50%
VCR⁰ percentile: 50-100%
value certainty gain (zscore)
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
0-50%
50-100%
|ΔVR⁰| percentile

Trading mental effort for confidence: Appendix

1. On the approximation accuracy of the expected confidence gain

The MCD model relies on the system's ability to anticipate the benefit of allocating resources to the decision process. Given the mathematical expression of choice confidence (cf. Equation 4 in the main text), this reduces to finding an analytical approximation to the following expression:

$$\bar{P} = E\left[s\left(\lambda|x|\right)\right] \tag{A1}$$

where $x \to s(x) = 1/1+\mathrm{e}^{-x}$ is the sigmoid mapping, λ is an arbitrary constant, and the expectation is taken under the Gaussian distribution of $x : N\left(\mu,\sigma^2\right)$, whose mean and variance are μ and σ^2, respectively.

Note that the absolute value mapping $x \to |x|$ follows a folded normal distribution, whose first two moments $E\left[|x|\right]$ and $V\left[|x|\right]$ have known expressions:

$$\begin{cases} E\left[|x|\right] = \sigma\sqrt{\dfrac{2}{\pi}}\exp\left(-\dfrac{|\mu|^2}{2\sigma^2}\right) + \mu\left(2\times s\left(\dfrac{\pi\mu}{\sigma\sqrt{3}}\right)-1\right) \\ V\left[|x|\right] = \mu^2+\sigma^2-E\left[|x|\right]^2 \end{cases} \tag{A2}$$

where the first line relies on a moment-matching approximation to the cumulative normal distribution function (Daunizeau, 2017a). This allows us to derive the following analytical approximation to Equation A1:

$$\bar{P} \approx s\left(\frac{E\left[|x|\right]}{\sqrt{\dfrac{1}{\lambda^2}+aV\left[|x|\right]}}\right) \tag{A3}$$

where setting $a \approx 3/\pi^2$ makes this approximation tight (Daunizeau, 2017a).

The quality of this approximation can be evaluated by drawing samples of $x:\ N\left(\mu,\sigma^2\right)$, and comparing the Monte-Carlo average of $s\left(\lambda\left|x\right|\right)$ with the expression given in Equation A3. This is summarized in Appendix-Figure 1 below, where the range of variation for the moments of x were set as follows: $\mu \in \left[-4,4\right]$ and $\sigma^2 \in \left[0,4\right]$.

===== Insert Appendix-Figure 1 here. =====

One can see that the error rarely exceeds 5%, across the whole range of moments $\left\{\mu,\sigma^2\right\}$ of the parent distribution. This is how tight the analytic approximation of the expected confidence gain (Equation 9 in the main text) is.

2. On the impact of model parameters for the MCD model

To begin with, note that the properties of the metacognitive control of decisions (in terms of effort allocation and/or confidence) actually depend on the demand for resources, which is itself determined by prior value representations (or, more properly, by the prior uncertainty σ^0 and the absolute means' difference $\left|\Delta\mu^0\right|$). Now, the way the MCD-optimal control responds to the resource demand is determined by effort efficacy and unitary cost parameters. In addition, MCD-optimal confidence may not trivially follow resource allocation, because it may be overcompensated by choice difficulty.

First, recall that the amount $\hat{z}$ of allocated resources maximizes the EVC:

$$\hat{z} = \arg\max_z \left[R \times \bar{P}_c\left(z\right) - \alpha z\right] \tag{A4}$$

where $\overline{P}_c(z)$ is given in Equation 9 in the main text. According to the implicit function theorem, the derivatives of $\hat{z}$ w.r.t. σ^0 and $\left|\Delta\mu^0\right|$ are given by (Gould et al., 2016):

$$\begin{cases} \dfrac{\partial \hat{z}}{\partial\left|\Delta\mu^0\right|} = -\dfrac{\left.\dfrac{\partial^2 \overline{P}_c(z)}{\partial\left|\Delta\mu^0\right|\partial z}\right|_{z=\hat{z}}}{\left.\dfrac{\partial^2 \overline{P}_c(z)}{\partial\left|\Delta\mu^0\right|^2}\right|_{z=\hat{z}}} \\ \dfrac{\partial \hat{z}}{\partial\sigma^0} = -\dfrac{\left.\dfrac{\partial^2 \overline{P}_c(z)}{\partial\sigma^0\partial z}\right|_{z=\hat{z}}}{\left.\dfrac{\partial^2 \overline{P}_c(z)}{\partial\sigma^{02}}\right|_{z=\hat{z}}} \end{cases} \tag{A5}$$

The double derivatives in Equations A5 are not trivial to obtain.

First, the gradient $\partial\overline{P}_c(z)/\partial\left|\Delta\mu^0\right|$ of choice confidence w.r.t. $\left|\Delta\mu^0\right|$ writes:

$$\begin{aligned} \frac{\partial\overline{P}_c(z)}{\partial\left|\Delta\mu^0\right|} &= \frac{\partial\overline{P}_c(z)}{\partial E\left[\left|\Delta\mu\right|\left|z\right.\right]}\frac{\partial E\left[\left|\Delta\mu\right|\left|z\right.\right]}{\partial\left|\Delta\mu^0\right|} + \frac{\partial\overline{P}_c(z)}{\partial V\left[\left|\Delta\mu\right|\left|z\right.\right]}\frac{\partial V\left[\left|\Delta\mu\right|\left|z\right.\right]}{\partial\left|\Delta\mu^0\right|} \\ &= 3K(z)\left(\left(2\sigma(z)+2\gamma z+\left|\Delta\mu^0\right|^2\right)\frac{\partial E\left[\left|\Delta\mu\right|\left|z\right.\right]}{\partial\left|\Delta\mu^0\right|} - \left|\Delta\mu^0\right| E\left[\left|\Delta\mu\right|\left|z\right.\right]\right) \end{aligned} \tag{A6}$$

where $K(z)\geq 0$ is given by:

$$K(z) = \frac{\pi\overline{P}_c(z)\left(1-\overline{P}_c(z)\right)}{\left(6\sigma(z)+3V\left[\left|\Delta\mu\right|\left|z\right.\right]\right)^{\frac{3}{2}}} \tag{A7}$$

Note that the gradient $\partial E\left[\left|\Delta\mu\right|\left|z\right.\right]/\partial\left|\Delta\mu^0\right|\geq 0$ in Equation A6 can be obtained analytically from Equation 7 in the main text. However, we refrain from doing this, because it is clear that deriving the right-hand term of Equation A6 w.r.t. both σ^0 and z will not bring any simple insight regarding the impact of $\left|\Delta\mu^0\right|$ onto $\hat{z}$.

Also, although the gradient $\partial\bar{P}_c(\hat{z})/\partial\sigma^0$ of choice confidence wr.t. σ^0 takes a much more concise form:

$$\begin{aligned}\frac{\partial\bar{P}_c(z)}{\partial\sigma^0} &= \frac{\partial\bar{P}_c(z)}{\partial\sigma(z)}\frac{\partial\sigma(z)}{\partial\sigma^0} \\ &= -\frac{3K(z)E\left[\left|\Delta\mu\right|\middle|z\right]}{\left(1+\beta z\sigma^0\right)^2}\end{aligned} \tag{A8}$$

it still remains tedious to derive its expression with respect to both σ^0 and z. This is why we opt for separating the respective effects of type #1 and type #2 efficacies.

First, let us ask what would be the MCD-optimal effort $\hat{z}$ and confidence $\bar{P}_c(\hat{z})$ when $\gamma=0$, i.e. if the only effect of allocating resources is to increase the precision of value representations. We call this the "β-effect". In this case, $E\left[\left|\Delta\mu\right|\middle|z\right]=\left|\Delta\mu^0\right|$ and $V\left[\left|\Delta\mu\right|\middle|z\right]=0$ irrespective of z. This greatly simplifies Equations A6, A7 and A8:

$$\begin{aligned}\left.\frac{\partial\bar{P}_c(z)}{\partial\left|\Delta\mu^0\right|}\right|_{\gamma=0} &= 6K(z)\sigma(z) \\ \left.\frac{\partial\bar{P}_c(z)}{\partial\sigma^0}\right|_{\gamma=0} &= -\frac{3K(z)\left|\Delta\mu^0\right|}{\left(1+\beta z\sigma^0\right)^2} \\ \left.K(z)\right|_{\gamma=0} &= \frac{\pi\bar{P}_c(z)\left(1-\bar{P}_c(z)\right)}{\left(6\sigma(z)\right)^{\frac{3}{2}}}\end{aligned} \tag{A9}$$

Inserting Equation A9 back into Equation A5 now yields:

$$\begin{cases} \left.\dfrac{\partial \hat{z}}{\partial\left|\Delta\mu^0\right|}\right|_{\gamma=0} = \dfrac{\beta K(\hat{z})\sigma(\hat{z}) - \left.\dfrac{\partial K(z)}{\partial z}\right|_{z=\hat{z}}}{\left.\dfrac{\partial K(z)}{\partial\left|\Delta\mu^0\right|}\right|_{z=\hat{z}}} \\ \left.\dfrac{\partial \hat{z}}{\partial \sigma^0}\right|_{\gamma=0} = \dfrac{2K(\hat{z})\beta\sigma^0 - \left.\dfrac{\partial K(z)}{\partial z}\right|_{z=\hat{z}}}{2K(\hat{z})\beta\hat{z} - \left.\dfrac{\partial K(z)}{\partial \sigma^0}\right|_{z=\hat{z}}} \end{cases} \tag{A10}$$

Now the sign of the gradients of $\hat{z}$ w.r.t. σ^0 and $\left|\Delta\mu^0\right|$ are driven by the numerators of Equation A10 because all partial derivatives of $K(z)$ have unambiguous signs:

$$\begin{aligned} \left.\frac{\partial K(z)}{\partial\left|\Delta\mu^0\right|}\right|_{\gamma=0} &= \frac{6\pi\left(1-2\bar{P}_c(z)\right)K(z)}{\left(6\sigma(z)\right)^{\frac{1}{2}}} \geq 0 \\ \left.\frac{\partial K(z)}{\partial \sigma^0}\right|_{\gamma=0} &= -\frac{\pi}{\left(1+\beta z\sigma^0\right)^2\left(6\sigma(z)\right)^{\frac{3}{2}}}\left(6\left(1-2\bar{P}_c(z)\right)K(z)\left|\Delta\mu^0\right| + \frac{\bar{P}_c(z)\left(1-\bar{P}_c(z)\right)}{4\sigma(z)^2}\right) \leq 0 \\ \left.\frac{\partial K(z)}{\partial z}\right|_{\gamma=0} &= \beta K(z)\sigma(z)\left(\frac{1}{4} + \frac{6\pi\left(1-2\bar{P}_c(z)\right)\left|\Delta\mu^0\right|}{\left(6\sigma(z)\right)^{\frac{3}{2}}}\right) \geq 0 \end{aligned} \tag{A11}$$

Replacing the expression for $\partial K(z)/\partial z$ in Equation A11 into Equation A10 now yields:

$$\begin{cases} \left.\dfrac{\partial \hat{z}}{\partial\left|\Delta\mu^0\right|}\right|_{\gamma=0} \propto 3\beta K(\hat{z})\sigma(\hat{z})\left(\dfrac{1}{4} - \dfrac{2\pi\left(1-2\bar{P}_c(\hat{z})\right)\left|\Delta\mu^0\right|}{\left(6\sigma(\hat{z})\right)^{\frac{3}{2}}}\right) \\ \left.\dfrac{\partial \hat{z}}{\partial \sigma^0}\right|_{\gamma=0} \propto \beta K(\hat{z})\left(2\sigma^0 - \dfrac{\sigma(\hat{z})}{4} - \dfrac{\pi\left(1-2\bar{P}_c(\hat{z})\right)\left|\Delta\mu^0\right|}{\sqrt{6\sigma(\hat{z})}}\right) \end{cases} \tag{A12}$$

At the limit $\left|\Delta\mu^0\right| \to 0$, then: $\partial\hat{z}/\partial\left|\Delta\mu^0\right| \geq 0$ and $\partial\hat{z}/\partial\sigma^0 \geq 0$. However, one can see from Equation A12 that there may be a critical value for $\left|\Delta\mu^0\right|$, above which the gradient $\partial\hat{z}/\partial\left|\Delta\mu^0\right|$ will eventually become negative. This means that the amount of allocated resources will behave as a bell-shaped function of $\left|\Delta\mu^0\right|$. This may not be the case along the σ^0 direction though, because $\sigma^0 \geq \sigma(z)$ and the last term in the brackets shrinks as σ^0 increases.

Similar derivations eventually yield expressions for the gradients of MCD-optimal confidence:

$$\begin{cases} \left.\dfrac{d\bar{P}_c(\hat{z})}{d\left|\Delta\mu^0\right|}\right|_{\gamma=0} = 3K(\hat{z})\sigma(\hat{z})\left(2+\beta\left|\Delta\mu^0\right|\sigma(\hat{z})\dfrac{\partial\hat{z}}{\partial\left|\Delta\mu^0\right|}\right) \\ \left.\dfrac{d\bar{P}_c(\hat{z})}{d\sigma^0}\right|_{\gamma=0} = 6K(\hat{z})\left|\Delta\mu^0\right|\left(\beta\sigma(\hat{z})^2\dfrac{\partial\hat{z}}{\partial\sigma^0}-\dfrac{1}{\left(1+\beta\hat{z}\sigma^0\right)^2}\right) \end{cases} \quad \text{(A13)}$$

Equation A13 implies that, under moderate type #1 efficacy ($\beta \approx 0$), MCD-optimal confidence decreases when $\left|\Delta\mu^0\right|$ decreases and/or when σ^0 increases, irrespective of the amount $\hat{z}$ of allocated resources. In other terms, variations in choice confidence are dominated by variations in the discriminability of prior value representations.

This analysis is exemplified on Appendix-Figure 2 below, which summarizes the β-effect, in terms of how MCD-optimal resource allocation and choice confidence depend upon $\left|\Delta\mu^0\right|$ and σ^0.

===== Insert Appendix-Figure 2 here. =====

One can see that, overall, increasing the prior variance σ^0 increases the resource demand, which eventually increases the MCD-optimal allocated effort $\hat{z}$. This, however, does not overcompensate for the loss of confidence incurred when increasing the prior variance. This is why the MCD-optimal confidence $\bar{P}_c(\hat{z})$ decreases with the prior variance σ^0. Note that, for the same reason, the MCD-optimal confidence increases with the absolute prior means' difference $\left|\Delta\mu^0\right|$.

Now the impact of the absolute prior means' difference $\left|\Delta\mu^0\right|$ on $\hat{z}$ is less trivial. In brief, when $\left|\Delta\mu^0\right|$ is high, the MCD-optimal allocated effort $\hat{z}$ decreases when $\left|\Delta\mu^0\right|$ increases. This is due to the fact that the resource demand decreases with $\left|\Delta\mu^0\right|$. However, there is a critical value for $\left|\Delta\mu^0\right|$, below which the MCD-optimal allocated effort $\hat{z}$ *increases* with $\left|\Delta\mu^0\right|$. This is because, although the resource demand still increases when $\left|\Delta\mu^0\right|$ decreases, the cost of allocating resources overcompensates the gain in confidence. For such difficult decisions, the system does not follow the demand anymore, and progressively de-motivates the allocation of resources as $\left|\Delta\mu^0\right|$ continues to decrease. In brief, the amount $\hat{z}$ of allocated resources decreases away from a "sweet spot", which is the absolute prior means' difference that yields the maximal confidence gain per effort unit. Critically, the position of this sweet spot along the $\left|\Delta\mu^0\right|$ dimension decreases with β and increases with α. This is because confidence gain increases, by definition, with effort efficacy, whereas it becomes more costly when α increases.

Second, let us ask what would be the MCD-optimal effort $\hat{z}$ and confidence $\overline{P}_c\left(\hat{z}\right)$ when $\beta = 0$, i.e. if the only effect of allocating resources is to perturb the value difference. The ensuing "γ -effect" is depicted on Appendix-Figure 3 below.

===== Insert Appendix-Figure 3 here. =====

In brief, the overall picture is reversed, with a few minor differences. One can see that increasing the absolute prior means' difference $\left|\Delta\mu^0\right|$ decreases the resource demand, which eventually decreases the MCD-optimal allocated effort $\hat{z}$. This can decrease confidence, if γ is high enough to overcompensate the effect of variations in $\left|\Delta\mu^0\right|$. When no effort is allocated however, confidence is driven by $\left|\Delta\mu^0\right|$, i.e. it becomes an increasing function of $\left|\Delta\mu^0\right|$. In contrast, variations in the prior variance σ^0 always overcompensate the ensuing changes in effort, which is why confidence always decreases with σ^0. In addition, the amount $\hat{z}$ of allocated resources decreases away from a sweet prior variance spot, which is the prior variance σ^0 that yields the maximal confidence gain per effort unit. Critically, the position of this sweet spot increases with γ and decreases with α, for reasons similar to the β-effect.

Now one can ask what happens in the presence of both the β-effect and the γ-effect. If the effort unitary cost α is high enough, the MCD-optimal effort allocation is essentially the superposition of both effects. This means that there are two "sweet spots": one around some value of $\left|\Delta\mu^0\right|$ at high σ^0 (β-effect) and one around some value of σ^0 at high $\left|\Delta\mu^0\right|$ (γ-effect). If the effort unitary cost α decreases, then the position of the β-sweet spot increases and that of the β-sweet spot decreases, until they effectively merge together. This is exemplified on Appendix-Figure 4 below.

===== Insert Appendix-Figure 4 here. =====

One can see that, somewhat paradoxically, the effort response is now much simpler. In brief, the MCD-optimal effort allocation $\hat{z}$ increases with the prior variance σ^0 and decreases with the absolute prior means' difference $\left|\Delta\mu^0\right|$. The landscape of the ensuing MCD-optimal confidence level $\bar{P}_c\left(\hat{z}\right)$ is slightly less trivial, but globally, it can be thought of as increasing with

$\left|\Delta\mu^0\right|$ and decreasing with σ^0. Here again, this is because variations in $\left|\Delta\mu^0\right|$ and/or σ^0 almost always overcompensate the ensuing effects of changes in allocated effort.

3. On MCD parameter estimation

Let y_t be a 6x1 vector composed of measured choice confidence, spreading of alternatives, value certainty gain, change of mind, response time, and subjective effort rating on trial t. Let u_t be a 4x1 vector, whose two first entries are composed of pre-choice value difference (ΔVR^0) and average value certainty (VCR^0) ratings, and whose two last entries encode consequential and penalized trials. Finally, let φ be the set of unknown MCD parameters (i.e. intrinsic effort cost α and effort efficacies β and γ), augmented with condition-effect parameters and affine transform parameters (see below). From a statistical perspective, the MCD model then reduces to the following observation equation:

$$\bar{y}_t = g\left(\varphi, u_t\right) + \varepsilon_t \qquad \text{(A14)}$$

where $\bar{y}$ denotes data that have been z-scored across trials, ε_t are model residuals, and the observation mapping $g\left(\varphi, u_t\right)$ is given by:

$$g(\varphi,u_t)=\begin{bmatrix} a_1+b_1\times s\left(\dfrac{\pi E\left[|\Delta\mu|\,\|\,\hat{z}\right]}{\sqrt{3\left(\dfrac{2}{1/\sigma^0+\beta\hat{z}}+V\left[|\Delta\mu|\,\|\,\hat{z}\right]\right)}}\right) \\ a_2+b_2\times\sqrt{\dfrac{\gamma\hat{z}}{\pi}}\exp\left(-\dfrac{|\Delta\mu^0|^2}{4\gamma\hat{z}}\right) \\ a_3+b_3\times s\left(-\dfrac{\pi|\Delta\mu^0|}{\sqrt{6\gamma\hat{z}}}\right) \\ a_4+b_4\times\beta\hat{z} \\ a_5+b_5\times\hat{z} \\ a_6+b_6\times\hat{z} \end{bmatrix} \qquad \text{(A15)}$$

where $E\left[|\Delta\mu|\,\|\,\hat{z}\right]$ and $V\left[|\Delta\mu|\,\|\,\hat{z}\right]$ depend upon γ (see Equations 7 and 8 in the main text). In Equation A15, $a_{1:6}$ and $b_{1:6}$ are the unknown offset and slope parameters of the (nuisance) affine transform on MCD outputs. Note that when fitting the MCD model to empirical data, theoretical pre-choice value difference and value certainty ratings are replaced by their empirical proxies, i.e. $\Delta\mu^0 \approx \Delta\mathrm{VR}^0$ and $1/\sigma^0 \approx \mathrm{VCR}^0$. In turn, given MCD parameters, Equations A14-A15 predict trial-by-trial variations in choice confidence, spreading of alternatives, value certainty gain, change of mind, response time, and subjective effort rating from variations in prior moments of value representations. We note that Equation A15 does not yet include condition-specific effects. As we will see, it will be easier to complete the definition of model parameters φ once we have explained the variational Laplace scheme for parameter estimation.

Recall that the variational Laplace scheme is an iterative algorithm that indirectly optimizes an approximation to both the model evidence $p(y|m,u)$ and the posterior density

$p\left(\varphi|y,m,u\right)$, where m is the so-called generative model (i.e., the set of assumptions that are required for inference). The key trick is to decompose the log model evidence into:

$$\ln p\left(y|m,u\right)=F\left(q\right)+D_{KL}\left(q\left(\varphi\right);p\left(\varphi|y,m,u\right)\right), \quad \text{(A16)}$$

where $q\left(\varphi\right)$ is any arbitrary density over the model parameters, D_{KL} is the Kullback-Leibler divergence and the so-called *free energy* $F\left(q\right)$, defined as:

$$F\left(q\right)=\left\langle \ln p\left(\varphi|m\right)+\ln p\left(y|\varphi,m,u\right)\right\rangle_q+S\left(q\right), \quad \text{(A17)}$$

where $S\left(q\right)$ is the Shannon entropy of q and the expectation $\langle g\rangle_q$ is taken under q.

From equation A16, maximizing the functional $F\left(q\right)$ w.r.t. q indirectly minimizes the Kullback-Leibler divergence between $q\left(\varphi\right)$ and the exact posterior $p\left(\varphi|y,m\right)$. This decomposition is complete in the sense that if $q\left(\varphi\right)=p\left(\varphi|y,m\right)$, then $F\left(q\right)=\ln p\left(y|m\right)$.

The variational Laplace algorithm iteratively maximizes the free energy $F\left(q\right)$ under simplifying assumptions (see below) about the functional form of q, rendering q an approximate posterior density over model parameters and $F\left(q\right)$ an approximate log model evidence (Daunizeau, 2017b; Friston et al., 2007). The free energy optimization is then made with respect to the sufficient statistics of q, which makes the algorithm generic, quick and efficient.

Under normal i.i.d. model residuals (i.e. ε_t : $N\left(0,1/\lambda\right)$), the likelihood function writes:

$$\begin{aligned} p\left(y|\varphi,\lambda,m,u\right) &= \prod_t p\left(y_t|\varphi,\lambda,m,u_t\right) \\ &= \prod_t N\left(g\left(\varphi,u_t\right),\frac{1}{\lambda}I\right) \end{aligned} \quad \text{(A18)}$$

where λ is the residuals' precision or inverse variance hyperparameter and the observation mapping $g(\varphi,u_t)$ is given in Equation A15.

We also use Gaussian priors $p(\varphi|m)=N(\eta_0,\Sigma_0)$ for model parameters and gamma priors for precision hyperparameters $p(\lambda|m)=Ga(\varpi_0,\kappa_0)$.

In what follows, we derive the variational Laplace algorithm under a "mean-field" separability assumption between parameters and hyperparameters, i.e.: $q(\varphi,\lambda)=q(\varphi)q(\lambda)$. We will see that this eventually yields a Gaussian posterior density $q(\varphi)\approx N(\eta,\Sigma)$ on model parameters, and a Gamma posterior density $q(\lambda)=Ga(\varpi,\kappa)$ on the precision hyperparameter.

First, let us note that, under the Laplace approximation, the free energy bound on the log-model evidence can be written as:

$$\begin{aligned} F(q) &= \langle I(\varphi)\rangle_{q(\varphi)} + S(q(\varphi)) + S(q(\lambda)) \\ &\approx I(\eta) + \frac{1}{2}\ln|\Sigma| + \frac{n_\varphi}{2}\ln 2\pi + \varpi - \ln\kappa + \log\Gamma(\varpi) + (1-\varpi)\psi(\varpi) \end{aligned} \quad \text{(A19)}$$

where n_φ is the number of parameters, $\Gamma(g)$ is the gamma function, $\psi(g)$ is the digamma function, and $I(\varphi)$ is defined as:

$$I(\varphi) = \left\langle \log p(\varphi|m) + \log p(y|\varphi,\lambda,m,u) + \log p(\lambda|m) \right\rangle_{q(\lambda)} \quad \text{(A20)}$$

Given the Gamma posterior $q(\lambda)$ on the precision hyperparameter, $I(\varphi)$ can be simply expressed as follows:

$$I(\varphi) = -\frac{1}{2}(\varphi-\eta_0)^T \Sigma_0^{-1}(\varphi-\eta_0) - \frac{\langle\lambda\rangle}{2}\sum_t (y_t - g(\varphi,u_t))^T (y_t - g(\varphi,u_t)) \quad \text{(A21)}$$

where we have ignored the terms that do not depend upon φ, and $\langle\lambda\rangle = E\left[\lambda \middle| y,m\right] = \varpi/\kappa$ is the posterior mean of the data precision hyperparameter λ.

The variational Laplace update rule for the approximate posterior density $q(\varphi)$ on model parameters now simply reduces to an update rule for its sufficient statistics:

$$q(\varphi) \approx N(\eta,\Sigma): \begin{cases} \eta = \arg\max_{\varphi} I(\varphi) \\ \Sigma = -\left[\left. \dfrac{\partial^2 I}{\partial\varphi^2} \right|_{\eta} \right]^{-1} \end{cases} \quad \text{(A22)}$$

In Equation A22, the first-order moment η of $q(\varphi)$ is obtained from the following Gauss-Newton iterative gradient ascent scheme:

$$\eta \leftarrow \eta - \left[\left. \frac{\partial^2 I}{\partial\varphi^2} \right|_{\eta} \right]^{-1} \left. \frac{\partial I}{\partial\varphi} \right|_{\eta} \quad \text{(A23)}$$

where the gradient and Hessians of $I(\varphi)$ are given by:

$$\begin{aligned} \frac{\partial I}{\partial\varphi} &= \Sigma_0^{-1}(\eta_0 - \varphi) + \langle\lambda\rangle \frac{\partial g}{\partial\varphi}^T \sum_t \left(y_t - g(\varphi,u_t) \right) \\ \frac{\partial^2 I}{\partial\varphi^2} &\approx -\Sigma_0^{-1} - \langle\lambda\rangle \sum_t \frac{\partial g}{\partial\varphi}^T \frac{\partial g}{\partial\varphi} \end{aligned} \quad \text{(A24)}$$

At convergence of the above gradient ascent, the approximate posterior density $q(\varphi)$ on the precision hyperparameter is updated as follows:

$$q(\lambda) = Ga(\varpi,\kappa): \begin{cases} \varpi = \varpi_0 + 3n_t - 1 \\ \kappa = \kappa_0 + \dfrac{1}{2} \sum_t \left(y_t - g(\eta,u_t) \right)^T \left(y_t - g(\eta,u_t) \right) + tr\left[\left. \dfrac{\partial g}{\partial\varphi} \right|_{\eta}^T \left. \dfrac{\partial g}{\partial\varphi} \right|_{\eta} \Sigma^{-1} \right] \end{cases} \quad \text{(A25)}$$

where n_t is the number of trials.

The variational Laplace scheme alternates between Equations A22 and A25 iteratively until convergence of the free energy.

Now, let us complete the definition of the model parameter vector $\varphi = \varphi_{1:17}$.

First, note that effort efficiency parameters are necessarily positive. Enforcing this constraint can be done using the following simple change of variable in Equation A15: $\beta = \exp(\varphi_1)$ and $\gamma = \exp(\varphi_2)$. In other words, $\varphi_{1:2}$ effectively measure efficiency parameters in log-space. Second, recall that we want to insert condition-specific effects in the model. More precisely, we expect “consequential” decisions to be more important than “neutral” ones, and “penalized” decisions effectively include an extraneous cost-of-time term. One can model the former condition effect by making R (cf. Equation 2 in the main text) sensitive to whether the decision is consequential ($u^{(c)} = 1$) or not ($u^{(c)} = 0$), i.e.: $R_t = \exp\left(\varphi_3 u_t^{(c)}\right)$, where t indexes trials, and φ_3 is the unknown weight of consequential choices on decision importance. This parameterization makes decision importance necessarily positive, and forces non-consequential trials to act as reference choices (in the sense that their decision importance is set to 1). We proxy the latter condition effect by making the effort unitary cost a function of whether the decision is penalized ($u^{(p)} = 1$) or not ($u^{(p)} = 0$), i.e.: $\alpha_t = \exp\left(\varphi_4 + \varphi_5 u_t^{(p)}\right)$, where φ_4 is the unknown intrinsic effort cost (in log-space), and φ_5 is the unknown weight of penalized choices on effort cost. The remaining parameters $\varphi_{6:17}$ lump the offsets ($a_{1:6}$) and log-slopes ($\log b_{1:6}$: this enforces a positivity constraint on slope parameters) of the affine transform.

Finally, we set the prior probability density functions on model parameters and hyperparameters as follows:

- $p(\varphi_i|m) = N(0,10^2)\ \forall i$, i.e. the prior mean of model parameters is $\eta_0 = 0$ and their prior variance is $\Sigma_0 = 10^2 \times I$.
- $p(\lambda|m) = Ga(1,1)$. Since the data has been z-scored prior to model inversion, this ensures that the prior and likelihood components of $I(\varphi)$ are balanced when the variational Laplace algorithm starts.

This completes the description of the variational Laplace approach to MCD inversion. For more details, we refer the interested reader to the existing literature on variational approaches to approximate Bayesian inference (Beal, 2003; Daunizeau, 2017b; Friston et al., 2007). We note that the above variational Laplace approach is implemented in the opensource VBA toolbox (Daunizeau et al., 2014).

In what follows, we use Monte-Carlo numerical simulations to evaluate the ability of this approach to recover MCD parameters. Our parameter recovery analyses proceed as follows. First, we sample a set of model parameters φ under a standard i.i.d. normal distribution. Here, we refer to φ_{ij} as i^{th} element of φ at the j^{th} Monte-Carlo simulation. Second, for each of these parameter set φ_{gj}, we simulate a series of N=100 decision trials according to Equations A14-A15 above (under random prior moments of value representations). Note that we set the variance of model residuals (ε in Equation A14) to match the average correlation between MCD predictions and empirical data (about 20%, see Figure 4 in the main text). We also used the same rate of neutral, consequential, and penalized choices as in our experiment. Third, we fit the model to the resulting simulated data (after z-scoring) and extract parameter estimates η_{gj} (at convergence of the variational Laplace approach). We repeat these three steps 1000 times, yielding a series of 1000 simulated parameter sets, and their corresponding 1000 estimated parameters sets. Should $\eta_{gj} \approx \varphi_{gj}\ \forall j$, then parameter recovery would be perfect. Appendix-Figure 14 below compares simulated and estimated parameters to each other across Monte-Carlo simulations. Note that we only report recovery results for $\varphi_{1:5}$, since we do not care about nuisance affine transform parameters.

We also quantify pairwise non-identifiability issues, which arise when the estimation method confuses two parameters with each other. We do this using so-called "recovery matrices", which summarize whether variations (across the 1000 Monte-Carlo simulations) in estimated parameters faithfully capture variations in simulated parameters. We first z-score simulated and estimated parameters across Monte-Carlo simulations. We then regress each estimated parameter against all simulated parameters through the following multiple linear regression model:

$$\eta_{ij} = \sum_{i'=1}^{5} \theta_{ii'} \varphi_{i'j} + \varsigma_{ij} \qquad \text{(A26)}$$

where $\theta_{ii'}$ are regression weights, and ς_{ij} are regression residuals. Here, regression weights are partial correlation coefficients between simulated and estimated parameters (across Monte-Carlo simulations). More precisely, $\theta_{ii'}$ quantifies the impact that variations of the simulated parameter $\varphi_{i'g}$ have on variations of the estimated parameter η_{ig}, conditional on all other simulated parameters. Would parameters be perfectly identifiable, then $\theta_{ii} \approx 1$ and $\theta_{ii'} \approx 0\ \forall i' \neq i$. Pairwise non-identifiability issues arise when $\theta_{ii'} \neq 0$. In other words, the regression model in Equation A26 effectively decomposes the observed variability in the series of estimated parameter η_{ig} into “correct variations” that are induced by variations in the corresponding simulated parameter φ_{ig}, and “incorrect variations” that are induced by the remaining simulated parameters $\varphi_{i'g}$ (with $i' \neq i$). This analysis is then summarized in terms of "recovery matrices", which simply report the squared regression weights $\theta_{ii'}^2$ for each simulated parameter (see right panel of Appendix-Figure 5 below).

===== Insert Appendix-Figure 5 here. =====

One can see that parameter recovery is far from perfect. This is in fact expected, given the high amount of simulation noise. However, no parameter estimate exhibits any noticeable estimation bias, i.e. estimation error is non-systematic and directly results from limited data reliability. Recovery matrices provides further quantitative insight regarding the accuracy of parameter estimation.

First, variability in all parameter estimates is mostly driven by variability in the corresponding simulated parameter (amount of “correct variability”: φ_1: 5.3%, φ_3: 17.4%, φ_4: 22.1%, φ_5: 22.7%, to be compared with “incorrect variability” – see below), except for type #1 efficacy (φ_2: 0.3%). The latter estimate is thus comparatively much less efficient than other MCD parameters. This is because $\beta = \exp(\varphi_2)$ only has a limited impact on MCD outputs. Second, there are no strong non-identifiability issues (total amount of "incorrect invariability" is always below 2.7%, even when including nuisance affine transform parameters $\varphi_{6:17}$), except for type #2 effort efficacy. In particular, the latter estimate may be partly confused with intrinsic effort cost (amount of “incorrect variability” driven by φ_1: 1.6%).

Having said this, the reliability of MCD parameter recovery is globally much weaker than in the ideal case, where data is not polluted with simulation noise (the amount of "correct variability" in this case, is higher than 95% for all parameters – results not shown). This means that acquiring data of higher quality and/or quantity may significantly improve inference on MCD parameters.

We note that the weak identifiability of type #1 effort efficacy (β) does not imply that some dependent variables will be less well predicted/postdicted than others. Recall that β indirectly influences all dependent variables, through its impact on the optimal amount of allocated resources. Therefore, all dependent variables provide information about β. Importantly, some dependent variables are more useful than others for estimating β. If empirical measures of these variables become unreliable (e.g., because they are very noisy), then β will not be identifiable. However, the reverse is not true. In fact, in our recovery analysis, we found no difference in postdiction accuracy across dependent variables. Now, the question of whether weak β identifiability may explain (out-of-sample) prediction errors regarding the impact of MCD input variables (such as ΔVR0) on dependent variables is more subtle. This is because, by construction, MCD parameters control the way MCD input variables eventually influence dependent variables. As one can see from the analytical derivations in section 2 of this Appendix, the impact of input variables on MCD dependent variables (in particular, the optimal amount of allocated resources) depends upon whether β dominates effort efficacy (cf. "β-effect") or not (cf. "γ-effect"). For example, if β dominates, then the relationship between ΔVR^0 and effort is bell-shaped (cf. Figure S6), whereas it is monotonic if β=0 (cf. Figure S7). This means that estimation errors on β may confuse the predicted relationship between input variables and MCD dependent variables.

4. Data descriptive statistics and sanity checks

Recall that we collect value ratings and value certainty ratings both before and after the choice session. We did this for the purpose of validating specific predictions of the MCD model (in particular: choice-induced preference changes: see Figure 10 in the main text). It turns out this also enables us to assess the test-retest reliability of both value and value certainty ratings. We found that both ratings were significantly reproducible (value: mean correlation=0.88, s.e.m.=0.01, $p < 0.001$, value certainty: mean correlation=0.37, s.e.m.=0.04, $p < 0.001$).

We also checked whether choices were consistent with pre-choice ratings. For each participant, we thus preformed a logistic regression of choices against the difference in value ratings. We found that the balanced prediction accuracy was beyond chance level (mean accuracy=0.68, s.e.m.=0.01, $p<0.001$).

5. Does choice confidence moderate the relationship between choice and pre-choice value ratings?

Previous studies regarding confidence in value-base choices showed that choice confidence moderates choice prediction accuracy (De Martino et al., 2013). We thus split our logistic regression of choices into high- and low-confidence trials, and tested whether higher confidence was consistently associated with increased choice accuracy. A random effect analysis showed that the regression slopes were significantly higher for high- than for low-confidence trials (mean slope difference=0.14, s.e.m.=0.03, p<0.001). For the sake of completeness, the impact of choice confidence on the slope of the logistic regression (of choice onto the difference in pre-choice value ratings) is shown on Appendix-Figure 6 below.

===== Insert Appendix-Figure 6 here. =====

These results clearly replicate the findings of De Martino and colleagues (2013), which were interpreted with a race model variant of the accumulation-to-bound principle. We note, however, that this effect is also predicted by the MCD model. Here, variations in both (i) the prediction accuracy of choice from pre-choice value ratings, and (ii) choice confidence, are driven by variations in resource allocation. In brief, the expected magnitude of the perturbation of value representations increases with the amount of allocated resources. This eventually increases the probability of a change of mind. However, although more resources are allocated to the decision, this does not overcompensate for decision difficulty, and thus choice confidence decreases. Thus, low-confidence choices will be those choices that are more likely to be associated with a change of mind. We note that the anti-correlation between choice confidence and change of mind can be seen by comparing Figures 7 and 8 in the main text.

6. How do choice confidence, difference in pre-choice value ratings, and response time relate to each other?

In the main text, we show that trial-by-trial variation in choice confidence is concurrently explained by both pre-choice value and value certainty ratings. Here, we reproduce previous findings relating choice confidence to both absolute value difference ΔVR^0 and response time (De Martino et al., 2013). First, for each participant, we regressed response time concurrently against both $|\Delta VR^0|$ and choice confidence. A random effect analysis showed that both have a significant main effect on response time (ΔVR^0: mean GLM beta=-0.016, s.e.m.=0.003, $p<0.001$; choice confidence: mean GLM beta=-0.014, s.e.m.=0.002; $p<0.001$), without any two-way interaction ($p=0.133$). This analysis is summarized in Appendix-Figure 7 below, together with the full three-way relationship between $|\Delta VR^0|$, confidence and response time.

===== Insert Appendix-Figure 7 here. =====

In brief, confidence increases with the absolute value difference and decreases with response time. This effect is also predicted by the MCD model, for reasons identical to the explanation of the relationship between confidence and choice accuracy (see above). Recall that, overall, an increase in choice difficulty is expected to yield an increase in response time and a decrease in choice confidence. This would produce the same data pattern as Appendix-Figure 7, although the causal relationships implicit in this data representation is partially incongruent with the computational mechanisms underlying MCD.

7. Do post-choice ratings better predict choice and choice confidence than pre-choice ratings?

The MCD model assumes that value representations are modified during the decision process, until the MCD-optimal amount of resources is met. This eventually triggers the decision, whose properties (i.e., which alternative option is eventually preferred, and with which confidence level) then reflects the modified value representations. If post-choice ratings are reports of modified value representations at the time when the choice is triggered, then choice and its associated confidence level should be better predicted with post-choice ratings than with pre-choice ratings. In what follows, we test this prediction.

In Section 4 of this Appendix, we report the result of a logistic regression of choice against pre-choice value ratings (see also Appendix-Figure 6). We performed the same regression analysis, but this time against post-choice value ratings. For each subject, we then measured the ensuing predictive power (here, in terms of balanced accuracy or BA) for both pre-choice and post-choice ratings. The main text also features the result of a multiple linear regression of choice confidence ratings onto $|\Delta VR^0|$ and VCR^0 (cf. Figure 8 in the main text). Again, we performed the same regression, this time against post-choice ratings. For each subject, we then measured the ensuing predictive power (here, in terms of percentage of explained variance or R^2) for both pre-choice and post-choice ratings.

A simple random effect analysis shows that the predictive power of post-choice ratings is significantly higher than that of pre-choice ratings, both for choice (mean difference in BA=7%, s.e.m.=0.01, $p<0.001$) and choice confidence (mean difference in R^2=3%, s.e.m.=0.01, $p=0.004$).

8. Analysis of eye-tracking data

We first checked whether pupil dilation positively correlates with participants' subjective effort ratings. We epoched the pupil size data into trial-by-trial time series, and temporally co-registered the epochs either at stimulus onset (starting 1.5 seconds before the stimulus onset and lasting 5 seconds) or at choice response (starting 3.5 seconds before the choice response and lasting 5 seconds). Data was baseline-corrected at stimulus onset. For each participant, we then regressed, at each time point during the decision, pupil size onto effort ratings (across trials).

Time series of regression coefficients were then reported at the group level, and tested for statistical significance (correction for multiple comparison was performed using random field theory 1D-RFT). Appendix-Figure 8 below summarizes this analysis, in terms of the baseline-corrected time series of regression coefficients.

===== Insert Appendix-Figure 8 here. =====

We found that the correlation between subjective effort ratings and pupil dilation became significant from 500ms after stimulus onset onwards. Note that, using the same approach, we found a negative correlation between pupil dilation and pre-choice absolute value difference $|\Delta VR^0|$. However, this relationship disappeared when we entered both $|\Delta VR^0|$ and effort into the same regression model.

Our eye-tracking data also allowed us to ascertain which item was being gazed at for each point in peristimulus time (during decisions). Using the choice responses, we classified each time point as a gaze at the (to be) chosen item or at the (to be) rejected item. We then derived, for each decision, the ratio of time spent gazing at chosen/rejected items versus the total duration of the decision (between stimulus onset and choice response). The difference between these two gaze ratios measures the overt attentional bias towards the chosen item. We refer to this as the gaze bias. Consistent with previous studies, we found that chosen items were gazed at more than rejected items (mean gaze bias=0.02, s.e.m.=0.01, p=0.067). However, we also found that this effect was in fact limited to low effort choices. Appendix-Figure 9 below shows the gaze bias for low- and high-effort trials, based on a median-split of subjective effort.

===== Insert Appendix-Figure 9 here. =====

We found that there was a significant gaze bias for low effort choices (mean gaze ratio difference=0.033, s.e.m.=0.013, p=0.009), but not for high effort choices (mean gaze ratio difference=0.002, s.e.m.=0.014, p=0.453). A potential trivial explanation for the fact that the gaze bias is large for low effort trials is that these are the trials where participants immediately recognize their favorite option, which attracts their attention. More interesting is the fact that the gaze bias is null for high effort trials. This may be taken as evidence for the fact that, on average, people allocate the same amount of (attentional) resources to both options. This is important, because we use this simplifying assumption in our MCD model derivations.

9. Comparison with evidence-accumulation (DDM) models

In the main text, we evaluate the accuracy of the MCD model predictions, without considering alternative computational scenarios. Here, we report results of a model-based data analysis that relies on the standard drift-diffusion decision or DDM model for value-based decision making (De Martino et al., 2013; Lopez-Persem et al., 2016; Milosavljevic et al., 2010; Ratcliff et al., 2016; Tajima et al., 2016).

In brief, DDMs tie together decision outcomes and response times by assuming that decisions are triggered once the accumulated evidence in favor of a particular option has reached a predefined threshold or bound (Ratcliff and McKoon, 2008; Ratcliff et al., 2016). Importantly here, evidence accumulation has two components: a drift term that quantifies the strength of evidence and a random diffusion term that captures some form of neural perturbation of evidence accumulation. The latter term allows choice outcomes to deviate from otherwise deterministic, evidence-driven, decisions.

Importantly, standard DDMs do not predict choice confidence, spreading of alternatives, value certainty gain, or subjective effort ratings. This is because these concepts have no straightforward definition under the standard DDM. However, DDMs can be used to make out-of-sample trial-by-trial predictions of, for example, decision outcomes, from parameter estimates obtained with response times alone. This enables a straightforward comparison of MCD and DDM frameworks, in terms of the accuracy of RT "postdictions" and change of mind out-of-sample prediction. Here, we also make sure both models rely on the same inputs: namely, pre-choice value (ΔVR^0) and value certainty (VCR^0) ratings as well as information about task conditions.

The simplest DDM variant includes the following set of five unknown parameters: the drift rate v, the bound's height b, the standard deviation of the diffusion term σ, the initial decision bias x_0, and the non-decision time T_{nd}. Given these model parameters, the expected response time (conditional on the decision outcome) is given by (Srivastava et al., 2016):

$$E\left[RT \middle| o, v, x_0, b, \sigma, T_{nd}\right] = \frac{b}{v}\left(2\coth\left(\frac{2vb}{\sigma^2}\right) - \left(1 + o\frac{x_0}{b}\right)\coth\left(\left(1 + o\frac{x_0}{b}\right)\frac{vb}{\sigma^2}\right)\right) + T_{nd} \qquad \text{(A27)}$$

where $o \in \{-1, 1\}$ is the decision outcome. One can then evaluate Equation A27 at each trial, given its corresponding set of DDM parameters. In particular, if one knows how, for example, drift rates vary over trials, then one can predict the ensuing expected RT variations. In typical applications to value-based decision making, drift rates are set proportional to the difference ΔVR^0 in value ratings (De Martino et al., 2013; Krajbich et al., 2010; Lopez-Persem et al., 2016; Milosavljevic et al., 2010). One can then define a likelihood function for observed response times from the following observation equation: $RT = E\left[RT \middle| o, v, x_0, b, \sigma, T_{nd}\right] + \varepsilon$, where ε are trial-by-trial DDM residuals. The variational Laplace treatment of the ensuing generative model then yields estimates of the remaining DDM parameters.

Out-of-sample predictions of change of mind (i.e., decision errors) can then be derived from DDM parameter estimates (Bogacz et al., 2006):

$$\begin{aligned} Q_{DDM} &= P\left(sign(o) \neq sign(v) \middle| v, b, \sigma, x_0\right) \\ &= \frac{1}{1+\exp\left(\dfrac{2vb}{\sigma^2}\right)} - \frac{1-\exp\left(\dfrac{2\left|v x_0\right|}{\sigma^2}\right)}{\exp\left(\dfrac{2vb}{\sigma^2}\right)-\exp\left(\dfrac{-2vb}{\sigma^2}\right)} \end{aligned} \quad \text{(A28)}$$

where Q_{DDM} is the DDM equivalent to the probability $Q(\hat{z})$ of a change of mind under the MCD model (see Equation 14 in the main text).

Here, we use two modified variants of the standard DDM for value-based decisions. In all of these variants, we allow the DDM system to change its speed-accuracy tradeoff according to whether the decision is consequential ($u^{(c)}=1$) or not ($u^{(c)}=0$), and/or "penalized" ($u^{(p)}=1$) or not ($u^{(p)}=0$). This is done by enabling the decision bound to vary over trials, i.e.: $b_t \equiv \exp\left(b^{(0)}+b^{(c)}u_t^{(c)}+b^{(p)}u_t^{(p)}\right)$, where t indexes trials. Here, $b^{(0)}$, $b^{(c)}$ and $b^{(p)}$ are unknown parameters that quantify the bound's height of "neutral" decisions, and the strength of "consequential" and "penalized" condition effects, respectively. The exponential mapping is used for imposing a positivity constraint on the resulting bound (see section 8 above). One might then expect that $b^{(c)}>0$ and $b^{(p)}<0$, i.e. "consequential" decisions demand more evidence than "neutral" ones, whereas "penalized" decisions favor speed over accuracy.

The two DDM variants then differ in terms of how pre-choice value certainty is taken into account (Lee and Usher, 2020):

- DDM1: at each trial, the drift rate is set to the affine-transformed certainty-weighted value difference, i.e. $\nu_t \equiv \nu^{(0)} + \nu^{(s)} \times VCR_t^0 \times \Delta VR_t^0$, where $\nu^{(0)}$ and $\nu^{(s)}$ are unknown parameters that control the offset and slope of the affine transform, respectively. Here, the strength of evidence in favor of a given alternative option is measured in terms of a signal-to-noise ratio on value. Note that the diffusion standard deviation σ is kept fixed across trials.
- DDM2: at each trial, the drift rate is set to the affine-transformed value difference, i.e. $\nu_t \equiv \nu^{(0)} + \nu^{(s)} \times \Delta VR_t^0$, and the diffusion standard deviation is allowed to vary over trials with value certainty ratings: $\sigma_t \equiv \exp\left(\sigma^{(0)} - \exp\left(\sigma^{(1)}\right) \times VCR_t^0\right)$. Here, $\sigma^{(0)}$ and $\sigma^{(1)}$ are unknown parameters that quantify the fixed and varying components of the diffusion standard deviation, respectively. In this parameterization, value representations that are more certain will be signaled more reliably. Note that the statistical complexity of DDM2 is higher than that of DDM1 (one additional unknown parameter).

For each subject and each DDM variant, we estimate unknown parameters from RT data alone using Equation A27, and derive out-of-sample predictions for changes of mind using Equation A28. We then measure the accuracy of trial-by-trial RT postdictions and out-of-sample change of mind predictions, in terms of the correlation between observed and predicted/postdicted variables. We also perform the exact same analysis under the MCD model (this is slightly different from the analysis reported in the main text, because only RT data is included in model fitting here).

To begin with, we compare the accuracy of RT postdictions, which is summarized in Appendix-Figure 10 below.

===== Insert Appendix-Figure 10 here. =====

One can see that the RT postdiction accuracy of both DDMs is higher than that of the MCD model. In fact, one-sample paired t-tests on the difference between DDM and MCD within-subject accuracy scores show that this comparison is statistically significant (DDM1: mean accuracy difference=12.3.%, s.e.m.=2.6%, $p<10^{-3}$; DDM2: mean accuracy difference=10.5%, s.e.m.=2.6%, $p<10^{-3}$; two-sided t-tests). In addition, one can see that DDM1 accurately captures variations in RT data that are induced by ΔVR^0 and VCR^0. However, DDM2 is unable to reproduce the impact of VCR^0 (cf. wrong effect direction). This is because, in DDM2, as value certainty ratings increase and the diffusion standard deviation decreases, the probability that DDM bounds are hit sooner decreases (hence prolonging RT on average). These results reproduce recent investigations of the impact of value certainty ratings on DDM predictions (Lee and Usher, 2020).

Now, Appendix-Figure 11 below summarizes the accuracy of out-of-sample change of mind predictions.

===== Insert Appendix-Figure 11 here. =====

It turns out that the MCD model exhibits the highest accuracy of out-of-sample change of mind predictions. One-sample paired t-tests on the difference between DDM and MCD within-subject accuracy scores show that this comparison reaches statistical significance for both DDM1 (mean accuracy difference=-5%, s.e.m.=2.4%, p=0.046; two-sided t-test) and DDM2 (mean accuracy difference=-9.9%, s.e.m.=3.4%, p=0.006; two-sided t-test). One can also see that neither DDM variant accurately predicts the effects of ΔVR^0 and VCR^0.

In brief, the DDM framework might be better than the MCD model at capturing trial-by-trial variations in RT data. This may not be surprising, given the longstanding success of the DDM on this issue (Ratcliff et al., 2016). The result of this comparison, however, depends upon how the DDM is parameterized (cf. wrong effect direction of VCR^0 for DDM2). More importantly, in our context, DDMs make poor out-of-sample predictions on decision outcomes, at least when compared to the MCD model. For the purpose of predicting decision-related variables from effort-related variables, one would thus favor the MCD framework.

10. Accounting for saturating γ-effect

When deriving the MCD model, we considered a linear γ-effect, i.e. we assumed that the variance of the perturbation $\delta(z)$ of value representation modes increases linearly with the amount z of allocated resources (cf. Equation 6 in the main text). However, one might argue that the marginal impact of effort on the variance of $\delta(z)$ may decrease as further resources are allocated to the decision. In other terms, the magnitude of the perturbation (per unit of resources) that one might expect when no resources have yet been allocated may be much higher than when most resources have already been allocated. In turn, Equation 6 would be replaced by:

$$\begin{aligned} \mu_i(z) &= \mu_i^0 + \delta_i \\ \delta_i &: N\left(0, f(z,\gamma)\right) \end{aligned} \quad \text{(A29)}$$

where the variance $f(z,\gamma)$ of the modes' perturbations would be a saturating function of z, e.g.:

$$f(z,\gamma) = \gamma_1\left(1-\exp(-\gamma_2 z)\right) \quad \text{(A30)}$$

where γ_1 is the maximum or plateau variance that perturbations can exhibit and γ_2 is the decay rate towards the plateau variance.

It turns out that this does not change the mathematical derivations of the MCD model, i.e. model predictions still follow Equations 9-14 in the main text, having replaced γz with $f(z,\gamma)$ everywhere.

Model simulations with this modified MCD model show no qualitative difference from its simpler variant (linear γ-effect), across a wide range of $\gamma_{1,2}$ parameters. Having said this, the modified MCD model is in principle more flexible than its simpler variant, and may thus exhibit additional explanatory power. We thus performed a formal statistical model comparison to evaluate the potential advantage of considering saturating γ-effects. In brief, we performed the same within-subject analysis as with the simpler MCD variant (see main text). We then measured the accuracy of model postdictions on each dependent variable, and performed a random-effect group-level Bayesian model comparison (Rigoux et al., 2014; Stephan et al., 2009). The results of this comparison are summarized on Appendix-Figure 12 below:

===== Insert Appendix-Figure 12 here. =====

First, one can see that considering saturating γ-effects does not provide any meaningful advantage in terms of MCD postdiction accuracy. Second, Bayesian model section clearly favors the simpler (linear γ-effect) MCD variant (linear efficacy: estimated model frequency=84.4% ±5.5%, exceedance probability=1, protected exceedance probability= 0.89). We note that other variants of the MCD model may be proposed, with similar modifications (e.g., nonlinear effort costs, non-Gaussian – skewed – value representations). Preliminary simulations seem to confirm that such modifications would not change the qualitative nature of MCD predictions. In other terms, the MCD model may be quite robust to these kinds of assumptions. Note that these modifications would necessarily increase the statistical complexity of the model (by inserting additional unknown parameters). Therefore, the limited reliability of behavioral data (such as we report here) may not afford subtle deviations to the simple MCD model variant we evaluate here.

11. Comparing MCD and model-free postdiction accuracy

The MCD model provides quantitative predictions for both effort-related and decision-related variables, from estimates of three native parameters (effort unitary cost and two types of effort efficacy), which control all dependent variables. However, the model prediction accuracy is not perfect, and one may wonder what is the added value of MCD compared to model-free analyses. To begin with, recall that one cannot make out-of-sample predictions in a model-free manner (e.g., there is nothing one can learn about effort-related variables from regressions of decision-related variables on ΔVR^0 and VCR^0). In contrast, a remarkable feature of model-based analyses is that training the model on some subset of variables is enough to make out-of-sample predictions on other (yet unseen) variables. In this context, MCD-based analyses show that variations in response times, subjective effort ratings, changes of mind, spreading of alternatives, choice confidence and precision gain can be predicted from each other under a small subset of modeling assumptions.

Having said this, model-free analyses can be used to provide a reference for the accuracy of MCD postdictions. For example, one may regress each dependent variable onto ΔVR^0, VCR^0, and indicator variables of experimental conditions (whether or not the choice is “consequential” and/or “penalized”), and measure the correlation between observed and postdicted variables.

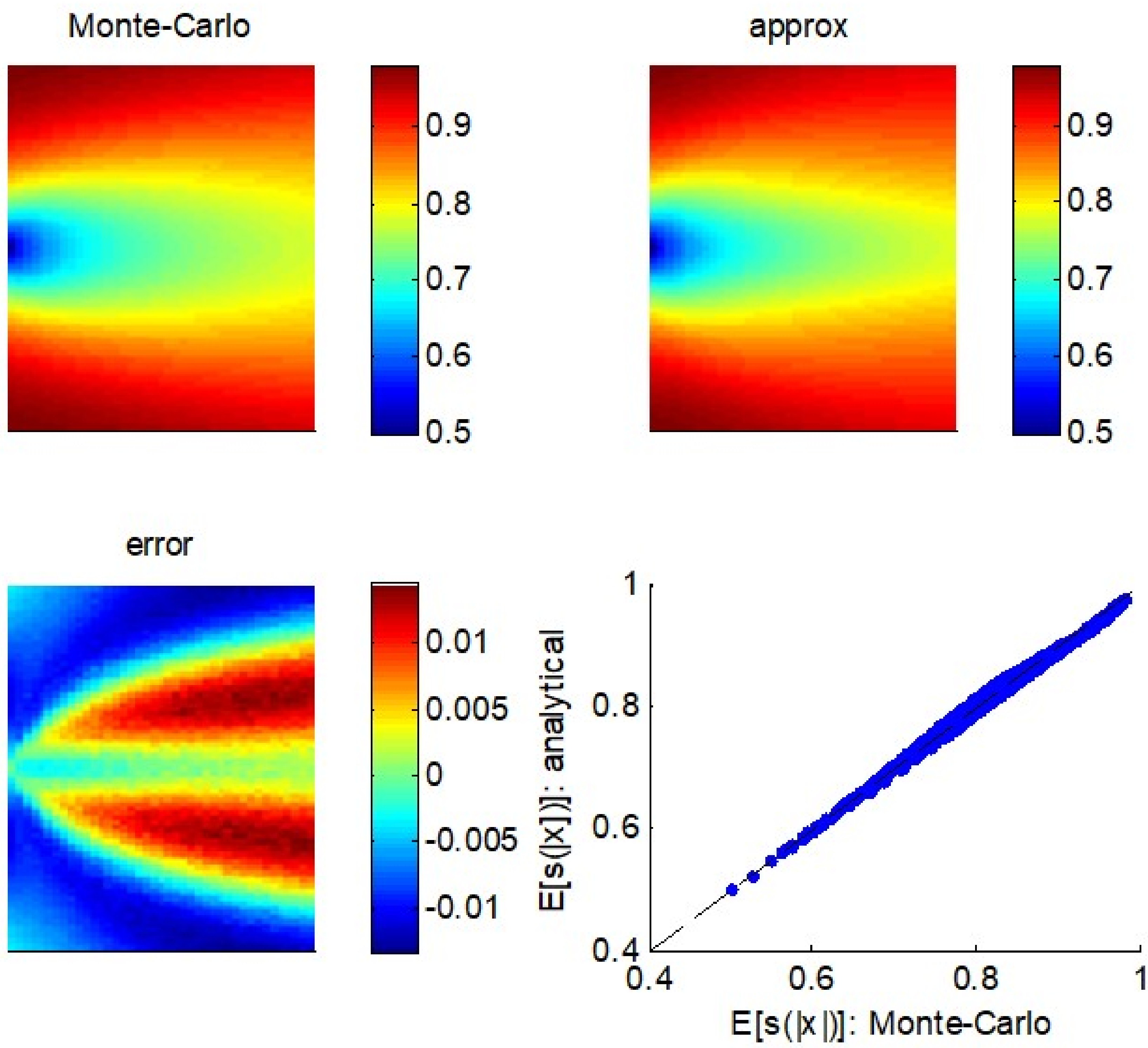
This provides a benchmark against which MCD postdiction accuracy can be evaluated. To enable a fair statistical comparison, we re-performed MCD model fits, this time fitting each dependent variable one by one (leaving the others out). In what follows, we refer to this as “MCD 1-variable fits”. The results of this analysis are summarized in Appendix-Figure 13 below:

===== Insert Appendix-Figure 13 here. =====

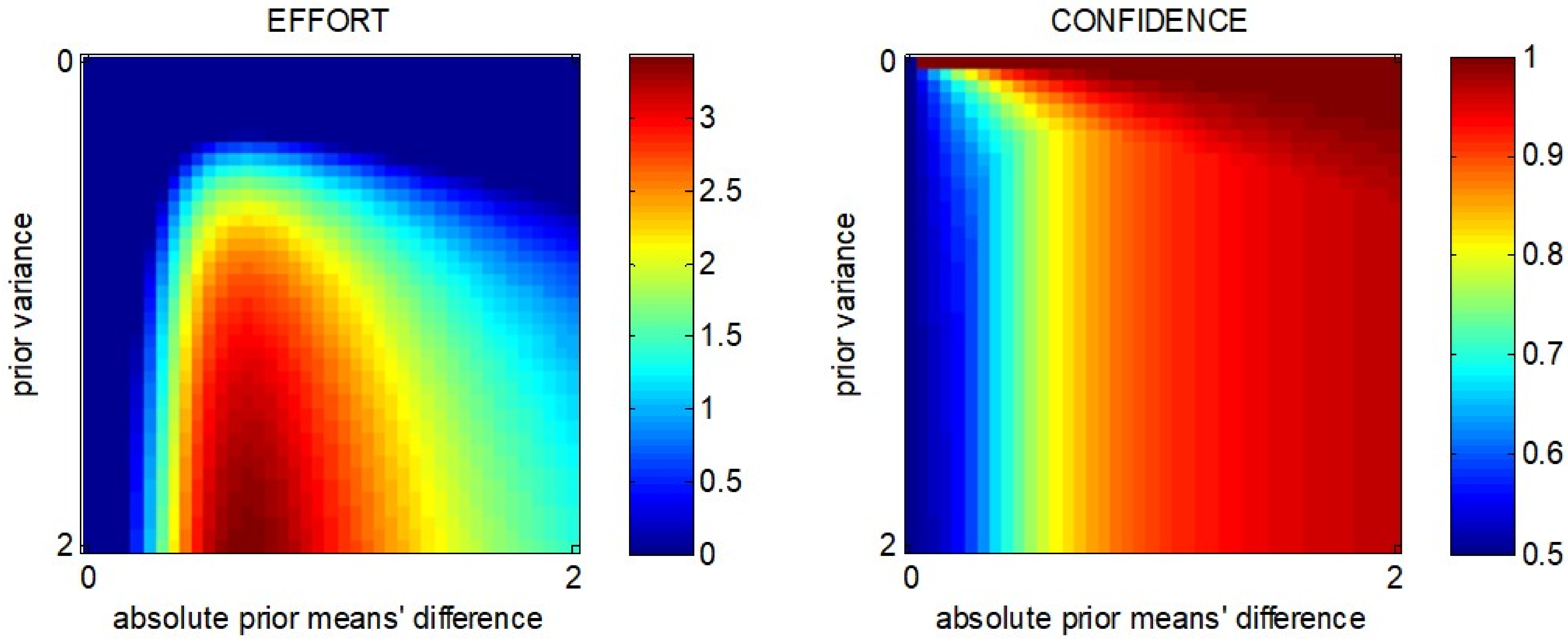
As expected, MCD 1-variable fits have better postdiction accuracy than the MCD “full-data” fit. This is because the latter approach attempts to explain all dependent variables with the same parameter set, which requires finding a compromise between all dependent variables.

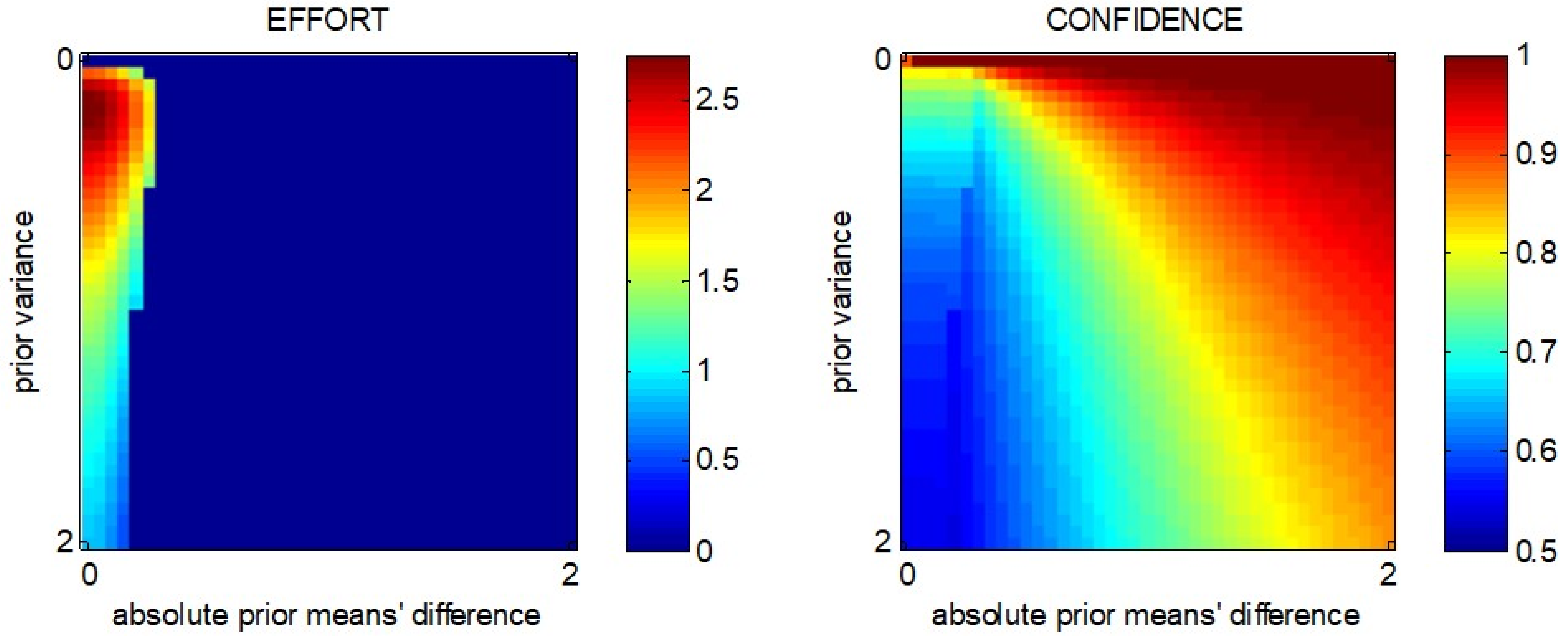
Now, model-free regressions seem to show globally better postdiction accuracy than MCD 1-variable fits: on average, the MCD model captures about 81% of the variance explained using linear regressions. However, the postdiction accuracy difference is only significant for effort-related variables (RT: $p=0.0002$, subjective effort rating: $p=0.0007$), but not for decision-related variables (choice confidence: $p=0.06$, spreading of alternatives: $p=0.28$, change of mind: $p=0.24$) except certainty gain ($p<10^{-4}$).

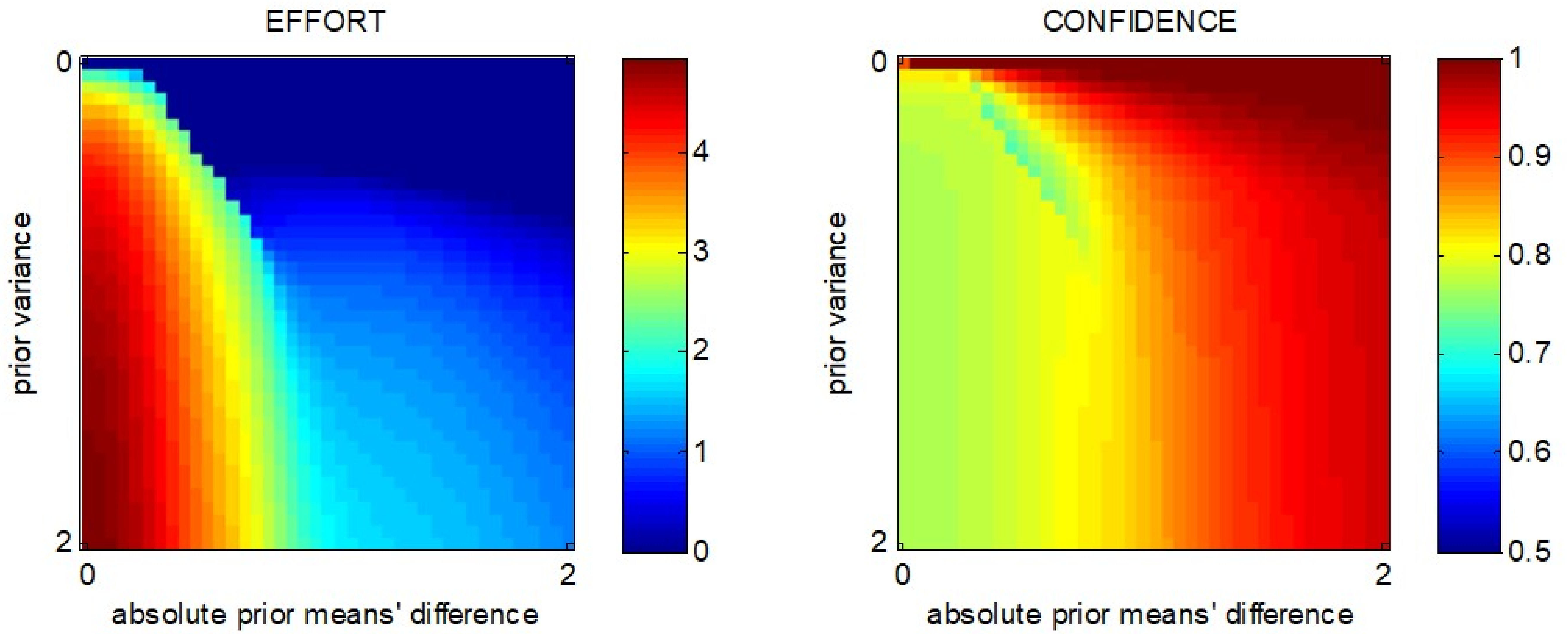
A likely explanation here is that the MCD model includes constraints that prevent 1-variable fits from matching the model-free postdiction accuracy level. In turn, one may want to extend the MCD model with the aim of relaxing these constraints. Having said this, these constraints necessarily derive from the modeling assumptions that enable the MCD model to make out-of-sample predictions. We comment on this and related issues in the Discussion section of the main text.

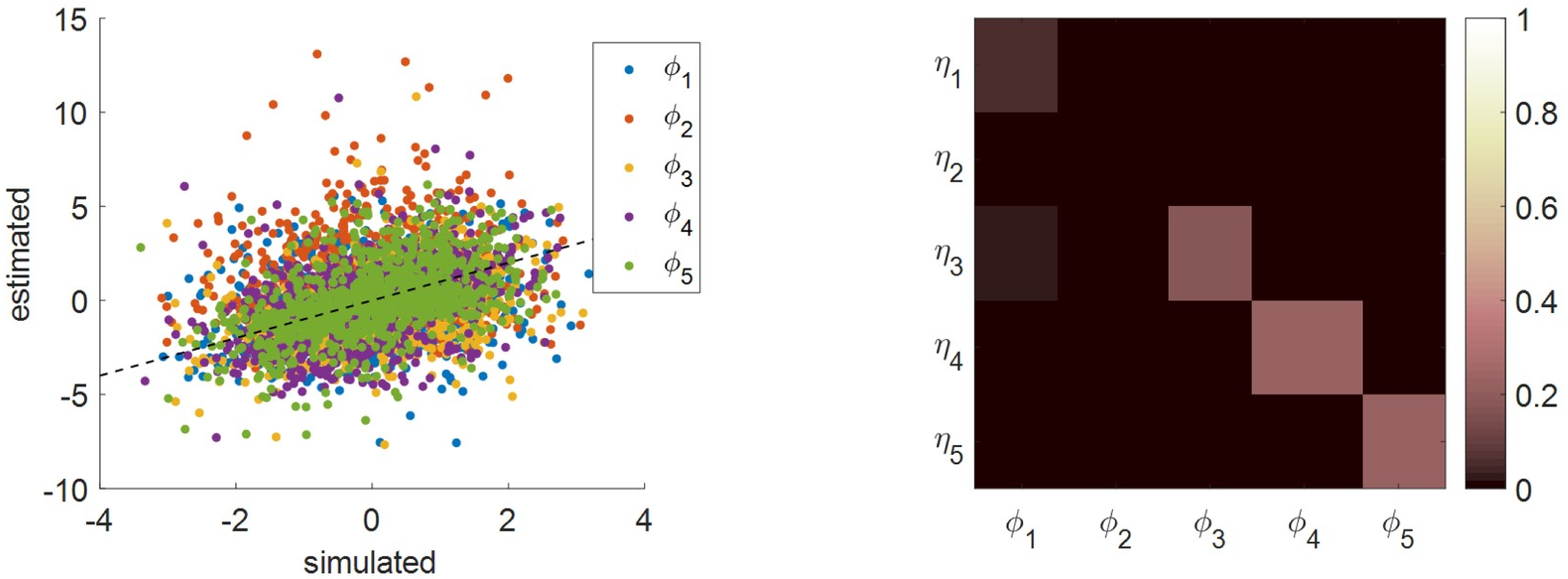


Monte-Carlo
approx
0.9
0.8
0.7
0.6
0.5
error
0.01
0.005
0
-0.005
-0.01
E[s(|x|)]: analytical
E[s(|x|)]: Monte-Carlo
1
0.8
0.6
0.4

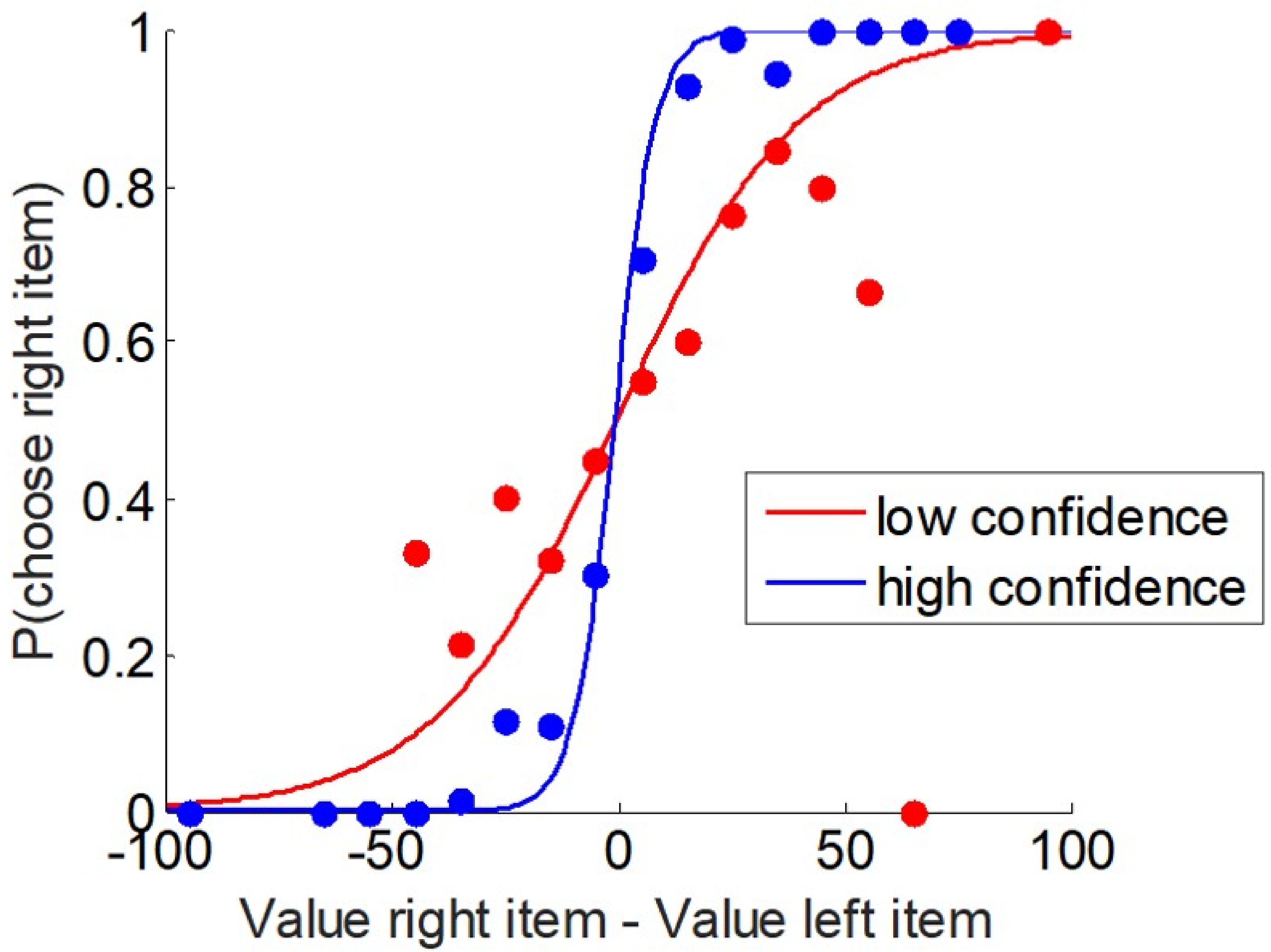

EFFORT
CONFIDENCE
prior variance
absolute prior means' difference
0
2
0.5
1
1.5
2.5
3
0.6
0.7
0.8
0.9


EFFORT
0
2
prior variance
0
2
absolute prior means' difference
2.5
2
1.5
1
0.5
0
CONFIDENCE
0
2
prior variance
0
2
absolute prior means' difference
1
0.9
0.8
0.7
0.6
0.5

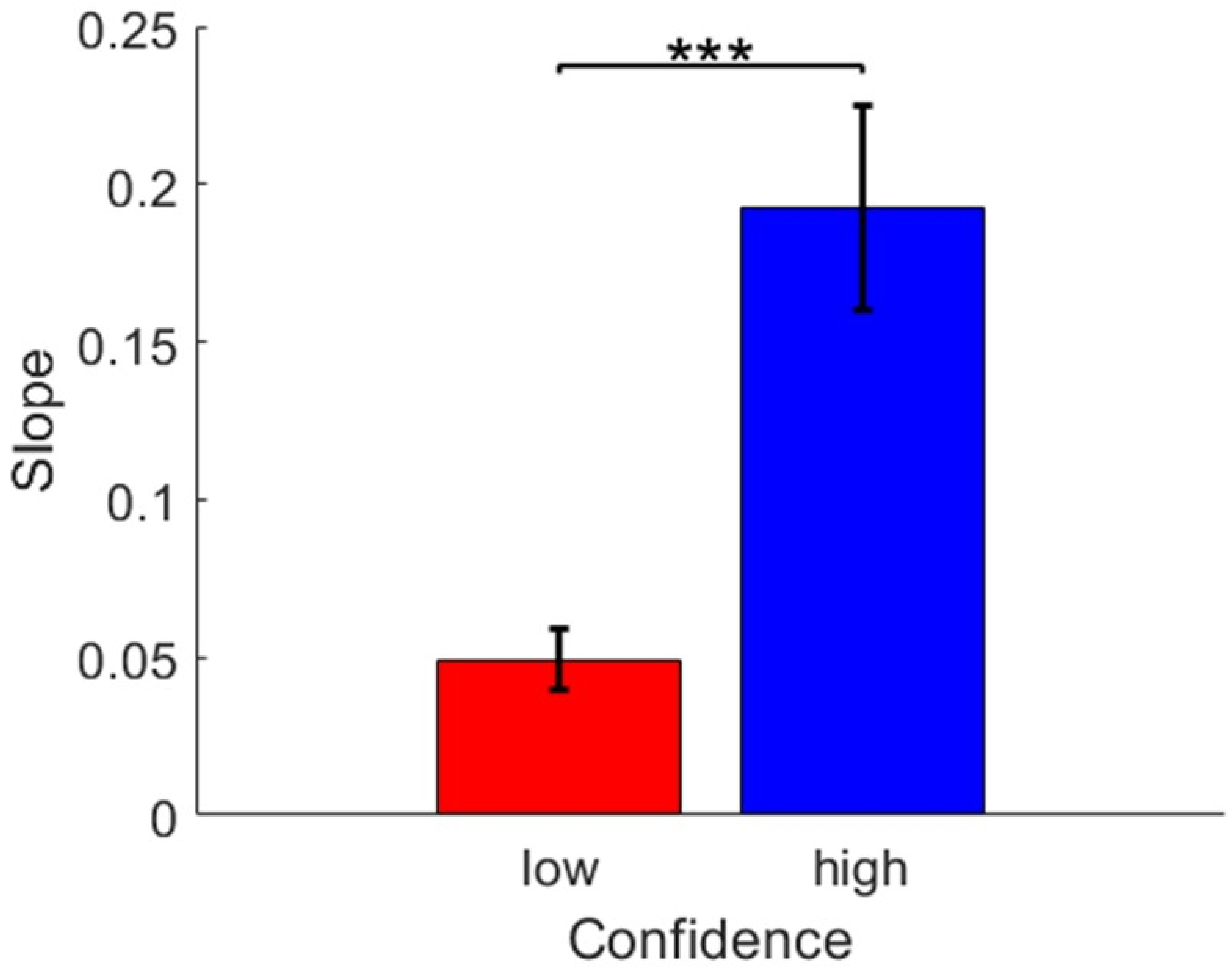

EFFORT
prior variance
absolute prior means' difference
0
2
0
2
0
1
2
3
4
CONFIDENCE
prior variance
absolute prior means' difference
0
2
0
2
0.5
0.6
0.7
0.8
0.9
1



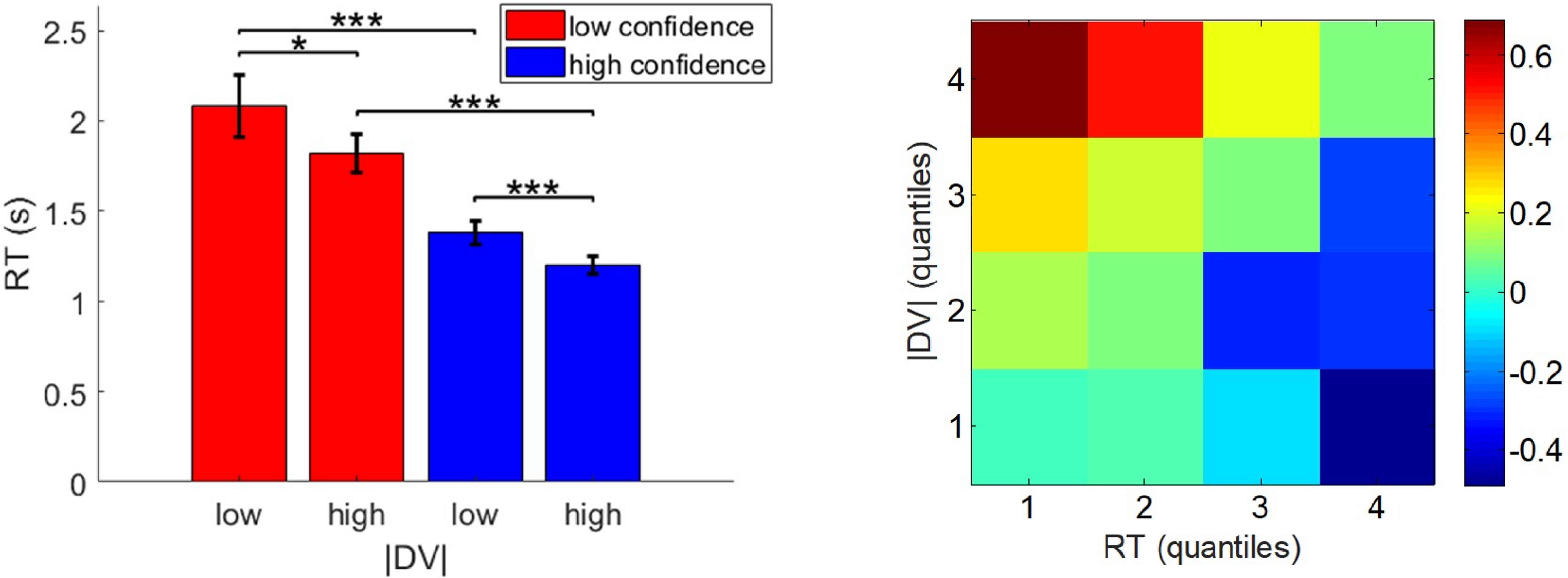
15
10
5
0
-5
-10
-4
-2
0
2
4
estimated
simulated
ϕ_1
ϕ_2
ϕ_3
ϕ_4
ϕ_5
η_1
η_2
η_3
η_4
η_5
ϕ_1
ϕ_2
ϕ_3
ϕ_4
ϕ_5
1
0.8
0.6
0.4
0.2
0



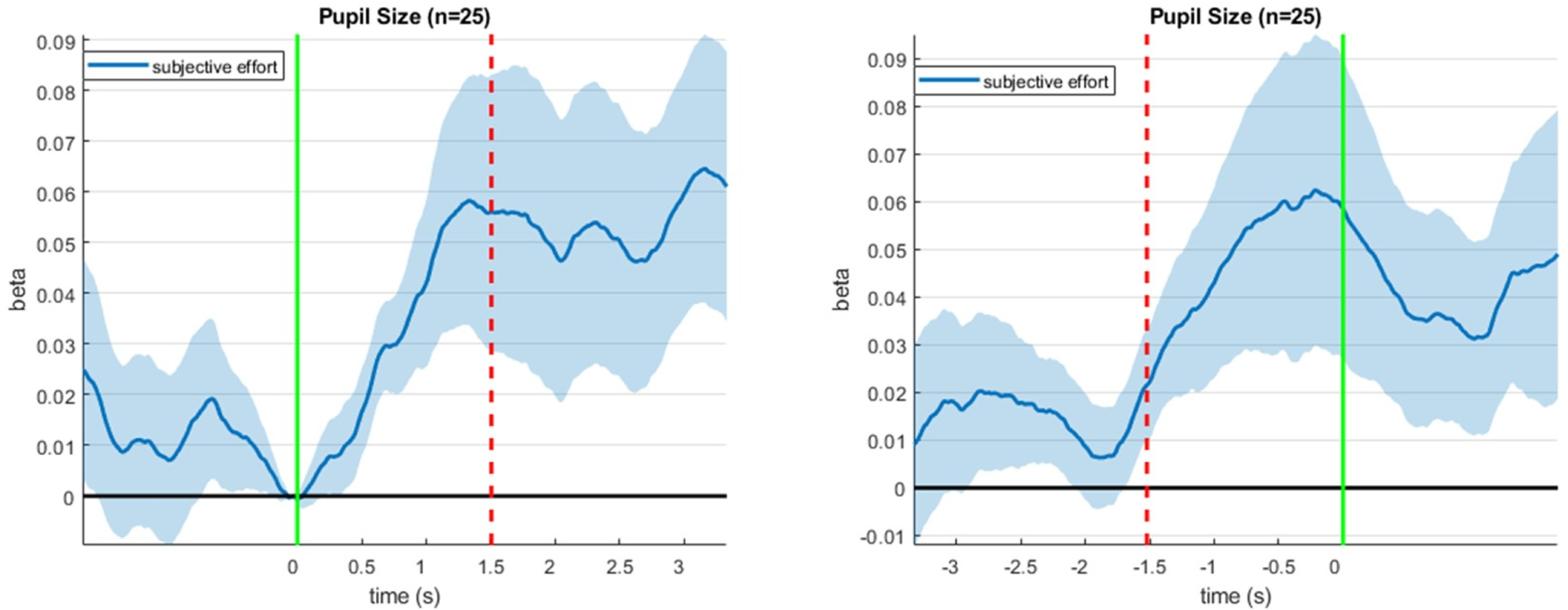
1
0.8
0.6
0.4
0.2
0
P(choose right item)
-100
-50
0
50
100
Value right item - Value left item
low confidence
high confidence



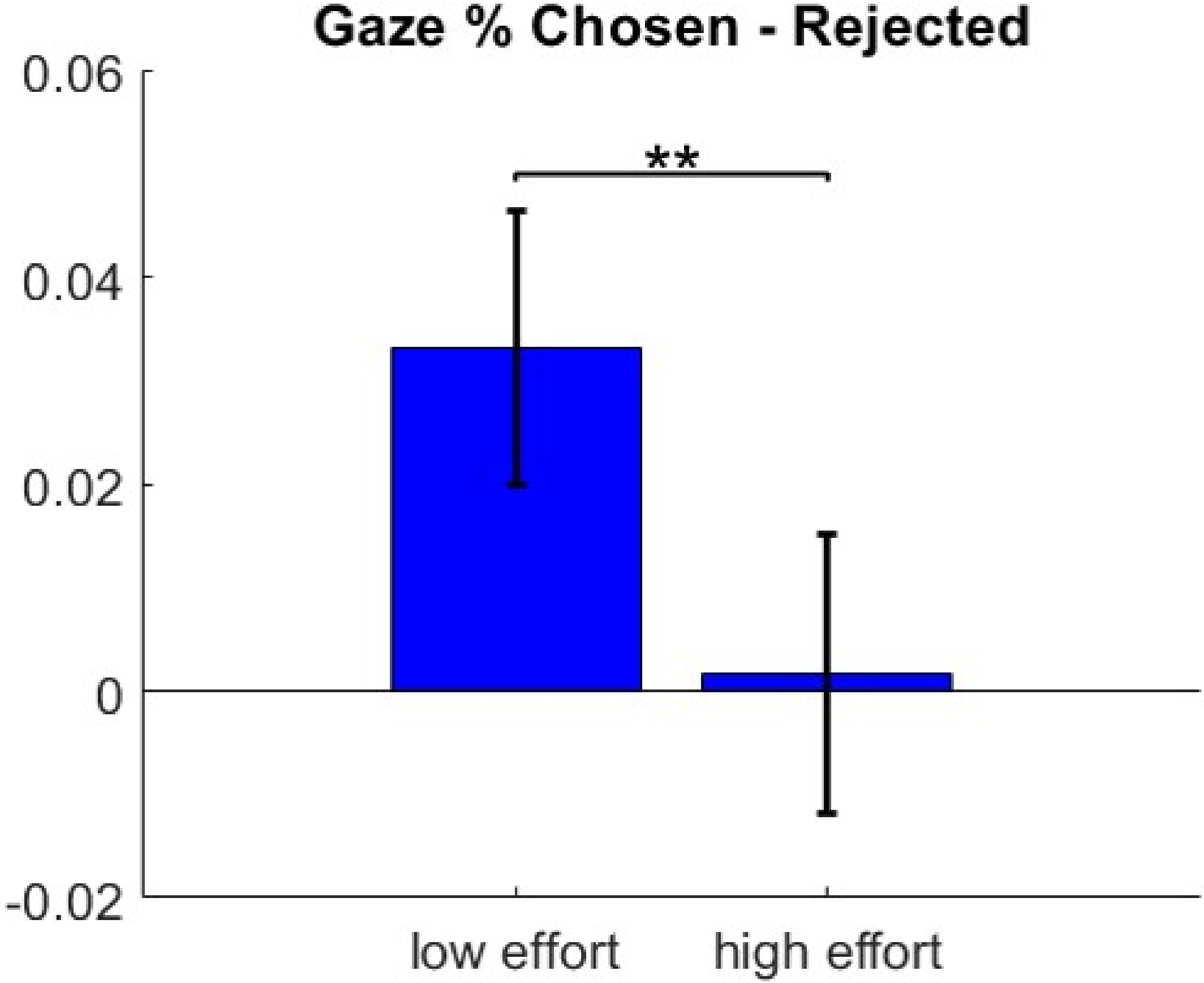
0.25
0.2
0.15
0.1
0.05
0
Slope
***
low
high
Confidence



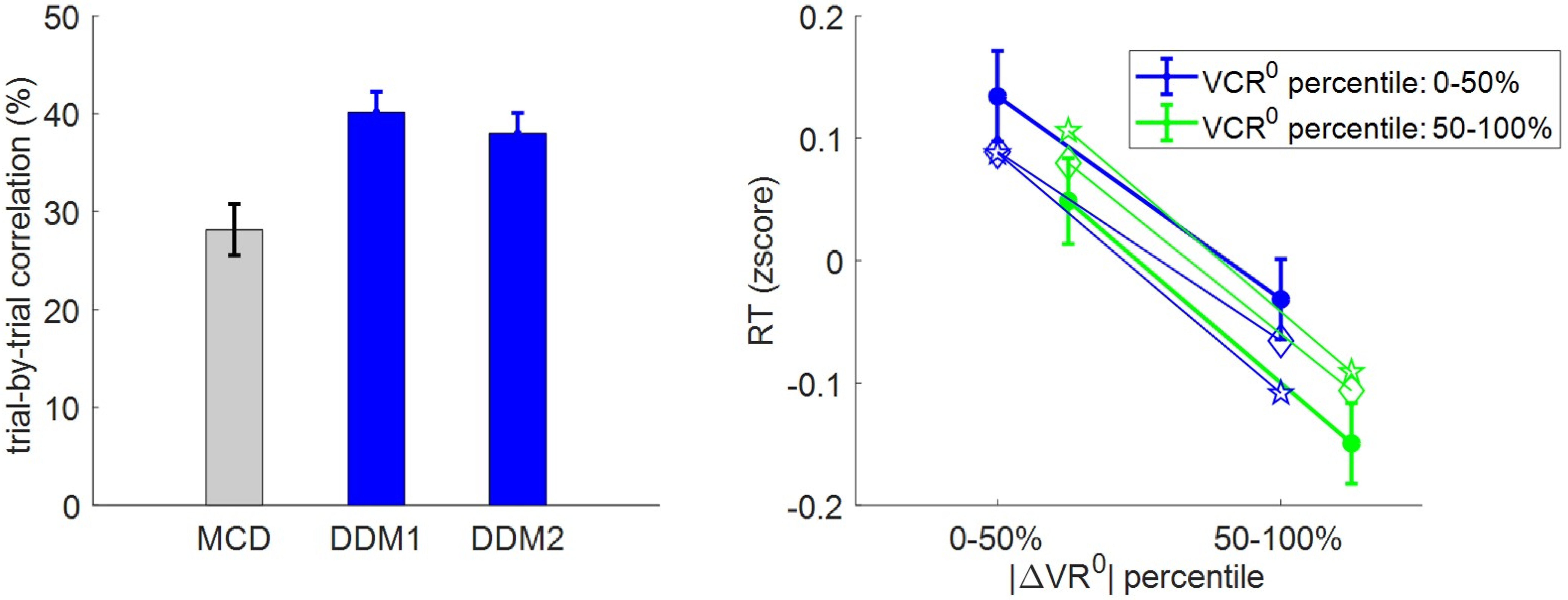
low confidence
high confidence
***
*
***
***
RT (s)
2.5
2
1.5
1
0.5
0
low
high
low
high
|DV|
|DV| (quantiles)
4
3
2
1
RT (quantiles)
0.6
0.4
0.2
0
-0.2
-0.4



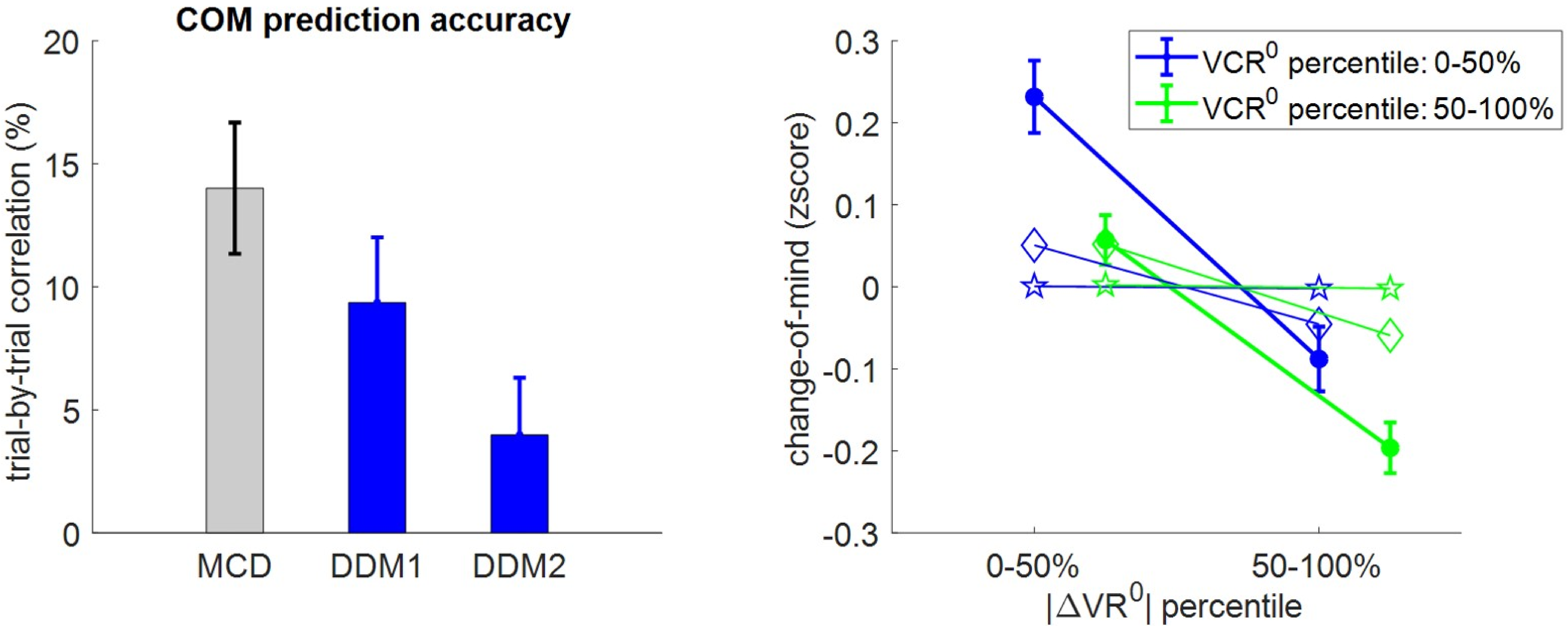
Pupil Size (n=25)
subjective effort
beta
time (s)
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
0.5
1
1.5
2
2.5
3
Pupil Size (n=25)
subjective effort
beta
time (s)
-0.01
-3
-2.5
-2
-1.5
-1
-0.5



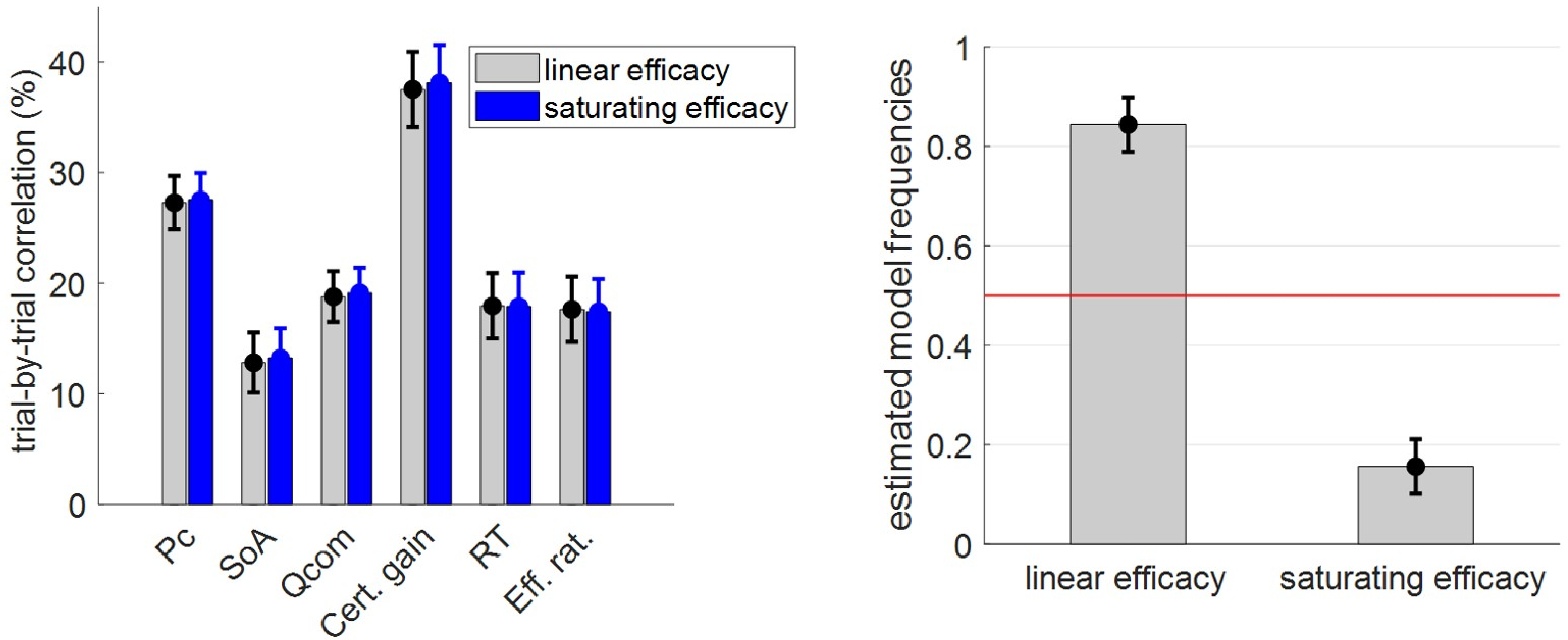
Gaze % Chosen - Rejected
0.06
0.04
0.02
0
-0.02
**
low effort
high effort



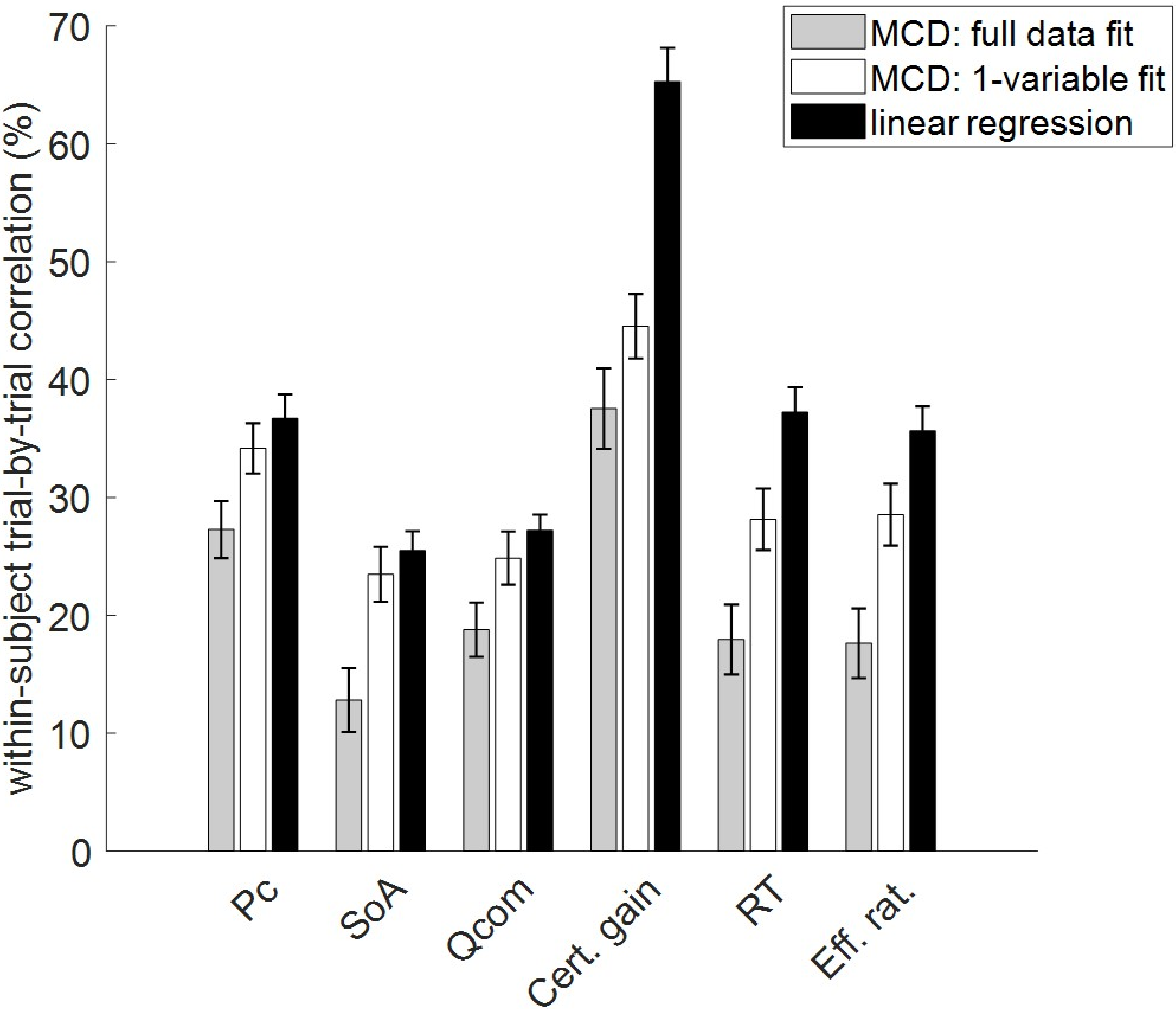
50
40
30
20
10
0
trial-by-trial correlation (%)
MCD
DDM1
DDM2
0.2
0.1
0
-0.1
-0.2
RT (zscore)
VCR⁰ percentile: 0-50%
VCR⁰ percentile: 50-100%
0-50%
50-100%
|ΔVR⁰| percentile



COM prediction accuracy
trial-by-trial correlation (%)
20
15
10
5
0
MCD
DDM1
DDM2
change-of-mind (zscore)
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
VCR0 percentile: 0-50%
VCR0 percentile: 50-100%
0-50%
50-100%
|ΔVR0| percentile



trial-by-trial correlation (%)
0
10
20
30
40
linear efficacy
saturating efficacy
Pc
SoA
Qcom
Cert. gain
RT
Eff. rat.
estimated model frequencies
0
0.2
0.4
0.6
0.8
1
linear efficacy
saturating efficacy



MCD: full data fit
MCD: 1-variable fit
linear regression
within-subject trial-by-trial correlation (%)
0
10
20
30
40
50
60
70
Pc
SoA
Qcom
Cert. gain
RT
Eff. rat.