



**HAL**  
open science

# The metacognitive control of decision-making

Douglas Lee

► **To cite this version:**

Douglas Lee. The metacognitive control of decision-making. *Neurons and Cognition [q-bio.NC]*. Sorbonne Université, 2019. English. NNT : 2019SORUS525 . tel-03030121

**HAL Id: tel-03030121**

**<https://theses.hal.science/tel-03030121>**

Submitted on 29 Nov 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



ed3c



Inserm



# Sorbonne Université

Ecole doctorale Cerveau Cognition Comportement

*Institut du Cerveau et de la Moelle Epinière / Motivation Brain Behavior team*

## The Metacognitive Control of Decision-Making

Présentée par

**Douglas LEE**

Thèse de doctorat de Neurosciences Cognitives

Dirigée par Jean DAUNIZEAU

Présentée et soutenue publiquement le 06 Décembre 2019

Devant un jury composé de :

M. Jean DAUNIZEAU	Sorbonne Université	Directeur de Thèse
M. Christian RUFF	University of Zurich	Rapporteur
M. Vincent DeGARDELLE	Panthéon-Sorbonne Université	Rapporteur
M. Falk LIEDER	Max Planck Institute	Examinateur
Mme. Liane SCHMIDT	Sorbonne Université	Examinatrice
M. Mehdi KHAMASSI	Sorbonne Université	Examinateur
M. Valentin WYART	Ecole Normale Supérieure	Membre Invité

## Acknowledgments

First, I would like to thank the researchers who kindly agreed to be part of my jury and to do me the honor of evaluating my work: Falk Lieder, Liane Schmidt, and Mehdi Khamasi, with a special thank you to Christian Ruff and Vincent de Gardelle, who accepted to review this dissertation. I would also like to thank Valentin Wyart, for agreeing to serve as a guest member in order to satisfy the jury selection requirements of my doctoral school. Finally, I thank Vincent and Valentin again for serving on my annual “thesis review committee” and offering feedback that helped me to progress on my project scientifically, as well as offered psychological comfort in knowing that external reviewers approved of the work I was doing.

I gratefully acknowledge Labex Bio-Psy for providing me with the funding to finance my research and time spent working towards my PhD. I would like to personally thank Brigitte Bouchard for always helping out with funding and administrative issues. I am honored to have been supported by the Labex Bio-Psy, without which this work could not have been completed.

My thanks also go to the ICM support teams, and in particular the PRISME platform, managed by the very helpful Karim N’Diaye and Gilles Rautureau. Without this platform and the help of these guys, my behavioral experiments would have been very difficult to conduct properly. Karim was always ready to help with recruiting, planning, or whatever else might have been needed. Gilles was always ready to help with any technical issues that might arise... as they often did!

I am particularly grateful to my thesis supervisor, Jean Daunizeau, for giving me the opportunity to work together. Jean had faith in my abilities from the very beginning. Although sometimes it felt like he perhaps placed too much faith in me, when I sometimes doubted myself and my decision to become a researcher, I think he knew what he was doing. The struggle to persist even when times are tough and it seems that it would be easier to just give up eventually leads to progress and development. Jean made sure to listen to me when I had concerns, to support me when I asked for it, and to let me suffer (in a good, pedagogical way). I won’t say that it was always pleasant, but I do value the experience that I have had with Jean as my mentor. He taught me a lot, both directly and indirectly. Above all, I truly appreciate that he gave me the opportunity to work with him and join the lab, even though I came from an entirely different academic background and had very limited scientific experience. I dare say that Jean’s gamble paid off, as we have produced some good work together, and collectively (with the team and other colleagues) had many hours of stimulating scientific conversation.

I am also grateful to Mathias Pessiglione, for his part in inviting me and welcoming me to the team. Mathias was also very supportive (during lab meetings, for example) of my development. He was always receptive of my ideas, yet always ready to challenge them. I am also grateful to Mathias for all of the times that he supported me with a letter or other form of recommendation (in addition to those of Jean). I am certain that the personal recommendations that I received from both Jean and Mathias helped me tremendously in meeting and discussing with other top researchers in the field, and that has been a wonderful way for me to learn.

I am grateful to the other PIs for their never-ending support and encouragement, as well as for their intellectual insight. Thanks Sebastien Bouret, Fabien Vinckier, and Raphael LeBouc!

One of the best parts of my PhD experience has been the MBB team. Scientifically, the team is excellent. Lab meetings for me were a treat, and random hallway discussions were a further delight. The lab organization and sheer size provided me with a great opportunity for scientific growth. But, beyond the science, being part of the team was a social joy. Each labmate had a very different personality, but that was part of made it such a nice experience to get to know and work alongside them. Though the composition of the team was always in flux (as is always the case in academia), everyone was always there to offer support when needed. I will look back very fondly on my days (years!) spent with the MBB, with a special place in my memories for each individual member that I was lucky enough to meet along the way. Here are just a few of the special thanks that I want to commemorate: Bastien—he was the first to make an effort to befriend me and get to know me, and a great help with both scientific and social matters; Alizee—she was basically my go-to person for help when I was new to the team, and she was always willing to be of assistance; Emma—she consistently offered great scientific advice, practical help, and pleasant social interaction; Quentin—our discussions were always very pleasant, no matter what the topic (and it didn't hurt that we have a similar style of thinking!); Antonius—I think everyone would agree that having Toni on the team is an all-around positive, but he has also been there to support me personally, for which I am very grateful.

Finally, I want to thank my family (my mom and my little sister, in particular). I can't claim that they have always (or ever?) understood the choices that I make in life, and moving to France to pursue a PhD in cognitive neuroscience surely was not something any of them would have ever predicted! Nevertheless, they have always made clear that I have their love and support, no matter what choices I make. I know they are all (including my dad and my brothers) very proud of me.

# TABLE OF CONTENTS

Chapter 1: Background.....	2
1. Subjective Value and Preferences.....	3
2. Preference Reversal.....	8
3. Constructed Preferences.....	23
4. Choice-Induced Preference Change.....	29
5. Multiple Option Decisions.....	37
6. Computational Mechanisms.....	39
Chapter 2: The Metacognitive Control of Decision-Making Model.....	55
1. Working hypothesis.....	56
2. Computational model.....	60
3. Impact of Input Variables on Predicted Output Variables.....	71
4. The online metacognitive control of decision-making.....	81
Chapter 3: Behavioral / Computational Studies.....	85
1. Choosing what we like versus liking what we choose.....	86
2. Trading Effort for Confidence.....	111
3. Implementation of Allocated Resources.....	179
4. Multi-Option EVC.....	203
Chapter 4: General Discussion.....	238
Conclusion.....	243
Bibliography.....	244
Appendix: Unresolved Issues.....	258

# **Chapter 1**

## **Background**

Decision-making is a fundamental component at the very core of the human experience. Every action that we take is the result of a decision made; every thought that we make provides input for decisions yet to be made. It is therefore tempting to claim that the ability to make “good” decisions, howsoever defined, is a critical trait for an archetypal member of the species to possess. Upon accepting such a claim, it should follow that one accept the importance of the study of the human decision-making process. Yet though this logic might be easy to accept, it is far more difficult to recognize good decision-making and distinguish it from bad decision-making. The principal reason for this is the subjective nature of the drivers of the decision process. Decision-making could be defined as “the process of identifying and choosing alternatives based on the values, preferences and beliefs of the decision-maker” [<https://en.wikipedia.org/wiki/Decision-making>]. Thus, modern decision-making researchers often focus on so-called *subjective value* as the key driver of the decision process.

### **Subjective Value and Preferences**

According to the Oxford English dictionary, the term *value* is principally defined as “the regard that something is held to deserve; the importance, worth, or usefulness of something.” It is readily apparent from this definition that value is at the core of decision-making. All else equal, a decision maker will always seek to choose the option that is the most useful to him. Moreover, value can be considered from an evolutionary perspective, where the value of something derives from the extent to which it facilitates survival or genetic propagation. From this perspective, it becomes clear why the brain would be equipped with a motivational system that determines the values of things according to their contribution to adaptive fitness. However, what is not explicitly declared in either of these definitions of value is the fact that it need not be universal. That is, the value that something holds in the estimation of one individual need not be the same

as the value estimation of another individual. In this regard, value is not something that exists outside of the mind of an individual decision maker. Instead, value can be thought of as an individual's subjective assessment, or feeling, of how much he likes or wants a certain decision option. This *subjective value* is based on the decision maker's information about the option itself, and on his beliefs about how such an option might benefit him if he were to choose it. The set of information and beliefs used to establish subjective value comes from the personal experience of the decision maker (e.g., observation of the environment, memory of past observations). The impression that a decision maker has about how valuable an option is, therefore, can be felt with more or less certainty, depending on the quantity and quality (e.g., reliability, consistency) of the information used to establish it. Of note, subjective value is something that can evolve over time as new information is obtained, old information is forgotten, or personal goals change.

Subjective value theory advances the idea that decision options (e.g., actions or tangible goods) do not hold any inherent value. Rather, their "value" is entirely determined by the degree to which the decision maker believes they will help him achieve his personal goals, along with the relative importance of those goals in the life of the decision maker. In this way, each individual decision maker might value the decision options in a way that differs tremendously from those of another decision maker. In this case, all decision makers would nevertheless be "correct" about the value of the options, regardless of how dramatically they differed in opinion, provided that the values represented the true subjective beliefs of the decision makers. The ordering of a set of options based on an individual decision maker's subjective valuation is referred to as that individual's *preferences* over that set. This concept of personal preferences is central to the study of decision-making in the fields of psychology, economics, and neuroscience.

Whatever the discipline, experimental research on decision-making requires a clear observation of the preferences of the decision makers. Perhaps the most traditional method of eliciting such

preferences is by survey, wherein the researcher directly asks the decision makers to state how strongly they feel about each particular decision option. This typically takes the form of a question such as, “How much do you like this item?” to which the decision makers respond according to their subjective assessment. The range of possible responses is generally bounded on both ends for practical reasons, and the granularity of the rating scale is determined by the researcher (e.g., 1 to 10 by increments of 1 or 0.1). This method has historically been the standard used by psychology researchers, and it remains prevalent in modern research. Strictly speaking, this method does not actually solicit preferences, merely subjective valuations. However, the link between valuations and preferences is direct and obvious, and implicit preferences can thus be inferred from the ratings themselves. To be clear, a stated valuation of Option 1 as 7 out of 10 and Option 2 as 6 out of 10 can be taken as a stated preference for Option 1 over Option 2 without much argument. Note that some researchers prefer to ask the decision makers to rank the options instead of rating them, making the inference of preferences perhaps even more straightforward.

Even though it might seem that the connection between self-reported valuations and inferred preferences cannot be questioned, some researchers nevertheless take issue with the implicit nature of preferences solicited in this way (i.e., the decision maker never actually declares his preferences, they must be inferred by the researcher). Furthermore, some question the accuracy of the self-reported valuations themselves, skeptical that the decision makers can adequately access their internal value estimates and translate them precisely onto the experimental rating scale. For these reasons, some researchers prefer to elicit preferences directly and explicitly through the use of a two-alternative forced-choice setup. Here, decision makers are presented with pairs of options and asked to choose which of the pair they prefer. As the response to this sort of question is equivalent to the statement, “I prefer Option 1 over Option 2,” the preference is explicit, leaving no room for it to be questioned as such. This method is especially popular

with economics researchers, and has been since the introduction of what they call *Revealed Preference Theory* [Samuelson, 1938]. The basic idea is that the preferences that an individual holds are hidden inside his mind and not observable, perhaps even by his own self. By choosing one option over another, however, the decision maker reveals his preference (the assumption is that a decision maker will always choose the option that he prefers). By revealing a collection of pairwise preferences, the preferences over the entire decision set can be established. This method can thus be quite useful for establishing preference ordinally, though it is not able to establish them cardinally (as the rating method is).

#### *Neural Correlates of Subjective Value*

A wide gamut of research has exposed a variety of neural correlates of subjective value. One team of researchers used functional magnetic resonance imaging (fMRI) to record brain activity while participants placed monetary bids for the right to eat different food items after the experiment [Plassmann, O'Doherty & Rangel, 2007]. The results showed that levels of activity in the medial orbitofrontal cortex (mOFC) and in the dorsolateral prefrontal cortex (dlPFC) correlated with willingness to pay (WTP), an economic framing of subjective value. Another team used fMRI to record brain activity during an intertemporal choice task (i.e., one option was larger than the other, but required a delay before receipt) [Kable & Glimcher, 2007]. The results of this study showed that neural activity in the ventral striatum (VS), medial prefrontal cortex (mPFC), and posterior cingulate cortex (PCC) all correlated with the revealed subjective value of the choice options. It has also been shown (using fMRI) that activity in the VS during market exchange (buying or selling) of lottery tickets indexed the subjective values of the options [DeMartino et al, 2009]. Here, the subjective values represented distortions in the mathematical expected values caused by the so-called *endowment effect*, in which people who already own a choice option tend to appraise it at a higher value than those who do not [Knetsch, 1989].

Developing beyond simple representation of subjective value, one study demonstrated that the concept of subject value was encoded similarly in the brain regardless of what sort of options were under consideration [Chib et al, 2009]. In this study, activity in the ventromedial prefrontal cortex (vmPFC) correlated with subjective valuations similarly across a variety of choice tasks (food, nonfood consumables, monetary gambles), which suggests that the brain encodes a “common currency” capable of encoding value for different categories [Chib et al, 2009]. Taking this a step further, later work showed that distinct brain regions encode subjective value for distinct types of choice option, which potentially inform a unified valuation network that translates the distinct values into a common currency to enable directly comparison and enable choice across reward types [Levy & Glimcher, 2011; see Levy & Glimcher, 2012 for a meta-analysis]. Here, food values were encoded by the hypothalamic region, monetary values were encoded by the PCC, and the common currency was encoded by the vmPFC. Another study confirmed the common encoding of value in the vmPFC and also found it in the VS, while simultaneously identifying a network that signals the salience of the option (e.g., the magnitude of its subjective value, whether positive or negative) [Bartra, McGuire & Kable, 2013]. Later research expanded the common valuation network to include the PCC along with the vmPFC and VS, and also identified a posterior-to-anterior gradient of value representations corresponding to concrete-to-abstract rewards [Clithero & Rangel, 2013]. The activity patterns in this network were consistent whether analyzed with respect to decision onset or conclusion, as well as for both monetary and primary (e.g., food) rewards. Taken together, the above findings leave little doubt that there is a brain system whose apparent job is to construct values, and that this same brain system is recruited whether the decision maker assesses the value of a single option or when states his preference between alternative options. While this does not provide evidence that subjective values actually conform to the goals of the decision maker, it does demonstrate that the theory of subjective value is realistic and biologically plausible.

## Preference Reversal

In principle, observed preferences should be identical regardless of whether they are inferred by ratings or revealed by choices, as both methods are believed to capture the same underlying mental representation of relative value. Indeed, this expectation holds in general. Countless studies have demonstrated a strong correlation between rating preferences (e.g., Option 1 rating > Option 2 rating) and choice preferences (e.g., Option 1 chosen versus Option 2). This correlation is so natural and irrefutable that it is often used as a sanity check in modern experimental studies, to verify that the rating and choice data were collected properly. However, it has long been demonstrated that the correlation does not always hold. To be clear, this means that a decision maker might sometimes assign a higher value rating to Option 1 than to Option 2, yet choose Option 2 versus Option 1. Furthermore, violations of the correlation between rating and choice have been shown to appear in systematic ways, ruling out experimental noise as the sole explanatory factor. Such violations have been dubbed *preference reversals* (i.e., the preferences of a decision maker reversed between the rating and choice tasks), and have spawned a vast body of literature in their own right.

Reports of preference reversal first entered the literature with reference to decision making under risk [Lichtenstein & Slovic, 1971]. The decision options in this domain are typically a relatively low payoff amount with a relatively high payoff probability (Option P) versus a relatively high payoff amount with a relatively low payoff probability (Option \$). The complementary probability was assigned to a nominal loss in all trials. The concept of risk here refers to the chance that a decision maker might lose money, although it would similarly apply to the possibility of not actually receiving the chosen payoff in the case that there were no possibility of loss. This study showed that classroom participants who chose Option P often rated Option \$ as being more valuable, inconsistencies that “violate every risky decision model” [Lichtenstein & Slovic, 1971]. The authors replicated their findings with experienced gamblers

in a Las Vegas casino, suggesting that the effect was not caused by artificial experimental circumstances [Lichtenstein & Slovic, 1973]. Soon after the initial paper was published, other researchers started to replicate the findings as well, proving that the phenomenon was real and persistent [Lindman 1971].

Hoping to discredit the findings of the aforementioned psychology studies, a team of economists conducted a particularly thorough series of experiments incorporating specific procedural variations that they believed would preclude the observed effect [Grether & Plott, 1979]. The authors began by examining a set of hypotheses based on potentially relevant economic theory, psychological theory, and experimental methods bias. These hypotheses are summarized below.

#### *Theory 1: Mis-specified Incentives*

The basic idea is that hypothetical choices, choices between options with imaginary payoffs, choices for which the payoff structure is ill-defined, et cetera, are choices that yield either no consequences or consequences of no importance to the decision maker. Preference reversal demonstrated via this type of decision setup might not be informative about real consequential decisions.

In their study, Grether and Plott always included real monetary payoffs, which were clearly communicated to participants before the experiments began. For choice, one trial was selected randomly (using a bingo cage) at the end of the experiment, and the option that they had selected on that trial was carried out (using a roulette wheel set up with probability of winning matching the chosen option). For rating, one trial was selected randomly and a BDM auction was implemented [Becker, DeGroot & Marschak, 1964]. In this BDM auction, a random number was generated within the same range that was available for the participant to value the options. If the number was larger than the value rating, the option was implemented (i.e., the gamble

was carried out); if smaller, the participant received a payoff equal to the randomly-generated number. The BDM auction is commonly accepted by economists as a way to ensure that value ratings actually reveal true preferences.

### *Theory 2: Income Effects*

Previous research has shown that there is an income effect on the attitudes towards risk that people exhibit. Specifically, the higher the income or wealth a person has, the more likely he is to seek risk (or more reasonably, the less likely he is to avoid risk). Grether and Plott suggested that the nature of the design of the previous studies might have created an income effect that would explain the alleged preference reversals. Their idea was as follows. The gambles that participants selected would only be enacted at the end of the experiment. The mathematical expected value of each gamble was a positive amount, and so as participants progressed through the experiment by selecting additional gambles, they formed some sort of mental account of a cumulative net gain that they expected to receive at the end of the experiment. Because the value estimates for each individual option were solicited after all choices had already been made, this expected gain might have somehow played the role of a real income, thus causing the participants to overvalue the gambles relative to their initial beliefs. This would be an income effect, because the overvaluation would in reality be an under-devaluation of the payoff amount due to the risk involved.

Grether and Plott made one simple change to the experimental design in order to prevent an income effect. Rather than enact all selected gambles, they randomly selected one trial at the end of the experiment and enacted only that trial. This way, participants should not have formed any expectation of increasing wealth that grew over the course of the experiment.

### *Theory 3: Indifference*

Previous studies required participants to select one option from each pair as their preferred; they did not allow for indifference (i.e., each option preferred equally). Grether and Plott thought that the alleged preference reversals could simply have been the result of a “systematic resolution of indifference on the part of subjects forced to record a preference.”

In their study, Grether and Plott allowed participants to respond “I don’t care” when asked to choose which option they preferred on each trial. If one of these trials was selected to be enacted, a coin was tossed to determine the choice.

### *Theory 4: Strategic Responses*

In their day-to-day lives, people often experience trading situations in some form or another. Such experience generally leads people to realize that trading (including buying and selling, which is merely trading that involves money) usually involves some degree of negotiation. In any negotiation, one is more likely to achieve a satisfactory end result if one starts at a point of relative advantage. Thus, a seller will have interest in proposing a high sale price and a buyer will have interest in proposing a low purchase price. Behavior in line with this theoretical observation develops almost automatically with experience. Though not necessarily consciously aware of it, people use this behavior strategically as they seek to improve their position in life. Grether and Plott suggest that this sort of automatic strategic behavior was a driving force in the alleged preference reversal, because the value ratings in the original experiment were solicited using sale price, which would bias the ratings upward and thus mimic a preference reversal. They note that in the next experiment, value ratings were solicited using purchase price, and the preference reversal effect was significantly lower.

To preclude this sort of strategic “market-type” behavior, Grether and Plott included a new control condition. In this condition, participants were not asked to name a buying or selling

price for the options, but rather to give “the exact dollar amount such that [they were] indifferent between the bet and the amount of money.” In this way, participants would have no foreseeable motivation to bias their responses in either direction.

#### *Theory 5: Probabilities*

The concept of probability is not always easy for people to grasp, hence they rely on their own subjective impression of what a specific probability might mean when they take it into consideration. Because subjective impressions are prone to instability, the subjective probabilities that participants used to make their choices or value ratings could change over the course of the experiment. This would clearly lead to inconsistencies in responses. Specifically, inconsistencies that arose in a systematic way could appear as preference reversals.

To eliminate the subjective aspect of probability, Grether and Plott made all probabilities operational by using a bingo cage as the random device used for enacting the gambles selected to be paid out. The gamble options were presented to participants in the form “if the number drawn is less than or equal to  $n$ , you lose  $\$x$ , and if the number drawn is greater than  $n$ , you win  $\$y$ .” Thus, participants had a concrete understanding of the probabilities that were presented in each option and did not need to rely on a subjective understanding.

#### *Theory 6: Elimination by Aspect*

A general theory of choice based on a covert iterative elimination process [Tversky, 1972] was presented by Grether and Plott as a potential contender for explaining preference reversal. This theory states that a decision maker views each option as a set of aspects. At each stage in the decision process, the decision maker randomly selects an aspect to focus on (with probability proportional to its subjective importance), and eliminates from consideration all options that do not include that selected aspect. While it is not clear how this would work in the case of the gambles discussed above, Grether and Plott proceed to dismiss it on theoretical grounds.

### *Theory 7: Lexicographic Semi-order*

This model was proposed by Tversky [1969] to demonstrate how binary choices could cycle in a predictable fashion. The basic idea is that a decision maker compares options based on their attributes, but measures the attributes in a coarse manner such that options with “close enough” attribute measurements are considered as being equal. If true, one might expect violations of transitivity. For example, a decision maker could declare by a series of choices that option  $W > X$ ,  $X > Y$ , and  $Y > Z$ , even if the “true” value of  $X > W$ ,  $Y > X$ , and  $Z > Y$ , provided that the differences in value within each of these pairs was too small to be detected. In this case, the same decision maker might choose option Z over W (a non-transitive choice), because the differential in value between those two options surpassed the threshold of detectability. Grether and Plott noted that this model could not account for the observed preference reversals that they were interested in, because no long chains of binary comparisons were involved, and because no small magnitude differences between options were present.

### *Theory 8: Information Processing—Decision Costs*

The basic idea here is that if a decision maker processed all available information about each option, such as its precise attribute composition and the importance of each attribute to the goals of the decision maker, he would have an accurate estimate of its subjective value. However, because exhaustively processing the available information is somehow disagreeable to the decision maker, he will tend to limit the amount of processing he undertakes before finalizing his choice. A simple choice rule that is often followed is to form a preliminary estimate (or “anchor”) of the value of an option by assessing it on the “most prominent” aspect; assessing the other aspects of the option allow the decision maker to adjust its subjective value upward or downward from the anchor. Empirical psychological studies have shown that adjustments are generally inadequate, such that the anchor has a disproportionately high influence on the

final decision. Grether and Plott claim that during the choice task, participants would focus on the probability as the most prominent aspect, but that during the rating task, they would focus on the amount. If true, this would lead to an apparent preference reversal.

Grether and Plott believed that the order in which participants performed the choice and ratings tasks could alter the anchors used and thus impact the relative preferences. To control for this possibility, they asked participants to perform half of the choices before the value ratings, and half after. Their claim was that “once the subject has ‘invested’ in a rule which yields a precise dollar value, then he/she would tend to use it repeatedly when the opportunity arises.” They thus expected that the apparent preference reversal effect would be obviated for the choices that were made after the ratings.

*Theory 9: Information Processing—Response Mode and Easy Justification*

According to Lichtenstein & Slovic [1971], “variations in response mode cause fundamental changes in the way people process information, and thus alter the resulting decisions.” They believe that a decision maker seeks to implement systematic procedures that allow for choices to be made fairly quickly and satisfactorily, while simultaneously being easy to justify (to one’s self and to others). The anchoring and adjustment process described above is proposed as the specific mechanism that general achieves these goals. Here, the only difference is that “the particular dimension used as an anchor is postulated to be a function of the context in which a decision is being made” [Grether & Plott, 1979]. Another way of explaining this could be that people have “true” preferences that are stable, but that they report distortions of these preferences that differ according to the context in which the report takes place. For example, certain words or situational details can serve as cues that draw the attention of the decision maker to one particular aspect over another, thus establishing an anchor that varies across

context. Grether and Plott note that this theory was in fact consistent with all empirical observations to date at the time of their study.

#### *Theory 10: Confusion and Misunderstanding*

If experimental participants do not clearly and completely understand the tasks they are to perform, the results might not be meaningful. For this reason, all participants in the aforementioned studies were instructed and trained before the experiment, and asked to repeat their choices during the course of the experiment. Confusion should therefore not have played a role in the responses. However, participants in some studies [Lindman, 1971] showed a sort of learning effect in which their preferences seemed to evolve across time. The existence of a learning effect implies that there was not complete certainty in the minds of the participants at the beginning of the experiment, which could have caused apparent preference reversals that disappeared over time. Furthermore, some subjects made some obvious errors, such as valuing the gamble option at a higher amount than the maximum payoff. However, this more likely suggests insufficient attention rather than confusion.

To take maximum precaution against confusion, Grether and Plott used instructions that were rather elaborate, including: “an explanation about why it was in the subjects' best interest to reveal their true reservation prices; a practice gamble; a demonstration of the procedures; and a written test. The correct answers to the test were discussed and subjects' questions were answered.”

#### *Theory 11: Frequency Low*

Here, Grether and Plott point out that the preference reversal phenomenon would be uninteresting if it occurred only infrequently or with a low percentage of the population. However, since previous studies reported consistent preference reversals in up to 73% of the participants, the authors decided that the low frequency claim was not relevant.

### *Theory 12: Unsophisticated Subjects*

Here, Grether and Plott point out that the test population should be representative of the general population in order for the results to be generalizable. The authors decided that the participants they would use in their studies met this criteria.

### *Theory 13: The Experimenters Were Psychologists.*

Test participants are often curious about the purpose of the experiment in which they are participating. It is an unfortunate reality that participants are often suspicious about psychology studies, as they have been known to include elements of deception. Such curiosity and especially suspicion can cause participants to respond differently than they naturally would, which could render the results less useful for the researcher.

Because Grether and Plott are economists, their replication of the experiments automatically addresses their concern about psychology studies.

### *Confirmation that Preference Reversal is Real*

In spite of their extensive theoretical considerations and experimental controls, the results of Grether and Plott's own study clearly demonstrated the preference reversal effect. The authors, in their own words, were perplexed. They did, however, offer one small clue about where future research might probe. The authors were not able to control for Theory 9, and thus not able to eliminate it as a possible cause of preference reversal. Thus, it remained theoretically possible that the apparent reversal of preferences is actually an imprecise calculation of preferences. In this case, it would be the presumably excessive cost of comprehensive information processing that leads to imperfect preference reports. The reports would be imperfect in different ways, depending on the precision to which the information about the different aspects had been

processed, which itself would depend on the context-driven allocation of attention to the different aspects.

The work of Grether and Plott was followed up by researchers who felt that the control conditions should have been even more rigorous than they already were [Pommerehne, Schneider & Zweifel, 1982; Reilly, 1982; reviewed in Slovic & Lichtenstein, 1983]. Pommerehne et al increased the magnitude of the payoffs compared to previous studies, and also added variation to the difference in expected value of the option gambles within choice pairs. In Reilly's study, the experiments were conducted in small groups where participants could interact with the experimenter for clarification purposes; the potential money to be won was physically present on the desks of the participants, for to increase saliency; and the expected value of each option was calculated for the participants, along with a description of the concept of expected value. Nevertheless, both of these studies showed a substantial portion of preference reversals. Various other economic studies in the ensuing years led to similar results [see Tversky, Slovic & Kahneman, 1990, for a review]. At the same time, psychological studies demonstrated the preference reversal effect outside the realm of basic financial economics, including decisions about job applicants, consumer products, and savings plans [Tversky, Sattath & Slovic, 1988].

#### *Preference Reversal as violations of choice axioms*

Once the robustness of the preference reversal phenomena had been demonstrated to the point that it had become universally accepted across multiple disciplines, researchers began to try to explain it rather than disregard it. Three classes of models attempting to explain preference reversal were introduced, each based on the violation of one specific axiom of the theory of preference-based choice (preference theory) [Von Neumann & Morgenstern, 1947; Tversky, Slovic & Kahneman, 1990].

## *Transitivity*

One of the most important foundations of preference theory is the Axiom of Transitivity. This axiom states that preferences of options within a set will form an ordering such that for any options X, Y, Z in a set, the relationships  $X > Y$  and  $Y > Z$  necessarily imply that  $X > Z$ . If it were ever the case that a decision maker preferred X over Y, Y over Z, and Z over X, this would be a violation of transitivity. In a more simple two-alternative choice like those described in the studies above, revealing that Option P is better than Option \$ (by choosing it from the pair), then revealing that Option \$ is better than Option P (by assigning it a higher value rating), could also be evidence of a covert violation of transitivity among the set of options, and many authors have interpreted preference reversal in this way. To address this, nontransitive choice models were developed, notably Regret Theory [Loomes & Sugden, 1982, 1983]. Here, the subjective value of an option is calculated as the expected value of that option across all possible states of the world in which it could be experienced. When part of a choice pair, the value of the alternative option now alters the value of the considered option. Specifically, for each possible state, the difference in the anticipated value of the considered and alternative options (in that state) is scaled by a regret factor and added to the independent value of the option. These modified subjective values are then incorporated into the expectation across states, and the option with the higher expected modified value is chosen. Although the authors were able to experimentally demonstrate the relationship between intransitivity and preference reversal [Loomes, Starmer & Sugden, 1989], further studies demonstrated that the majority of the experimentally-observed preference reversals were not due to violations of transitivity [Tversky, Slovic & Kahneman, 1990].

### *Response Invariance*

The principal of response invariance holds that the preferences that an individual reveals should be identical regardless of the method used to reveal them. Situations that do not respect response invariance are said to include response bias. Theory 9 from Grether and Plott is essentially a response bias theory. It does not claim that preferences are nonexistent or volatile, merely that their externalization via self-report can be biased by response context. To examine the ability of this type of theory to explain preference reversals, response bias models were developed. The *Contingent Tradeoff Models* assert that preference reversal is caused by an over-weighting of the aspect of the options that is “consistent” with the rating scale [Tversky, Sattath & Slovic, 1988]. In experimental task that the authors used, pricing monetary gambles, the consistent dimension was the financial amount. *Expression Theory* asserts that there are three stages in the decision-making process—encoding of option features, evaluation of how option features map to value, and expression of value—and that different preference elicitation methods result in different expression [Goldstein & Einhorn, 1987]. The experimental results that these authors obtained supported their respective models, although they admitted that it was unclear how well they might be able to explain other types of preference reversals (i.e., options other than risky financial offers).

### *Independence*

The Axiom of Independence states that preferences between options must be independent of context. That is, a preference for one specific option over another must not change if additional alternatives are added to the choice comparison set. If this axiom did not hold, it could be reasonable to expect that the subjective value that a decision maker assigned to an option might be different whether it was assessed in isolation (rating task) or in tandem with another option (choice task). Generalized (non-independent) utility theories [Holt, 1986; Karni & Safra, 1987;

Segal, 1988] were introduced to try to explain preference reversals through the framework of independence violations, although no formal model was established to fully explain the phenomenon.

#### *Assessing the alternative explanations*

A study by Tversky, Slovic, and Kahneman [1990] attempted to assess the three alternative interpretations outlined above to determine which, if any, was correct. To distinguish between failures of transitivity and failures of response invariance, the authors extended the original experimental design to include cash amounts to be compared to the gamble options. Thus, the new option triples would include Option P, Option \$, and Option C (cash), and choices between all pairs within each triple would be made. Importantly, the cash amount for each triple was set to be in between the participants' value ratings for Option P and Option \$. The authors focused on trials where Option P had been chosen over Option \$, but the value of Option \$ had been rated higher than that of Option P, using the Option C choices to classify response patterns. Choices of Option \$ over Option C and Option C over Option P indicated intransitivity, since that suggests  $P > \$$  but simultaneously  $\$ > C > P$ . Choices of Option C over Option \$ and Option C over Option P indicated "over-pricing" of Option \$ (response bias), since that suggests  $C > \$$  but simultaneously  $\$ > C$  (from the ratings).

The authors in this study also included a special payoff structure that allowed them to test the axiom of independence, while remaining incentive-compatible. Experimental groups were sometimes not paid at all for their responses, sometimes paid based on the previously established BDM procedure, and sometimes paid based on what the authors refer to as the *ordinal payoff scheme* [Tversky, Slovic & Kahneman, 1990]. Under this scheme, a pair of options is randomly selected at the end of the experiment. The value ratings of the options are compared, and whichever option was rated higher is enacted. This method uses ratings to

establish within-pair orderings, but not necessarily cash-equivalent valuations that would rely on expected utility theory. Within-pair orderings based on choice or on rating should be identical, so any situation where this was not the case would be classified as a preference reversal. A generalized (non-independent) utility model would rationalize these preference reversals by assuming that the decision maker preferred a random choice over a choice of either option. However, this would explain only random preference reversals, not the systematic ones observed empirically [Tversky, Slovic & Kahneman, 1990].

The results provided by Tversky, Slovic, and Kahneman [1990] eliminated the independence and transitivity explanations of preference reversals, while suggesting that failure of response invariance was the root cause of the phenomenon. With respect to independence, the mere observation that preference reversals were equally abundant in the ordinal payoff scheme group (where payoffs were not contingent on participants' responses) as in the BDM payoff group serves to rule it out as a causal factor. Recall that the ordinal payoff scheme was devised to relax the independence axiom, whereas the BDM payoff scheme relies on that axiom. With respect to transitivity, only 10% of observed preference reversals were consistent with the diagnostic pattern indicative of transitivity violation (see above). Therefore, transitivity was excluded as an important causal factor. The majority of preference reversals were consistent with the diagnostic pattern indicative of a failure of procedure invariance, specifically the overpricing of Option \$. Another group of researchers came to the same conclusion based on their similar study [Bostic, Hernstein & Luce, 1990]. Indeed, this had already been demonstrated even before the preference reversal phenomenon had been exposed [Slovic & Lichtenstein, 1968]. In that study, regression analysis showed that participants relied primarily on the probability dimension when making their choices but relied primarily on the amount dimension when making their value ratings, which is a clear violation of procedure invariance.

### *Scale Compatibility*

The authors in the study above interpret the overpricing of Option \$ as an effect of *scale compatibility*. Here, the word *scale* refers to the unit label assigned to a particular measurement. For example, payoff probability is measured on a likelihood scale and payoff amount is measured on a monetary scale. Because the task soliciting value ratings requires responses measured on a monetary scale, payoff amount will become more salient (as it uses the same scale) and thus receive more attention than payoff probability. This distortion in attention allocation is purportedly what causes the overpricing and ensuing preference reversals. This explanation is in line with another study that demonstrated the way in which participants use the anchoring and adjustment strategy when determining value ratings [Schkade & Johnson, 1989]. These authors showed that anchoring was heavily influenced by scale compatibility. They further showed that the frequency of preference reversal was correlated with anchoring and adjusting, as well as with differential attention (looking time) given to the information about payoff amount versus probability. Beyond risky choice, there is also data supporting the scale compatibility hypothesis for options with more or less temporally distant payoffs [Tversky, Slovic & Kahneman, 1990].

Some researchers have proposed that the weight given to an aspect of an option during valuation (or choice, when relevant) is enhanced by its compatibility with the response scale [Slovic, Griffin & Tversky, 1990]. The rationale behind this hypothesis is twofold: 1) additional mental operations are needed to map from one scale into another, increasing effort and potential for error; 2) response mode tends to focus attention on the compatible features of an option [Tversky & Thaler, 1990]. Slovic, Griffin, and Tversky [1990] illustrated this by showing that participants placed greater weight on scale compatible information when predicting either the market values of a set of companies (monetary scale) or the ranking of the companies by market value (rank scale). Further support for the idea that information processing demand for an aspect

biases attention allocation away from that aspect was provided in a separate line of work [Johnson, Payne & Bettman, 1988]. An additional prediction of this hypothesis is that preference reversals will be lower when there is no scale compatibility. Slovic, Griffin, and Tversky [1990] showed a reduction of preference reversals by almost 50% when they used non-monetary (e.g., movie tickets, dinner vouchers) versus monetary payoffs.

As summarized by Tversky and Thaler [1990], the findings described above “are in contrast to the standard economic formulation of choice which assumes that, in the presence of complete information, people act as if they could look up their preferences in a book, and respond to situations accordingly: choose the item most preferred; pay up to the value of an item to obtain it; sell an item if offered more than its value; and so on.” This suggests that there does not exist a set of pre-defined preferences (or values) in the mind of a decision maker, but rather such preferences are constructed during the decision process. It further suggests that the manner in which the preferences are constructed is different in different decision contexts.

### **Constructed Preferences**

Research on the construction of preferences during decision-making or judgment tasks has since developed in its own right. The existing research on preference reversals had provided convincing evidence that the core issue that led to the observed phenomenon was the failure of the principle of invariance (specifically, procedure invariance). Other work not only confirmed that procedure invariance does not hold, but additionally showed that description invariance also does not hold (the so-called *framing effect*) [Slovic, Fischhoff & Lichtenstein, 1980; Kahneman & Tversky, 1979; Tversky & Kahneman, 1981]. Having seemingly located the source of preference reversal, researchers now sought to figure out the cognitive mechanisms behind that source.

The process of preference construction can be thought of as an interaction between the properties of the decision-making system and the properties of the decision task [Payne, Bettman & Johnson, 1992; Slovic, 1995]. Task properties include procedure, description, and context, and have been discussed above. System properties include information stored in memory, decision strategies (heuristics) available, cognitive resource capacity, and personal goals. In particular, a decision maker can simultaneously seek to fulfil multiple goals (e.g., maximize accuracy, minimize effort, maximize emotional state, minimize complexity) [Bettman, Luce & Payne, 1998]. Conflicting goals can be prioritized differently in different contexts, which could understandably cause altered preference construction.

The construction of value can also be broken down into two components: one stable, one unstable [Payne et al, 1999]. The stable component is effectively the measurements of the various aspects (or attributes) of an option. These could be measured more or less precisely or accurately, but they are nevertheless constant across context. The unstable component is effectively the importance (or weights) of the attributes to the decision maker. These weights are highly subjective, difficult to calculate, and clearly context-dependent. Constructed preferences would be based on constructed values, which would amount to the weighted sum of attribute measurements.

An extensive body of research has now shown that preferences are calculated during the decision process, and that they are sensitive to context [Warren, McGraw & Van Boven, 2010]. In other words, preferences are constructed “on demand”. Moreover, research has shown that preferences are calculated to different degrees of precision and accuracy, dependent on the various goals, cognitive constraints, and prior experience of the decision maker [Warren, McGraw & Van Boven, 2010]. Goals affecting the degree of calculation include: need to make accurate and justifiable decisions [Sanbonmatsu & Fazio, 1990]; need for cognition (i.e., a tendency to engage in and enjoy activities that require thinking) [Mantel & Kardes, 1999]; need

to avoid effort [Chaiken, 1987]; and need for closure (i.e., a desire for a firm answer to a question and an aversion toward ambiguity) [Kruglanski & Webster, 1996]. Calculation precision and/or accuracy can be diminished by cognitive constraints such as: time pressure [Dhar, Nowlis & Sherman, 2000; Payne, Bettman, Coupey & Johnson, 1992; Dijksterhuis, 2004]; depleted cognitive control [Pocheptsova et al, 2009]; distraction [Albarracin & Wyer, 2001]. Calculation precision and/or accuracy can also be diminished by experience due to: familiarity [Feldman & Lynch, 1988; Hastie & Park, 1986]; repetition [Schneider & Shiffrin, 1977; Fazio et al, 1986]; and expertise [Shanteau, 1988; Kramer, 2007].

Although one could argue that it is now clear that preferences are constructed at the time of a decision or judgment, it remains unclear why this might be so. More accurately, the part that remains partially unresolved is why the construction process does not always lead to the same expressed preferences. Preference construction is tantamount to information processing for the purpose of advising a decision. When information input into a decision system leads to systematically different outputs (i.e., not merely random variation), it means that the information must have been processed differently. There are two possible explanations for how this could happen: either the system performs different operations on the full set of information, or the system processes different subsets of the information. Because the goal of the preference construction system should be identical regardless of elicitation method, it does not seem likely that the first explanation would be true. Therefore, it is more likely true that the decision system processes less than the entirety of information that is theoretically available, and that the collection of information that is processed is not always the same. Various accounts of this conclusion, along with theories about its implications, have been proposed, with no consensus yet arising about which should be accepted.

*Bounded rationality*

The modern study of decision-making began with economic theory based on fully rational decision making agents, that is, agents who seek to optimize some objective (e.g., maximize value acquisition) in a mathematically optimal manner. However, psychology research repeatedly demonstrated that human decision makers very often do not behave rationally. Herbert Simon implored the economists to incorporate the psychology findings into their models of decision-making, specifically suggesting that the tenet of perfect rationality be replaced with what he called *bounded rationality* [Simon, 1955, 1956, 1957]. This concept preserved the notion that decision makers behave rationally, only it constrained the degree of rationality that a human decision maker could have at its disposal [March, 1978]. The main constraints used to summarize the theory of bounded rationality are limits on: available information; cognitive capacity; time [Simon, 1982]. This summary could be compacted even further, noting that all three constraints are in effect the same—a limit on the ability to process information. Thus, the very definition of bounded rationality implies that the human decision system will process information incompletely. Although it does not directly explain why different information will be processed during separate instances of similar decisions, it requires only a minor assumption that constraints are always fluctuating in order to arrive at a satisfactory explanation.

Convincing as it may be, bounded rationality is not sufficient to explain apparent deficiencies in human decision-making. Numerous studies have shown that people often make choice errors even when there are no obvious constraints on time, cognitive resources, or available information. Daniel Kahneman and Amos Tversky illustrated this with numerous studies under the umbrella of their “heuristics and biases” program [Tversky & Kahneman, 1974; Kahneman, Slovic & Tversky, 1982; Tversky & Kahneman, 1989; Tversky & Kahneman, 1991; Gilovich, Griffin & Kahneman, 2002]. The main message of these researchers was that the human decision-making system is prone to using heuristics to solve the problems it faces, rather than

calculating optimal solutions; these heuristics in turn lead to systematic and predictable biases in behavior that can cause decrements in the well-being of the decision maker. An opposing school of thought, the “fast and frugal” camp led by Gerd Gigerenzer and Peter Todd, argued that the usage of decision heuristics can actually lead to better performance than more sophisticated methods that rely on perfect integration of all relevant information [Gigerenzer & Goldstein, 1996; Gigerenzer & Todd, 1999; Gigerenzer & Selten, 2002]. To this day, there remains some debate as to whether heuristic decision-making is dysfunctional or adaptive, but everyone seems to agree on a common fundamental truth—variations in decision behavior result from variations in the information set under consideration, with or without active constraints on information processing.

#### *Dual-process theory*

One of the most important ideas proliferated by the work of Kahneman and Tversky is that the human decision system is in fact made of two separate subsystems (System 1 and System 2), each operating independently of the other [Kahneman, 2011]. System 2 is responsible for the deliberate style of decision-making that calculates based on all available resources (and is thus subject to constraints and bounded rationality); System 1 handles decisions quickly and often automatically, refraining from heavy information processing. It is therefore System 1 that accounts for those times when a decision maker neglects relevant information, even when he does not need to. Kahneman and Tversky believed that this was the default response mode, and that most decisions were made in this way. However, they were not able to provide an explanation as to why the System 1 versus System 2 conflict exists, nor as to when and how System 2 will become active and override System 1. Despite the abundance of studies that refer to this sort of *dual-process theory* of decision-making, no one has yet proposed a satisfactory answer to these questions.

### *Rational inattention*

Whereas the idea that people sometimes “think” and sometimes do not (in reference to System 2 vs System 1) is seductive, it should not overshadow the possibility that the human decision system might operate in a more continuous fashion. Here, the question would not be whether to think or not, but how much to think about a given decision before making a response. In other words, the contextual circumstances of a particular decision might encourage the system to process more or less information compared to the same situation in a different context or to a different decision. The underlying justification for this idea is that processing information is costly, and that costs are something that people seek to avoid. This idea has been introduced in economics research, under the designation *rational inattention* [Sims, 2003; Caplin & Martin, 2014; Caplin & Dean, 2015]. According to this line of work, decision makers acquire and incorporate additional information until they have reached some satisficing level of utility [Simon, 1955]. In a series of choice experiments in which information about option values was presented in a disintegrated format, participants were asked to provide their momentary preferences at random times throughout the trials [Caplin & Dean, 2011; Caplin, Dean & Martin, 2011]. The results show that, as predicted, the participants incorporated information incrementally across time and used it to construct their preferences. Furthermore, the final responses that the participants gave at the end of each trial suggest that they used a strategy to stop incorporating additional information after a certain point. This subjective information accumulation threshold can be viewed as the point at which processing additional information was thought to be more costly than beneficial [Caplin, Dean & Martin, 2011; Gabaix et al, 2006]. The data also show that changes in decision context (e.g., increased complexity, additional options) caused participants to adjust their accumulation thresholds in predictable ways. In effect, rational inattention describes self-imposed bounded rationality. That is, even in situations where there are no inherently limiting constraints, the decision system itself creates

ad hoc virtual constraints to account for the costly nature of information processing. The end result is identical to that caused by “naturally” bounded rationality.

The picture that emerges from the above line of work can be summarized as follows. People make decisions (e.g., choices or judgments) based on subjective valuations and preferences; the preferences are constructed at the moment of the decision; constraints on information processing (natural or virtual) affect preference construction differently for different types of decisions; this induces different mechanisms of preference construction that do not always yield identical results. In other words, value ratings and choices might be inconsistent even if the decision system behaves rationally, due to information being processed differently by divergent task algorithms. However, this picture neglects other important known facts. So far, we have focused on the question of how preferences influence decisions. An entirely separate line of research exists that claims to show exactly the opposite—decisions influence preferences. Many researchers on decision-making ignore this work, but the evidence is substantial and the implications important. If preferences for an option can be altered by the mere act of choosing it, decision makers could become biased away from (subjectively) optimal behavior. The adaptive fitness of such a bias is unclear. Thus, in order to more completely understand the topic of decision-making, one must include the study of choice-induced preference change.

### **Choice-Induced Preference Change**

In 1956, Jack Brehm introduced the Free-Choice Paradigm to the world of experimental psychology. In his study, Brehm presented participants with a collection of consumer goods and asked them to rate each item qualitatively (on a scale with eight intervals labeled “extremely desirable”, ..., “definitely not at all desirable”). The participants were allowed to take as long as they wanted to inspect the items (average of 15 minutes), which were all located together on

the same table. The participants were then presented with a pair of items from the set, which they were told had been selected randomly but were in fact chosen to control for decision difficulty (e.g., very similarly-rated ratings for the high-difficulty condition). Whichever item the participants chose from their pair was given to them as compensation for participating in the research study. The participants were then asked to re-rate each of the items, one at a time and with the actual items out of sight. The results of this study revealed that chosen items were rated as more desirable (than initially) and rejected items were rated as less desirable, and also that the closer the initial ratings of the choice pair items (higher dissonance / choice difficulty), the greater the magnitude of rating change. These results were interpreted as empirical support for the theory of cognitive dissonance [Festinger, 1962]. According to this theory, a decision maker required to choose one option from a set of equally-valued options will experience cognitive dissonance, or psychological discomfort, regardless of which option he chooses. For example, the decision maker might regret having accepted the less attractive features of the chosen option or having foregone the more attractive features of the rejected option. Or, he might simply feel unconfident about having selected the best option. According to Festinger's theory, the decision maker makes a post-choice adjustment to his preferences such that the chosen option is now valued higher than it was before the choice, and the rejected option is now valued lower. Through such adjustments, the decision maker resolves the dissonance that had been created by the choice. This theory thus provided a descriptive social psychological account of the choice-induced preference change (CIPC) phenomenon, but provided no insight to its purported cognitive mechanisms or what possible role it might play in adaptive fitness.

For more than 50 years, researchers replicated the FCP study and philosophically debated the putative role of cognitive dissonance and the associated phenomenon of spreading of alternatives. The effect was even demonstrated in small children and capuchin monkeys [Egan, Santos & Bloom, 2007; Egan, Bloom & Santos, 2010]. One key difference in the children and

monkey studies, however, was that agents would select an alternative without knowing precisely what item they would eventually receive. For example, children would pick a container that held multiple items, and were then presented with whichever non-visible item happened to be at the top of the container. Monkeys on the other hand, would search through a box to find a non-visible item, while being tricked into thinking that there were multiple items inside. This line of research suggested that “blind choices” (i.e., choices made without clear prior knowledge about which option was being selected) could also cause CIPC. However, a blind choice could not possibly create cognitive dissonance in the traditional sense, because a necessary condition for such dissonance to occur is that the decision-maker feels personally responsible for the selection that he made. Evidence from later studies [Sharot, Velasquez & Dolan, 2010; Nakamura & Kawabata, 2013] more directly demonstrated CIPC in blind choice situations. In these studies, participants performed the FCP tasks, but during the choice section, the options were not actually displayed (nonsense strings of symbols were displayed instead of vacation destination names). Participants were told that the option identities were masked, and that their choices would be driven by unconscious perception. Upon choosing, the option that they had selected (unwittingly) was revealed to them. Critically here, subjects were led to believe in the power of subliminal decision-making (i.e., deception was used to induce a sense of responsibility for their choices). The results showed significant spreading of alternatives. Sharot et al contrasted this with a control condition in which the computer made the choice, which did not show any significant spreading. In each of these studies, the decision makers could not possibly contemplate the options before selecting one (because they did not know what they were), and thus could not regret their choices in the way suggested above. In spite of this, CIPC still occurred. These results demonstrated that choice-induced preference change might occur automatically, without the need for value-based calculations. However, the results

do not speak to the possibility of additional CIPC that might be more directly related to the subjective valuation process.

*Is choice-induced preference change a reliable empirical construct?*

In 2010, Chen and Risen highlighted a methodological flaw in traditional FCP studies, which had a major impact on the field. In their paper, the authors provided a detailed mathematical explanation for how apparent spreading of alternatives could occur without being caused by choice. The idea is that ratings provide imperfect (i.e., noisy) information about participants' true preferences. Choices, too, provide imperfect information about preferences—a chosen item is more likely to be truly preferred than a rejected item. So, in a scenario where a participant is faced with a choice between two items that were rated identically (or similarly), the choice provides additional information regarding which of the items was truly valued higher (which would be revealed if all rating noise was somehow removed). Here, the first ratings are considered as random draws from probability distributions (which overlap in the case of similar ratings). Sometimes these random draws will be close to the means of their respective distributions, but sometimes they will be closer to the tails. In these latter situations, “regression to the mean” will cause subsequent random draws to be (on average) closer to the true distribution means. Chen and Risen argued that this regression to the mean effect could explain the spreading of alternatives, which would then reduce to a mere statistical artifact. To test their claim, the authors first implemented a novel experimental variant of the FCP, in which items were rated, then rated again, then chosen between. In this condition, choice cannot possibly cause preference change, because it occurs after all re-ratings have been completed. However, post-rating choices would still hold information about preferences, and could thus still be used to predict spread. The authors demonstrated this experimentally (i.e., spreading was observed in the Rating-Rating-Choice condition). However, in Chen and Risen's second experiment (same study), the degree of spreading was more than double in the standard Rating-Choice-

Rating (RCR) condition compared to the control Rating-Rating-Choice (RRC) condition. This suggests that while the statistical artifact pointed out by the authors may be valid, it is not enough to void the long-lasting assertions about cognitive dissonance reduction. Moreover, numerous studies have since reported higher spreading in the RCR condition versus the RRC [Voigt et al, 2018; Salti et al, 2014; Coppin et al, 2012, 2014; Sharot et al, 2012]. A previous study showed spreading with amnesic decision-makers (which Chen and Risen mentioned might mimic the RRC condition, since choices that were not remembered might not be able to cause preference change) [Lieberman et al, 2001]. Another study showed that spreading was indeed linked to episodic memory, where spreading only occurred for choice pairs for which the choice outcome had been remembered [Chammat et al, 2017]. This series of studies strengthened the empirical bases of choice-induced preference change, and revived the interest for identifying its underlying cognitive mechanism.

#### *Is cognitive dissonance reduction adaptive?*

The idea that a decision maker should alter his preferences after a choice as a means of reducing cognitive dissonance is an idea that, *a priori*, would seem to have low adaptive fitness. Logically, choices should be made based on value, and value should be determined based on meaningful information. Distortions to value such as those caused by the purported cognitive dissonance reduction would only decrease the optimality of choices, one might be tempted to claim. The Action-Based Model of Dissonance [Harmon-Jones, Harmon-Jones & Levy, 2015; Harmon-Jones, Amodio & Harmon-Jones, 2009] provides an interesting account of a possible adaptive value of cognitive dissonance reduction. This model is founded on the idea that decision outcomes are closely linked with action tendencies. That is to say, the main purpose of making a decision is to select which action to perform. Thus, a choice should naturally be associated with a preparation to perform the selected action. The action-based model proposes that discrepant cognitions (e.g., having equivalent cravings for two snack options but explicitly

choosing to eat one of them) have the “potential to interfere with effective and unconflicted action” [Harmon-Jones, Amodio & Harmon-Jones, 2009]. Reducing dissonance (e.g., by adjusting option valuation) thus facilitates efficient performance (i.e., following through with the chosen option). In line with the original theory of cognitive dissonance, the negative feeling associated with dissonance is essentially an emotional cue for the decision maker to adjust his beliefs (or actions).

Another possibility is that the cognitive dissonance reduction actually takes place *before* the choice is made, rather than after. The basic idea presented by the Action-Based Model of Dissonance would be preserved, only the timing of the alleged preference revision would be different. In this case, the preference change might actually be *instrumental* to the decision process, rather than reactive. The logic is that when a decision maker is faced with two options that he cannot distinguish in terms of value, he feels cognitive dissonance. However, rather than choosing randomly and later adjusting his preferences, he might instead adjust the preferences for the very purpose of enabling him to choose one option over the other. This sort of preference adjustment would be tantamount to preference construction at the time of decision. So, even though it might at first seem to be a minor point, the timing of the preference change is actual of critical importance for understanding whether it is a cognitive bias or an adaptive feature that is central to rational decision-making. This is something that has only recently begun to be explored, and will be discussed further below.

#### *Neural Correlates of Choice-Induced Preference Change*

Recent neuroimaging findings lend support to the CIPC phenomenon. Izuma et al [2010] first showed that the striatum encoded subjective value of options in a free-choice paradigm task with snack foods. The authors then showed that the change in value ratings from pre- to post-choice correlated with the change in striatal activity. Beyond subjective value and choice-

induced changes therein, this study also investigated the neural correlates of the purported cognitive dissonance itself. Using their own measure of dissonance, the cognitive dissonance index (CDI), the authors showed that activity in the dACC correlated with CDI during the post-choice rating task. The CDI for each option is defined (post-choice) as a function of its choice (chosen or rejected) and pre-choice preference rating:

$$CDI = Ch * (9 - rating) + (1 - Ch) * rating$$

$$Ch \in \{0,1\}$$

Because CDI is simply a contrast between a binary choice signal (i.e., chosen versus rejected) and a continuous value signal, the observed dACC activity could perhaps alternatively be interpreted as detection of conflicting information, which the dACC is already known to perform [e.g., Weissman et al, 2003, 2004; Egner et al, 2007]. There was also a correlation between CDI and dlPFC activity, which was interpreted as conflict resolution, in line with the literature [e.g., Mansouri, Buckley & Tanaka, 2007; Oehrn, 2014]. This conflict detection and resolution was intended characterize the CIPC—detection when realizing that the option had been rejected while simultaneously rated highly, resolution by adjusting the value rating downward to make it seem more reasonable that it would have been rejected. However, the authors did not show a link between CDI and CIPC, nor between neural activity and CIPC, so their interpretation still remains partially unsupported by the data. In addition, this study does not elucidate when the covert value reassessments occurred (e.g., during or after the choice).

A later study attempted to show that the CIPC occurred after the choice by using electroencephalography (EEG) to record event-related potentials (ERPs) in the fronto-central region [Colosio et al, 2017]. In this study, the authors used the difference in ERP magnitude immediately following response for difficult versus easy choices as a proxy for a cognitive dissonance signal, and showed that this signal correlated with the eventual CIPC. The authors

interpreted the ERP as error-related negativity (ERN), akin to an error detection signal [e.g., Holroyd & Coles, 2002; Yeung, Botvinik & Cohen, 2004; Gehring et al, 2018]. Their idea was that the error-like signal informed the decision maker that he had made a choice that was inconsistent with his previous value ratings, which would encourage him to later revise those ratings. But the ERN is not always interpreted as an error detection signal. It has also been shown to represent choice confidence [Boldt & Yeung, 2015]. Here, the difference between the ERN for difficult and easy choices would demonstrate low versus high confidence choices, which would offer no information about the timing of the CIPC. Unfortunately, Colosio et al only reported average measures and showed correlations across subjects. Without trial-by-trial contrasts within subject, there is no clear evidence regarding the timing of the CIPC (i.e., during or after the choice).

Although there is not yet a consensus as to whether CIPC occurs during or after a decision, several studies have shown that neural activity *during* choice deliberation correlates with and predicts CIPC. One study showed that increased activity in the right inferior frontal gyrus, the medial fronto-parietal regions, and the ventral striatum, as well as decreased activity in the anterior insula, during the decision phase of the experiment positively correlated with eventual observed CIPC [Jarcho, Berkman, & Lieberman, 2010]. This suggests that option re-evaluation might occur at the time of the decision. Another study showed that activity in the dorsal anterior cingulate cortex (dACC) and the left anterior insula (left aINS) was distinguishable for difficult versus easy choices, and that the activity of these regions correlated with activity in the posterior cingulate cortex (PCC) during choice [Kitayama et al, 2013]. The intensity of PCC activity, in turn, predicted the magnitude of the CIPC on a trial-by-trial basis. This suggests that choice difficulty caused cognitive dissonance, which was resolved by preference change (CIPC). The involvement of the dACC and the aINS at the time of deliberation in the CIPC phenomenon had previously been demonstrated in a cognitive dissonance study using an induced-compliance

paradigm [Van Veen et al, 2009]. Unsurprisingly, these studies have demonstrated that the activity, during the decision, of key regions of the brain's valuation and cognitive control systems are involved in CIPC.

However, other studies have provided contradictory evidence. Harmon-Jones et al [2008] manipulated left frontal cortical activity in participants *after* they had made difficult choices. Results showed that decreased activity (via neurofeedback) caused lower CIPC, and increased activity (via action-oriented mindset induction) caused higher CIPC. Though the spatial specificity of the regions involved was low, the timing of the manipulations was such that the results seem to support post-choice dissonance reduction as the cause of the CIPC. Another study [Mengarelli et al, 2013] used transcranial direct current stimulation (tDCS) to downregulate the activity of the left dorsolateral prefrontal cortex (dlPFC) *after* choice and observed decreased CIPC. It was later shown that using transcranial magnetic stimulation (TMS) to downregulate activity in the posterior medial frontal cortex (pmMFC) after choice also led to decreased CIPC [Izuma et al, 2015]. Although compatible with post-choice CIPC, such causal effects could instead be due to the post-choice disturbance of value representations that resulted from intra-choice CIPC.

### **Multiple Option Decisions**

Everything described thus far has been in reference to decisions between two, and only two, options. But in real life, we are often faced with decisions where the number of possible options is more than just two. These multi-option decisions are the focus of a growing body of work, whose early successes date back to the empirical demonstration of specific forms of preference reversals, in particular: the so-called *decoy effects*. Here, decision makers often reverse their preferences between two items when a third option is offered as an alternative. This behavior

violates the norm of *independence of irrelevant alternatives*, also known as the *independence axiom* [Arrow, 1950]. This norm states that an individual's preference between two options must remain unaltered regardless of what other options might be included in the set of alternatives. Three principal decoy effects have been exposed in the literature: the "attraction", "compromise", and "similarity" effects. The "attraction effect" induces a decision maker to increase his relative preference for option A over option B when option C is made available, where C is similar but clearly inferior to A [Huber, Payne & Puto, 1982; Ratneshwar, Shocker & Stewart, 1987]. The "compromise effect" increases the preference for A over B when C is such that  $C > A > B$  in some ways, but  $C < A < B$  in other ways [Simonson, 1989]. The idea is that A can now be thought of as a compromise between C and B. The "similarity effect" increases the preference for A over B when C is similar to B in terms of both overall value and qualitative features [Tversky, 1972; Huber, Payne & Puto, 1982]. Fascinating as these decoy effects may be, they cannot (to our current knowledge) be explained by theories of simple value-based decision making. Their explanation relies on models of attribute-based decision making, which is beyond the scope of this current work.

But multi-alternative decisions have other properties that disclose more general aspects of the covert mechanisms underlying decision making. First among these is Hick's law, which states that response time is an increasing function of number of decision options [Hick, 1952]. It has also been shown that choice stochasticity is greater for choice sets that contain three options rather than two. This suggests that the probability that the decision maker will choose the option that he had previously rated highest will continue to decrease as more alternative options are added. One prominent explanation for this phenomenon is the theory of divisive normalization [Louie, Khaw & Glimcher, 2013]. This theory, inspired from basic research on the physiology of neurons, states that the firing rate of a neuron representing the value of a particular decision option will be normalized by the sum of the firing rates of the pools of neurons representing the

value of the other alternatives in the option set. This satisfies the constraint that the level of total neural activity remains stable across decisions. This theory explains the observation that as the value of the “distractor” option (i.e., the one with the lowest subjective value rating) increases, the choice stochasticity (between the other options) also increases. This is because an increasing distractor value will require increasing neural activity for its representation, which in turn will decrease the neural activity for the other options. This directly results in a decrease in the difference of the neural activity of the best and second best options, which causes an increase in choice stochasticity as the option values become less distinguishable.

An alternative theory, which does not rely on neural coding dynamics, is that distractor decision options capture attention from the decision maker and thus increase stochasticity [Gluth, Spektor & Rieskamp, 2018]. The basic idea is that when the values of the first and second options are similar, the decision is difficult and thus requires attention to choose accurately between them. However, when a third option steals some attention from the other options, the reduction in attention for the “relevant” options necessarily decreases the probability that the best option will be chosen. Furthermore, the higher the value of the distractor, the more attention it will capture, the more interference it will cause in the choice between the relevant options.

We note that, to the best of our knowledge, no studies exist to date that examine choice-induced preference change with option sets greater than two.

## **COMPUTATIONAL MECHANISMS**

Until now, we have seen that people make decisions based on subjective preferences that are imperfectly and differentially constructed during the decision process. We have also seen that the preference construction process seems to incorporate information related to previous decisions involving the same option. However, it is not yet clear through what cognitive

mechanisms either of these processes takes place. Many models have been proposed to account for preference construction during a choice task, which could perhaps be generalized to include any decision task. To our knowledge, no such quantitative model exists to account for the choice-induced preference change phenomenon.

### *Signal detection theory*

The earliest and most simple computational model of decision making that propagated throughout the field is based on *signal detection theory*. This theory was initially developed to provide a means to measure the differentiability of information-bearing patterns (signal or stimulus) from random patterns that distract from the information (noise) [Marcum, 1960]. Psychologists later adapted the theory to examine behavior in the two-alternative forced choice task, in which case the option with a lower value rating plays the role of the distractor. The basic set-up of this type of model is that each option is represented by a Gaussian distribution, in order to capture both the estimates of subjective value and the certainty of the estimates. One of the key variables of this model is the sensitivity index, or  $d'$  (“d prime”), which measures how different the two distributions are (inversely proportional to their region of overlap). The other key variable is the criterion, measures the pre-choice bias of the decision maker to respond to one option or the other. During a choice, the decision maker is assumed to take a random sample from each distribution and use that to determine his preference (and thus, his choice). The larger the value of  $d'$ , the more likely the decision maker is to choose the option with the “true” higher value (i.e, the higher mean). The smaller the  $d'$ , the more likely the decision maker is to make an “error” (or preference reversal). In an unbiased situation, the criterion would be set at the intersection of the probability distributions. If, for example, the decision maker believed that one type of response (the “default”) was more often the best one in his current environment, he would shift his criterion away from the mean of that option’s distribution. In this case, the balance of evidence in favor of the non-default option over the default would need

to be much higher in order for the decision maker to select that option. The signal detection theoretic model of decision making did a good job of accounting for choice accuracy, also known as the probability of choosing correctly (or consistently with pre-choice ratings). Although providing a strong foundation for computational models of decision making, signal detection theory was not able to account for anything other than accuracy.

### *Drift-diffusion models*

Models based upon signal detection theory were later extended to account for another quantitative feature of decision processes, namely: decision time [Pike, 1973]. Most of these models share a common underlying principle, which is referred to as *evidence accumulation to bound*. As the term suggests, these models are built on the idea that a decision system “accumulates evidence” in favor of each of the options until the cumulative evidence in favor of one option passes a bound. Here, evidence is basically information, accumulation is basically cognitive processing, and bound is basically a lower limit for the amount of subjective proof that an option must offer before the decision maker will voluntarily declare it as preferred. The two main classes of model that have proven popular among researchers are so-called *drift-diffusion* and *race* models, in their standard formats as well as numerous variants.

The drift-diffusion model is built around a single accumulator variable, which represents the balance of evidence in favor of one option versus the other. The variable is unidimensional, with an initial value of zero. Thus, as evidence accumulates over time, an instantaneous value greater than zero represents option 1 being favored, while an instantaneous value less than zero represents option 2 being favored. At each time step ( $t$ ), a new sample of evidence ( $e$ ) is drawn from a probability distribution for each option, and the difference in the values of the new evidence for the options ( $i=1,2$ ) is added to the accumulator ( $x$ ).

$$e_{it} \sim N(\mu_i, \sigma_i)$$

$$x_t = x_{t-1} + e_{1t} - e_{2t}$$

The evidence is modeled as a random variable to account for noise that interferes with the decision process. This can come from sources external to the system (e.g., perceptual impediments, ambiguous context) or internal (e.g., incongruent memories, inconsistent neural activity). The process unfolds in time until the accumulated evidence crosses a threshold set at a pre-determined distance from the starting point (in both the positive and negative directions), and the favored option at that moment is selected as the chosen option. The key components of this model are thus: drift rate, which is the mean of the difference in evidence for option 1 versus option 2; diffusion rate, which is the variance of the difference in evidence (though it is sometimes modeled as intrinsic noise in the decision system); threshold, which sets the cumulative target for required evidence. Note that one of these parameters must be fixed in order for the model to be identifiable. The choice of which parameter to fix is arbitrary, and the chosen parameter is often set to one for simplicity.

The drift-diffusion model (DDM) was first introduced as a tool for representing cognitive processes in a study of memory retrieval [Ratcliff, 1978]. Not long after, it was discovered that the DDM could be used to describe two-alternative forced-choice decisions, which are the standard for experimental work on decision-making [Ratcliff & Rouder, 1998]. Usage of the DDM has since become widespread by researchers in psychology, neuroscience, and even economics [Ratcliff et al, 2016]. However, not all researchers prefer to use the DDM in its original form. Below is a sample list of the most common adjustments to the basic model, which allow for across trial variability. Starting points for the accumulator that differ from zero have been used to represent default expectations, where the decision maker believes that one option is more likely to be preferred even before decision deliberation begins [Ratcliff & Rouder, 1998; Ratcliff, Van Zandt & McKoon, 1999]. Some researchers like to include so-called *non-*

*decision time*, which represents a delay between option presentation and start of deliberation, thought to capture the time it takes for the options to be perceived and recognized by the decision system [Ratcliff & Tuerlinckx, 2002]. *Collapsing bounds*, which simply refers to thresholds that decrease over time, are often added so as to include the idea that the decision maker will become ever more eager to respond as time passes, due to inherent costs of information processing [Ratcliff, Thapar & McKoon, 2001, 2003, 2004; Thapar, Ratcliff & McKoon, 2003]. Recently, a new branch of DDM, the attentional DDM (aDDM) has started to be used. This version of the model includes an extra parameter that increases the weight of the evidence for the option currently being gazed upon at each point in time within a decision trial [Krajbich, Armel & Rangel, 2010]; it has been extended to multi-option decisions [Krajbich & Rangel, 2011]. This model allows attentional shifts and durations to be purely random, and hence requires eye gaze measurements to make (post-hoc) quantitative predictions. At least one recent attempt has been made to enhance the aDDM by assigning attention to each option based on its relative value [Gluth, Spektor & Rieskamp, 2018]. This so-called Mutual Inhibition with Value-based Attentional Capture (MIVAC) model is basically the same as the original aDDM, except that the probability of gazing at a particular option at any time point during deliberation increases with that option's estimated value (at that time point).

The popularity of the DDM throughout modern decision-making research is understandable. This simple model does a good job accounting for choice data from standard two-alternative forced-choice decision tasks in a wide variety of research paradigms [Ratcliff & McKoon, 2008]. In particular, this model accounts well for choice accuracy, response time averages for correct and incorrect responses, and response time distributions for correct and incorrect responses. Furthermore, it is nicely suited to demonstrate the so-called *speed-accuracy tradeoff* [Bogacz et al, 2006]. The speed-accuracy tradeoff is a common phenomenon wherein a decision maker can either invest more time in the deliberation process and expect to achieve more

accurate results, or forego higher accuracy in order to respond more quickly. The threshold in the DDM effectively establishes this tradeoff. Due to noise in the accumulation process, slower responses will be more accurate as the uncorrelated noise averages out. Higher thresholds yield longer response times, and therefore greater accuracy. Bogacz et al [2006] show that among a variety of more complex models, the DDM is the one that implements the statistically optimal algorithm (highest accuracy given speed, fastest speed given accuracy).

The DDM is sometimes criticized as being no more than a tool to fit experimental data, with no generative capacity for predictive modeling. However, it has been shown that the DDM can be exactly equivalent to a generative model based on Bayesian inference equations [Bitzer et al, 2014]. Using a geometric framework, the DDM has been shown to guarantee optimality (minimal response time for a desired level of accuracy) even with an arbitrarily large option set [Kvam, 2019]. It has also been shown that the DDM with Bayesian updating of accumulated evidence and collapsing bounds represents the optimal decision policy for a series of sequential decisions within a fixed time period [Tajima, Drugowitsch & Pouget, 2016]; this holds for multi-option decisions as well [Tajima et al, 2019].

### *Race models*

The race model is similar in design to the DDM, except for one key distinction—this model is based on the dynamics of two separate accumulator variables, each representing the cumulative sum of evidence for one of the options. The sequential sampling aspect of the race model is identical to that of the DDM. However, the race accumulators follow a simpler rule—each accumulator moves in direct accordance with its own information samples. Here, the decision process terminates once either of the accumulators surpasses the pre-determined threshold, and that accumulator which crosses first determines which option is chosen. The race model was first introduced as a tool to represent cognitive processing during perceptual decision tasks

[Vickers, 1979]. This seminal variant of the race model did not, however, accurately match experimental data. Fundamentally, for difficult decisions where both options have very high and similar values, the race model predicts very fast responses, which is at odds with empirical data. For this reason, later variants of the race model relax the assumption of independence between the accumulator variables, and consider situations of mutual inhibition instead.

The *Leaky, Competing Accumulator (LCA) Model* includes a separate evidence accumulator for each decision option, but augments the basic race model with “leakage” and “mutual inhibition” [Usher & McClelland, 2001]. The leakage term degrades the strength of accumulated evidence as it ages. With such a mechanism, the accumulators (if left unbounded) will eventually reach an equilibrium point where the inflow of new evidence is balanced by the decay of old evidence. The mutual inhibition between the accumulators causes a gain in evidence for one option to simultaneously decrement the evidence for the other option (note that this model is valid for any number of options). As more evidence samples are accumulated, the best option tends to dominate, and its cumulative activity represents its value relative to the other option. In this way, the LCA model can be thought of as a hybrid between a race model and a DDM—it offers independent accumulators (as in a race model) and a threshold based on relative evidence (as in a DDM). This type of model has been shown to be as successful in describing empirical data as the DDM, yet it has not been adopted by nearly as many researchers in the field, presumably because it is computationally intensive to implement [Ratcliff & Smith, 2004; Ratcliff et al, 2005; Ratcliff & McKoon, 2008].

More recently, a group of researchers has introduced their own version of the race model, with the aim of predicting *choice confidence* reports in addition to choices and response times [DeMartino et al, 2013]. The term choice confidence refers to the decision maker’s subjective belief that the option he chose (or is preparing to choose) is indeed the one that he prefers, an important psychological variable that had not been considered by previous models. Like the

LCA model, this model also relies mutual inhibition between accumulators, although here the degree of inhibition is asymmetric and time-variant. In this model, as in the standard race model, there are separate (but equivalent) bounds for each accumulator. The main contribution of this model is its ability to measure choice confidence. Here, confidence is represented by the distance between accumulators at the end of a given trial (i.e., when one stream of evidence reaches the threshold). The DeMartino et al race model provides a novel computational perspective on confidence, yet it stands to be improved upon. This is because the model predicts that choices that are made quickly due to substantial noise in the decision system will achieve higher levels of confidence, which does not seem reasonable.

#### *Limitations of evidence accumulation models*

In spite of its merits, we believe that these kinds of models are of only limited use for the ever-advancing field of decision-making research. While they do perform well with respect to choice accuracy and response time, they have little to say about other features of decision processes, such as choice confidence, subjective effort, choice-induced preference change. In particular, they do not account for systematic decision-by-decision variations in decision features. This is because they somehow assert that such variations are driven by the stochastic component of the model, which remains, by construction, unpredictable.

In addition, some of the core components of these models can be called into question. One such component is the threshold. To our knowledge, no one has yet offered an explanation as to how the thresholds could be established on a decision-by-decision basis based on cognitive/computational principles. The magnitude of the threshold is generally determined by fitting data, and stable across trials. One notable exception is the model of Tajima et al, under the additional assumption that a decision maker will reassess the reward rate of the environment before beginning a new decision, which would result in a trial-by-trial threshold adjustment.

However, this would force confidence levels to correlate with previous decisions, as opposed to the properties of the current decision to be made.

Another potentially problematic component of these models is that they do not account for the fact that preferences are known to be constructed (and/or modified) dynamically during the decision process itself. As a reminder, evidence accumulation models take the value of each option as direct input to the decision process. Although overlooked by proponents of evidence accumulation models, this implies that options' values were already calculated before the decision, and stay stable throughout the decision process. Yet the entire point of the original work on constructed preference was to explain the widely-recognized fact that preferences established by rating were not the same as those established by choice. With evidence accumulation models, rating-based and choice-based preferences are clearly not independent, and we believe that this should be cited as a theoretical flaw in their design. This important oversight was pointed out in a recent study, and the authors provided a solution based on efficient coding (see below) and Bayesian decoding principles [Polania, Woodford & Ruff, 2019]. It remains to be seen if the recommendations of these authors will become common practice in the field.

Note that the notion of choice-induced preference change *does* allow that option values can be known before a decision, but it too does not allow that they be stationary. Specifically, values will increase for items that are chosen and decrease for items that are rejected. This possibility is something that is excluded from every known variant of evidence accumulation models.

### *Alternative models*

In addition to the popular families of model described above, there have also been several alternative models of decision-making proposed in recent years. Some of the more interesting

models include those related to an alleged urgency signal to respond, neural network variants of the DDM, and “efficient coding” models.

The “urgency-gating” class of model was proposed as an alternative to the more common sequential sampling models (i.e., DDM and race models) [Ditterich, 2006; Churchland, Kiani & Shadlen, 2008]. In particular, early proponents of this type of model were not satisfied with the fact that in most studies that tested evidence accumulation to bound models, within-trial information was almost always constant [Cisek, Puskas & El-Murr, 2009]. Because of this, the evolving dynamics described by such models (i.e., evidence accumulation) could instead be evidence of a growing signal related to the urgency of making a choice. Indeed, in the case of constant evidence, the two models are mathematically equivalent and hence impossible to compare [Cisek, Puskas & El-Murr, 2009]. When evidence varies across time (within a trial), however, the models make different predictions. Effectively, the urgency-gating model predicts that a decision will be more heavily influenced by late evidence relative to early, due to the model’s increasing gain function that scales the incoming evidence samples. Although some authors have shown results in favor of the urgency-gating model [Thura et al, 2012; Carland et al, 2015], others have invalidated it [Evans et al, 2017]. Other authors have discovered situations in which the evidence accumulation and urgency-gating models can be more easily contrasted, and found mixed results [Hawkins et al, 2015]. More empirical data is needed before a consensus can be made about the validity of the urgency-gating class of model.

Other prominent models of decision-making follow a similar logic as evidence accumulation models, but are framed from the perspective of neural activity. These “neural network” variants of sequential sampling models have been developed as researchers have started to look for neural evidence that corresponds to the various components of the different sequential sampling models, as behavioral data alone is often insufficient to distinguish models [see O’Connell et al, 2018, for a review]. Neural network models start from the premise that the decision apparatus

in the brain is composed of distinct pools of neurons, each of which serves a different purpose during a decision. The first layer of the neural network contains a separate pool of neurons for each decision option. Input to this layer comes directly from the stimuli (or more accurately, from other brain regions responsible for encoding the perception of the stimuli). The input (i.e., evidence) for each option is represented by a Gaussian random variable whose mean is equal to the value of that option and whose variance characterizes neural encoding noise. The inflow of the evidence stream follows a Poisson distribution. This input excites the activity of the respective pools of neurons, causing the firing rates to increase in proportion to the value of the evidence. From this point, the neural network models diverge into two primary classes. In the first class of neural network model, each pool of neurons in the first layer has two output connections [Wang, 2002; Wong & Wang, 2006]. One of these is a recurrent excitatory connection (i.e., to and from the same pool), which sustains persistent activity and thus allows for working memory and evidence accumulation over time. The other is an inhibitory connection to the pool representing the value of the other option(s). This creates a sort of competition between the neural pools. Eventually, the circuit dynamics will arrive at an attractor state, in which the two pool maintains persistent activity—one elevated, one suppressed. The option whose pool maintains elevated activity will be the option that is chosen as preferred. This model has successfully accounted for empirical findings in a variety of perceptual and value-based decision tasks [see Wang, 2008, for a review].

In the second class of neural network model, the first layer of neuron pools (i.e., those representing the values of each decision option) feeds into a second layer, where a pool of inhibitory neurons resides [Louie, Grattan & Glimcher, 2011]. Each pool in the first layer has an excitatory connection to the pool in the second layer. This causes activity in the second layer to increase when activity in *any* of the first layer pools increases. The second layer of neurons also feeds back into the first layer, with an inhibitory connection to each of the pools. This

circuit causes the activity in each of the first layer pools to be dampened in proportion to the cumulative activity across the first layer. According to this model, the value signal for an option will depend on the value signal for the alternative option(s), supporting the theory that value-based decision-making is context dependent. The authors demonstrated that both behavioral and physiological experimental data could be well accounted for by their model. Nevertheless, further exploration of this type of model is needed before

The Wang model can be reduced so that it is mathematically identical to the DDM, but otherwise, its dynamics enable it to account for the empirical observation that response times for error choices (i.e., those where the higher valued option was not chosen) are typically longer than for correct choices [Wong & Wang, 2006]. This is something that the basic DDM cannot predict, because the response time distributions for correct and error trials are always identical [Luce, 1986]. Allowing for trial-by-trial variability in the drift rate and/or starting point resolves this discrepancy [Ratcliff & Rouder, 1998]. Although not necessarily intuitive why these parameters would vary across a set of identical decisions, it would be reasonable to assume that it might be something that comes into play during a series of sequential decisions. This is something that is starting to be examined in the field, but it is beyond the scope of this current project.

The Glimcher model, on the other hand, quantitatively diverges from the DDM. In particular, any model based on this sort of divisive normalization will predict larger response times and higher error rates for decisions where options have high value versus low. This is because for a similar level of difficulty (e.g., value difference in an economic decision), highly valued options will cause high levels of inhibition, which will diminish the difference in the value signals. This will imply a lower drift rate in a DDM, which, in turn, will predict slower responses and a higher likelihood of choosing erroneously. However, empirical data shows exactly the opposite—response times are typically smaller, and both accuracy and choice confidence

higher, when the values of the options are higher. In spite of the prevalence of this observation, there is currently no consensus as to the cognitive mechanisms at work.

Although it is not technically a model, it is worth mentioning here the principle of efficient coding. Put simply, this principle holds that the manner in which the brain encodes information (e.g., value) should be efficient, in the sense that mental representations should be constructed using the minimum amount of neural activity possible [Bhui & Gershman, 2018]. The rationale behind this is two-fold. First, neural activity is metabolically costly. It has been estimated that the brain accounts for 20% of resting oxygen consumption in humans, principally for signaling purposes [Laughlin, 2001], and that only 1% of the neurons in the brain can be substantially active concurrently [Lennie, 2003]. This provides a strong argument for why natural selection would have favored an efficient coding scheme. Second, information in the environment is highly structured and correlated, and thus redundant [Louie & Glimcher, 2012]. For this reason, neural coding that was not efficient would be unnecessarily wasteful, and thus detrimental to the adaptive fitness of the decision maker.

The divisive normalization model described above was built on the efficient coding principle. The authors refer to divisive normalization as a spatial aspect of efficient coding, wherein different pools of neurons (i.e., neurons in different spatial locations) adapt their firing rates in order to most efficiently distinguish the option values, given the inherent constraints of the system [Louie & Glimcher, 2012]. These same authors also suggest a temporal aspect of efficient coding, which they refer to as *range adaptation*. Here, the efficient coding adapts across time according to the evolving local (i.e., short-term) statistics of the environment, such as the mean, variance, and range of recent rewards. Range adaptation has also been supported by empirical data [Khaw, Glimcher & Louie, 2017; Louie & Glimcher, 2012; Kobayashi et al, 2010; Padoa-Schioppa, 2009].

The efficient coding principle has also been used as the foundation for a recent model that attempts to provide a unified framework to link empirically-observed effects of choice variability, biases, and confidence [Polania, Woodford & Ruff, 2019]. This model maximizes information transmission (within the decision system) by optimally allocating value representation resources according to the system's prior knowledge about the distribution of valuable options in its environment. Specifically, efficient coding explains that option values that lie within a zone of high density (i.e., many other options have values in the vicinity) will be encoded with higher precision than those in a zone of low density. This means that the explicit ratings should be more or less noisy depending on their location on the rating scale. This information can then be used to estimate the "true" underlying values of the options via Bayesian inference, where such values will be those that maximize the likelihood of the observed ratings. The combination of efficient coding and Bayesian decoding used in this model makes an interesting new prediction, which the authors refer to as a repulsion bias. This bias causes ratings for values that are close to the peak of the prior distribution (of rewards expected in the environment) to be repulsed from that peak. Ratings for values far from the peak will be attracted to it, as with classical Bayesian updating. Finally, this model predicts that choices between options in a high-density value zone will be made with more confidence, due to the higher fidelity of their value representations. Higher confidence for extreme ratings (high or low), as has previously been demonstrated [Lebreton et al, 2015], is also permitted in this model. Each of the predictions of this model (i.e., variability, bias, and confidence) was validated experimentally [Polania, Woodford & Ruff, 2019]. In addition to its theoretical merit and empirical validation, this model also provides a practical means for researchers to improve their own modeling and analyses. The most common models of choice (see above) take as the starting point of the decision process the previously-elicited value ratings, even though this ignores the likely possibility that the observed ratings are themselves the result of a (perhaps similar)

cognitive process. The model presented in this study offers a solution to this problem, where the “true” subjective values can be inferred from the ratings and used to account for the choices, instead of the ratings themselves [Polania, Woodford & Ruff, 2019].

### *Interim Conclusion*

Until now, we have learned that people make decisions in accordance to their personal preferences, which are often based on the value of the options available for consideration. Value in general refers to the extent to which a decision option will improve the decision maker’s life in some way. It is subjective, in the sense that it is agent-specific, assessed based on the individual circumstances and beliefs of the agent making the decision. Yet value is not something that exists in its own right, stored away somewhere in the memory of the agent. Value is determined on an ad hoc basis, wherein two options that at first glance appear identical can actually be assessed as having different values. This is because the way in which the value of the considered options, or perhaps the ordering of preferences among them, will be constructed can change according to context. Indeed, preferences between options can be reversed, even if the agent behaves rationally. The ability of the agent to process all information perfectly is impaired by inherent constraints, biological, physical, or other. Moreover, the costs involved with information processing will cause the agent to focus attention on some pieces of information and ignore others, even without any inherent constraints. A variety of computational models have been put forth to illustrate the dynamics of a simple decision process, and they seem to do a good job of summarizing the relationship between deliberation time and choice accuracy. Yet their ability to account for features of subjective value-based decision-making, such as preference changes or confidence about the choice, implores for a new approach that offers a more complete account of the decision process. With this current

project, we hope to provide such an approach. In what follows, we first present our thoughts that fundamentally guide our work, followed by a mathematical model that we believe describes an algorithmic implementation of simple value-based decision-making that could potential be used by the brain. Although we acknowledge that it is far from a self-contained solution, we hope that it will prove useful for future research in our shared quest to understand what happen in the mind of a decision maker during deliberation.

## **Chapter 2**

### **The Metacognitive Control of Decision-Making Model**

## WORKING HYPOTHESIS

One question that motivates our research is: why are the choices we make not always consistent with our stated preferences? The superficial answer to this seems to be: because our preferences are constructed at the moment of decision, when pondering about which of the alternative options that compose the choice set we most like and/or want. This invites us to ask the following question: why do we carefully ponder some decisions, but not others? Decisions permeate every aspect of our lives—from long-term decisions such as which career to pursue, to short-term decisions such as what to eat on a given day, to immediate decisions such as how to interpret something seen—but the amount of effort that we put into different decisions varies tremendously. Theoretically, thinking hard should always lead to an optimal choice, provided that we have sufficient information about the state of the world and about personal preferences. Yet, we often fail to even make use of all of the information that *is* available to us. We tend to rely on habitual behavior and decision heuristics (instead of more analytical cognitive processes), which can lead to systematic biases and judgment errors [Kahneman, 2011]. It has been proposed that we are forced to neglect decision-relevant information (rather than making “purely rational” decisions that incorporate all available information) by inherent constraints on human rationality (e.g., imperfect knowledge, insufficient time, inadequate processing power) that prevent us from being able to unequivocally determine which option is the best one [Simon, 1955]. However, an alternative perspective is that the brain has a preference for *efficiency* over accuracy [Thorngate, 1980]. Without necessarily being aware of it, we constantly make performance trade-offs between obtaining precision and conserving energy (i.e., limiting mental effort). Such trade-offs often lead us to make decisions without deep contemplation, even when constraints are not tight. In this light, effortless decision making (even with its associated biases) could be viewed as an evolutionary adaptation, rather than a

flaw that somehow managed to evade the filter of evolution. Thus remains an important open question—how do we determine how much mental effort to invest in a decision task?

Our working hypothesis is that mental effort allocation, or cognitive control during decision making, is strategically deployed. That is to say, the amount of executive resources (e.g., attention, working memory) that a decision maker (DM) will allocate to a decision task will be proportional to the expected value (benefit minus cost) of the allocation. It is common knowledge that increased effort can lead to enhanced performance on a wide variety of tasks, and also that reward can incentivize increased effort. Importantly, many studies have empirically demonstrated that investing mental effort is costly [Shenhav et al, 2017]. The benefit of cognitive control in a decision-making task is that it increases the likelihood that the DM will choose the (subjectively-defined) best option; the cost is directly inherent in the exertion of cognitive control itself. The brain should conduct a sort of cost-benefit analysis of the mental resource investment in an automatic and optimal way, or the expected value of cognitive control [Shenhav, Botvinick & Cohen, 2013]. The selected level of cognitive control should thus represent the best trade-off between effort and performance.

An important element of decision making is the concept of choice confidence, or the DM's subjective feeling that the chosen (or potentially chosen) option is indeed the best one. The nature of this confidence is contentious in current scientific debate. Some theories purport that confidence is merely an afterthought of the decision task, or perhaps a quantitative assessment of the relative strength of the chosen option versus the rejected option [DeMartino et al, 2013]. Others maintain that confidence could be useful for the DM, but only if he were to encounter similar decisions in the future [Folke et al, 2017]. In contrast to these ideas, it is our firm contention that choice confidence is itself instrumental to the decision task at hand. A DM will be inclined to resist making a choice that he does not feel confident about. If the best choice is obvious, the decision will be easy and confidence high, and the DM will make the choice

without further thought. Otherwise, the DM will prefer to invest effort into the task in order to increase his confidence until it reaches a minimum threshold level. This threshold will be different for each decision, determined in part by decision difficulty and the importance of that decision to the various goals that the DM might have. In effect, the DM is selecting the amount of confidence to purchase, with effort being the currency of payment. The brain of the DM will monitor the instantaneous confidence level throughout deliberation. As long as confidence remains below the threshold level, specific to the current decision and determined by a cost-benefit arbitrage, the DM will continue to invest additional mental effort. This effort leads directly to supplementary information processing, which provides higher precision and potentially greater differentiation to the mental representations of the decision options. This in turn causes a boost in confidence, which will either eventually surpass the threshold or stabilize at some sub-threshold level and encourage the DM to cease deliberation. Critically, this process also allows for changes of mind, where the DM eventually chooses the option that he was initially inclined to reject, because he has modified his representation of the options' values while pondering the decision. This is a critical difference with evidence accumulation models, for which values are left unchanged by the decision process. In turn, these models consider all decisions that do not match the reported values as errors. Note that, under our framework, monitoring confidence would serve no purpose if the DM could not change his mind (and was somehow forced to choose according to his initial uncertain preference). But if the DM is allowed to change his mind during deliberation, then optimizing the effort-for-confidence tradeoff becomes unambiguously instrumental. In other words, we ponder decisions because we consider the possibility of changing our mind when contemplating alternative options.

To expound our idea, we developed a computational model to demonstrate the mechanics of the mental effort allocation process. Our model relies on three main components that we believe drive the process: decision importance, decision difficulty, and the costly nature of effort. A

decision whose consequences are expected to have a greater impact on the DM's life will be more important, and thus motivate a higher level of resource allocation. A decision whose options are comparable in value will be more difficult, and thus require more resources in order for the DM to be able to distinguish the options and make the best choice. Finally, it is crucial to note that mental effort allocation is costly and thus provides a disincentive for the DM to simply expend maximal effort on every decision task that he is ever faced with. The technical details of our model are provided below.

In brief, our model attempts to capture the case-by-case control of mental effort allocation during decision-making tasks. As elucidated below, it borrows from the theoretical framework of "Expected Value of Control" (EVC), which was originally proposed to account for the impact of motivation on cognitive control [Shenhav, Botvinick & Cohen, 2013]. It offers a simple yet robust explanation for when people will be inclined to invest more versus less mental resources in a decision-making task, and what the effects of this investment on several key decision variables should look like. This will become more clear below, when we present the details of the model.

In contrast to evidence accumulation models, ours is in the spirit of the literature on constructed preferences [Tversky & Thaler, 1990; Payne, Bettman & Johnson, 1993; Lichtenstein & Slovic, 2006], and is aligned with empirical results on choice-induced preference change. Recall that constructed preferences demonstrates that decision makers often make choices that are inconsistent with their previously-stated preferences. This observation has been explained by suggesting that preferences are not merely retrieved from a stable database in the decision maker's memory, but rather calculated during the choice task. This is the starting point of our model, which then proposes the mechanisms through which the process of the preference calculations is controlled. Recall that choice-induced preference change demonstrates that preferences are not stable [see Harmon-Jones & Harmon-Jones, 2007; Izuma & Murayama,

2013 for reviews]. Here, the mere act of choosing one option over another tends to cause increased valuations for the chosen items and decreased valuations for the rejected items. As will be explained below, our model can explain such choice-induced preference changes under the assumption that valuation changes occur *before* the choice is finalized, not *after*.

## COMPUTATIONAL MODEL

The aim of this model is to describe the cognitive mechanisms that are engaged during a subjective value-based choice task between two options. The agent will choose the option that it believes to be better (i.e., the one with higher expected subjective value), but only if its confidence (i.e., subjective probability that the option to be selected is the best one) reaches a target threshold. At any point during the deliberation process, if the agent's instantaneous confidence level is sub-threshold, it will incrementally invest mental effort in order to process additional information in hopes of achieving greater confidence. Upon reaching either a satisficing level of confidence or exhausting the full amount of resources that he is willing to allocate to the task, the agent will finalize its choice.

In what follows, we propose a simple mathematical model of how cognitive resources are allocated in the context of value-based decision-making. In brief, we assume that this type of motivational control of decision-making is composed of two elementary cognitive processes: the subjective evaluation of choice items (how much we like or want things) and the monitoring of choice confidence (how certain we are that the currently-preferred option is better than the alternatives). The model describes how mental effort "fuels" these two processes, through the allocation of costly cognitive resources that eventually change the representations of item values and their precisions. The critical questions we address are whether and how expended

cognitive effort is restricted when making decisions, a phenomenon that at times leads to irrational, biased behavior and/or errors.

Our complete model can be thought of as being composed of two separate parts, or phases, which we will reveal below. The first part can be labeled the “allocation phase”. It is during this phase that the DM’s brain establishes a “budget” for how much effort the DM should be willing to invest in a specific task (which might be different from what the DM *does* invest, when all is said and done). The purpose of this initial phase is to set a threshold for effort allocation (the optimal upper bound), along with another threshold for the minimal satisficing choice confidence level (which will equal the expected level of choice confidence given that the maximal effort allocation is reached). These thresholds will constrain the dynamics of the second part of the model.

The core of our computational model, in its simplest form, is the most basic of classical economic models—a direct cost-benefit comparison resulting in a net expected value. The expected value (EV) here refers to the overall gain that a DM will expect to achieve as the result of having invested effort into a specific decision-making task. Effort investment enhances performance but carries a cost. The EV is thus comprised of a benefit term (B) and a cost term (C):

$$EV = E[B - C] \tag{1}$$

where  $E[.]$  refers to an expectation that accounts for predictable stochastic influences that ensue from allocating resources (details below). The benefit term can be broken down into the reward (R) for choosing the best option times the probability of choosing the best option ( $P_c$ ):

$$B = R * P_c \tag{2}$$

Note that the reward scaling factor R captures the importance of making a "good" decision, whereas  $P_c$  really measures one's confidence in making a "good" decision. The true benefit of

resource allocation, however, is not the expected confidence itself, but rather the expected *gain* in confidence (i.e., expected confidence after resource allocation minus expected confidence before resource allocation). Specific to the notion of cognitive control, the expected value of control (EVC), as well as each of its components, will be a function of the amount of cognitive resources invested in the task (which we label  $z$ ):

$$EVC(z) = R * (P_c(z) - P_c(0)) - C(z) \quad (3)$$

The only requirement for  $C$  is that it be monotonically increasing (in this paper, we assume it is linear for simplicity and without loss of generality). The EVC equation can thus be simplified as follows:

$$EVC(z) = R * (P_c(z) - P_c(0)) - \alpha z \quad (4)$$

where  $\alpha$  is simply the cost of effort per unit of resources.

To derive an expression for  $P_c$ , we will refer to a decision task of the simplest form—binary choice, singular dimension of comparison. For each option, we represent the DM's *pre-choice* subjective value ( $V$ ) as a normally distributed random variable with mean equal to the perceived value and variance equal to the uncertainty of that value:

$$V_i \sim N(\mu_i, \sigma_i); i \in \{1,2\} \quad (5)$$

In other words, the DM is endowed with a probabilistic representation of option values, whose mode  $\mu$  measures the strength of the subjective feeling of liking/wanting the option, and whose width  $\sigma$  measures how uncertain that feeling is (see Figure 1). As we will see below, this probabilistic representation evolves along with the allocation of resources during the decision process.

When making a decision, however, it is not the individual option values that are important, but rather the difference between them. We thus examine this difference ( $\delta V$ ) as a normally

distributed random variable with mean equal to the difference of the option means and variance equal to the sum of the option variances:

$$\delta V_{i-j} \sim N(\mu_i - \mu_j, \sqrt{\sigma_i^2 + \sigma_j^2}); i, j \in \{1, 2\} \quad (6)$$

A correct choice is defined as choosing the option with the highest value:

$$\text{correct choice} \stackrel{\text{def}}{=} \begin{cases} \text{option } i, & \text{if } \delta V_{i-j} > 0 \\ \text{option } j, & \text{otherwise} \end{cases} \quad (7)$$

Subjective uncertainty regarding option values eventually translates into uncertainty regarding what the "correct choice" is. This is summarized in terms of the probability  $P_c$  of "committing to the correct choice":

$$P_c = \begin{cases} 1 - cdf_{\delta V_{i-j}}, & \text{if } E[\delta V_{i-j}] > 0 \\ cdf_{\delta V_{i-j}}, & \text{otherwise} \end{cases} \quad (8)$$

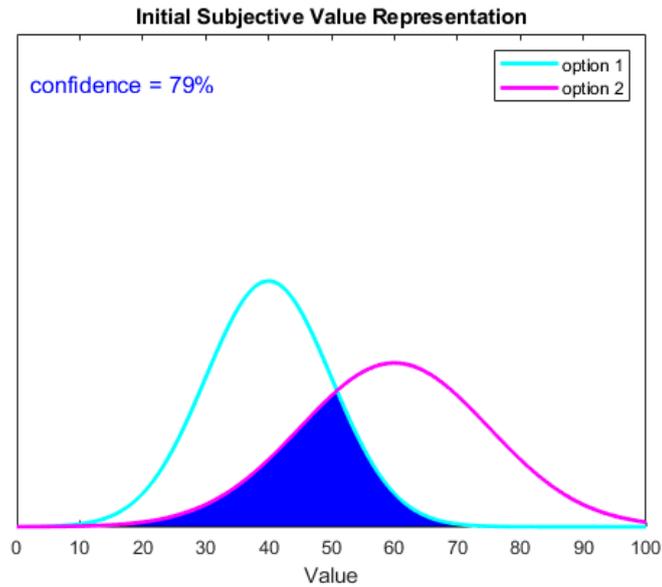
where  $cdf$  is the cumulative density function of the corresponding normal distribution, evaluated at  $\delta V_{i-j} = 0$ . Critically,  $P_c$  can be viewed as measuring subjective *choice confidence*, in the sense that confidence stems from the belief that the currently preferred option is better than the other (see Figure 1).

As long as the choice that is considered aligns with the preference (i.e., the sign of  $\delta V_{i-j}$ ), the assignment of any particular option to either  $i$  or  $j$  is arbitrary but has a symmetrical impact on  $P_c$ , which yields the following simplification of the above equation:

$$P_c = 1 - cdf_{|\delta V_{i-j}|} \quad (9)$$

$$\approx \frac{1}{1 + e^{\frac{-\pi |\mu_i - \mu_j|}{\sqrt{3(\sigma_i^2 + \sigma_j^2)}}}} \quad (10)$$

where the second line derives from a simple analytical approximation to the normal cumulative density function [Daunizeau, 2017].



**Figure 1:** The two Gaussian curves represent the values of two hypothetical choice options. Here, the mean of each distribution represents the decision maker’s subjective estimate of how much he appreciates that option; the variance of each distribution represents the decision maker’s feeling of uncertainty about his estimate. When comparing the two options, the region of overlap between the curves represents the decision maker’s lack of confidence about knowing which option he prefers.

As discussed above, choice confidence (i.e.,  $P_c$ ) will be instrumental in the effort allocation process. If initial confidence (with zero or some other baseline level of effort) is already high enough, the DM will not be motivated to invest additional effort, because any further increase in choice confidence would be overcompensated by its induced subjective cost. Otherwise, the DM will invest mental effort with the goal of increasing  $P_c$  to the point where it reaches a minimum (subjective) threshold level. In our model, effort investment impacts  $P_c$  in two ways: (i) via an increase in the precision of  $\delta V$ ; and (ii) via a change in the first-order moment of  $\delta V$  that allows for the possibility of reversing the DM's initial preference. This follows from the fact that resources are allocated for the purpose of an information-gathering process (regarding item values), which changes the DM's probabilistic representations of choice option values by both increasing their precisions and perturbing (in an unpredictable manner) their centers of mass.

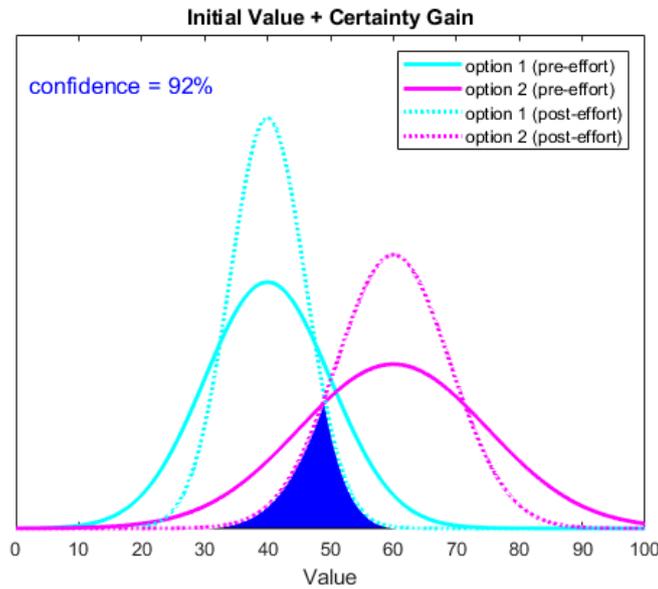
The more information the DM takes into consideration, the greater is the precision of his value estimates. Thus, the precision  $\phi$  of  $\delta V$  after effort is the initial precision plus an additive term that is a function of the amount of allocated resources (herein linear, for simplicity and without loss of generality). The magnitude of the increase in precision is scaled by an efficiency parameter that we call  $\beta$ :

$$\phi \stackrel{\text{def}}{=} \frac{1}{\sigma_i^2 + \sigma_j^2} \quad (11)$$

$$\phi \xrightarrow{z} \phi + \beta z \quad (12)$$

The increase in precision causes an increase in  $P_c$  (see Figure 2):

$$\frac{1}{1 + e^{-\pi \sqrt{\frac{\phi + \beta z}{3}} |\mu_i - \mu_j|}} > \frac{1}{1 + e^{-\pi \sqrt{\frac{\phi}{3}} |\mu_i - \mu_j|}} \quad (13)$$



**Figure 2:** Through mental effort, the decision maker can gain confidence about his preferences even if his initial beliefs do not change. Here, the option value curves become narrower as the decision maker invests mental resources to process additional information and lower his uncertainty. This in turn causes the region of overlap to shrink, indicating a gain in choice confidence.

Allocating resources also implies greater potential for changes in the DM's value estimates. These changes are equally likely to be positive or negative, with a range proportional to the amount of allocated resources. A mathematically convenient way to capture this phenomenon is via a stochastic perturbation term on the center of mass  $\mu$  of value representations. We thus include this in our model as a normally-distributed random variable with mean zero and variance equal to effort scaled by an efficiency parameter that we call  $\gamma$ :

$$|\delta\mu| \xrightarrow{z} |\delta\mu + \Delta| \quad (14)$$

$$\Delta(z) \sim N(0, \gamma z) \quad (15)$$

A congruent change (increase) in  $\delta V$  causes an increase in  $P_c$  (see Figure 3):

$$\frac{1}{1+e^{-\pi\sqrt{\frac{\Phi}{3}}|\mu_i-\mu_j+\Delta|}} > \frac{1}{1+e^{-\pi\sqrt{\frac{\Phi}{3}}|\mu_i-\mu_j|}} \text{ if } |\mu_i - \mu_j + \Delta| > |\mu_i - \mu_j| \quad (16)$$

but any incongruent change in  $\delta V$  (i.e.,  $|\mu_i-\mu_j+\Delta| < |\mu_i-\mu_j|$ ) can lead to a decrease in  $P_c$ .

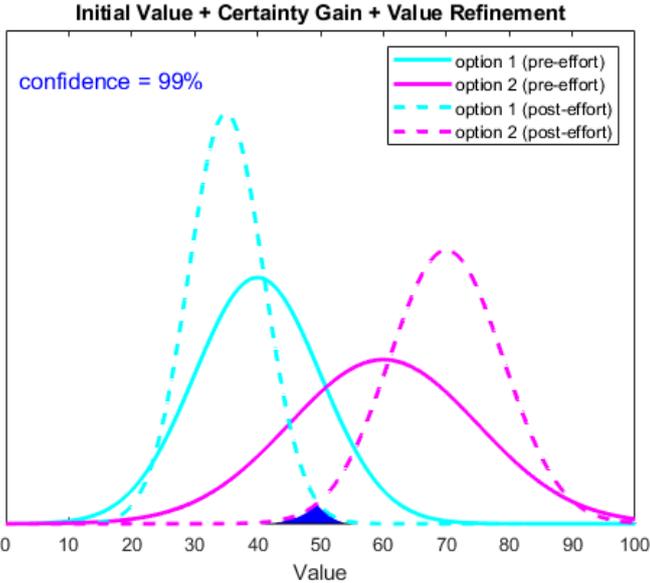
Nevertheless, the impact of  $\Delta$  on the *expected* confidence  $E[P_c]$  is always beneficial, meaning allocating resources always increases the expected confidence. To see this, first note that the expected absolute difference between the two modes after resource allocation  $E[|\mu_i-\mu_j+\Delta|]$  is given by:

$$E \left[ |\mu_i - \mu_j + \Delta| \mid z \right] = \sqrt{\frac{2\gamma z}{\pi}} e^{\left(\frac{-|\mu_i-\mu_j|^2}{2\gamma z}\right)} + |\mu_i - \mu_j| * \text{erf} \left( \frac{|\mu_i-\mu_j|}{\sqrt{2\gamma z}} \right) \quad (17)$$

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \quad (18)$$

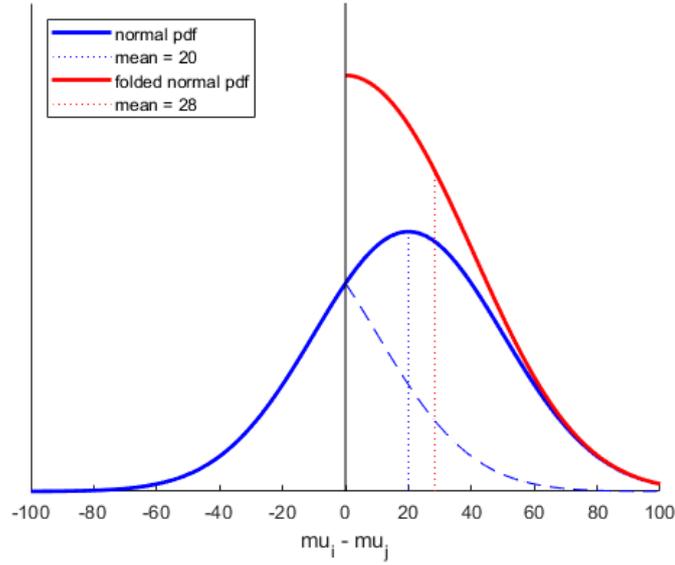
where we have simply used the expression for the first-order moment of the so-called “folded normal distribution”. Critically, the expected value of a folded normal distribution is always greater than the expected value of the associated ordinary normal distribution (see Figure 4). Hence, the DM will be encouraged to invest resources to allow for perturbations (i.e., potential changes of mind) even though he knows this could sometimes work against him (because on

average, the effect will be beneficial). Here, we would want to pass  $\delta V$  through the  $P_c$  function before taking its expectation, but for mathematical simplification, we instead rely on the approximation  $E[P_c(|\mu_i - \mu_j + \Delta|)] \approx P_c(E[|\mu_i - \mu_j + \Delta|])$ .



**Figure 3:** Through mental effort, the decision maker can adjust gain confidence about his preferences by changing his initial beliefs. Here, the option value curves shift apart as the decision maker considers information incongruent with his initial value estimates. This in turn causes the region of overlap to shrink, indicating a gain in choice confidence.

A crucial feature of the stochastic perturbation term of our model is the fact that it allows for the possibility that the DM will change his mind, eventually choosing the option that he was initially inclined to reject. The probability of changing one’s mind is simply the probability that the perturbation causes the initial ordering of option values to reverse. The fact that this probability is non-zero is important because otherwise, the outcome of a choice should always be identical with or without effort, and therefore it would never be rational to allocate any resources to the task.



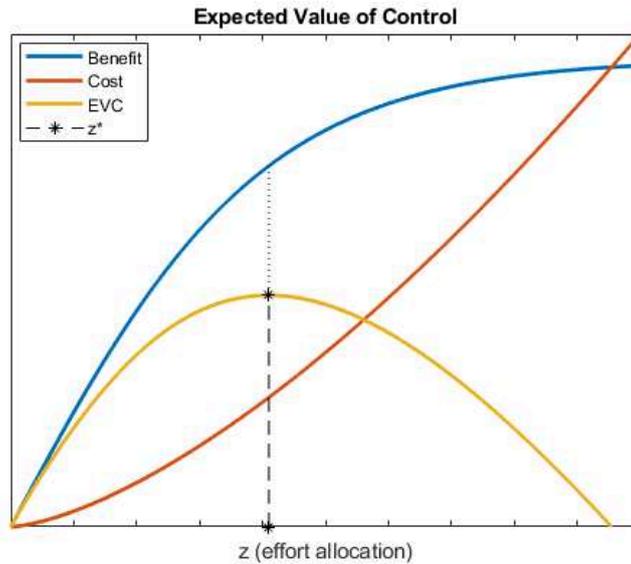
**Figure 4:** The mean of the “folded” Normal probability density function is always strictly greater than the mean of its corresponding standard Normal pdf. Here, the portion of the standard pdf that lies in the negative domain is “folded” back into the positive domain and added to the density already found there.

For any particular decision task at hand, the DM will select the optimal level of effort allocation ( $z^*$ ) by maximizing the EVC:

$$z^* = \underset{z}{\operatorname{argmax}} EVC \quad (19)$$

where the expressions for the expected impact of resource allocation on the distribution of  $\delta V$  has been inserted into the mathematical definition of the expected choice confidence in the EVC. Because the expected choice confidence ( $E[P_c]$ ) and the cost of effort ( $C$ ) are both monotonically increasing functions of  $z$ , the optimal effort allocation ( $z^*$ ) will be the level at which the marginal cost equals the marginal benefit, beyond which the cost would outweigh the benefit (see Figure 5). Note that the EVC-optimal level of expected choice confidence ( $P_c^*$ ) is obtained by inserting  $z^*$  into the  $E[P_c]$  formula:

$$P_c^* \approx \frac{1}{1 + e^{-\pi \sqrt{\frac{\phi + \beta z^*}{3}} E[|\mu_i - \mu_j + \Delta| | z^*]}} \quad (20)$$



**Figure 5:** The benefit of mental effort is a monotonically increasing, concave function of resource allocation. The cost of mental effort is also monotonically increasing with resource allocation. The expected value of control is simply the difference between the benefit and cost functions. Here, the maximum value of the EVC function indicates the EVC-optimal amount of resource allocation ( $z^*$ ).

Now the model predicts the EVC-optimal amount of allocated resources  $z^*$  as a function of three important features of the decision task (initial difficulty  $|\delta\mu|$ , initial precision  $\phi$ , and incentive  $R$ ), and three parameters that capture some idiosyncrasies specific to the DM (the subjective effort cost of resource allocation  $\alpha$ , and the efficiency of effort investment regarding either precision gain or value revision,  $\beta$  and  $\gamma$ ). Note that the model also allows us to predict how the reported choice confidence level, the probability of changing one’s mind, the change in rating confidence, and the average spreading of alternatives should vary as initial difficulty, initial precision, and incentive vary.

The second part of our model can be labeled the “implementation phase”. It is during this phase that the resources are actually allocated, initiating the cognitive processes that enable the reassessment of value representations. Because it does not make sense to say that one could simply expend a pre-allocated amount of effort all at once, it is reasonable for us to claim that effort is invested *incrementally*. (This might become more clear if one thinks of investing

mental effort in terms of processing an additional bit of information, activating an additional neuron, or some other such discrete but fine-grained depiction.) The DM will continue to invest additional resources until either of the two aforementioned thresholds is reached, at which point he will make a choice. The threshold for maximal effort allocation is merely  $z^*$ , and the threshold for satisficing choice confidence is merely  $P_c^* = E[P_c|z^*]$ .

Computationally, the implementation phase of our model can be represented as a drift-diffusion process that captures the evolution of choice confidence across time. Here, we are interested in how the process evolves across resource investment rather than time. The initial location of the process for any given decision will be the probability of making the correct choice without investing any effort, or  $P_c(0)$ . Across total effort investment ( $z$ ), the location will depend on incremental random drifts in the valuations of the options, along with incremental gains in the precisions of said valuations. The confidence follows a random walk because the stochastic perturbation term will impact the instantaneous value representations (more precisely, their modes) (see Figure 6).

$$|d\mu_0| = d\mu(z = 0) \tag{21}$$

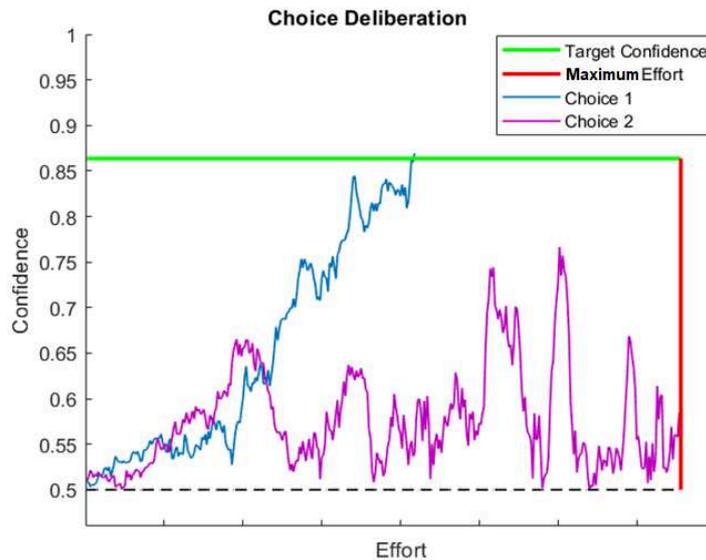
$$|d\mu_z| = |d\mu_{z-1} + \Delta|, \Delta \sim N(0, \gamma), z > 0 \tag{22}$$

$$\phi_0 = \phi(z = 0) \tag{23}$$

$$\phi_z = \phi_{z-1} + \beta, z > 0 \tag{24}$$

$$P_z = \frac{1}{1 + e^{-\pi \sqrt{\frac{\phi_z}{3}} |d\mu_z|}} \tag{25}$$

In concurrence with the evolution of choice confidence, a resource monitor will tally the accumulation of invested resources. The DM will continue to invest additional increments of resources until one of the thresholds is reached by its respective accumulator process ( $z = z^*$  or  $P \geq P^*$ ).



**Figure 6:** The implementation phase of cognitive resource allocation during a decision can be represented by a drift-diffusion process. Here, the two curves represent two individual decision trials where the DM initially had identical beliefs about the options. The starting point (i.e., with zero effort) shows the default level of choice confidence if the DM declines to deliberate, which is based solely on his uncertain prior beliefs. Through effort, the allocated resources allow for additional information to be processed, which in turn impacts how confident the DM feels about the imminent choice. The horizontal threshold represents the confidence level that the DM seeks to achieve prior to committing to a choice. The vertical threshold represents the maximum amount of resources that the DM is willing to allocate before terminating deliberation with a choice. Here, Choice 1 represents a situation where the information that the DM processed was helpful early on, allowing him to reach a high level of confidence and stop deliberating sooner than anticipated. Choice 2 represents a situation where the processed information was conflicting and unhelpful, and the DM ended up making his choice with a lower-than-desired level of confidence.

## Impact of Input Variables on Predicted Output Variables

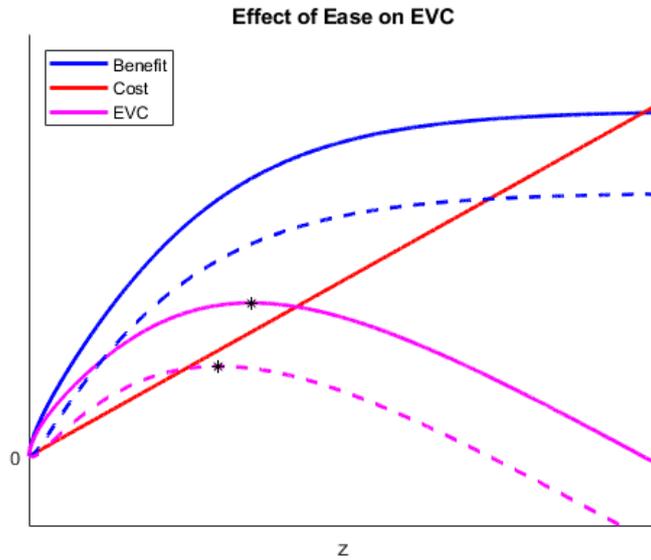
Our model makes some clear behavioral predictions dependent on a small set of input variables (i.e., distance, certainty) and parameters (i.e., alpha, beta, gamma). Below, we briefly describe the model predictions as each of the different input factors is varied.

### *Choice Ease*

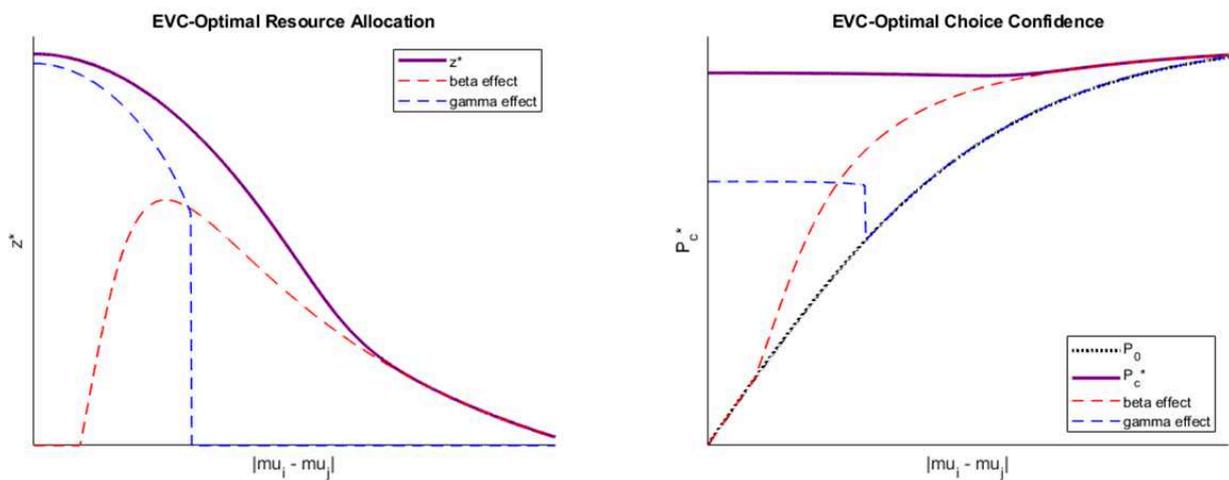
The most important driver of our model is what we call choice ease. Here, ease is defined as the distinguishability of the options in a choice task. The primary component of ease is what we call *value distance*, or the absolute difference between the modes of the value representations of the two choice options. The second component of choice ease is what we call

*value certainty*, which is the level of certitude that a DM has about his valuations of different options. An increase in either distance or certainty will make the two options more distinguishable. Logically, the more distinct the values of the options are, the easier the choice will be for the decision maker. Our model follows this same logic. As such, an increase in ease (e.g., replacing one choice option with an option whose subjective value estimate is further from that of the comparison option, or whose value estimate is more certain) is expected to cause a decrease in allocated mental resources (depending on  $\beta$  and  $\gamma$  parameter values—see below for more details). This is because the cost function for effort allocation remains unchanged, while the benefit function pointwise decreases (see Figure 7). This is so because the core of the benefit is choice confidence, and with an easier choice, the initial (i.e., pre-deliberation) confidence is higher. Higher initial confidence implies lower potential confidence gain via resource allocation. The result is that the maximum of the EVC function (i.e., the optimal resource allocation) occurs with fewer allocated resources.

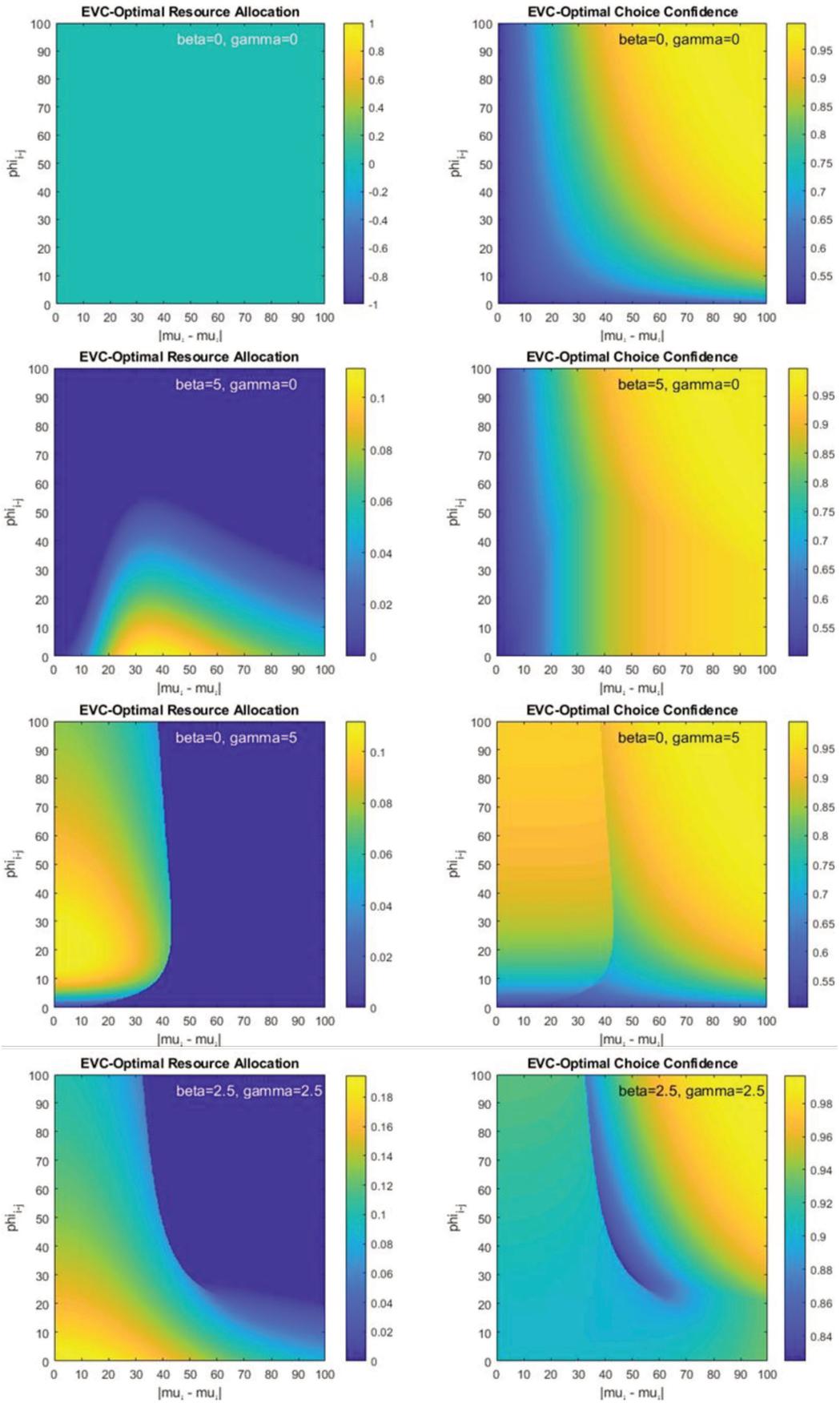
The effect of choice ease on resource allocation is monotonically decreasing (again, depending on  $\beta$  and  $\gamma$  parameter values). Here we separate the effects of value distance and value certainty for illustrative purposes (see Figure 8). Interestingly, the predicted effect of choice ease on the EVC-optimal expected choice confidence is less straightforward. In our model, choice confidence is effectively the same as choice ease, defined as the distinguishability of the options. Thus, an increase in ease should yield an increase in EVC-optimal expected confidence. Yet the impact of resource allocation on confidence can distort the relationship between ease and EVC-optimal expected confidence, as the EVC-optimal resource allocation is itself a function of ease. This slightly non-trivial relationship is depicted in Figure 9 below.



**Figure 7:** The EVC-optimal level of resource allocation ( $z^*$ ) is the point where the EVC curve is maximized. Here, we show the effect of a lower versus higher initial level of choice ease. The solid curves represent a lower level, the dashed curves a higher level. Here, we see that an increase in ease causes a decrease in benefit (i.e., expected gain in confidence), which leads to a decrease in  $z^*$ .



**Figure 8:** The purple curves illustrates the monotonically decreasing relationship between value distance (a key component of choice ease) and the EVC-optimal levels of resource allocation ( $z^*$ ) (left plot) and expected choice confidence ( $z^*$ ) (right plot). As mentioned above, this relationship depends on the beta and gamma parameters. Here, the separate effects of each parameter are illustrated. The “beta effect” shows what the relationships would look like if gamma were not a factor. The “gamma effect” shows what the relationships would look like if beta were not a factor. Note that there is also an interaction effect between beta and gamma, which helps to blend the two effects together more smoothly when both parameters are factors.

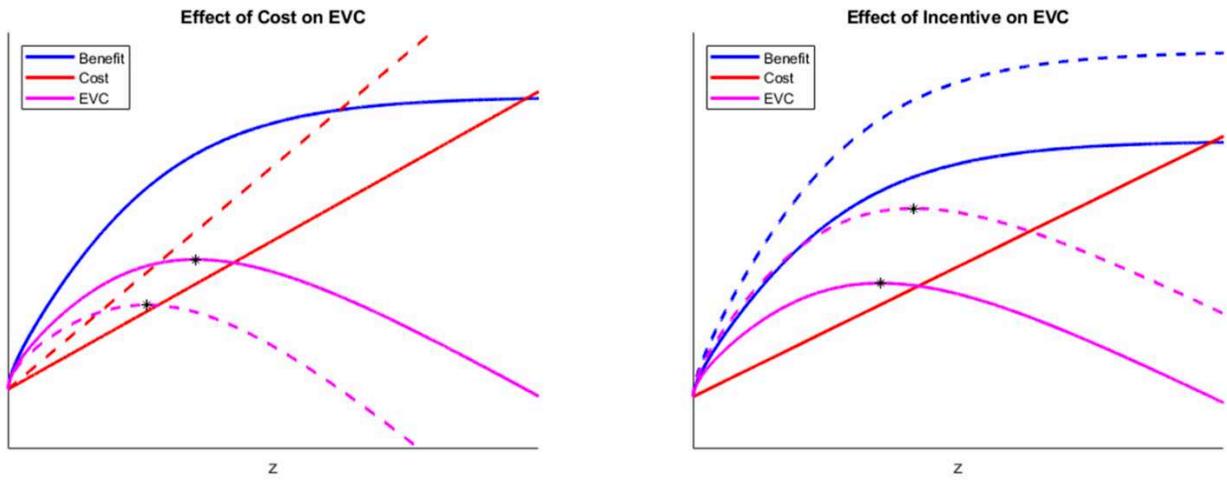


**Figure 9:** Here, the beta and gamma effects are illustrated further. Each plot shows the relationship between choice ease and the key output variables of our model. Ease is broken down into value difference (horizontal axes) and value certainty (vertical axes). The left plots show EVC-optimal resource allocation, the right plots show EVC-optimal expected choice confidence. The first row shows the situation when both beta and gamma are equal to zero. The second row shows the beta effect. The third row shows the gamma effect. The bottom row shows a typical situation, where both beta and gamma are factors.

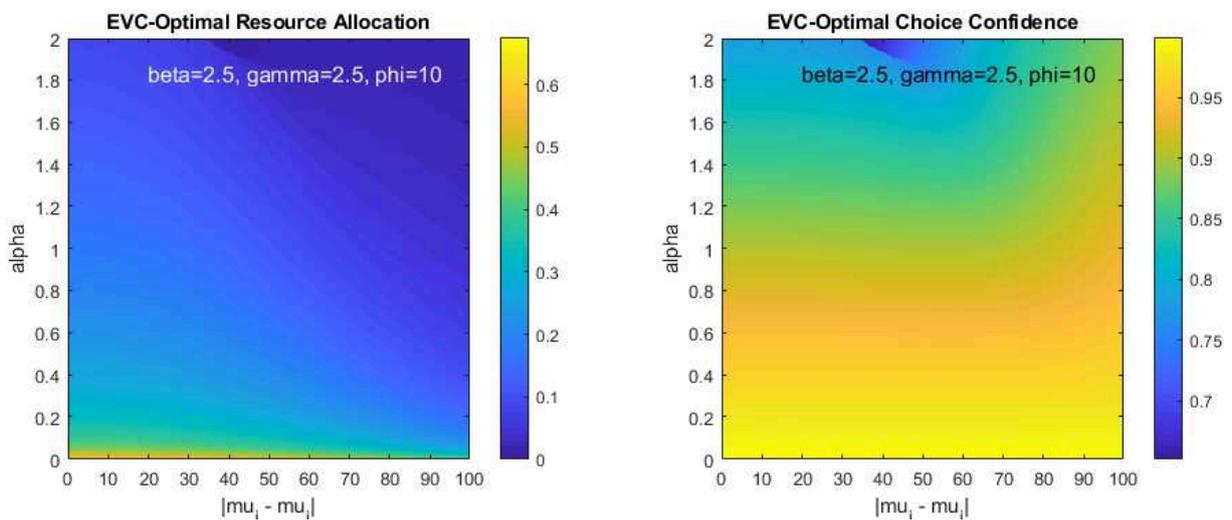
### *Cost-Benefit Tradeoff*

In our model, the subjective effort cost of cognitive resource investment is characterized by the *alpha* parameter, and the subjective benefit of confidence gain is characterized by the R term. These subjective costs and benefits are inherent properties of the decision maker (i.e., they quantify subjective traits), but they could act in conjunction with externally imposed incentives and/or constraints. Intuitively, an increase in the cost of resource allocation will encourage the decision maker to allocate fewer resources to a given choice task. In our model, this would be captured as a pointwise increase in the cost function, with no change in the benefit function, resulting in a lower optimal resource allocation level (see Figure 10). On the other hand, an increase in the benefit term (e.g., an increased decision importance) will encourage the decision maker to allocate more resources. In our model, this would be captured by a pointwise increase in the benefit function, with no change to the cost function, resulting in a higher optimal resource allocation level (see Figure 10).

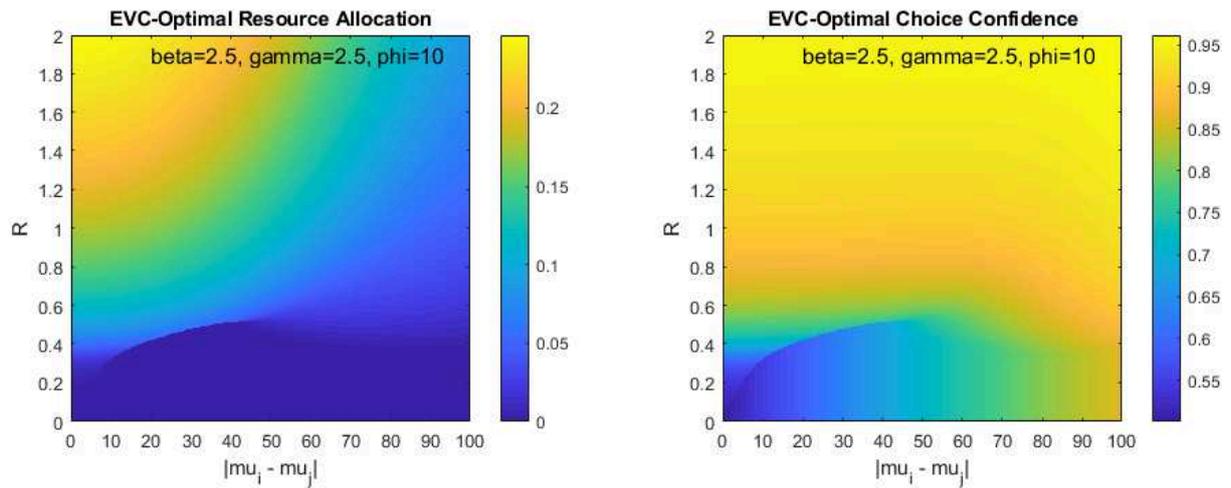
The impact of alpha on resource allocation is monotonically decreasing (see Figure 11). Alpha impacts EVC-optimal expected confidence only indirectly through the relationship between effort and confidence. As the relationship between resource allocation and confidence is monotonically increasing, the impact of alpha on EVC-optimal expected confidence follows its impact on resource allocation (i.e., monotonically decreasing) (see Figure 11). In contrast, the impact of incentive (R) on both EVC-optimal resource allocation and expected confidence is monotonically increasing (see Figure 12).



**Figure 10:** Here, we show the effect of a lower versus higher cost of effort (left plot) and incentive for confidence (right plot). The solid curves represent a lower level, the dashed curves a higher level. Here, we see that an increase in cost directly leads to a decrease in  $z^*$ . Conversely, an increase in incentive causes an increase in benefit, which leads to an increase in  $z^*$ .

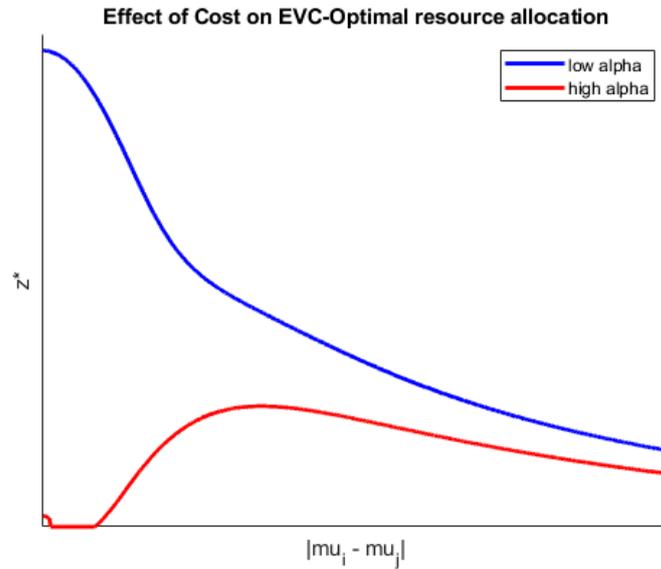


**Figure 11:** The impact of effort cost ( $\alpha$ ) on EVC-optimal resource allocation (left plot) and choice confidence (right plot) is monotonically decreasing.



**Figure 12:** The impact of incentive ( $R$ ) on EVC-optimal resource allocation (left plot) and choice confidence (right plot) is monotonically increasing.

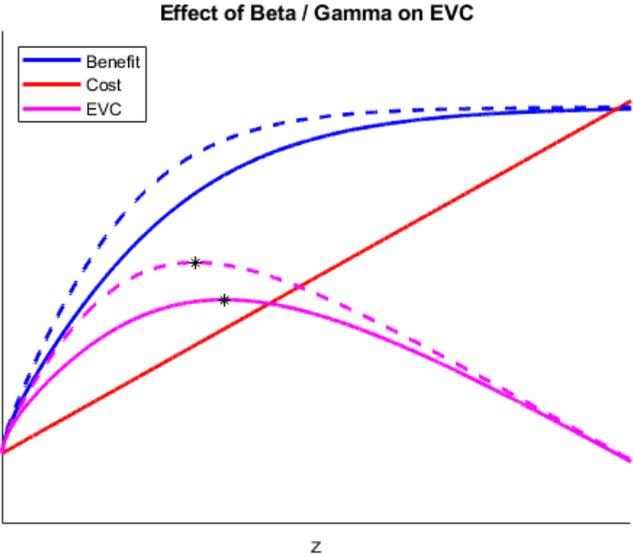
Depending on the magnitude of the beta and gamma parameters, changes in either alpha or incentive could lead to qualitative changes in the relationship between choice ease and EVC-optimal resource allocation (see Figure 13). For example, with a low incentive, resources are not allocated for the most difficult choices, only for those of sufficient ease. In this *beta-dominant regime*, the probability of changing one's mind is rather low, and the main effect of resource allocation is to feel more comfortable with one's choice. With a high incentive, however, the decision maker allocates more resources when choice difficulty increases. In this *gamma-dominant regime*, the probability of changing one's mind is higher, thus making deliberation all the more valuable.



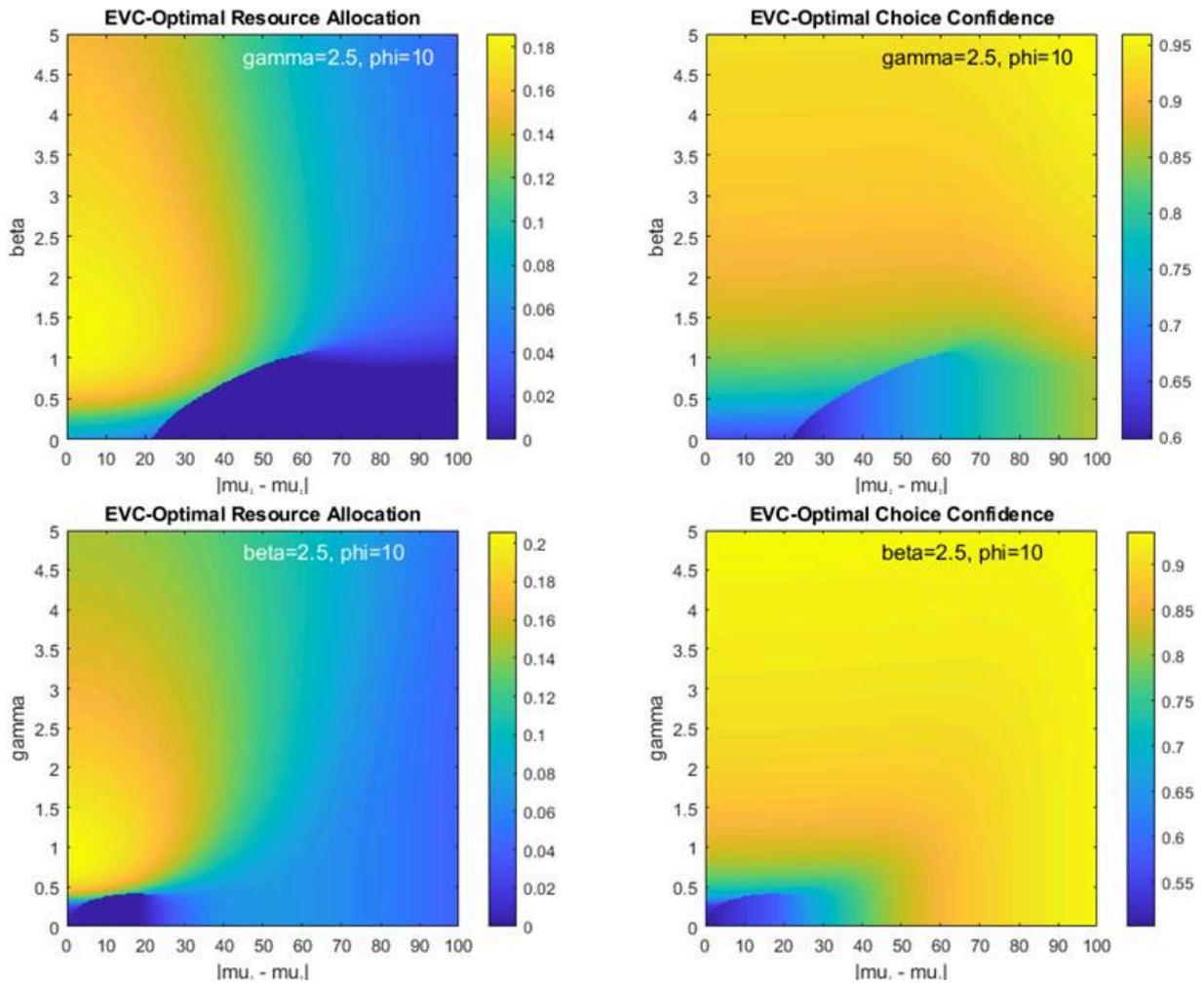
**Figure 13:** Depending on effort cost ( $\alpha$ ), our model can predict qualitatively different DM behavior. Here, we show that in a high cost situation, the DM to allocate resources for relatively easy choices, but not for difficult ones. In a low cost situation, all else equal, the DM will allocate more resources as difficulty increases.

Now efficacy parameters enter the model through the expected benefit term (see Figure 14). Interestingly, the relationship between either of these parameters and EVC-optimal resource allocation is non-monotonic. Recall that, for small values, increases in the beta or gamma parameters can be thought of as an increase of the efficiency of the “beta effect” or the “gamma effect”, respectively. Small increases in efficiency will encourage greater resource allocation, because the expected benefit of allocating resources will overcompensate its cost. Further increases in efficiency, however, will encourage lesser resource allocation. This is because the expected benefit function is asymptotic at maximum plateau. Increasing the beta or gamma parameter beyond a certain point will allow this plateau to be approached faster, thus reducing the amount of resources whose expected benefit will be outweighed by its expected cost. Note that the effect of an increase in beta or gamma on EVC-optimal expected confidence, on the other hand, will be monotonically increasing (see Figure 15). (See Figure 16 for an illustration

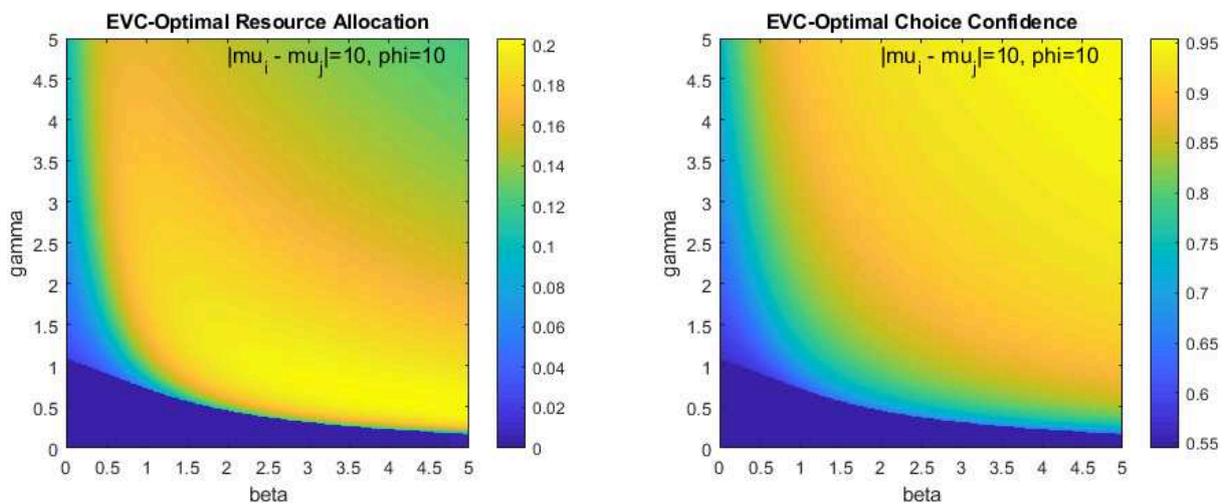
of how EVC-optimal resource allocation and expected confidence change as both beta and gamma change.)



**Figure 14:** Here, we show the effect of a lower versus higher cognitive resource efficiency (either beta or gamma). The solid curves represent a lower level, the dashed curves a higher level. Here, we see that an increase in efficiency leads to a decrease in  $z^*$ .



**Figure 15:** For low levels of  $\beta$  or  $\gamma$ , a small increase will lead to an increase in EVC-optimal resource allocation (left plot) and expected confidence (right plot). However, further increases beyond a certain point will lead to a decrease in resource allocation and an increase in expected confidence.



**Figure 16:** As with either parameter individually, when beta and gamma change together, the effect on EVC-optimal is non-monotonic. The effect on EVC-optimal expected confidence remains monotonically increasing.

### **The online metacognitive control of decision-making**

The motivational decision control described above focuses on the issue of identifying the optimal amount of cognitive resources to allocate during a decision task, given that the prior preference of the DM is uncertain. Importantly, the DM may change his mind while reassessing the value of the alternative options, meaning that the preferred option may eventually be different from the prior preference. Such change of mind is due entirely to the stochastic perturbation term, which can eventually reverse the prior preference. This implies that the opportunity to change one's mind depends upon the amount of resources allocated to the decision, which is itself determined by both the properties of the prior value representations and the effort efficacy and cost parameters. Nevertheless, evaluating the probability to change one's mind requires to consider the online implementation of decision control.

Under the above premises, decision control could simply be realized by monitoring online the amount of allocated resources, and stopping when reaching the EVC-optimal level. Another reasonable solution would be to monitor confidence, which evolves randomly under the influence of the stochastic perturbation term, and stop when reaching the EVC-optimal expected level. Hence, we suggest that the monitoring system stops whenever either of these two criteria is reached. In such a scenario, confidence is central for deciding whether or not to continue to invest resources in the decision. This is why we refer to this as a form of *metacognitive* control of decision-making.

Now, it turns out that this metacognitive control architecture is close to a drift-diffusion decision model (DDM) with a specific form of collapsing bounds. This can be seen as follows. Let  $b(z)$

be a sigmoidal function of allocated resources, whose inflexion point is the resource budget ( $z^*$ ), and whose magnitude is  $P_c^*$ , i.e.:

$$b(z) = \left( E[P_c(z^*)] - \frac{1}{2} \right) * \text{sigmoid}(\lambda(z^* - z)) + \frac{1}{2} \quad (26)$$

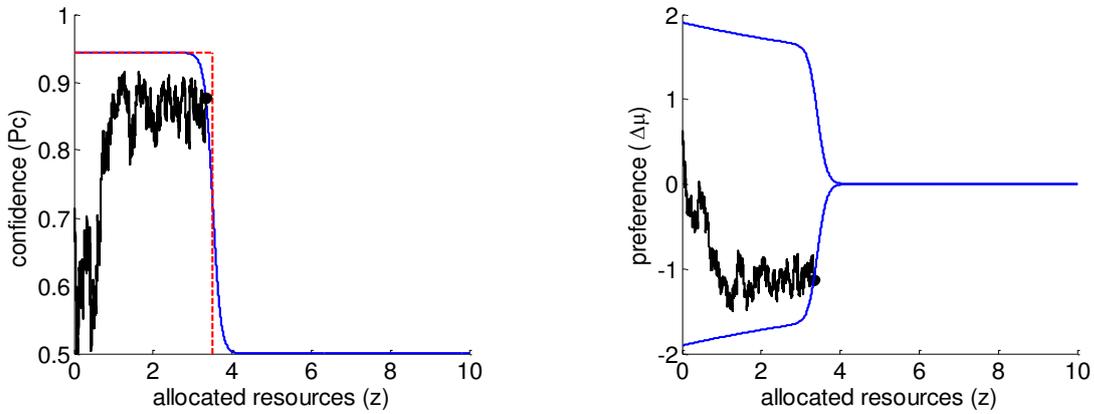
where  $\lambda$  controls the sigmoid's slope. If used as a resource-dependent bound for confidence, one can think of  $b(z)$  as an approximation to the double-threshold scenario, which is recovered exactly at the limit  $\lambda \rightarrow \infty$ . Having said this, one may relax this constraint and consider softer confidence bounding. In this context,  $\lambda$  effectively controls the stiffness of the motivational control of decision-making. Now this type of online control yields outputs identical to the following DDM: the DDM's state variable (i.e., accumulated evidence) would be the instantaneous preference  $\Delta\mu(z)$ , whose stochastic evolution would follow the standard Brownian motion form; a decision would be triggered whenever  $\Delta\mu(z)$  reaches either of the positive/negative collapsing bounds:

$$\Delta_{\pm} = \frac{\pm\sqrt{6} * \log\left(\frac{1}{b(z)} - 1\right)}{\pi\sqrt{\theta_0 + \beta z}} \quad (27)$$

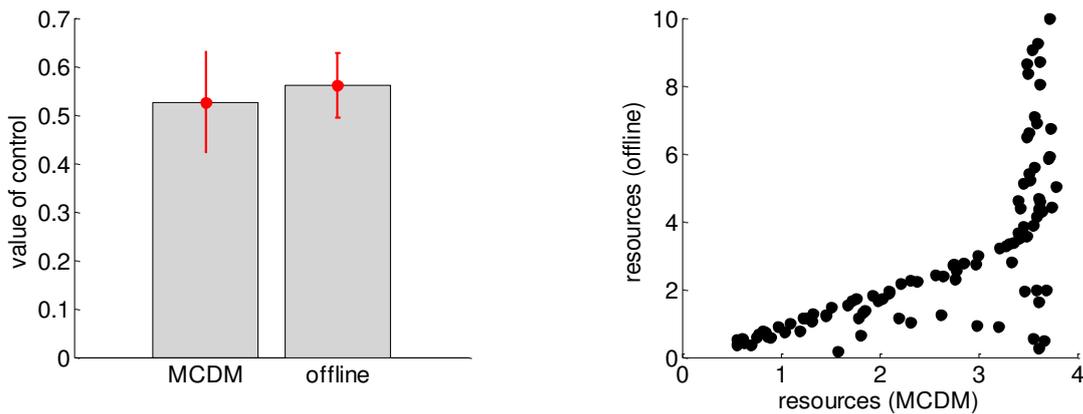
A sample path of both "drifting-to-bound" decision processes (confidence versus preference) is exemplified below (see Figure 17). Note that, in this example, everything is kept identical, including the underlying stochastic perturbation term.

Now one may also ask how well such online metacognitive control of decision-making performs. We can do this as follows. First, we identify the threshold  $\Delta^*$  that should be applied to the system sample paths  $\Delta\mu(z)$  to maximize the expected value of control (where the expectation is taken under Monte-Carlo simulations of the system's sample paths). We call this the "offline" threshold, because it can only be identified after having sampled all paths. Then we compare the distributions of effective value of control (i.e. the discounted value  $RP_c(z) - \alpha z$  of allocating the amount of resources  $z$ ), when one applies the threshold  $\Delta^*$  or the collapsing

boundaries  $\Delta_{\pm}$ . One can also ask whether the amounts of resources that would be allocated are similar. This is exemplified on below (see Figure 18).



**Figure 17: A sample path example of the metacognitive control of decision-making.** Left: confidence (black line, y-axis) is plotted as a function of the amount of allocated resources (x-axis). The soft confidence bound  $b(z)$  is shown in blue (here, we used  $\lambda \approx 20$ ), and the underlying double-threshold in red. Right: instantaneous preference  $\Delta\mu(z)$  (black line, y-axis) is plotted as a function of the amount of allocated resources (x-axis). The ensuing collapsing boundaries  $\Delta_{\pm}$  are shown in blue. Note that their shape depends upon the control stiffness parameter. In particular, when the control stiffness becomes weak (e.g.,  $\lambda < 1$ ), the collapsing boundaries look like exponentially decaying functions of  $z$ .



**Figure 18: Efficiency of the metacognitive control of decision-making.** Left: the sample average value of control ( $\pm 1$  standard deviation) for both MCDM and offline thresholding strategies are shown. Right: the amount resources allocated using MCDM (x-axis) is plotted against the amount of resources allocated using the offline strategy (y-axis), for each sample path of the decision process. This picture generalizes to different settings of effort efficacies, unitary cost and control stiffness.

One can see that both strategies yield very similar outcomes, both in terms of the average value of control and in terms of pseudo trial-by-trial variations in the amount of allocated resources. This is important because it provides construct validity to the metacognitive control of decision making, which essentially yields EVC-optimal resource allocation. Having said this, one may be tempted to conclude that the metacognitive control of decision-making is essentially similar to some basic form of DDM. The critical difference, however, is that such metacognitive control adapts the bounds on a decision-by-decision basis, in order to perform an EVC-optimal resource allocation for each distinct decision. Note that this still allows for inter-individual differences, which would be captured by inter-individual differences in the effort efficacy, effort cost and control stiffness parameters.

## **Chapter 3**

### **Behavioral / Computational Studies**

### **Study 1: Choosing What We Like versus Liking What We Choose**

In order to better demonstrate our model and to validate it with empirical data, we conducted a series of behavioral experiments. With Study 1, summarized in the journal article, “Choosing what we like versus liking what we choose: how choice-induced preference change might actually be instrumental to decision-making,” we show evidence that preferences are constructed at the time of choice, and that these constructed preferences often differ from those constructed in a rating task. Furthermore, we show that preferences constructed during a choice task in proximity to a rating task are more accurate and persistent than those constructed during a rating task alone. Critically, we provide evidence that the more accurate preferences are constructed during the choice task, rather than after it or during the subsequent rating task. We do this by clearly demonstrating the link between preference change observed with the options in each choice and self-reported confidence levels in the same choices. Study 1 thus provides a solid theoretical foundation for us to introduce our model in more detail in Study 2.

**Choosing what we like vs liking what we choose:**

**How choice-induced preference change might actually be instrumental to decision-making.**

Douglas Lee <sup>1,2,3</sup>, Jean Daunizeau\* <sup>1,2,3</sup>

<sup>1</sup> Sorbonne University, Paris, France

<sup>2</sup> Institut du Cerveau et de la Moelle épinière, Paris, France

<sup>3</sup> INSERM UMR S975

Address for correspondence:

Jean Daunizeau

Motivation, Brain and Behavior Group

Brain and Spine Institute

47, boulevard de l'Hôpital, 75013, Paris, France

Tel: +33 1 57 27 47 19

E-mail: [jean.daunizeau@gmail.com](mailto:jean.daunizeau@gmail.com)

Keywords: value, cognitive dissonance, metacognition, choice confidence

## ABSTRACT

For more than 60 years, it has been known that people report higher (lower) subjective values for items after having selected (rejected) them during a choice task. This phenomenon is coined “choice-induced preference change” or CIPC, and its established interpretation is that of “cognitive dissonance” theory. In brief, if people feel uneasy about their choice, they later convince themselves, albeit not always consciously, that the chosen (rejected) item was actually better (worse) than they had originally estimated. While this might make sense from a pragmatic psychological standpoint, it is challenging from a theoretical evolutionary perspective. This is because such a cognitive mechanism might yield irrational biases, whose adaptive fitness would be unclear. In this work, we assume that CIPC is mostly driven by the refinement of option value representations that occurs *during* (and not *after*) difficult choices. This makes CIPC the epiphenomenal outcome of a cognitive process that is instrumental to the decision. Critically, our hypothesis implies novel predictions about how observed CIPC should relate to two specific meta-cognitive processes, namely: choice confidence and subjective certainty regarding pre-choice value judgments. We test these predictions in a behavioral experiment where participants rate the subjective value of food items both before and after choosing between equally valued items; we augment this traditional design with reports of choice confidence and subjective certainty about value judgments. The results confirm our predictions and provide evidence against the standard post-choice cognitive dissonance reduction explanation. We then discuss the relevance of our work in the context of the existing debate regarding the putative cognitive mechanisms underlying cognitive dissonance reduction.

## INTRODUCTION

The causal relationship between choices and subjective values goes both ways. By definition, choices are overt expressions of subjective values, which is the basis of decision theory (Slovic, Fischhoff & Lichtenstein, 1977). However, one's choices also influence one's values, such that actions or items seem to acquire value simply because one has chosen them. Such "choice-induced preference change" (CIPC) has been repeatedly demonstrated via the so-called "free-choice paradigm" (Brehm, 1956). Here, people rate the pleasantness of (e.g., food) items both before and after choosing between pairs of equally pleasant items. Results show that the post-choice pleasantness ratings of chosen (rejected) options are typically higher (lower) than their pre-choice pleasantness ratings, which has traditionally been taken as empirical evidence for the existence of a "cognitive dissonance" reduction mechanism, triggered by difficult choices (see Harmon-Jones & Harmon-Jones, 2007; Izuma & Murayama, 2013 for reviews). An established variant of this interpretation is that people rationalize their choice *ex post facto* as they think along the lines of, "I chose (rejected) this item, so I must have liked it better (worse) than the other one," and hence adjust their internal values accordingly (Bem, 1967). Over the past decade, neuroimaging studies have demonstrated that the act of choosing between similarly-valued options causes changes in the brain's encoding of subjective values (Izuma et al, 2010; Voigt et al, 2018). This has lent neurobiological support to the theory, and cognitive dissonance reduction is now the popular explanation behind a broad variety of important irrational sociopsychological phenomena, ranging from, for example, post-vote political opinion changes (Beasley & Joslyn, 2001) to post-violence hostile attitude worsening (Acharya, Blackwell & Sen, 2015).

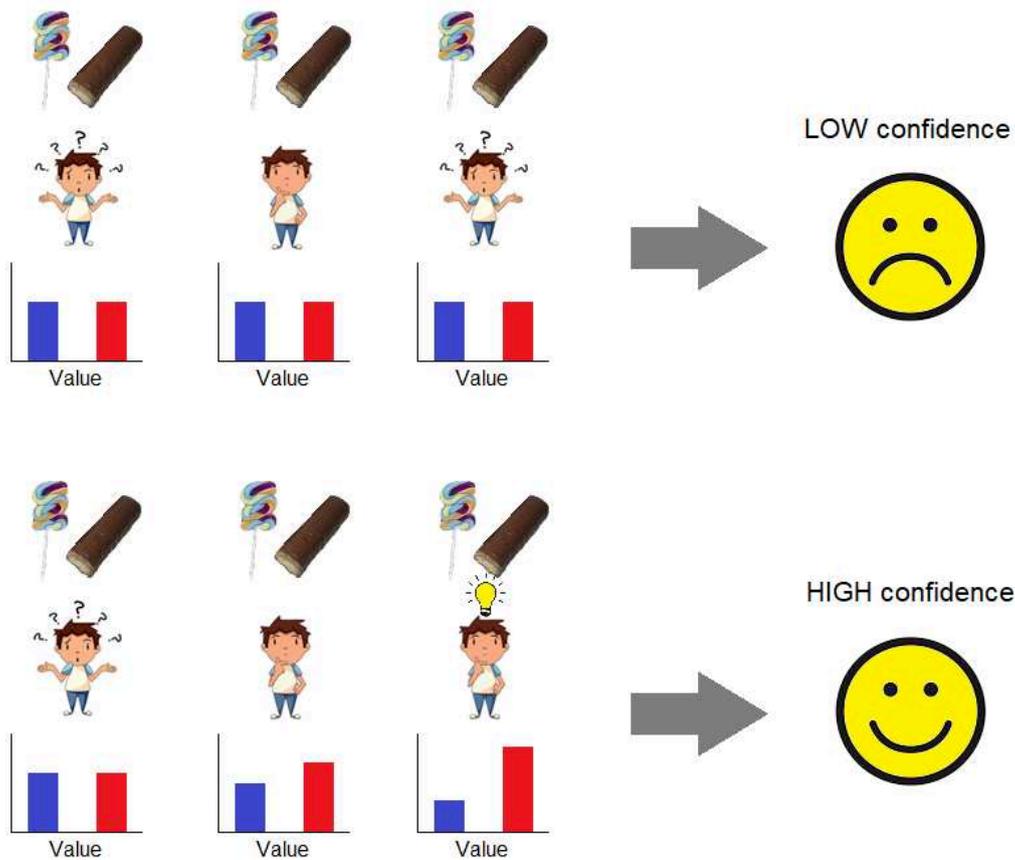
This is not to say, however, that the theory of cognitive dissonance reduction has remained unchallenged. The first issue is theoretical in essence. In brief, it is unclear why evolutionary pressure would have favored post-choice cognitive dissonance reduction mechanisms, given

that they could eventually induce irrational cognitive biases that have no apparent adaptive fitness (Akerlof and Dickens, 1982; Gilad et al., 1987; Perlovsky, 2013). For example, in the context of evidence-based decision making, standard cognitive dissonance theory predicts the appearance of confirmation and overconfidence biases. This is simply because weak beliefs should be reinforced by subsequent choices, despite the lack of any additional piece of evidence (Allahverdyan and Galstyan, 2014; Navajas et al., 2016). At least in principle, cognitive dissonance reduction may of course have other behavioral consequences that would overcompensate for the adverse selective pressure on cognitive biases. Now if it possessed such adaptive fitness, then it would undoubtedly be expressed in many other animal species. This is, however, an unresolved issue in the existing ethological literature (Egan, Santos & Bloom, 2007; West et al, 2010). Second, the main experimental demonstration of cognitive dissonance has also been challenged on statistical grounds. In 2010, Chen and Risen reported a methodological issue in the way CIPC had typically been measured and explained. The basic idea was that simple random variability in repeated value ratings could confound classical measures of CIPC in the context of the free-choice paradigm. The authors provided a detailed mathematical explanation for how such a statistical confound might eventually cause an apparent CIPC (see Chen & Risen, 2010 for details), and introduced a clever control condition. Here, both first and second value ratings are provided before any choice is ever made, thus precluding choice from causally influencing reported subjective values. Results show that significant CIPC occurs regardless of whether the choice is made before or after the second rating. Although this supports the validity of the authors' statistical criticism, subsequent studies also demonstrated that the magnitude of CIPC is significantly greater when the choice is made before the second value rating (Salti et al, 2014; Coppin et al, 2013; Coppin et al, 2012; Sharot et al, 2012). Taken together, the current theoretical and empirical bases of CIPC do not yet provide a straightforward portrait of why and how choice may influence subjective values.

Interestingly, recent neuroimaging evidence suggests that, in the context of typical free-choice paradigms, preference changes occur *during* the decision, not *after* it (Jarcho, Berkman, & Lieberman, 2010; Colosio et al, 2017). Recall that people are reluctant to make a choice that they are not confident about (De Martino et al, 2013). But envisioning competing possibilities during a choice provides a new context that highlights the unique aspects of the alternative options (Tversky & Thaler, 1990; Lichtenstein & Slovic, 2006; Warren, McGraw, & Van Boven, 2011). In turn, the act of choosing may change preferences by reappraising aspects of choice options that may not have been considered thoroughly before (Sharot et al., 2009). In other words, preferences may change during deliberation in order to facilitate the choice process (Simon, Stenstrom & Read, 2015; Simon, Krawczyk & Holyoak, 2004; Simon & Holyoak, 2002). This is important, because it allows for the possibility that preference changes may be instrumental for the process of decision making, which would resolve most theoretical concerns. This is the essence of our working hypothesis. We reason that decision difficulty drives people to reassess the values of the alternative options before committing to a particular choice. The ensuing refinement of internal value representations will eventually raise choice confidence enough to trigger the decision, which may or may not be aligned with pre-choice value ratings. Critically for our theory, the more difficult the decision, the more deliberation and potential reassessment of value representations, the more likely a change of mind and the related CIPC. This would make CIPC the epiphenomenal outcome of a cognitive process that is instrumental to the decision.

Importantly, our working hypothesis makes two original predictions that deviate from standard post-choice cognitive dissonance reduction theory. Recall that the magnitude of CIPC is known to increase with the absolute difference between pre-choice option values, which is typically taken as a proxy for choice difficulty (Izuma and Murayama, 2013; Sharot et al., 2012). We argue, however, that choice difficulty is better defined in terms of the similarity of value

representations. The critical difference is that value representations may be uncertain, i.e. the feeling of liking and/or wanting a given choice option may be imprecise. In other terms, subjective estimates of choice difficulty derive from both value difference and metacognitive judgments about value uncertainty. Our first prediction regards the impact of the latter. In particular, we predict that CIPC should increase with subjective uncertainty regarding pre-choice value judgment. This is because pre-choice value uncertainty raises choice difficulty, which calls for more value reassessment. In contrast, standard post-choice cognitive dissonance reduction theory predicts that post-choice dissonance should be highest when pre-choice values were *a priori* certainly equal, i.e. for choices with minimal value distance and maximal value certainty. Second, we predict that CIPC should positively correlate with choice confidence, controlling for the impact of decision difficulty. This is because, under our hypothesis, CIPC indirectly signals a successful improvement of choice confidence, due to reassessed values spreading apart. In contrast, standard dissonance reduction theory would posit that choices made with low confidence should trigger the strong aversive dissonance feelings that eventually lead to CIPC (see Figure 1). Note that we can rule out variants of Chen and Risen's statistical confound by showing that, if anything, evidence for both predictions is weaker when the choice is made after the second value rating.



**Figure 1:** Top: an example decision under the framework of post-choice cognitive dissonance reduction. Here, the DM is not sure which item to choose, because he estimates their values to be equal. He deliberates about the options, but in the end, he is still unsure, so he is forced to choose randomly. This lack of confidence about the choice causes dissonance, which is then resolved by adjusting his value estimates. Note that under this framework, lower confidence choices lead to higher CIPC (hence a negative correlation). Bottom: the same decision taken under the framework of pre-choice value refinement. Here, while the DM deliberates, he starts to realize that the candy bar is more valuable than the lollipop. Eventually, this value revision reaches the point where he can decisively make a choice, with high confidence. Note that under this framework, higher CIPC leads to higher confidence (hence a positive correlation).

## METHODS

We adapted the experimental design of Chen and Risen (2010), which includes two groups of participants. The so-called RCR (Rating, Choice, Rating) group of participants was asked to rate the value of a series of items both before and after making paired choices. In contrast, the RRC (Rating, Rating, Choice) group of participants rated the items twice before making the paired choices. As we will see, comparisons between the RCR and the RRC (control) group

serve to rule out variants of Chen and Risen's statistical confound. In our adapted experimental design, when evaluating subjective values, participants now also rated their subjective certainty regarding their value judgment. In addition, they also reported how confident they were when making paired choices. This allows us to assess the impact of both subjective uncertainty about value judgments and choice confidence on CIPC.

### ***Participants***

A total of 123 people participated in this study. The RCR group included 65 people (45 female; age: mean=29, stdev=9, min=19, max=53). The RRC group included 58 people (34 female; age: mean=33, stdev=11, min=18, max=55). All participants were native French speakers. Each participant was paid a flat rate of 12€ as compensation for one hour of time.

### ***Materials***

We wrote our experiment in Matlab, using the Psychophysics Toolbox extensions (Brainard, 1997). The experimental stimuli consisted of 108 digital images, each representing a distinct sweet snack item (including cookies, candies, and chocolates). Prior to the experiment, participants received written instructions about the sequence of tasks, including typical visual examples of rating and choice trials.

### ***Experimental Design***

The experiment was divided into three sections, following the classic Free-Choice Paradigm protocol: Rating #1, Choice, Rating #2 (RCR group) or Rating #1, Rating #2, Choice (RRC group). Note that only in the RCR group do Rating #1 and Rating #2 correspond to pre-choice and post-choice ratings. Participants underwent a brief training session prior to the main testing phase of the experiment. There was no time limit for the overall experiment, nor for the different

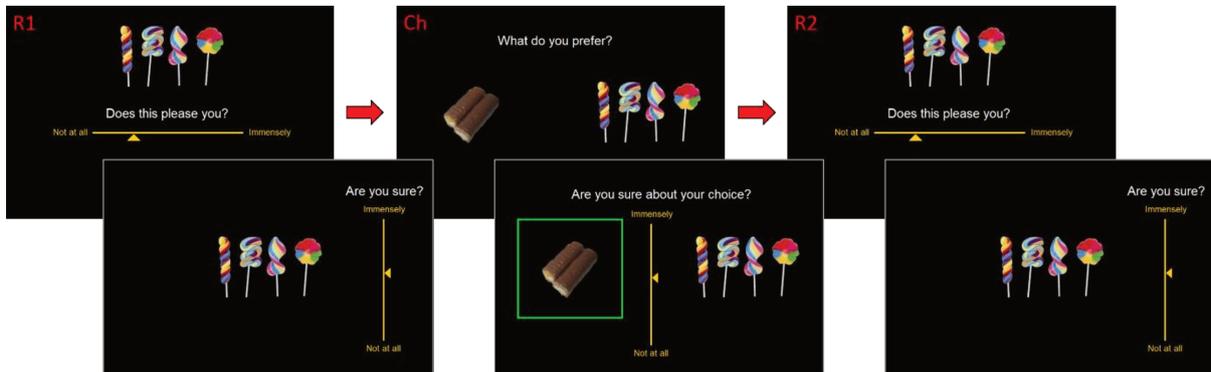
sections, nor for the individual trials. Within-trial event sequences are described below (see Figure 2).

Rating: Participants rated the stimulus items in terms of how much each item pleased them. The entire set of stimuli was presented to each participant, one at a time, in a random sequence (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, “Does this please you?” using a horizontal Likert scale (from “not at all” to “immensely”) to report their subjective valuation of the item. Participants then had to respond to the question, “Are you sure?” using a vertical Likert scale (from “not at all” to “immensely”) to indicate their level of subjective uncertainty regarding the preceding value judgment. At that time, the next trial began.

Choice: Participants chose between pairs of items in terms of which item they preferred. The entire set of stimuli was presented to each participant, one pair at a time, in a random sequence of pairs. Each item appeared in only one pair. The algorithm that created the choice pairs first sorted all items into 10 bins, then paired off (at least) half of the items within each bin, then paired off all remaining items across bins. This ensured that at least half of choices would be between items of similar subjective value (value rating difference  $< 1/10$  of the full rating scale, as shown in previous studies to cause CIPC), but that a substantial portion would be associated with greater value differences.

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, “What do you prefer?” using the left or right arrow key. Participants then had to respond to the question, “Are you sure about your choice?”

using a vertical Likert scale to report their level of confidence in the preceding choice. At that time, the next trial began.

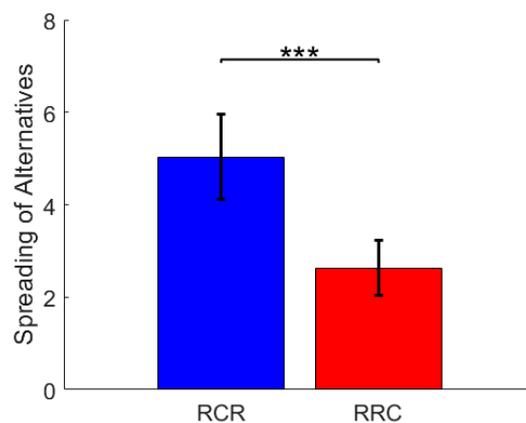


**Figure 2: Experiment Tasks:** participants individually estimated their subjective value of each option, along with their certainty about each estimation; next, they chose their preferred option from pairs of options, and provided their subjective confidence and effort estimates for each choice; finally, participants again estimated their subjective value and certainty for each option.

## RESULTS

Before testing our hypothesis (against both statistical confounds and standard post-choice cognitive dissonance reduction theory), we performed a number of simple data quality checks. First, we assessed the test-retest reliability of both value judgments and their associated certainty reports. For each participant, we thus measured the correlation between ratings #1 and #2 (across items). We found that both ratings were significantly reproducible (value ratings: correlation = 0.862, 95% CI [0.838, 0.886],  $p < 0.001$ ; certainty ratings: correlation = 0.472, 95% CI [0.409, 0.535],  $p < 0.001$ ). Second, we asked whether choices were consistent with value ratings #1. For each participant, we thus performed a logistic regression of paired choices against the difference in value ratings. We found that the balanced prediction accuracy was beyond chance level (group mean = 0.685, 95% CI [0.666, 0.703],  $p < 0.001$ ). Third, we checked that choice confidence increases both with the absolute value difference between paired items,

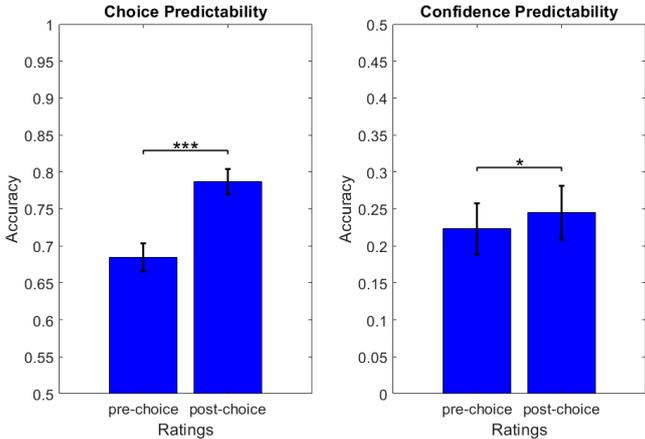
and with the mean certainty reports about value judgments (of the paired items). For each participant, we thus performed a multiple linear regression of choice confidence against absolute value difference and mean judgment certainty (ratings #1). A random effect analysis shows that both have a significant effect at the group level ( $R^2 = 0.223$ , 95% CI [0.188, 0.258]; absolute value difference: GLM beta = 9.046, 95% CI [7.841,  $\infty$ ],  $p < 0.001$ ; mean judgment certainty: GLM beta = 3.157, 95% CI [2.137,  $\infty$ ],  $p < 0.001$ ). Fourth, we asked whether we could reproduce previous findings that CIPC is higher in the RCR group than in the RRC group. For each participant, we thus measured the magnitude of CIPC in terms of the so-called "spreading of alternatives" (SoA), calculated as the mean difference in value rating gains between chosen and unchosen items ( $\text{SoA} = [\text{rating}\#2 - \text{rating}\#1]_{\text{chosen}} - [\text{rating}\#2 - \text{rating}\#1]_{\text{unchosen}}$ ). As expected, we found that SoA is significant in both groups (RCR group: SoA = 5.033, 95% CI [4.118, 5.949],  $p < 0.001$ ; RRC group: SoA = 2.635, 95% CI [2.047, 3.224],  $p < 0.001$ ). In addition, SoA is significantly higher in the RCR group than in the RRC group (SoA difference = 2.398,  $p < 0.001$ ) (Figure 3).



**Figure 3:** Mean and 95% CI for spreading of alternatives in the RCR and RRC groups

In what follows, and unless stated otherwise, we focus on the RCR group of participants. Recall that, under our hypothesis, the deliberation that takes place during the decision process is expected to cause a refinement of internal value representations until a target level of choice confidence is met and the decider commits to a choice. To begin with, we thus asked whether certainty about value judgments improved after the choice had been made. For each participant, we thus estimated the mean difference between post-choice and pre-choice certainty reports (across all items). A random effect analysis then shows that post-choice certainty reports are significantly higher than pre-choice certainty reports (certainty increase = 3.781, 95% CI [1.810, 5.752],  $p < 0.001$ ). This finding supports our claim but does not provide evidence for or against classical post-choice cognitive dissonance reduction theory. We then asked whether post-choice ratings better predict choice (and choice confidence) than pre-choice ratings. First, we performed another logistic regression of paired choices, this time against the difference in post-choice value ratings (ratings #2). The ensuing choice prediction accuracy is higher than with pre-choice value ratings (accuracy = 0.787, 95% CI [0.770, 0.804], accuracy gain = 0.103, 95% CI [0.082, 0.124],  $p < 0.001$ ) (Figure 4). Second, we regressed choice confidence, this time against post-choice absolute value difference and mean judgment certainty. The ensuing amount of explained variance in choice confidence reports is higher than with pre-choice ratings ( $R^2 = 0.245$ , 95% CI [0.209, 0.281],  $R^2$  gain = 0.022, 95% CI [0.001, 0.042],  $p = 0.02$ ) (Figure 4). When testing for the significance of differences in pre-choice and post-choice regression parameters, we found that this gain in explanatory power is more likely to be due to value ratings (GLM beta difference = 0.763, 95% CI [0.180,  $\infty$ ],  $p = 0.018$ ) than to certainty reports (GLM beta difference = 0.158, 95% CI [-0.451,  $\infty$ ],  $p = 0.34$ ). These results are important, because they validate basic requirements of our pre-choice CIPC hypothesis. However, they are equally compatible with both pre-choice and post-choice CIPC mechanisms. Next, we focus on testing the two specific predictions regarding the relationship between CIPC and meta-

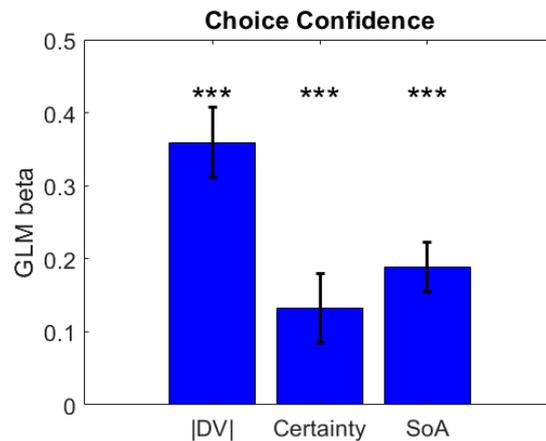
cognitive processes, which discriminate mechanisms of pre-choice value refinements from post-choice cognitive dissonance reduction.



**Figure 4:** Mean and 95% CI for choice predictability (left) and choice confidence predictability (right), based on pre- vs post-choice ratings

We now assess the statistical relationships between CIPC and both choice difficulty and choice confidence. For each participant, we performed a multiple linear regression of choice confidence onto absolute difference in pre-choice value ratings, mean pre-choice judgment certainty reports, and spreading of alternatives (Figure 5). The idea is that if CIPC does indeed occur during the choice (as opposed to after), it can effectively be re-labeled “gain in ease” and as such would play the same role as pre-choice ease (value difference/certainty). As expected, a random effect analysis on the ensuing parameter estimates shows that confidence significantly increases with the absolute difference in pre-choice value ratings (GLM beta = 0.359, 95% CI [0.311, 0.407],  $p < 0.001$ ) and with pre-choice judgment certainty (GLM beta = 0.133, 95% CI [0.085, 0.180],  $p < 0.001$ ) More importantly, we found that confidence also increases with spreading of alternatives (GLM beta = 0.189, 95% CI [0.154, 0.227],  $p < 0.001$ ). The latter

findings support our hypothesis, and are incompatible with classical post-choice cognitive dissonance reduction theory (cf. Figure 1 above).



**Figure 5:** Mean and 95% CI for GLM beta weights of absolute value difference, average value certainty (within a choice pair), spreading of alternatives on choice confidence

Finally, we aimed at ruling out statistical confounds. This can be done by showing that if the above statistical relationships exist in the RRC group, they should be significantly weaker than in the RCR group. We thus performed the above analyses on data acquired in participants from the RRC group, which we compared to the RCR group of participants using standard random effect analyses. First, the gain in choice prediction accuracy (from rating #1 to rating #2) is significantly higher in the RCR group than in the RRC group (accuracy gain difference = 0.0362,  $p=0.008$ ). Second, and most importantly, both the impact of absolute pre-choice value difference (GLM beta diff = -0.095,  $p=0.0039$ ) and choice confidence (GLM beta diff = 0.044,  $p=0.049$ ) on SoA are significantly higher in the RCR than in the RRC group. Note that some comparisons between the two groups turned out not to be significant (gain in subjective certainty regarding value judgments:  $p=0.17$ , gain in confidence prediction accuracy:  $p=0.58$ ,

impact of subjective value certainty on SoA:  $p=0.096$ ). Nevertheless, taken together, these findings are unlikely under a chance model of random variations in value ratings.

## **DISCUSSION**

In this work, we have presented empirical evidence that challenges standard interpretations of CIPC, in particular: post-choice cognitive dissonance reduction theory (and its self-perception variants). However, we would like to highlight that we do not dispute the concept of cognitive dissonance itself, nor even the idea that cognitive dissonance is the root cause of CIPC. In fact, our view only differs from standard post-choice cognitive dissonance reduction theory in one single aspect, namely: the temporal dynamics of the dissonance reduction. According to standard post-choice cognitive dissonance reduction theory, choices made with low confidence trigger strong aversive dissonance feelings that are resolved by retrospectively matching internal value representations to the choice. We would rather say that no choice commitment is made until internal value representation refinements allow choice confidence to reach a satisfying (non-aversive) level. This idea is a simple extension of the so-called Action-based model of cognitive dissonance (Harmon-Jones, Harmon-Jones, & Levy, 2015; Harmon-Jones, Amodio, & Harmon-Jones, 2009), and of recent attempts to formalize cognitive dissonance in terms of neurocognitive theories of predictive processing in the action-perception loop (Kaaronen, 2018). That cognitive dissonance reduction occurs during the decision process (and not after) is critical, however, because it is now endowed with a clear functional purpose (namely: improving decision accuracy).

Note that one may challenge our interpretation of the observed relationship between confidence and CIPC, in terms of evidence against post-choice CIPC. For example, one may argue that CIPC might occur after people commit to a choice, but before they get a feel for how confident

they are about that choice. This sounds paradoxical however, in the sense that experiencing cognitive dissonance in this context simply means feeling uneasy about one's choice, i.e. lacking confidence about it. In any case, this line of reasoning cannot apply to the observed impact of value certainty on CIPC. Recall that we probe metacognitive judgments about value certainty before the choice, using rating scales at the time when each item is presented (immediately after first-order value judgments). Therefore, the relationship between CIPC and value certainty that we disclose empirically cannot derive from metacognitive processes that occur after the choice has been made.

Our results apparently contradict the recent finding that CIPC only occurs for choices where the agent is later able to recall which option was chosen and which was rejected (Salti et al, 2014). This is because remembering choices has no causal role under our pre-choice value refinement hypothesis. This contrasts with standard cognitive dissonance reduction theory, where post-choice option re-evaluation requires the memory trace of relevant choices. This interpretation is compatible with the observation that, when re-evaluating items after the choice, activity in the hippocampus discriminates between remembered and non-remembered choices (Chammat et al, 2017). These two studies thus provide apparent evidence that CIPC occurs after the choice has been made. However, we contend that this theory remains unsupported until empirical evidence is found for memory traces of information that is critical for post-choice CIPC (namely: whether an option was chosen or rejected, what was the option's pre-choice value, and which option was the alternative during the relevant choice). In addition, the relationship between CIPC and memory might be confounded by choice difficulty. In brief, the more difficult a decision is, the more value reassessment it will eventually trigger, the more likely the agent is to remember his choice. Alternatively, post-choice reports of internal values may rely on slightly unstable episodic memory traces of intra-choice CIPC. The latter scenario is actually compatible with the fact that activity in the left dorsolateral prefrontal cortex (during

choice) predicts the magnitude of CIPC only when the choices are remembered (Voigt et al, 2018), and also with the intra-choice CIPC interpretation of the causal impact of post-choice activity perturbations (see below). In any case, either or both of these scenarios would explain why intra-choice CIPC might exhibit an apparent (non-causal) relationship with choice memory. Finally, note that the causal implication of memory is inconsistent with the assessment of amnesic patients, who exhibit normal CIPC despite severe deficits in choice memory (Liebermann et al, 2001).

Even though the cognitive architecture that underlies intra-choice cognitive dissonance reduction is yet to be disclosed, recent neuroimaging findings shed light on the question of whether CIPC occurs during or after the decision. On the one hand, a few brain stimulation studies suggest that perturbing brain activity *after the choice* (in particular: in the left dorsolateral and/or posterior medial frontal cortices) disrupts the observed CIPC (Mengarelli et al, 2013; Izuma et al, 2015). Although compatible with post-choice CIPC, such causal effects can be due to the post-choice disturbance of value representations that resulted from intra-choice CIPC. On the other hand, many recent studies show that brain activity measured *during the choice process* is predictive of the magnitude of CIPC (Colosio et al, 2017; Jarcho, Berkman, & Lieberman, 2010; Kitayama et al, 2013; van Veen et al, 2009, Voigt et al, 2018). Unsurprisingly, 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). Note that current neurocomputational theories of ACC suggest that it is involved in controlling how much mental effort should be allocated to a given task, based upon the derivation of the so-called "expected value of control" (Musslick et al., 2015; Shenhav et al., 2013). This is highly compatible with our results, under the assumption that pre-choice value reassessment is a controlled and effortful process that trades mental effort against choice confidence. We will pursue this computational scenario in subsequent publications.

Nevertheless, we consent that CIPC may be driven by both pre-choice value reassessment and post-choice cognitive dissonance reduction mechanisms. The quantitative contribution of the latter effect, however, may have been strongly overestimated. In our view, this is best demonstrated, though perhaps unintentionally, in the results of the “blind choice” study from Sharot and colleagues (Sharot et al., 2010). Here, participants rated items both before and after making a blind choice that could not be guided by pre-existing preferences (because the items were masked). Critically, although blind choice precludes any instrumental value refinement process, preferences were altered after the choice. Interestingly, the effect size is rather small, i.e. the ensuing CIPC magnitude was estimated to be around  $0.07 \pm 0.03$ . This is to be compared with the CIPC magnitude of the RCR and RRC conditions in two other studies by the same authors (Sharot et al., 2009, 2012), namely:  $0.38 \pm 0.08$  (RCR condition, with non-blind choices) and  $0.11 \pm 0.06$  (RRC condition, which was not included in the “blind choice” study). In other terms, CIPC under blind choice is smaller than the apparent CIPC that unfolds from the known statistical confounds of the free choice paradigm. Note that if this had not have been the case, then post-choice CIPC cognitive reduction effects would dominate and we would not have confirmed our predictions. Taken together, this series of studies clearly supports the notion that option re-evaluation occurs, at least partly, at the time of the decision.

## **CONCLUSION**

In conclusion, our results lend support to the hypothesis that choice-induced preference change is caused by an intra-choice refinement of option value representations that is motivated by difficult decisions, and undermine the established theory of post-choice cognitive dissonance resolution. We also demonstrate the relevance of meta-cognitive processes (cf. reports of choice confidence and certainty about value judgments) to choice-induced preference change. This contributes to moving forward the state of the 60-year-old research on the reciprocal influence between choice and subjective value.

## REFERENCES

- Acharya, A., Blackwell, M., and Sen, M. (2015). Explaining Attitudes from Behavior: A Cognitive Dissonance Approach (Harvard University, John F. Kennedy School of Government).
- Akerlof, G. A., & Dickens, W. T. (1982). The economic consequences of cognitive dissonance. *The American economic review*, 72(3), 307-319.
- Allahverdyan, A. E., & Galstyan, A. (2014). Opinion dynamics with confirmation bias. *PloS one*, 9(7), e99557.
- Beasley, Joslyn (2001). Cognitive Dissonance and Post-Decision Attitude Change in Six Presidential Elections. *Political Psychology*. 22 (3): 521–540.
- Bem, D. J. (1967). Self-Perception: An Alternative Interpretation of Cognitive Dissonance Phenomena. *Psychological Review*, 74, 183-200.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10, 433-436.
- Brehm, J. W. (1956). Postdecision changes in the desirability of alternatives. *The Journal of Abnormal and Social Psychology*, 52(3), 384.
- Chammat, M., El Karoui, I., Allali, S., Hagège, J., Lehongre, K., Hasboun, D., ... & Navarro, V. (2017). Cognitive dissonance resolution depends on episodic memory. *Scientific reports*, 7, 41320.
- Chen, M. K., & Risen, J. L. (2010). How choice affects and reflects preferences: revisiting the free-choice paradigm. *Journal of personality and social psychology*, 99(4), 573.
- Colosio, M., Shestakova, A., Nikulin, V. V., Blagovechtchenski, E., & Klucharev, V. (2017). Neural mechanisms of cognitive dissonance (revised): An EEG study. *Journal of Neuroscience*, 37(20), 5074-5083.
- Coppin, G., Delplanque, S., Bernard, C., Cekic, S., Porcherot, C., Cayeux, I., & Sander, D. (2014). Choice both affects and reflects preferences. *The Quarterly Journal of Experimental Psychology*, 67(7), 1415-1427.
- Coppin, G., Delplanque, S., Porcherot, C., Cayeux, I., & Sander, D. (2012). When flexibility is stable: implicit long-term shaping of olfactory preferences. *PloS one*, 7(6), e37857.
- De Martino, B., Fleming, S.M., Garrett, N., and Dolan, R. (2013). Confidence in value-based choice. *Nat Neurosci* 16, 105–110.
- Egan, L. C., Bloom, P., & Santos, L. R. (2010). Choice-induced preferences in the absence of choice: Evidence from a blind two choice paradigm with young children and capuchin monkeys. *Journal of Experimental Social Psychology*, 46(1), 204-207.
- Egan, L.C., Santos, L.R., and Bloom, P. (2007). The origins of cognitive dissonance evidence from children and monkeys. *Psychological Science* 18, 978–983.
- Gilad, B., Kaish, S., & Loeb, P. D. (1987). Cognitive dissonance and utility maximization: A general framework. *Journal of Economic Behavior & Organization*, 8(1), 61-73.

- Harmon-Jones, E., Harmon-Jones, C., & Levy, N. (2015). An action-based model of cognitive-dissonance processes. *Current Directions in Psychological Science*, 24(3), 184-189.
- Harmon-Jones, E., Amodio, D. M., & Harmon-Jones, C. (2009). Action-based model of dissonance: A review, integration, and expansion of conceptions of cognitive conflict. *Advances in experimental social psychology*, 41, 119-166.
- Harmon-Jones, E., Harmon-Jones, C., Fearn, M., Sigelman, J. D., & Johnson, P. (2008). Left frontal cortical activation and spreading of alternatives: tests of the action-based model of dissonance. *Journal of personality and social psychology*, 94(1), 1.
- Harmon-Jones, E., & Harmon-Jones, C. (2007). Cognitive dissonance theory after 50 years of development. *Zeitschrift für Sozialpsychologie*, 38(1), 7-16.
- Izuma, K., Akula, S., Murayama, K., Wu, D. A., Iacoboni, M., & Adolphs, R. (2015). A causal role for posterior medial frontal cortex in choice-induced preference change. *Journal of Neuroscience*, 35(8), 3598-3606.
- Izuma, K., & Murayama, K. (2013). Choice-induced preference change in the free-choice paradigm: a critical methodological review. *Frontiers in psychology*, 4, Art-No.
- Izuma, K., Matsumoto, M., Murayama, K., Samejima, K., Sadato, N., & Matsumoto, K. (2010). Neural correlates of cognitive dissonance and choice-induced preference change. *Proceedings of the National Academy of Sciences*, 107(51), 22014-22019.
- Jarcho, J. M., Berkman, E. T., & Lieberman, M. D. (2010). The neural basis of rationalization: cognitive dissonance reduction during decision-making. *Social cognitive and affective neuroscience*, 6(4), 460-467.
- Kaaronen, R.O. (2018). A Theory of Predictive Dissonance: Predictive Processing Presents a New Take on Cognitive Dissonance. *Front. Psychol.* 9.
- Kitayama, S., Chua, H. F., Tompson, S., & Han, S. (2013). Neural mechanisms of dissonance: An fMRI investigation of choice justification. *Neuroimage*, 69, 206-212.
- Lichtenstein, S., & Slovic, P. (Eds.). (2006). *The construction of preference*. Cambridge University Press.
- Lieberman, M. D., Ochsner, K. N., Gilbert, D. T., & Schacter, D. L. (2001). Do amnesics exhibit cognitive dissonance reduction? The role of explicit memory and attention in attitude change. *Psychological Science*, 12(2), 135-140.
- Mengarelli, F., Spoglianti, S., Avenanti, A., & Di Pellegrino, G. (2013). Cathodal tDCS over the left prefrontal cortex diminishes choice-induced preference change. *Cerebral Cortex*, 25(5), 1219-1227.
- Musslick, S., Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2015). A computational model of control allocation based on the Expected Value of Control. In *The 2nd multidisciplinary conference on reinforcement learning and decision making*.
- Navajas, J., Bahrami, B., & Latham, P. E. (2016). Post-decisional accounts of biases in confidence. *Current Opinion in Behavioral Sciences*, 11, 55-60.
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge University Press.

- Payne, J. W., Bettman, J. R., Schkade, D. A., Schwarz, N., & Gregory, R. (1999). Measuring constructed preferences: Towards a building code. In *Elicitation of preferences* (pp. 243-275). Springer, Dordrecht.
- Perlovsky, L. (2013). A challenge to human evolution—cognitive dissonance. *Frontiers in Psychology*, 4, 179.
- Risen, J. L., & Chen, M. K. (2010). How to Study Choice-Induced Attitude Change: Strategies for Fixing the Free-Choice Paradigm. *Social and Personality Psychology Compass*, 4(12), 1151-1164.
- Salti, M., El Karoui, I., Maillet, M., & Naccache, L. (2014). Cognitive dissonance resolution is related to episodic memory. *PloS one*, 9(9), e108579.
- Sharot, T., Fleming, S. M., Yu, X., Koster, R., & Dolan, R. J. (2012). Is choice-induced preference change long lasting?. *Psychological science*, 23(10), 1123-1129.
- Sharot, T., Velasquez, C. M., & Dolan, R. J. (2010). Do decisions shape preference? Evidence from blind choice. *Psychological science*, 21(9), 1231-1235.
- Sharot, T., De Martino, B., & Dolan, R. J. (2009). How choice reveals and shapes expected hedonic outcome. *Journal of Neuroscience*, 29(12), 3760-3765.
- Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron*, 79(2), 217-240.
- Simon, D., Stenstrom, D. M., & Read, S. J. (2015). The coherence effect: Blending cold and hot cognitions. *Journal of Personality and Social Psychology*, 109(3), 369.
- Simon, D., Krawczyk, D. C., & Holyoak, K. J. (2004). Construction of preferences by constraint satisfaction. *Psychological Science*, 15(5), 331-336.
- Simon, D., & Holyoak, K. J. (2002). Structural dynamics of cognition: From consistency theories to constraint satisfaction. *Personality and social psychology review*, 6(4), 283-294.
- Slovic, P., Fischhoff, B., & Lichtenstein, S. (1977). Behavioral decision theory. *Annual review of psychology*, 28(1), 1-39.
- Tversky, A., & Thaler, R. H. (1990). Anomalies: preference reversals. *Journal of Economic Perspectives*, 4(2), 201-211.
- Van Veen, V., Krug, M. K., Schooler, J. W., & Carter, C. S. (2009). Neural activity predicts attitude change in cognitive dissonance. *Nature neuroscience*, 12(11), 1469.
- Voigt, K., Murawski, C., Speer, S., & Bode, S. (2018). Hard decisions shape the neural coding of preferences. *bioRxiv*, 298406.
- Warren, C., McGraw, A. P., & Van Boven, L. (2011). Values and preferences: Defining preference construction. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2(2), 193-205.
- West, S., Jett, S.E., Beckman, T., and Vonk, J. (2010). The phylogenetic roots of cognitive dissonance. *J Comp Psychol* 124, 425–432.

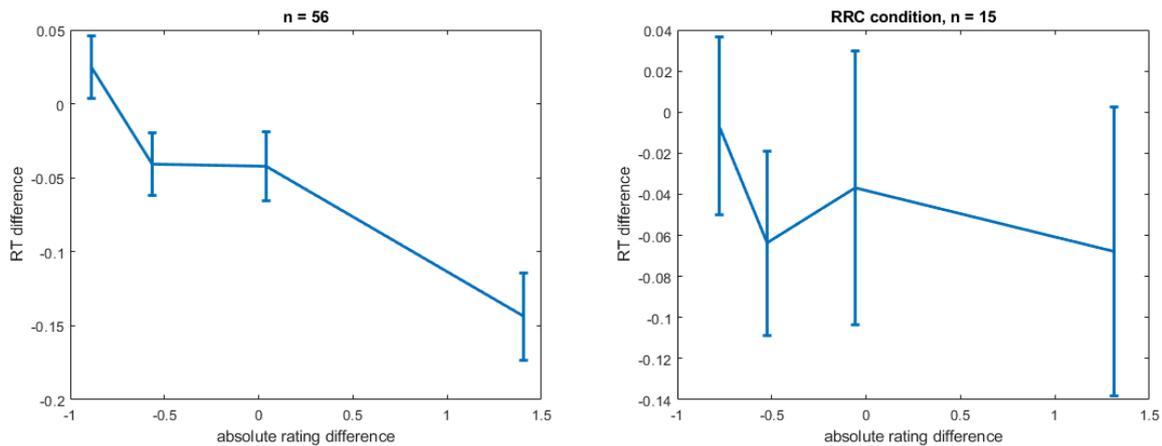
## Study 1 – Appendix

### *Negative Spreading of Alternatives*

Beyond the results described in our journal article, there were also a few other findings that could be of interest to our work. One such finding is related to the trial-by-trial variability of the choice-induced preference change. Our data show that choices yielded positive SoA *on average*, but that individual choices often yielded negative SoA. This is entirely consistent with our framework, which holds that information considered during deliberation could be either congruent or incongruent with previous beliefs, meaning that deliberating could sometimes cause the options to be *less* distinguishable (i.e., negative SoA). It is also consistent with the statistical artifact framework, under which subsequent ratings change randomly in either direction. However, this data is not consistent with the cognitive dissonance framework, where negative SoA should never occur, because that would potentially entail the *creation* of dissonance rather than its resolution.

### *Pre- vs Post-Choice Rating RT*

We examined the rating data to look for the potential impact of choice deliberation on post-versus pre-choice ratings. We found that response times were shorter for post-choice ratings versus pre-choice ratings (mean difference=0.984s, s.e.m.=0.137,  $p < 0.001$ ). Furthermore, this decrease in RT strongly negatively correlated with magnitude of change from pre- to post-choice ratings (beta=-0.079,  $p < 0.001$ ). This correlation was not present in the control condition, where choice deliberation occurred only after both rating tasks ( $p=0.931$ ).



### Rating change magnitude vs rating speed increase

This data provides one final indication that a DM might re-evaluate his subjective valuation of choice options during deliberation. We interpret the correlation between magnitude of the change in rating and decrease in response time to mean that at some point between the pre- and post-choice ratings (presumably during the choice task), the DM refined his evaluations of some of the items (i.e., those in difficult choice pairs). This refinement simultaneously resulted in changes in value estimates and increases in certainty about those estimates. The items with greater estimate changes would have undergone greater contemplation (on average), thus leading to greater certainty. We believe that the reduction in RT between the two rating tasks is indirect evidence of this increase in certainty. This would explain the observation that post-choice ratings were faster for all items, as well as the negative correlation between rating change and RT change. The competitor frameworks would not predict such relationships. Cognitive dissonance would predict the opposite pattern. If a DM were to re-evaluate items only at the time when asked to provide post-choice ratings, this should cause an *increase* in rating RT and a *positive* correlation between rating change and RT

### *Certainty versus Precision*

Our theory fundamentally relies on the concept of precision, for which we intended to use certainty as a proxy. However, it turns out that the mapping from certainty to precision is not

as clear or straightforward as we had hoped it would be. This first became obvious when we observed that even though pre- to post-choice rating certainty change had a positive mean, approximately half of items showed a *decrease* in certainty. In order for our theory to be more complete, we would need to decipher the mapping between certainty and precision. It is possible that participants interpreted our certainty variable as something more akin to *information congruency*. This would result in a similar sense of rating confidence, as our certainty variable was meant to capture, but would allow for both gains and losses of certainty with the processing of additional information. It could be interesting to try to disentangle these two different but closely related variables in future studies.

### *Memory and CIPC*

Recall that it has been shown that choice-induced preference change is stronger when the choices are explicitly remembered [Salti et al, 2014; Chammat et al, 2017]. This is taken as support for the claim that preferences are revised after the choice has already been made, through some later process. We believe, however, that we can provide an equally plausible and more ecological explanation. Under our framework, choices that are made with relatively low effort are those for whom the best option is apparent based on initial value estimates. Such easy choices will not be salient and will therefore be less memorable than difficult choices. On the contrary, difficult choices will call for attention as more information is processed and more accurate and precise value representations are formed. This increased attention and information processing will in turn lead both to the formation of more robust memories of the choice, and to stronger preference changes. Hence, the connection between memory and preference change could be due to a common cause, with no real link between the two.

## **Study 2: Trading Effort for Confidence**

With our Study 2, summarized in the journal article titled, “Trading effort for confidence: the motivational control of decision making,” we build on the theoretical foundation and empirical support that we provided in Study 1. Here we provide a full summary of our computational model, along with all relevant mathematical details. We move beyond a description of preference formation and its relationship with choice confidence, now introducing the concept of mental effort as the force that empowers the preference formation (which leads to confidence). In Study 2, we first validate our own model and then show that it accounts for the data in ways that the current leading models in the field are not capable of doing.

**Trading Mental Effort for Confidence:**

**The Metacognitive Control of Value-Based Decision-Making**

Douglas Lee<sup>1,2</sup>, Jean Daunizeau<sup>1,3</sup>

<sup>1</sup> Brain and Spine Institute (ICM), Paris, France

<sup>2</sup> Sorbonne University, Paris, France

<sup>3</sup> 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](mailto:jean.daunizeau@gmail.com)

Keywords:

Word count: 8684

## **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 the amount of cognitive resources that is deployed during a decision is controlled by an effort-confidence tradeoff. Importantly, the anticipated benefit of allocating resources varies in a decision-by-decision manner according to decision difficulty and importance. The ensuing MCD model predicts choices, decision time, subjective feeling of effort, choice confidence, and choice-induced preference change. As we will see, these predictions are critically different from accumulation-to-bound models of value-based decisions. We compare and test these predictions in a systematic manner, using a dedicated behavioral paradigm. Our results provides a mechanistic link between mental effort, choice confidence, and preference reversals, which suggests alternative interpretations of existing related neuroimaging findings.

## 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. This is why they are sometimes viewed as the outcome of inherent neurocognitive constraints on the brain's decision processes, e.g., limited attentional and/or mnemonic capacity (Giguère and Love, 2013; Lim et al., 2011; Marois and Ivanoff, 2005) or unreliable neural representations of decision-relevant information (Drugowitsch et al., 2016; Wang and Busemeyer, 2016; Wyart and Koechlin, 2016). 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;

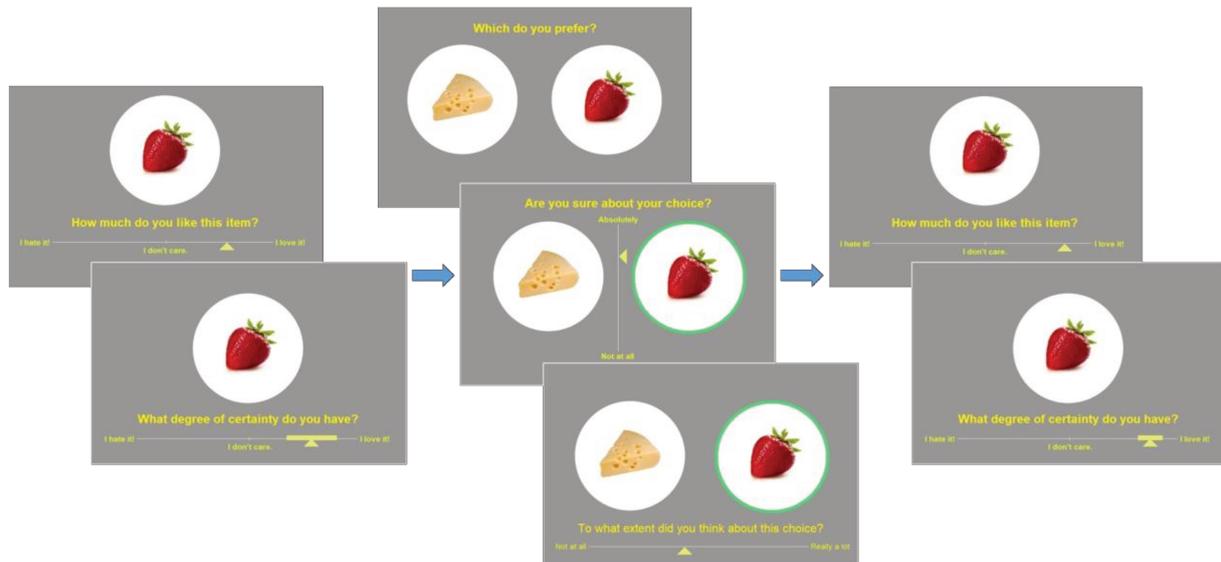
Ratcliff and McKoon, 2008; Ratcliff et al., 2016). Further computational work demonstrated that some variants of these models actually implement an optimal solution to speed-accuracy tradeoff problems (Ditterich, 2006; Drugowitsch et al., 2012). From a theoretical standpoint, this implies that accumulation-to-bound 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 proved 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., 2012) or default preferences (Lopez-Persem et al., 2016), lent empirical support to this type of interpretation. Further credit also came 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, despite the widespread use of these models in decision neuroscience, no evidence of a trial-by-trial accumulation signal has ever been observed in neural recordings in brain systems supporting value-based decisions. In fact, contradictory empirical evidence has even been recently reported in the context of perceptual decisions (Latimer et al.,

2015, 2017). In addition, accumulation-to-bound models neglect the possibility that people may reassess the value of alternative options during decisions (Slovic, 1995; Tversky and Thaler, 1990; Warren et al., 2011). For example, contemplating competing possibilities during a choice may highlight features of alternative options that may not have been considered thoroughly before (Sharot et al., 2010). Under this view, apparent preference reversals are not errors: they are deliberate changes of mind. Lastly, accumulation-to-bound models may make nonsensical predictions, in particular with respect to confidence (Lebreton et al., 2015). As we will show below, existing variants of these models that care about choice confidence (De Martino et al., 2013; Tajima et al., 2016) predict that choice confidence should decrease when the reliability of value signals increases! Here, we propose an alternative computational model of value-based decision-making that resolves most of these concerns.

We start with the premise that 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 reassess option values until they reach a satisfactory level of confidence about their preference. Such effortful mental deliberation engages 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. 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. As we will see, the MCD model makes clear quantitative predictions that differ from accumulation-to-bound models.

We test these predictions by asking participants to choose between pairs of food items, both before and after having reported their judgment about each item's value and their subjective certainty about value judgements. Note that we also measure choice confidence, decision time, and subjective effort for each decision (cf. Figure 1 below).



**Figure 1. 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).

The MCD model predicts choice, decision time, subjective effort, choice confidence, probability of changing one's mind, and choice-induced preference change on a decision-by-decision basis, 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 what follows, we present our main dual computational/behavioral results.

## RESULTS

First, we compare the MCD model to two established models of value-based decision making, namely: an optimal drift-diffusion model with collapsing bounds (Tajima et al., 2016) and a modified race model (De Martino et al., 2013). These two models use variants of the accumulation-to-bound principle, and they can make quantitative predictions regarding the impact of pre-choice value and value certainty ratings (cf. Supplementary Materials). Second, we test a few specific novel predictions that the MCD model makes and that have no analog under alternative frameworks. 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 the Supplementary Materials.

- **Comparing models of decision time and choice confidence**

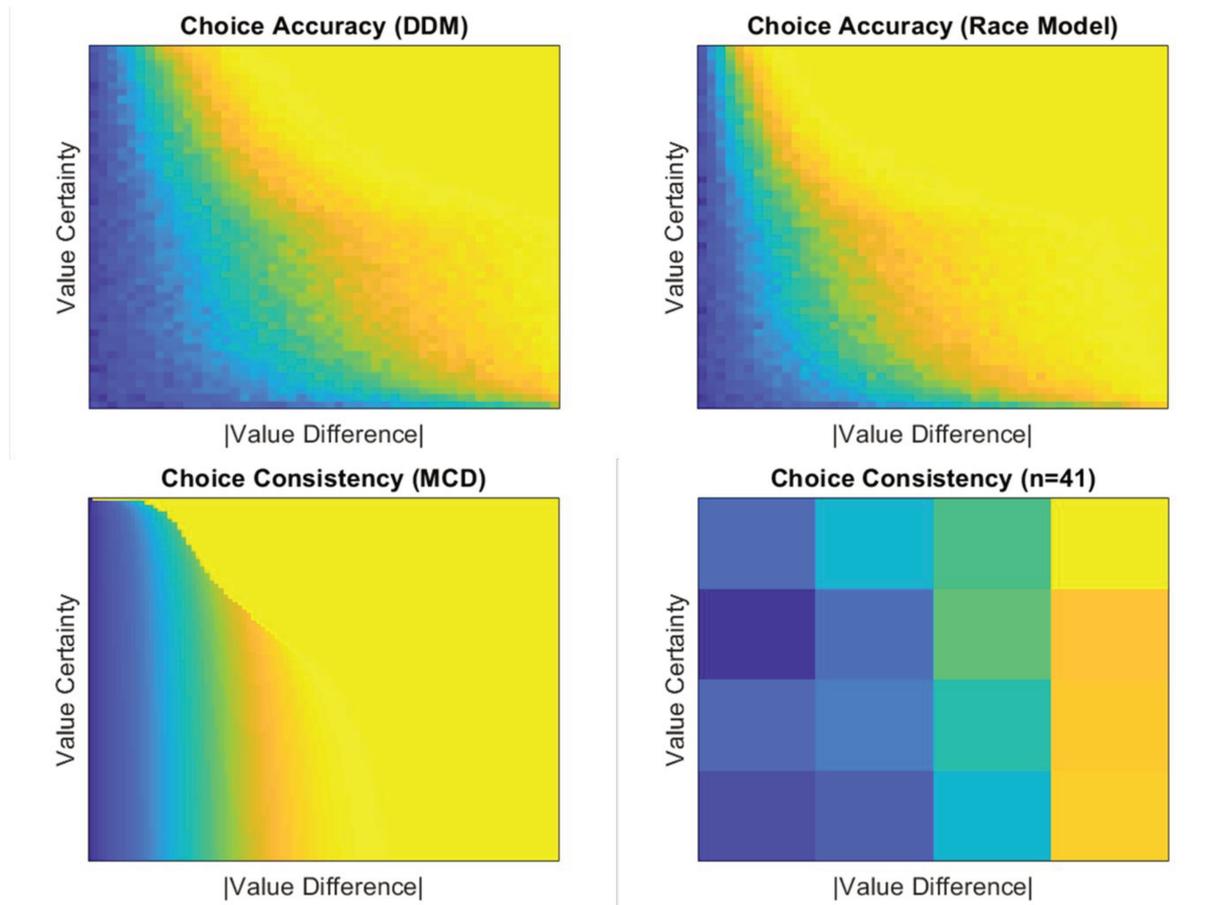
In what follows, we compare existing computational models of the relationship between choice, value, confidence, and decision time. At this point, suffice it to say that, under accumulation-to-bound models, value uncertainty ratings proxy the magnitude of the stochastic noise in the evidence accumulation process. In contrast, under the MCD model, they simply capture the precision of subjective value representations before the choice. As we will see, all models make rather similar predictions regarding the impact of value ratings. However, they disagree about the impact of value certainty ratings. We will now inspect the three-way relationships between pre-choice value and value certainty ratings and each choice feature (namely: prediction accuracy, decision time, and confidence). 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:  $VCR^0$ ). In each case, we will summarize the empirical data and the corresponding model prediction.

First, we checked how choice prediction accuracy relates to  $|\Delta VR^0|$  and  $VCR^0$ . Here, we measure choice accuracy in terms of the rate of choices that are congruent with preferences derived from pre-choice value ratings  $\Delta VR^0$ . Under accumulation-to-bound models, choice accuracy should increase with  $|\Delta VR^0|$ , and decrease with  $VCR^0$ . This is because the relative impact of stochastic noise on the decision decreases with choice ease, and its magnitude decreases with value certainty ratings. The MCD model makes the same prediction, but for a different reason. In brief, increasing  $|\Delta VR^0|$  and/or  $VCR^0$  will decrease the demand for effort, which implies that the probability of changing one's mind will be smaller. Figure 2 below shows all quantitative model predictions and summarizes the corresponding empirical data.

One can see that the data seem to conform to the models' predictions. To confirm this, we ran, for each participant, a multiple logistic regression of choice accuracy against  $|\Delta VR^0|$  and  $VCR^0$ . A random effect analysis shows that both have a significant positive effect at the group level ( $|\Delta VR^0|$ : mean GLM beta=0.17, s.e.m.=0.02,  $p<0.001$ ;  $VCR^0$ : mean GLM beta=0.07, s.e.m.=0.03,  $p=0.004$ ). Note that people make "inaccurate" choices either because they make mistakes or because they change their mind during the decision. In principle, we can discriminate between these two explanations because we can check whether "inaccurate" choices are congruent with post-choice value ratings (change of mind) or not (error). This is important, because accumulation-to-bound models do not allow for the possibility that value representations change during decisions (hence all "inaccurate" choices would be deemed "errors"). It turns out that, among "inaccurate" choices, mind changes are more frequent than errors (mean rate difference=2.3%, s.e.m.=0.01,  $p=0.032$ ). Note that analyses of mind

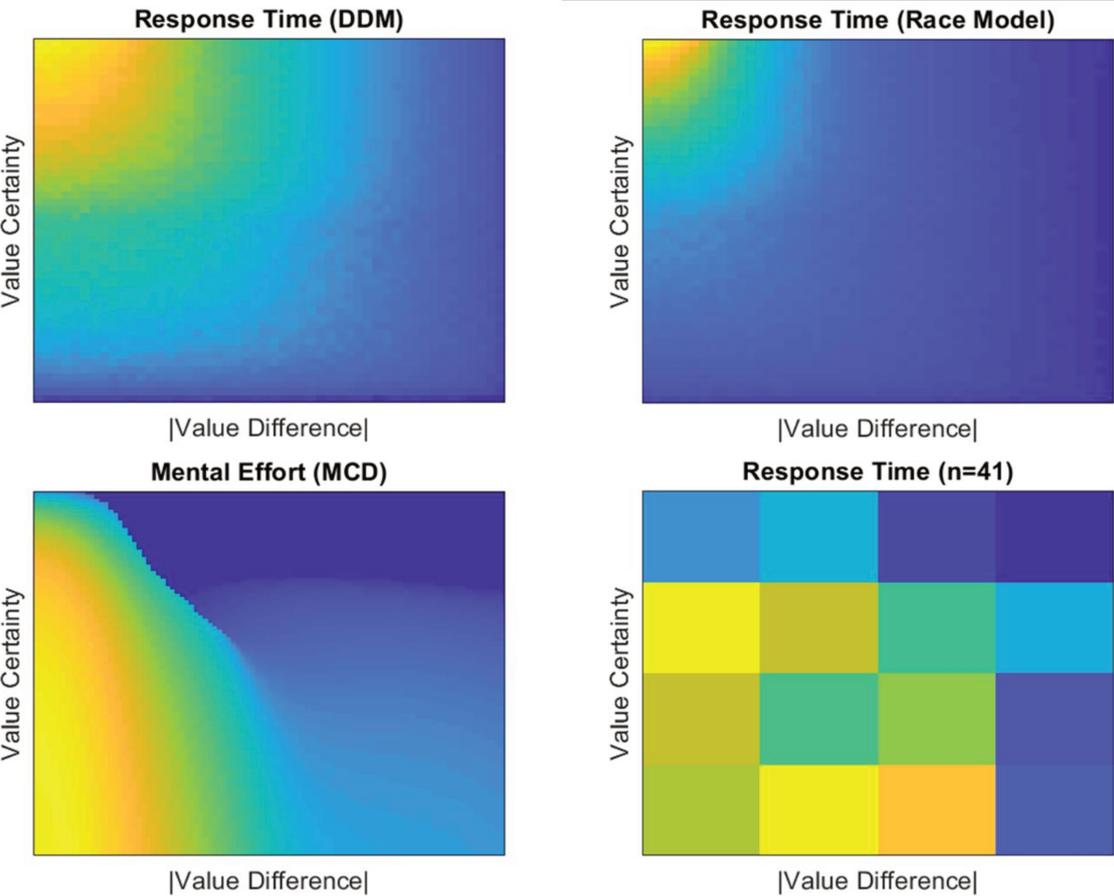
changes yield qualitatively identical results as choice accuracy (we refer the interested reader to Supplementary Material).



**Figure 2. Three-way relationship between choice accuracy, value, and value certainty.** Upper-left panel: prediction of the DDM model: choice accuracy (color code) is shown as a function of  $|\Delta VR^0|$  (x-axis) and  $CR^0$  (y-axis). Upper-right panel: prediction of the race model: same format. Lower-left panel: prediction of the MCD model: same format. Lower-right panel: empirical data: same format.

Second, we checked how decision time relates to first- and second-order pre-choice ratings. Under accumulation-to-bound models, decisions are triggered whenever the stochastic evidence accumulation process reaches a predefined threshold. Now, increasing  $\Delta VR^0$  effectively increases the drift rate, eventually decreasing the expected decision time. In addition, expected decision time increases with  $VCR^0$ , because the

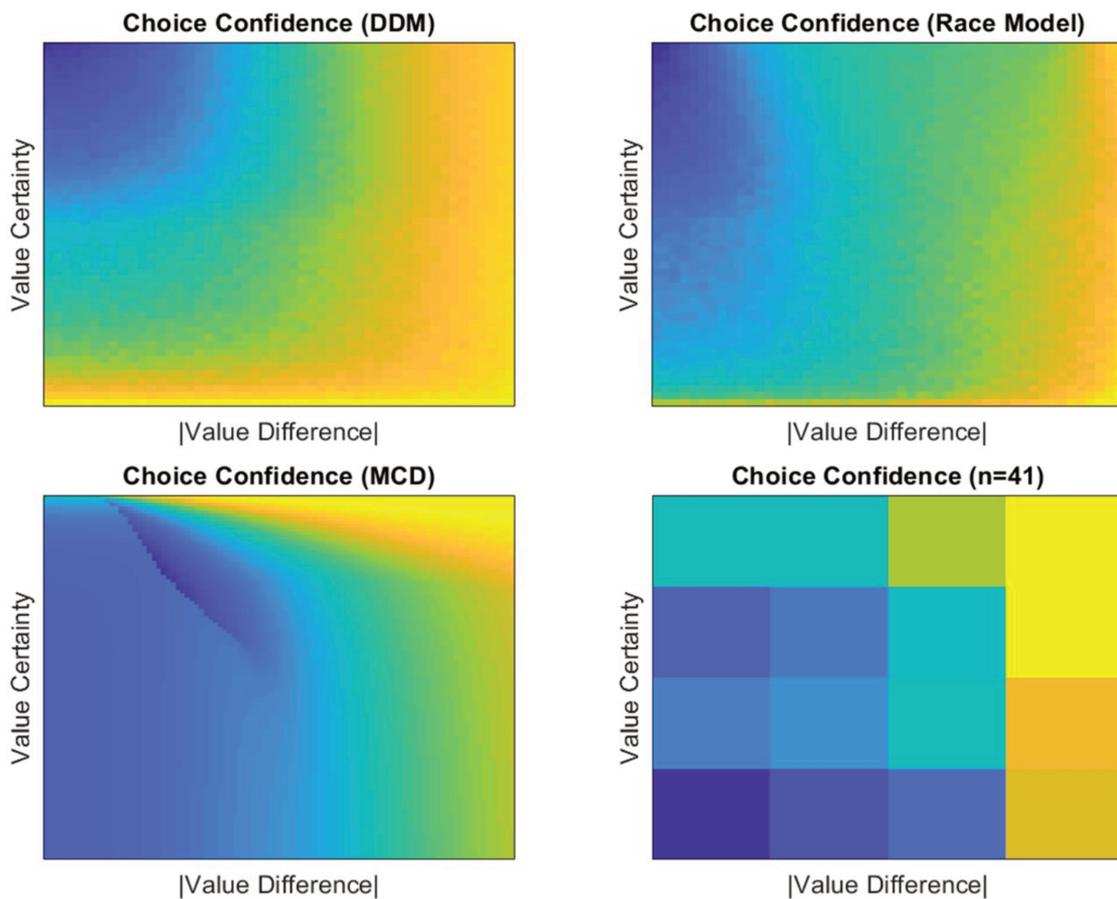
probability of an early bound hit decreases when the noise magnitude decreases. Under the MCD model, decision time can be thought of as a proxy for effort duration. Here, increasing  $|\Delta VR^0|$  and/or  $VCR^0$  will decrease the demand for effort, which will result in smaller expected decision time. In other words, the MCD model differs from accumulation-to-bound models with respect to the impact of  $VCR^0$  on decision time. Figure 3 below shows all quantitative model predictions and summarizes the corresponding empirical data.



**Figure 3. Three-way relationship between decision time, value, and value certainty.**  
 Upper-left panel: prediction of the DDM model: decision time (color code) is shown as a function of  $|\Delta VR^0|$  (x-axis) and  $CR^0$  (y-axis). Upper-right panel: prediction of the race model: same format. Lower-left panel: prediction of the MCD model: same format. Lower-right panel: empirical data: same format.

One can see that the decision time data behave as predicted by the MCD model. Here, we also ran, for each participant, a multiple logistic regression of decision times against  $|\Delta VR^0|$  and  $VCR^0$ . A random effect analysis shows that both have a significant and negative effect at the group level ( $|\Delta VR^0|$ : mean GLM beta=-0.13, s.e.m.=0.02,  $p<0.001$ ;  $CR^0$ : mean GLM beta=-0.06, s.e.m.=0.02,  $p=0.005$ ).

Third, we checked how choice confidence relates to  $|\Delta VR^0|$  and  $VCR^0$ . Under the DDM model, choice confidence is defined as the height of the optimal collapsing bound when it is hit. Because bounds are collapsing with decision time, confidence increases with  $|\Delta VR^0|$  and decreases with  $VCR^0$ . Under the race model, confidence is defined as the gap between the two value accumulators when the bound is hit. As with the DDM model, increasing  $|\Delta VR^0|$  trivially increases confidence. In addition, increasing  $VCR^0$  decreases the expected gap between the best and the worst value accumulators (Lebreton et al., 2015). Under the MCD model, confidence reflects the discriminability of value representations after optimal resource allocation. Critically, although more resources are allocated to the decision when either  $|\Delta VR^0|$  or  $VCR^0$  decrease, this does not overcompensate for decision difficulty, and thus choice confidence decreases. As before, Figure 4 below shows all quantitative model predictions and summarizes the corresponding empirical data.



**Figure 4. Three-way relationship between choice confidence, value, and value certainty.** Upper-left panel: prediction of the DDM model: choice confidence (color code) is shown as a function of  $|\Delta VR^0|$  (x-axis) and  $CR^0$  (y-axis). Upper-right panel: prediction of the race model: same format. Lower-left panel: prediction of the MCD model: same format. Lower-right panel: empirical data: same format.

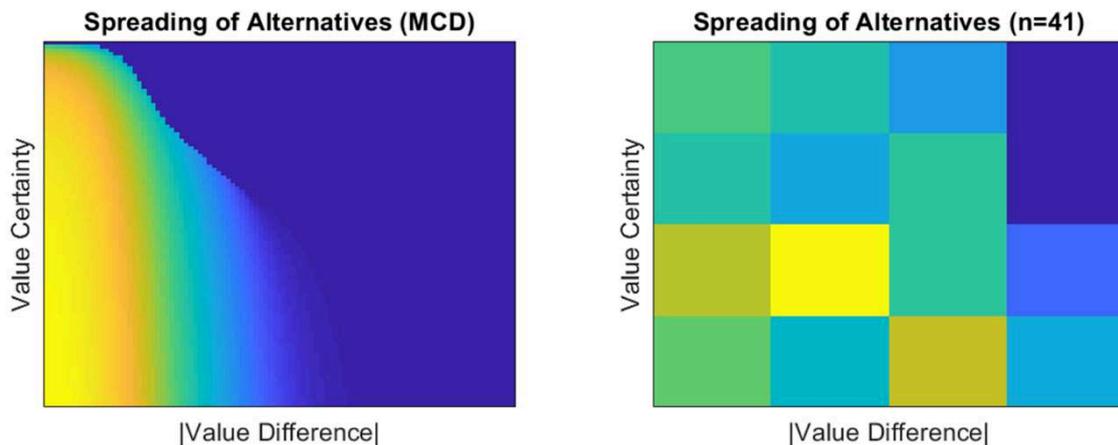
One can see that the choice confidence follows the MCD model predictions. Again, we ran, for each participant, a multiple logistic regression of confidence against  $|\Delta VR^0|$  and  $VCR^0$ . A random effect analysis shows that both have a significant and positive effect at the group level ( $|\Delta VR^0|$ : mean GLM beta=0.23, s.e.m.=0.02,  $p<0.001$ ;  $VCR^0$ : mean GLM beta=0.15, s.e.m.=0.03,  $p<0.001$ ). Note that this is unlikely to be a trivial consequence of peoples' decision time readout, since confidence is only mildly correlated with decision time (mean correlation=-0.32, s.e.m.=0.03,  $p<0.001$ ).

- **Subjective feeling of effort, choice-induced preference change, decision importance, and cost of time**

So far, we have provided evidence that choice confidence and decision time are better explained with the MCD model than with accumulation-to-bound models. In what follows, we will evaluate some additional quantitative predictions that are specific to the MCD model. The derivation of each of these predictions is detailed in the Model section below.

First, recall that MCD really is about the allocation of costly cognitive resources, i.e. mental effort, into the decision process. One may thus ask whether the subjective feeling of effort *per se* follows the MCD predictions. Recall that increasing  $|\Delta VR^0|$  and/or  $VCR^0$  will decrease the demand for mental resources, which will result in the decision being associated with a lower feeling of effort. To check this, we thus performed a multiple linear regression of subjective effort ratings against  $|\Delta VR^0|$  and  $VCR^0$ . A random effect analysis shows that both have a significant and negative effect at the group level ( $|\Delta VR^0|$ : mean GLM beta=-0.20, s.e.m.=0.03,  $p<0.001$ ;  $VCR^0$ : mean GLM beta=-0.05, s.e.m.=0.02,  $p=0.025$ ). A graphical summary of the data can be seen in the Supplementary Material.

Second, the MCD model predicts how value representations will be modified during the decision process. In particular, choice-induced preference change should globally follow the optimal effort allocation. More precisely, the reported value of alternative options should spread apart, and the expected spreading of alternatives should be decreasing with  $|\Delta VR^0|$  and  $VCR^0$ . Figure 5 below shows the model predictions and summarizes the corresponding empirical data.

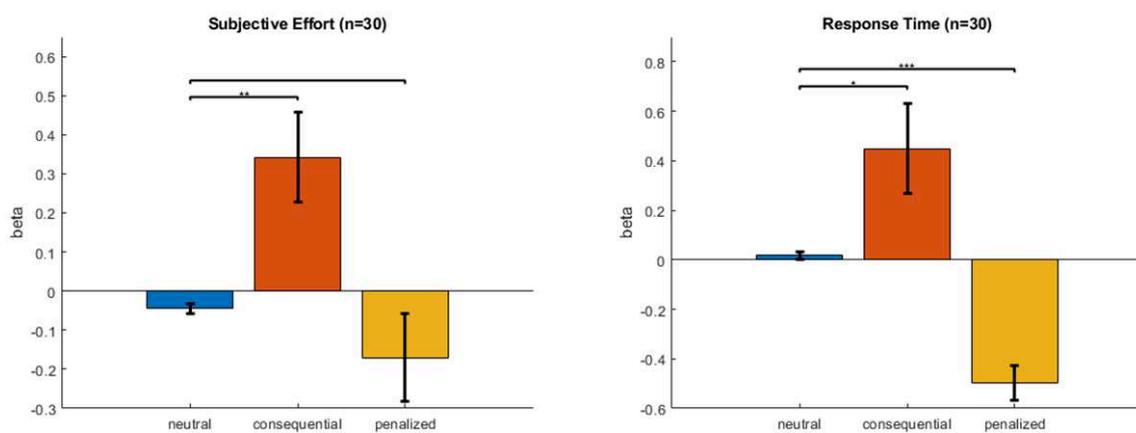


**Figure 5. Three-way relationship between choice-induced preference change, value, and value certainty.** Left panel: prediction of the MCD model: the spreading of alternatives (color code) is shown as a function of  $|\Delta VR^0|$  (x-axis) and  $VCR^0$  (y-axis). Right panel: empirical data: same format.

One can see that the spreading of alternatives follows the MCD model predictions. A random effect analysis confirms this, showing that both  $|\Delta VR^0|$  and  $CR^0$  have a significant negative effect at the group level ( $|\Delta VR^0|$ : mean GLM beta=-0.09, s.e.m.=0.03,  $p<0.001$ ;  $CR^0$ : mean GLM beta=-0.04, s.e.m.=0.02,  $p=0.027$ ). Note that this replicates our previous findings on choice-induced preference change (Lee and Daunizeau, 2019). In addition to expected changes in value ratings, the MCD model predicts that the precision of value representations should increase after the decision has been made (cf. " $\beta$ -effect" in the Supplementary Material). Indeed, post-choice value certainty ratings are significantly higher than pre-choice value certainty ratings (mean difference=1.34, s.e.m.=0.51,  $p=0.006$ ). Importantly, under the MCD model, post-choice ratings are simply reports of modified value representations at the time when the choice is triggered. Therefore, choice and its associated confidence level should be better predicted with post-choice ratings than with pre-choice ratings.

Indeed, we found that the predictive power of post-choice ratings is significantly higher than that of pre-choice ratings, both for choice (mean prediction accuracy difference=7%, s.e.m.=0.01,  $p<0.001$ ) and choice confidence (mean prediction accuracy difference=3%, s.e.m.=0.01,  $p=0.004$ ). Details regarding this analysis can be found in the Supplementary Material.

Third, the MCD model predicts that, all else being equal, effort increases with decision importance and decreases with costs. We checked the former prediction by asking participants to make a few decisions where they knew that the choice would be real, i.e. they would actually have to eat the chosen food item. We refer to these trials as "consequential" decisions. To check the latter prediction, we imposed a financial penalty that increases with decision time. These experimental manipulations are described in the Methods section. Figure 6 below shows subjective effort ratings and decision times for "neutral", "consequential" and "penalized" decisions, when controlling for  $|\Delta VR^0|$  and  $VCR^0$  (see the Supplementary Material for more details).



**Figure 6. Comparison of "neutral", "consequential", and "penalized" decisions.** Left: Mean (+/- s.e.m.) effort ratings are shown for "neutral" (blue), "consequential" (red) and "penalized" (yellow) decisions. Right: Mean (+/- s.e.m.) decision time (same format). Both datasets were corrected for  $|\Delta VR^0|$  and  $VCR^0$ .

One can see that subjective effort reports and decision times follow the MCD model predictions. More precisely, both subjective effort reports and decision times were significantly higher for "consequential" decisions than for "neutral" decisions (mean effort difference=0.39, s.e.m.=0.12,  $p=0.001$ ; mean decision time difference=0.43, s.e.m.=0.19,  $p=0.017$ ). In addition, decision times are significantly faster for "penalized" than for "neutral" decisions (mean decision time difference=-0.51, s.e.m.=0.08,  $p<0.001$ ). Note that although the difference in reported effort between "neutral" and "penalized" decisions does not reach statistical significance (mean effort difference=-0.13, s.e.m.=0.12,  $p=0.147$ ), it goes in the right direction.

## **METHODS**

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 épinière*). All participants were native French speakers. All participants were from the non-patient population with no reported history of psychiatric or neurological illness.

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. The experiment was conducted entirely in French.

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.

- **Participants**

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 each participant was paid a flat rate of 20€ as compensation for his time plus an average of 4€ as a bonus. One group of 11 participants was excluded from the cross-condition analysis only (see below), due to technical issues.

- **Materials**

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.

- **Procedure**

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 (Chen and Risen, 2010; Izuma and Murayama, 2013): pre-choice item ratings, choice, and post-choice item ratings (see Figure 1 above). There was no time limit for the overall experiment, nor for the different sections, nor for the individual trials. Item rating and choice sessions are described below.

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, "How certain are you 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.

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 decision 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.

Note: In the Results section, we refer to  $\Delta VR^0$  as the difference between pre-choice value ratings of items composing a choice set. Similarly,  $CVR^0$  is the average pre-choice value certainty ratings across items composing a choice set.

- **Conditions**

The choice section of the experiment included trials of three different conditions: "neutral" (60 trials), "consequential" (7 trials), and "penalized" (7 trials), which were randomly intermixed. Immediately prior to each "consequential" trial, participants were instructed that they would be required to eat, at the end of the experiment, a portion of the item that they were about to choose. Immediately prior to each "penalized" trial, participants were instructed that they would lose 0.20€ for each second that they would take to make their choice.

## **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 proactive anticipation of how these resources will perturb the internal representations of subjective values. As we will see, the computational properties of the MCD model are critically different from accumulation-to-bound models of value-based decision-making, which we briefly describe in the Supplementary Material.

### 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:

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

(1)

where the expectation accounts for predictable stochastic influences that ensue from allocating resources (this will be more clear 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 will be made more clear below,  $P_c(z)$  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 the decision maker may be unsure about how much he likes/wants 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 values take the form of Gaussian probability density functions, as follows:

$$p(V_i) = N(\mu_i, \sigma_i) \tag{3}$$

where  $\mu_i$  and  $\sigma_i$  are the mode and the variance of the probabilistic value representations, 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:

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

where 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).

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  $\sigma_i$  and the modes  $\mu_i$  of value representations will change, before having allocated the resources (i.e., without having processed the information). In brief, allocating resources essentially has two impacts: (i) it decreases the uncertainty  $\sigma_i$ , and (ii) it perturbs the modes  $\mu_i$  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. Under simple Bayesian belief update rules, this reduces to stating that the variance of a given probabilistic value representation decreases in proportion to the amount of allocated effort, i.e.:

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

where  $\sigma_i^0$  is the prior variance of the representation (before any effort has been allocated), and  $\beta$  controls the efficacy with which resources increase the precision of value representations. Formally speaking, Equation 5 has the form of a Bayesian update of the belief variance in a Gaussian-likelihood model, where the precision of the likelihood term is  $\beta z$ . More precisely,  $\beta$  is the precision increase that follows from allocating a unitary amount of resources  $z$ . In what follows, we will refer to  $\beta$  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  $i^{\text{th}}$  value representation, having allocated allocating 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.:

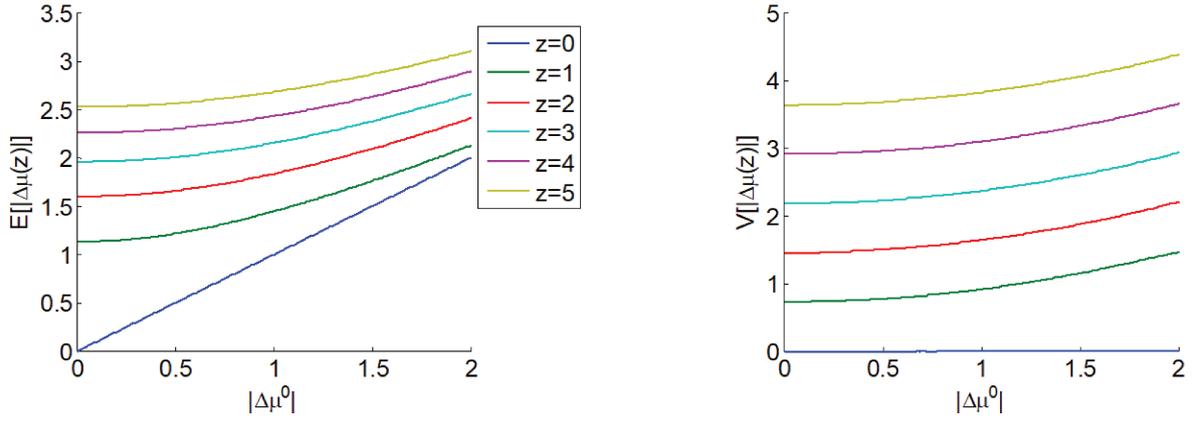
$$\begin{aligned} \mu_i(z) &= \mu_i^0 + \delta_i \\ \delta_i &\sim N(0, \gamma z) \end{aligned} \tag{6}$$

where  $\mu_i^0$  is the mode of the value representation before any effort has been allocated, and  $\gamma$  controls the relationship between the amount of allocated resources and the variance of the perturbation term  $\delta$ . The higher  $\gamma$ , the greater the expected perturbation of the mode for a given amount of allocated resources. In what follows, we will refer to  $\gamma$  as the "*type #2 effort efficacy*".

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  $\delta$ , 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:

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

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  $E[|\Delta\mu||z]$  depends upon both the absolute prior mean difference  $|\Delta\mu^0|$  and the amount of allocated resources  $z$ . This is depicted on Figure 7 below.



**Figure 7. The expected impact of allocated resources onto value representations.** Left: the expected absolute mean difference  $E[|\Delta\mu|z]$  (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: Variance  $V[|\Delta\mu|z]$  of the absolute mean difference ; same format.

One can see that  $E[|\Delta\mu|z] - |\Delta\mu^0|$  is always greater than 0 and increases with  $z$  (and if  $z = 0$ , then  $E[|\Delta\mu|z] = |\Delta\mu^0|$ ). 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  $V[|\Delta\mu|z]$  of the absolute means' difference is derived from the expression of the second-order moment of the corresponding folded normal distribution:

$$V[|\Delta\mu|z] = 2\gamma z + |\Delta\mu^0|^2 - E[|\Delta\mu|z]^2 \quad (8)$$

One can see on Figure 7 that  $V[|\Delta\mu||z]$  increases with the amount  $z$  of allocated resources (but if  $z = 0$ , then  $V[|\Delta\mu||z] = 0$ ).

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  $z$ :

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

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  $\beta$  and  $\gamma$ . 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 the Supplementary Material.

To complete the cost-benefit model, and without loss of generality, we will 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  $\alpha$  determines the effort cost of allocating a unitary amount of resources  $z$ . In what follows, we will refer to  $\alpha$  as the "effort unitary cost". We refer to  $\alpha$  as the "effort unitary cost".

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

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

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 the Supplementary Material.

Note: at this point, the MCD model is agnostic about what the allocated resource is. Empirically, we relate  $\hat{z}$  to two different measures, namely: decision time and the subjective feeling of effort. The former makes sense if one thinks of decision time in terms of effort duration, which increases the cumulative engagement of neurocognitive resources. The latter relies on the subjective cost incurred when deploying neurocognitive resources, which would be signaled by experiencing mental effort. We will comment on this in the Discussion section. Also, 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.  $z_1 \approx z_2 \triangleq z$ ). This simplifying assumption is justified by eye-tracking data (cf. Supplementary Material). Finally, we investigate the effect of decision importance by comparing effort and decision time in "neutral" versus "consequential" decisions (cf. Methods section).

## 2. Corollary predictions of the MCD model

In the previous section, we derived the MCD-optimal resource allocation, 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 on Figure 3 (and/or Figure S4 of the Supplementary Material), as a function of (empirical proxies for) the prior absolute difference between modes  $|\Delta\mu^0|$  and the prior certainty  $1/\sigma^0$  of value representations. But, this 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 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 invests and its associated confidence, on a decision-by-decision basis. This quantitative prediction is shown on Figure 4.

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.:

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

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  $\hat{z}$ . Note that, by definition, choice accuracy (i.e., congruence of choice and prior preference  $\Delta\mu^0$ ) is but  $1-Q(z)$ , which is shown on Figure 2.

Lastly, we can predict 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:

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

where  $\Delta\delta(z) \sim N(0, 2\gamma z)$  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

$\overline{SOA}(\hat{z})$ :

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

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  $\hat{z}$ .

In summary, the MCD model predicts, given the prior absolute difference between modes  $|\Delta\mu^0|$  and the prior certainty  $1/\sigma^0$  of value representations, choice accuracy, choice confidence, choice-induced preference change, decision time and/or subjective feelings of effort. Note that, when testing the decision-by-decision predictions of the MCD model, we use  $\Delta VR^0$  and  $CVR^0$  as empirical proxies for  $\Delta\mu^0$  and  $1/\sigma^0$ , respectively.

## DISCUSSION

In this work, we have presented a novel computational model of decision-making which explains the intricate relationships between choice accuracy, decision time, subjective effort, choice confidence, and choice-induced preference change. This model assumes that deciding between alternative options whose values are uncertain induces a demand for allocating cognitive resources to processing value-relevant information. Cognitive resource allocation then optimally trades effort for confidence, given the discriminability of prior value representations. Such metacognitive control of decisions

or MCD makes novel predictions that differ from standard accumulation-to-bound models of decision-making, including a drift-diffusion model that was proposed as an optimal policy for value-based decision making (Tajima et al., 2016). But, 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 decider'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 stationary nonetheless. This why the DDM cannot predict choice-induced preference changes. This critical distinction extends to other types of accumulation-to-bound models, including race models ( De Martino et al, 2013; Tajima et al, 2019).

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.,  $\Delta VR^0$ ) 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 the framework of the race model that we compared to the MCD model. 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 choices might rather be 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. 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 would like to discuss a few features of the MCD model. First, we did not specify what determines the reward component, which quantifies decision importance and acts as an effective 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 time of the decision (cf. partner choices), the breadth of its repercussions (cf. political decisions), or its instrumentality with respect to the achievement of superordinate goals (cf. 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 might also be influenced by the prior (intuitive/emotional/habitual) appraisal of option values. For example, we found that, all else equal, people spent much 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). Probing this type of influence will be the focus of forthcoming publications.

Second, our current version of the MCD model relies upon a simple variant of resource costs. We note that rendering the cost term nonlinear (e.g., quadratic) does not change the qualitative nature of the MCD model 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. For example, the cost of allocating attention to a given option may depend upon whether this option would be a priori chosen or not. This might eventually explain systematic decision biases and differences in decision times between default and non-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 decision time as a proxy for cognitive effort. In fact, as is evident from the comparison between “penalized” and “neutral” choices, imposing an external penalty cost on decision time reduces, as expected, the ensuing subjective effort. More generally, however, the dual optimization of effort dimensions might render the relationship between effort and decision time 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 feeling of effort is not strongly impacted (cf. Figure 6). In any case, the relationship between effort and decision time might depend upon the relative costs of effort duration and intensity, which might itself 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 dynamic 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, proactive-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.

## REFERENCES

- 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.
- 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.
- 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 rostral lateral 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 rostral lateral 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.
- 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.
- Harlé, K.M., and Sanfey, A.G. (2007). Incidental sadness biases social economic decisions in the Ultimatum Game. *Emot. Wash. DC* *7*, 876–881.
- 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*.
- Izuma, K., and Murayama, K. (2013). Choice-Induced Preference Change in the Free-Choice Paradigm: A Critical Methodological Review. *Front. Psychol.* *4*.
- Kahneman, D., Slovic, P., and Tversky, A. (1982). *Judgment Under Uncertainty: Heuristics and Biases* (Cambridge University Press).
- Krajibich, 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.
- Latimer, K.W., Yates, J.L., Meister, M.L.R., Huk, A.C., and Pillow, J.W. (2015). Single-trial Spike Trains in Parietal Cortex Reveal Discrete Steps During Decision-making. *Science* *349*, 184–187.
- Latimer, K.W., Huk, A.C., and Pillow, J.W. (2017). No cause for pause: new analyses of ramping and stepping dynamics in LIP (Rebuttal to Response to Reply to Comment on Latimer et al 2015). *BioRxiv* 160994.
- 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 Daunizeau, J. (2019). Choosing what we like vs liking what we choose: How choice-induced preference change might actually be instrumental to decision-making. *BioRxiv* 661116.
- 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.
- 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.
- Ozcinmer, 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., Ozcinmer, 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*.
- 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.
- Sharot, T., Velasquez, C.M., and Dolan, R.J. (2010). Do Decisions Shape Preference? Evidence From Blind Choice. *Psychol. Sci.* *21*, 1231–1235.
- Shenhav, A., & Karmarkar, U. R. (2019). Dissociable components of the reward circuit are involved in appraisal versus choice. *Scientific reports*, *9*(1), 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.
- 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.
- 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.

---

## **Trading Mental Effort for Confidence: Supplementary Material**

### **1. Data descriptive statistics and sanity checks**

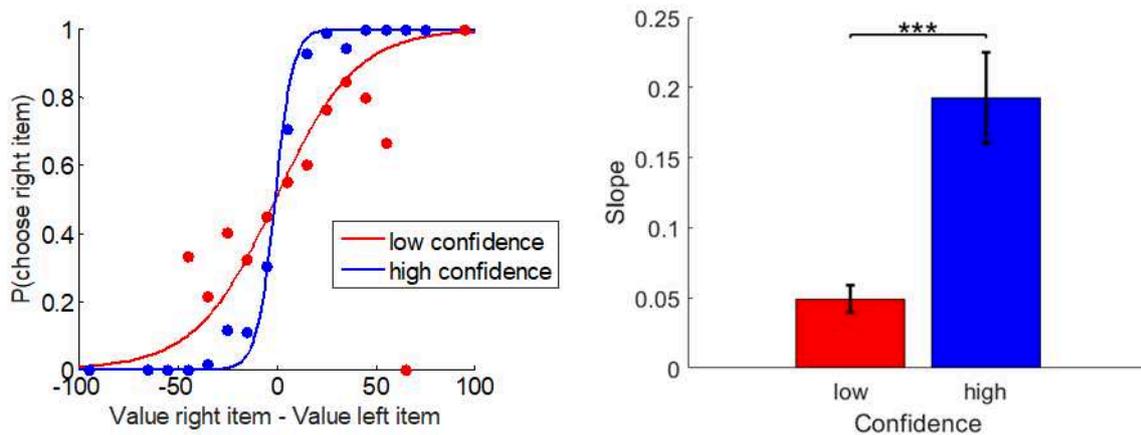
Recall that we collect value ratings and value certainty ratings both before and after the choice session. We did this for the purpose of validating specific predictions of the MCD model (in particular: choice-induced preference changes: see Figure 5 of the main text). It turns out this also enables 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 performed 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,  $p < 0.001$ ).

### **2. Does choice confidence moderate the relationship between choice and pre-choice value ratings?**

Previous studies regarding confidence in value-based 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 Figure S1 below.



**Figure S1. Relationship between choices, pre-choice value ratings and choice confidence.** Left: 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: the corresponding logistic regression slope (y-axis) is shown for both high- (blue) and low- (red) confidence trials (group means +/- s.e.m.).

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 degrades the prediction accuracy of choice from pre-choice value ratings (which have been changed during the decision process). However, although more resources are allocated to the decision, this does not

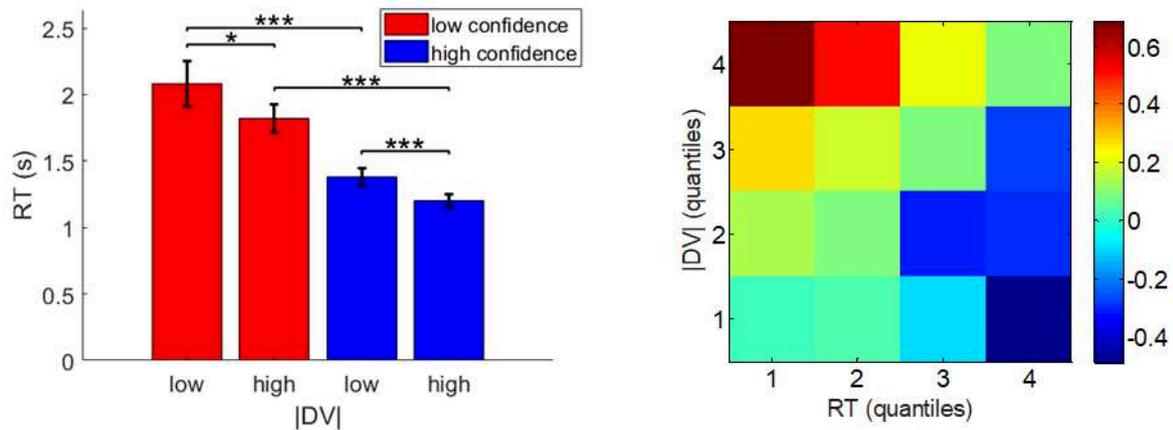
overcompensate for decision difficulty, and thus choice confidence decreases. Thus, low-confidence choices will be those choices that cannot be well predicted with pre-choice value ratings. We note that the anti-correlation between choice confidence and choice accuracy can be seen by comparing Figures 2 and 4 of the main text.

### **3. How do choice confidence, difference in pre-choice value ratings, and decision 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  $\Delta VR^0$  and decision time (De Martino et al., 2013). First, we regressed, for each participant, decision time concurrently against both  $|\Delta VR^0|$  and choice confidence. A random effect showed that both have a significant main effect on decision time ( $\Delta VR^0$ : mean GLM beta=-0.016, s.em.=0.003,  $p<0.001$ ; choice confidence: mean GLM beta=-0.014, s.em.=0.002;  $p<0.001$ ), without any two-way interaction ( $p=0.133$ ). This analysis is summarized in Figure S2 below, together with the full three-way relationship between  $|\Delta VR^0|$ , confidence and decision time.

In brief, confidence increases with the absolute value difference and decreases with decision 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 decision time and a decrease in choice confidence. This would produce the same data pattern as Figure S2, although the causal relationships implicit in this data

representation is partially incongruent with the computational mechanisms underlying MCD.



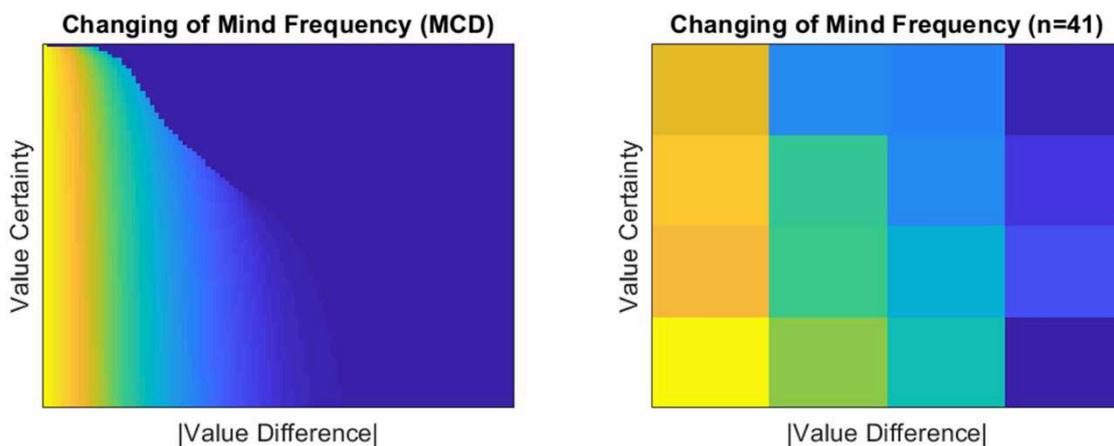
**Figure S2. Relationship between pre-choice value ratings, choice confidence, and decision time.** Left: decision time (y-axis) is plotted as a function of low- and high-  $|\Delta VR^0|$  (x-axis) for both low- (red) and high- (blue) confidence trials. Error bars represent s.e.m. Right: A heatmap of mean z-scored confidence is shown as a function of both decision time (x-axis) and  $|\Delta VR^0|$  (y-axis).

#### 4. Analysis of changes of mind

In the main manuscript, we show that choice accuracy increases with pre-choice value difference  $\Delta VR^0$  and pre-choice value certainty  $VCR^0$ . Recall that choice accuracy was defined in terms of the rate of choices that are congruent with preferences derived from pre-choice value ratings. Now, people make "inaccurate" choices either because they make mistakes or because they change their mind during the decision. In principle, we can discriminate between these two explanations because we can check whether "inaccurate" choices are congruent with post-choice value ratings (change of mind) or not (error). This is important, because accumulation-to-bound models do not allow for the possibility that value representations change during decisions. Hence all "inaccurate" choices would be deemed "errors", which are driven by stochastic noise in the evidence accumulation process. It turns out that most choices are "accurate"

(mean choice accuracy =73.3%, s.e.m.=1%), and less than half of the "inaccurate" choices are classified as "errors" (mean error rate=12%, s.e.m.=0.01), which is significantly less than "mind changes" (mean rate difference=2%, s.e.m.=0.01,  $p=0.032$ ). In addition, choice confidence and (post- versus pre-choice) value certainty gain were significantly higher for "changes of mind" than for "errors" (choice confidence: mean difference=13.7, s.e.m.=2.1,  $p<0.001$ ; value certainty gain: mean difference=2.6, s.e.m.=1.4,  $p=0.035$ ).

Thus, one may wonder what would be the impacts of both pre-choice value difference  $\Delta VR^0$  and pre-choice value certainty  $VCR^0$  on choice accuracy, if one were to remove "errors" from "inaccurate" choices. Figure S3 below shows both the predicted and measured three-way relationship between the probability of changing one's mind,  $\Delta VR^0$  and  $VCR^0$ .



**Figure S3. Relationship between the probability of changing one's mind, value ratings, and certainty ratings.** Left: Prediction under the MCD model: a heatmap of the probability of changing one's mind is shown as a function of both  $|\Delta VR^0|$  (x-axis) and  $VCR^0$  (y-axis). Right: Empirical data: same format.

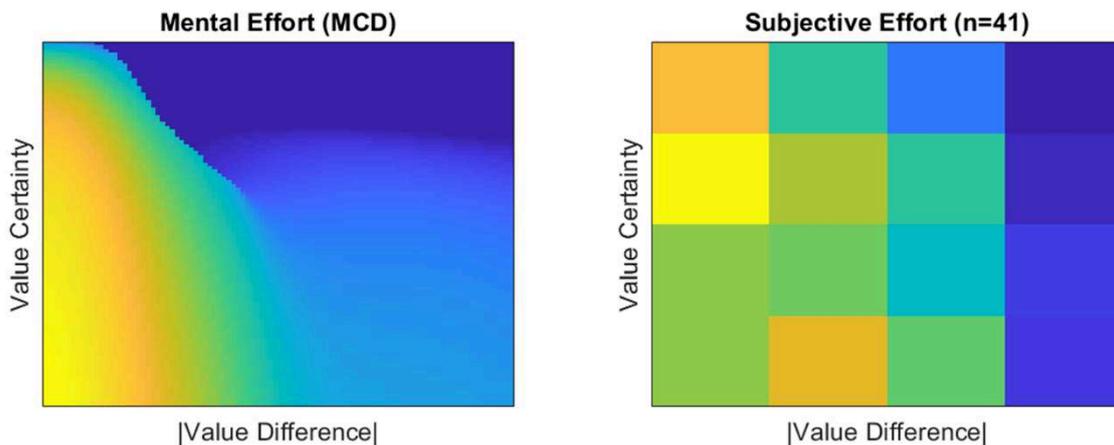
Recall that, under the MCD model, the probability of changing one's mind increases with the resource demand, which decreases when either  $|\Delta VR^0|$  or  $VCR^0$  increase. One can see that the data seem to conform to this prediction. To check this, we ran, for each participant, a multiple logistic regression of change of mind against  $|\Delta VR^0|$  and  $VCR^0$ . A random effect analysis shows that both have a significant and negative effect at the group level ( $\Delta VR^0$ : mean GLM beta=-0.16, s.e.m.=-0.02,  $p<0.001$ ;  $VCR^0$ : mean GLM beta=-0.08, s.e.m.=0.02,  $p<0.001$ ). These results are qualitatively similar to the analysis of choice accuracy (cf. Figure 2 in the main text).

## 5. Analysis of the subjective feeling of effort

In the main manuscript, we show that decision time decreases with pre-choice value difference  $\Delta VR^0$  and pre-choice value certainty  $VCR^0$ . The focus on decision time was motivated by the fact that all models could make quantitative—and thus comparable—predictions. In brief, we found that the effect of  $VCR^0$  on decision time was consistent with the MCD model, but not with accumulation-to-bound models. Now, under the MCD model, decision time is but a proxy for effort duration. Here, we ask whether the subjective feeling of effort *per se* follows the MCD model predictions. This is possible because we asked participants to rate how effortful each decision felt. Figure S4 below shows both the predicted and the measured three-way relationship between effort,  $|\Delta VR^0|$  and  $VCR^0$ .

One can see that the reported subjective feeling of effort closely matches model predictions. One may ask whether people's effort reports may be trivial post-choice read-outs of decision time and/or choice confidence. This, however, is unlikely, given that people's subjective effort is reducible neither to decision time (mean

correlation=0.39, s.e.m.=0.04), nor to choice confidence (mean correlation=-0.48, s.e.m.=, p=0.05).

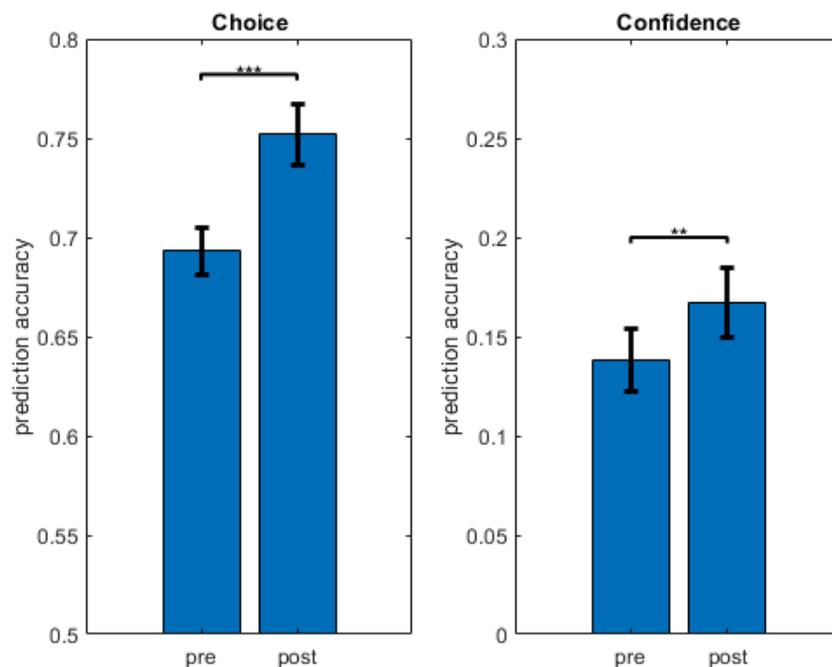


**Figure S4. Relationship between subjective effort, value ratings, and certainty ratings.** Left: Prediction under the MCD model: a heatmap of the MCD-optimal effort allocation is shown as a function of  $|\Delta VR^0|$  (x-axis) and  $VCR^0$  (y-axis). Note: this prediction is identical to Figure 3 in the main text (MCD model). Right: Empirical data: same format.

## 6. 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 the first section of this Supplementary Material, we report the result of a logistic regression of choice against pre-choice value ratings (see also Figure S1). We performed the same regression analysis, but this time against post-choice value ratings. Figure S5 below shows 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 4). Again, we performed the same regression, this time against post-choice ratings. Figure S5 below shows the ensuing predictive power (here, in terms of percentage of explained variance or  $R^2$ ) for both pre-choice and post-choice ratings.

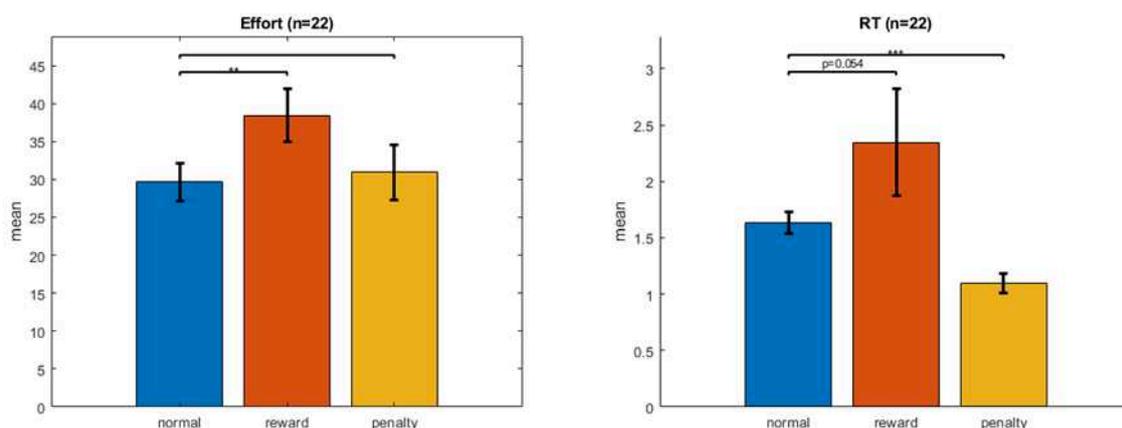


**Figure S5. Comparison of the predictive power of pre-choice versus post-choice ratings.** Left: Mean (+/- s.e.m.) BA of logistic regressions of choice against pre-choice (left) and post-choice (right) value ratings. Right: Mean (+/- s.e.m.)  $R^2$  of multiple linear regressions of choice confidence against pre-choice (left) and post-choice (right) 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,  $p=0.004$ ).

## 7. Cross-condition analysis: decision importance and cost of decision time

As featured in the main manuscript, we intermixed "neutral" trials with two specific sets of trials, in which we either manipulated decision importance (cf. "consequential" decisions) or the cost of decision time (cf. "penalized" decisions). Figure S6 below shows the mean subjective effort ratings and decision times for "neutral", "consequential" and "penalized" decisions.



**Figure S5. Comparison of "neutral", "consequential", and "penalized" decisions.** Left: Mean (+/- s.e.m.) effort rating are shown for "neutral" (blue), "consequential" (red) and "penalized" (yellow) decisions. Right: Mean (+/- s.e.m.) decision time (same format).

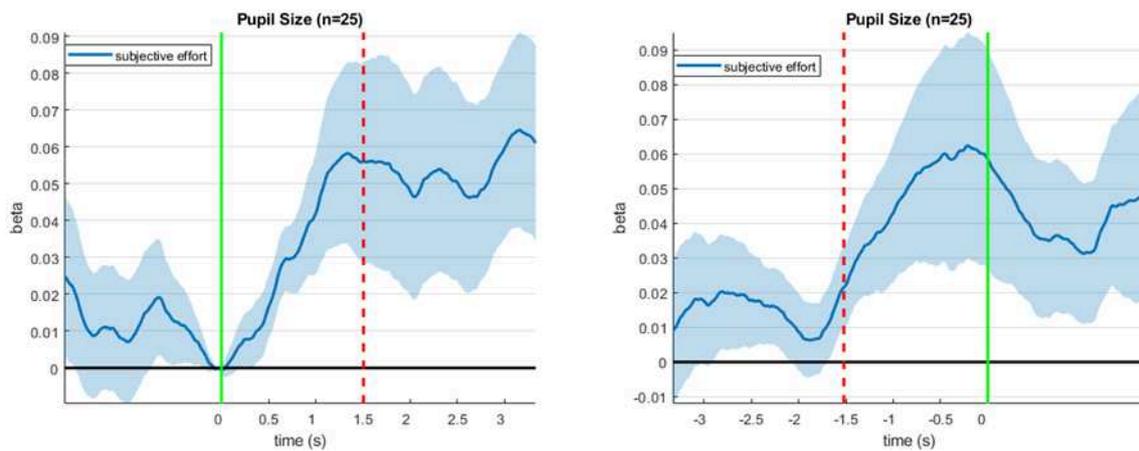
Overall, the data partially follows the model predictions. In particular, subjective effort and decision time are both significantly higher for "consequential" than for "neutral"

decisions (effort: mean difference=9.0, s.e.m.=2.2,  $p<0.001$ ; DT: mean difference=0.56, s.e.m.=0.32,  $p=0.043$ ). In addition, decision time is significantly lower for "penalized" than for "neutral" decision (mean DT difference=-0.46, s.e.m.=0.08,  $p<0.001$ ). However, there is no noticeable difference between reported efforts in "neutral" and "penalized" decisions (mean effort difference=0.6, s.e.m.=2.1,  $p=0.604$ ). This comparison, however, may be confounded by between-condition differences in  $\Delta VR^0$  or  $VCR^0$ . For each participant, we thus performed a multiple linear regression of effort and DT onto  $|\Delta VR^0|$  and  $VCR^0$ , including all types of trials. Corrected effort and DT can now be compared, after having removed the effects of  $|\Delta VR^0|$  and  $VCR^0$ . This is what Figure 6 of the main text shows. As one can see, the overall pattern is similar to Figure S5. As before, subjective effort and decision time are both significantly higher for "consequential" than for "neutral" decisions (effort: mean GLM beta difference=0.39, s.e.m.=0.12,  $p=0.001$ ; DT: mean GLM beta difference=0.43, s.e.m.=0.19,  $p=0.017$ ), and decision time is significantly lower for "penalized" than for "neutral" decisions (mean DT GLM beta difference=-0.51, s.e.m.=0.08,  $p<0.001$ ). Finally, the difference between reported efforts in "neutral" and "penalized" decisions is now almost significant (mean effort GLM beta difference=-0.13, s.e.m.=0.12,  $p=0.147$ ).

## **8. Analysis of eye-tracking data**

We first checked whether pupil dilation positively correlates with participants' reports of subjective effort. 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). Figure S6 below summarizes this analysis, in terms of the baseline-corrected time series of regression coefficients.

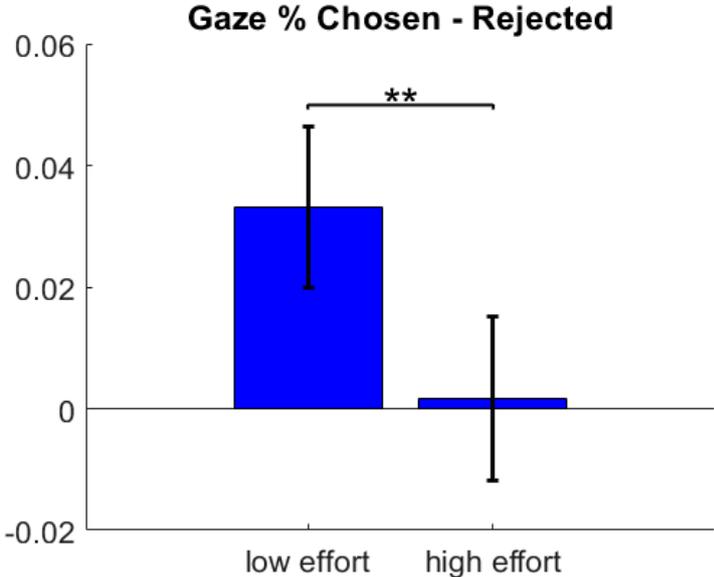


**Figure S6. Correlation between pupil size and reports of subjective effort during decision time.**

Left: Mean ( $\pm$  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: 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).

We found that the correlation between effort and pupil dilation was becoming 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/unchosen 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. Figure S7 below shows the gaze bias for low and high effort trials, based upon a median-split of subjective effort.



**Figure S7. 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.

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 on both items. This is important, because we use this simplifying assumption in our MCD model derivations.

### 9. On the accuracy of the predicted 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:

$$\bar{P} = E[s(\lambda|x)] \tag{S1}$$

where  $x \rightarrow s(x) = 1/(1 + e^{-x})$  is the sigmoid mapping,  $\lambda$  is an arbitrary constant, and the expectation is taken under the Gaussian distribution of  $x \sim N(\mu, \sigma^2)$ , whose mean and variance are  $\mu$  and  $\sigma^2$ , respectively.

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

$$\begin{cases} E[x] = \sigma \sqrt{\frac{2}{\pi}} \exp\left(-\frac{|\mu|^2}{2\sigma^2}\right) + \mu \left(2 \times s \left(\frac{\pi \mu}{\sigma \sqrt{3}}\right) - 1\right) \\ V[x] = \mu^2 + \sigma^2 - E[x]^2 \end{cases} \quad (\text{S2})$$

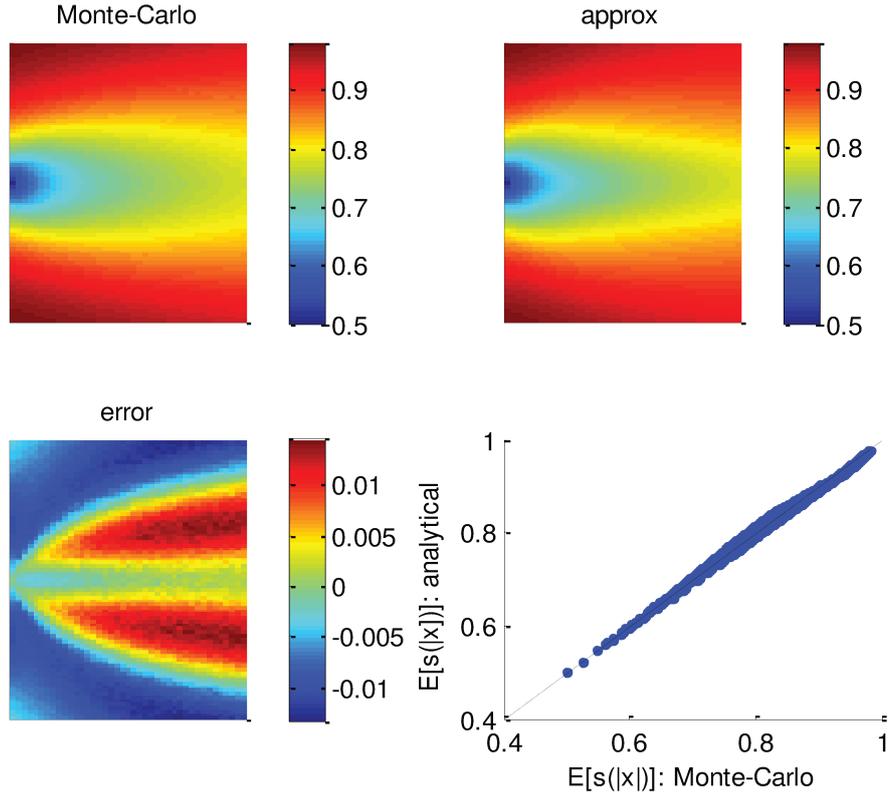
where the first line relies on a moment-matching approximation to the cumulative normal distribution function (Daunizeau, 2017). This allows us to derive the following analytical approximation to Equation S1:

$$\bar{P} \approx s \left( \frac{E[x]}{\sqrt{\frac{1}{\lambda^2} + aV[x]}} \right) \quad (\text{S3})$$

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

The quality of this approximation can be evaluated by drawing samples of  $x \sim N(\mu, \sigma^2)$ , and comparing the Monte-Carlo average of  $s(\lambda|x)$  with the expression given in Equation S3. This is summarized in Figure S8 below, where the range of variation for the moments of  $x$  were set as follows:  $\mu \in [-4, 4]$  and  $\sigma^2 \in [0, 4]$ .

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



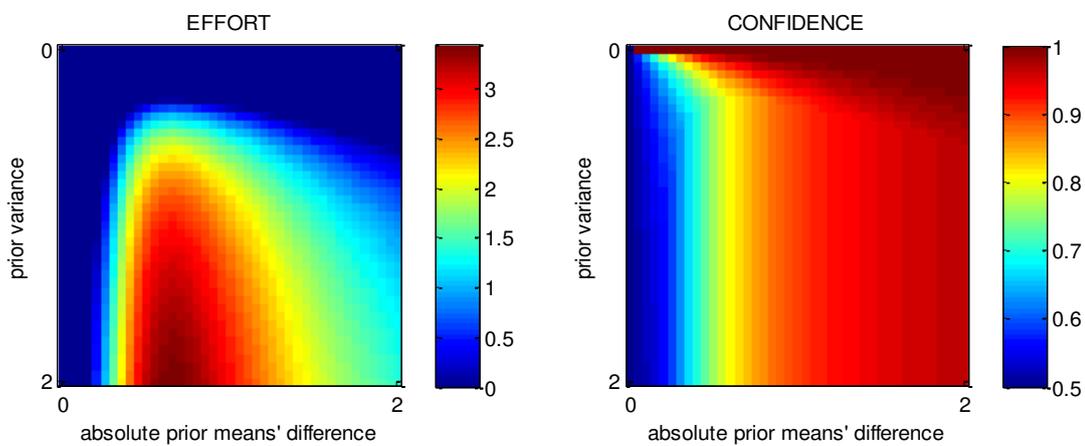
**Figure S8: Quality of the analytical approximation to  $\bar{P}$ .** Upper left panel: the Monte-Carlo estimate of  $\bar{P}$  (colour-coded) is shown as a function of both the mean  $\mu \in [-4, 4]$  (y-axis) and the variance  $\sigma^2 \in [0, 4]$  (x-axis) of the parent process  $x \sim N(\mu, \sigma^2)$ . Upper right panel: analytic approximation to  $\bar{P}$  as given by Equation S3 (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.

## 10. On the impact of model parameters for the MCD model

First, note that the properties of the metacognitive control of decisions (in terms of effort allocation and/or confidence) actually depends upon the demand for resources, which is itself determined by prior value representations. Now the way the MCD-optimal control responds to the resource demand (which is fully specified

by the prior uncertainty  $\sigma^0$  and the absolute means' difference  $|\Delta\mu^0|$ ) is determined by effort efficacy and unitary cost parameters.

Let us first 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 " $\beta$ -effect". It is depicted on Figure S9 below.

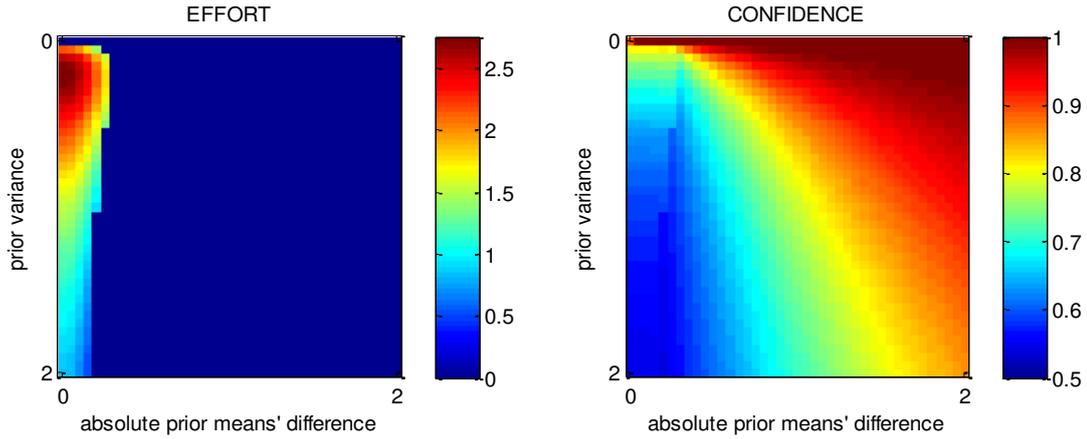


**Figure S9. The  $\beta$ -effect: MCD-optimal effort and confidence when effort has no impact on the value difference.** MCD-optimal effort (left) and confidence (right) are shown as a function of the absolute prior mean difference  $|\Delta\mu^0|$  (x-axis) and prior variance  $\sigma^0$  (y-axis).

One can see that, overall, increasing the prior variance  $\sigma^0$  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  $\bar{P}_c(\hat{z})$  always decreases with the prior variance  $\sigma^0$ . Note that, for the same reason, the MCD-optimal confidence always increases with the absolute prior means' difference  $|\Delta\mu^0|$ . Now the

impact of the absolute prior means' difference  $|\Delta\mu^0|$  on  $\hat{z}$  is less trivial. In brief, when  $|\Delta\mu^0|$  is high, the MCD-optimal allocated effort  $\hat{z}$  decreases with  $|\Delta\mu^0|$ . This is due to the fact that the resource demand decreases with  $|\Delta\mu^0|$ . However, if  $|\Delta\mu^0|$  decreases even more, it eventually reaches a critical point, below which the MCD-optimal allocated effort  $\hat{z}$  increases with  $|\Delta\mu^0|$ . This is because, although the resource demand still decreases with  $|\Delta\mu^0|$ , 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  $|\Delta\mu^0|$  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 decreases with  $\beta$  and increases with  $\alpha$ . This is because the confidence gain increases, by definition, with effort efficacy, whereas it becomes more costly when  $\alpha$  increases.

Let us now ask what would be the MCD-optimal effort  $\hat{z}$  and confidence  $\bar{P}_c(\hat{z})$  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 Figure S10 below.

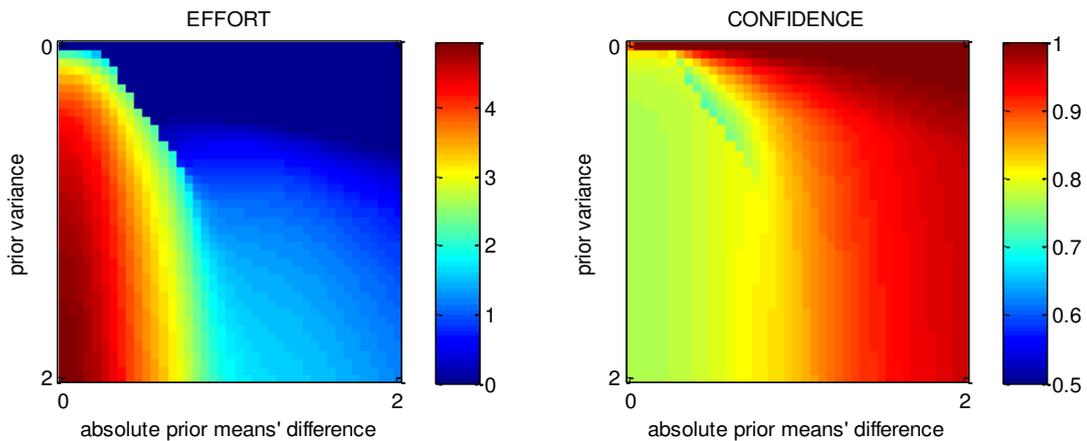


**Figure S10. The  $\gamma$ -effect: MCD-optimal effort and confidence when effort has no impact on value precision. Same format as Fig S9.**

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

Now one can ask what happens in the presence of both the  $\beta$ -effect and the  $\gamma$ -effect. If the effort unitary cost  $\alpha$  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  $|\Delta\mu^0|$  at high  $\sigma^0$  ( $\beta$ -effect) and one around some value of  $\sigma^0$  at high  $|\Delta\mu^0|$  ( $\gamma$ -effect). If the effort unitary cost  $\alpha$  decreases, then the position of the  $\beta$ -sweet spot increases and that of the  $\beta$ -sweet spot decreases, until they effectively merge together. This is exemplified on Figure S11 below.



**Figure S11. MCD-optimal effort and confidence when both types of effort efficacy are operant.**  
Same format as Fig S9.

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  $\sigma^0$  and decreases with the absolute prior means' difference  $|\Delta\mu^0|$ . The landscape of the ensuing MCD-optimal confidence level  $\bar{P}_c(\hat{z})$  is slightly less trivial, but globally, it can be thought of as increasing with  $|\Delta\mu^0|$  and decreasing with  $\sigma^0$ . Here again, this is because variations in  $|\Delta\mu^0|$  and/or  $\sigma^0$  almost always overcompensate the ensuing changes in allocated effort.

## 11. Accumulation-to-bound process models

In the main text, we compare the MCD model to two variants of accumulation-to-bound process models, namely: an optimal DDM with collapsing bounds (Tajima et al., 2016) and a modified race model (De Martino et al., 2013). We focus on these two models because both can make quantitative predictions regarding choice, value, decision time, and choice confidence.

Recall that DDMs essentially solve the problem of comparing uncertain values to make accurate choices as quickly as possible. Tajima and colleagues assume that, at the beginning of each trial, the two options have true but unknown values  $V_1$  and  $V_2$ , which the decision maker only indirectly accesses through some instantaneous noisy evidence  $x_j^t = V_j + \varepsilon_j^t$ , where  $t$  indexes time. Here,  $\varepsilon_j^t \sim N(0, \sigma)$  is some Gaussian random noise with variance  $\sigma$ , which partially masks the true values. The noise variance  $\sigma$  thus effectively controls the (un)reliability of the evidence signal  $x_j^t$ . The decision maker then progressively updates his/her posterior estimate  $\hat{V}_j^t$  by accumulating past evidence signals. When neglecting, for the sake of simplicity, the decision maker's prior belief about  $V_j$ , the decision maker's value estimate is given by:

$$\hat{V}_j^t \approx \frac{1}{t} \sum_{t'=1}^t x_j^{t'} = V_j + \frac{1}{t} \sum_{t'=1}^t \varepsilon_j^{t'}. \text{ As one can see, } \hat{V}_j^t \text{ fluctuates along with the noise, and}$$

behaves as a random walk that eventually converges towards  $V_j$ . Now the model also assumes that the decision maker pays a cost  $c$  per second of evidence accumulation. Optimal decision making then amounts to finding a policy that maximizes the expected discounted value  $\langle V_j | x_j^1, x_j^2, \dots, x_j^T \rangle - c \langle T \rangle$ . It turns out that the optimal policy is to wait until the estimated value difference  $\Delta \hat{V}^t = \hat{V}_1^t - \hat{V}_2^t$  eventually hits any of two (upper or

lower) collapsing bounds, at which point the decision maker commits to the corresponding choice. Setting these optimal collapsing bounds is done by numerically solving the so-called Bellman equation. In this work, we simply use the code written by Tajima and colleagues to perform their simulations. Let us now make two remarks on this model. First, the height of the bound at the time when it is hit measures choice confidence. This is essentially because choice confidence increases with  $\Delta\hat{V}$ .

The race model of De Martino and colleagues was specifically proposed to predict choice confidence in the context of value-based decisions. Here, separate decision variables  $\hat{V}_1^t$  and  $\hat{V}_2^t$  accumulate evidence for each option, with the decision being determined by which accumulator reaches the threshold first. At each time step, a new evidence sample  $x^t \sim N(\Delta V, \sigma)$  is drawn from a Gaussian distribution, which is a noisy measure of the value difference between the two options. The model then assumes that only the decision variable that benefits from the evidence sample is updated. For example, if  $x^t > 0$ , then  $\hat{V}_1^{t+1} = \hat{V}_1^t + x^t$  and  $\hat{V}_2^{t+1} = \hat{V}_2^t$ . This ensures that decision variables can only increase with accumulation time. Here, confidence is defined as the gap  $|\Delta\hat{V}^t|$  between the two variables when the bound is hit.

Critically, both models can make predictions about choice accuracy, decision time, and choice confidence from value ratings... and value certainty ratings (although this was never exploited before). This is because, under both models, the more uncertain people are about option values, the less reliable evidence signals will be, i.e. the higher the noise variance  $\sigma$  should be. As we see when simulating the model, increasing  $\sigma$  increases the probability of an early bound hit. For the DDM model, this implies that increasing  $\sigma$  increases choice confidence (because optimal bounds are collapsing

over decision time). In addition, under the race model, increasing  $\sigma$  increases the average gap between the two accumulators, eventually yielding the same prediction.

Lastly, let us highlight that one of the core assumption of both these models is that option values do not change during the decision. This is because the models focus on comparing option values, not on constructing them (option values are considered as inputs to the value comparison system). This effectively prevents them from being able to explain choice-induced preference change.

## Study 2 – Appendix

### *Effort proxies*

Experimental researchers often debate what should be the standard method to gauge mental effort allocation. Response time provides an objective measure, free from subjective distortions or the participants' inability to accurately gauge their allocated effort. But, RT can accumulate for reasons other than effort allocation, such as day-dreaming or other sorts of distraction, or even *lack* of effort. Self-reports of feelings of effort are less ambiguous, but subject to the unwanted influence of metacognitive limitations or even intentional manipulation. Even though each measure individual can be called into question, assessing both simultaneously might be the most satisfying solution (until something better is discovered). In our study, self-reported estimates of subjective effort positively correlated with response time, as expected (beta=0.437,  $p < 0.001$ ,  $r^2 = 0.191$ ). This suggests that RT is a fair, though not great, proxy for mental effort allocation, or that subjects have a decent, but not perfect, ability to gauge their effort exertion level. Either way, it is reassuring to find that both of these measures behaved similarly in every one of our analyses, which implies that they are indeed capturing the same information.

### *Model comparison against a “null” model*

Before fundamentally validating our model against established versions of the DDM and race models, we initially demonstrated that our effort-based depiction of the decision-making process better explains the data than a passive “null” model whose output is primarily driven by statistical noise. We designed such a model built on the idea that internal value signals are stable but noisy. As such, the ratings (both initial and final) for any particular item are drawn from a Gaussian distribution with mean equal to the true (hidden) value and variance equal to the inverse of the certainty (taken to represent precision). Choices are made according to samples drawn from the same distributions. Note that this “statistical noise model” contains no

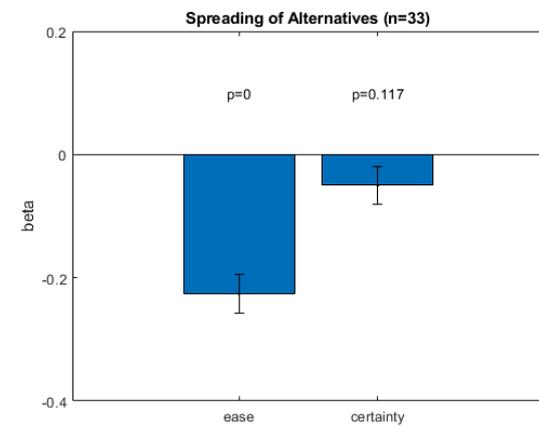
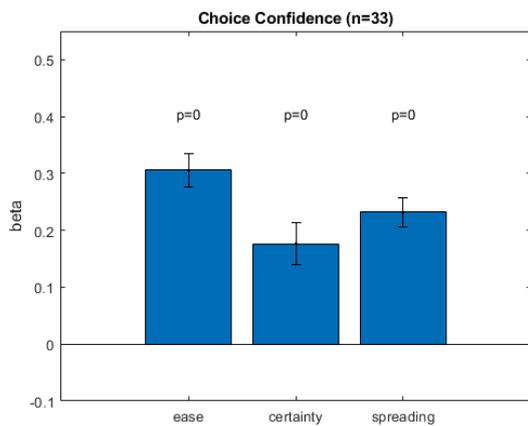
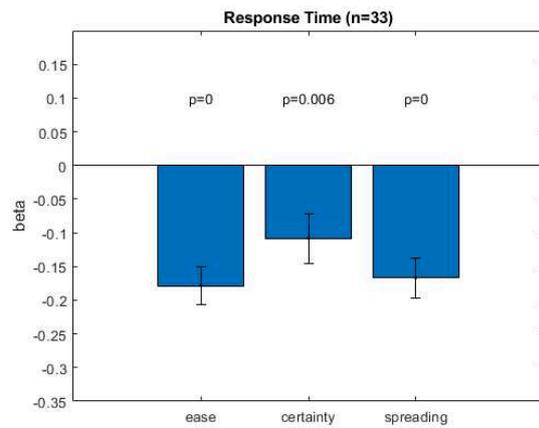
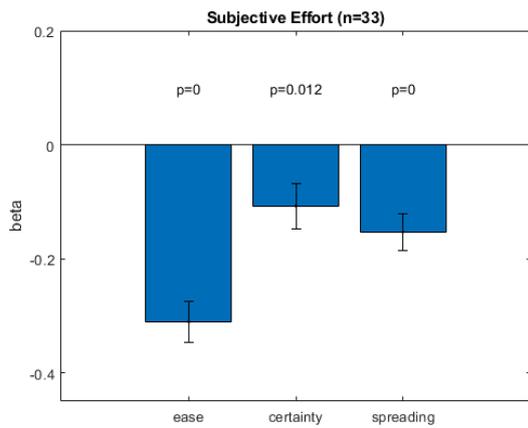
concept of either confidence or effort (or response time). One could calculate an objective probability of choosing correctly, which could be interpreted as something comparable to choice confidence. This probability could be approximated by a sigmoidal function of the absolute distance between the true values, scaled by precision. Effort could then be computed in proportion to the derivative of the confidence function, although we know of no principled reason to do so. This null model does predict spreading of alternatives, although here all apparent spreading is purely artifactual, resulting from the “regression to the mean” effect when repeatedly randomly sampling from Gaussian distributions (see Study 1 for a description of the statistical noise source of spreading). Results simulated using this null model show that it also predicts the basic data fairly well. However, the more interesting aspect of our model is not its ability to predict elementary relationships between independent and dependent variables, but rather the relationships between dependent variables (see below). The experimental data match all of these predictions. The null model, on the other hand, predicts no correlations between these dependent variables.

#### *Relations between effort, spreading of alternatives, and confidence*

Beyond the relationships between independent and dependent variables, our model can also make predictions about the relationships between the dependent variables (note: this only holds when including the implementation phase, which was excluded from the publication for our Study 2). Specifically, the model predicts a negative correlation between spreading of alternatives and effort, a positive correlation between spreading and confidence, and a (counter-intuitive) negative correlation between effort and confidence. To test these predictions in our experimental data, we used a general linear model (GLM) for each of our main dependent variables of interest (subjective effort, RT, choice confidence) with our main independent variables of interest (ease, certainty), as well as the spreading of alternatives (after controlling for the effect of ease and certainty), as regressors. The data confirm all predicted relationships

(SP-EF:  $\beta=-0.153$ ,  $p<0.001$ ; SP-RT:  $\beta=-0.167$ ,  $p<0.001$ ; SP-CC:  $\beta=0.232$ ,  $p<0.001$ ).

For the relationship between confidence and effort, we directly tested their correlations (EF-CC:  $\beta=-0.560$ ,  $p<0.001$ ; RT-CC:  $\beta=-0.365$ ,  $p<0.001$ ). The relationships survived after removing the impact of ease and certainty (EF-CC:  $\beta=-0.542$ ,  $p<0.001$ ; RT-CC:  $\beta=-0.345$ ,  $p<0.001$ ).



**Impact of choice ease and spreading of alternatives on effort and confidence**

### *Rating sum*

The data unexpectedly show a clear pattern with respect to the sum of the ratings of items in a choice pair: a negative correlation with response time and a positive correlation with choice confidence (RT:  $\beta = -0.138$ ,  $p < 0.001$ ; confidence:  $\beta = 0.303$ ,  $p < 0.001$ ). Our model makes no predictions with respect to rating sum, but perhaps it should. One possibility is that when options are both well-liked, the DM is tempted more strongly and thus chooses more quickly. Or, it could be that confidence relates not only to the probability of being correct, but also to the probability of not regretting one's choice. If both options are well-liked, there is little likelihood that choosing either option would lead to regret. In this way, confidence will get inflated, which in turn will allow the DM to stop deliberating sooner. These concepts merit further investigation, and should be considered in future modeling endeavors.

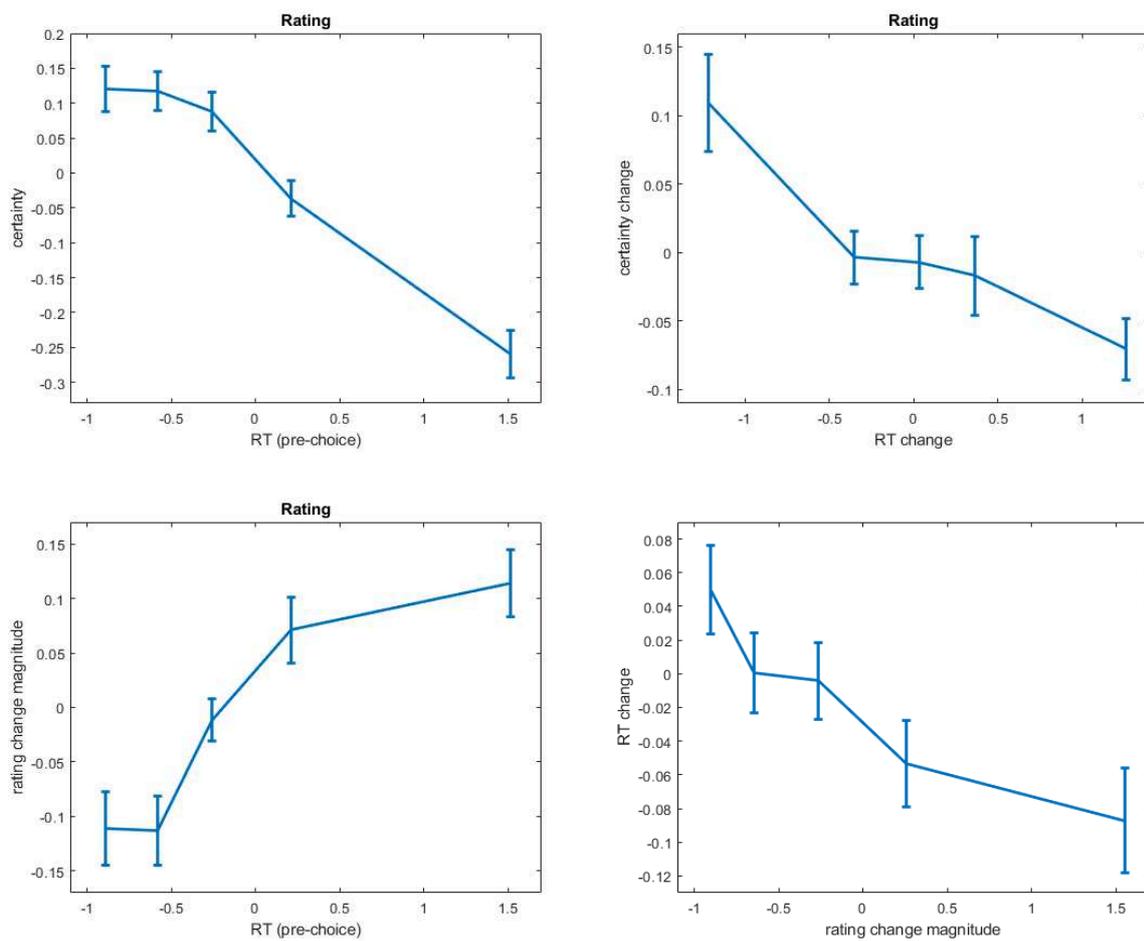
### *Familiarity*

At the end of our experiment, participants provided familiarity ratings for each item, using a slider scale to answer the question, "How well do you know this item. This familiarity variable shows the expected positive correlation with value certainty (mean correlation = 45%,  $p < 0.001$ ,  $r^2 = 0.198$ ). We also unexpectedly observed a positive correlation with value (mean correlation = 51%,  $p < 0.001$ ). Furthermore, we found that familiarity correlates negatively with both subjective effort and RT (effort: mean correlation = -16%,  $p = 0.001$ ; RT: mean correlation = -12%,  $p < 0.001$ ), and positively with choice confidence (mean correlation = 29%,  $p < 0.001$ ). Our model does not currently include the concept of familiarity, but given its apparent importance, this could be an interesting variable to explore further in future studies.

### *Ratings, certainty, and response time*

We also made some interesting observations with respect to the rating data. Specifically, we found a negative correlation between pre-choice rating RT and value certainty ( $\beta = -0.117$ ,

$p < 0.001$ ). Similarly, rating RT change correlated negatively with value certainty change ( $\beta = -0.065$ ,  $p < 0.001$ ). Pre-choice rating RT correlated positively with rating change magnitude ( $\beta = 0.084$ ,  $p < 0.001$ ). Rating change magnitude correlated negatively with rating RT change ( $\beta = -0.060$ ,  $p < 0.001$ ). Finally, across all subjects, 79% of items had lower post- vs pre-choice rating RT and ratings overall were significantly faster after the choice (mean pre-choice RT: 3.15s, mean post-choice RT: 2.47s,  $p < 0.001$ ).



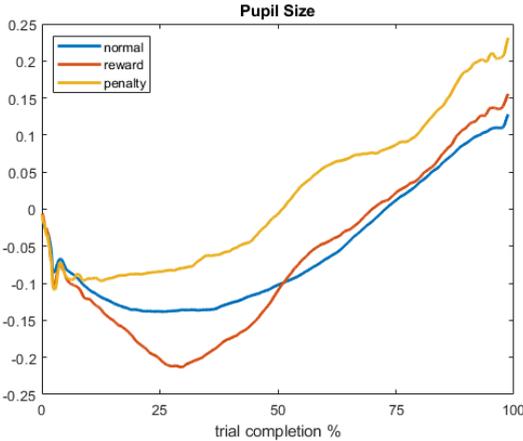
**Relationships between rating RT and rating certainty**

### *Pupil dilation across conditions*

The behavioral data fundamentally validate our model with respect to both the reward and cost manipulations. The interpretation of the pupil data is less straightforward. As our model predicts increased effort with increased reward, we expected pupil dilation to be highest in the reward condition. The pupil data thus validate our model with respect to the reward condition. As our model predicts decreased effort with increased cost, we might have expected dilation to be lowest in the penalty condition. The data would thus have seemed to contradict our model predictions (greater dilation versus ostensible predicted lesser dilation). However, we had hypothesized that dilation and viewing time are critical orthogonal components of gaze activity, which would suggest that the lower RT foreseen in the penalty condition could represent a substantial decrease in effort *duration* that was partially offset by an increase in effort *intensity* (evidenced by pupil dilation). If true, the pupil data would still be consistent with our model predictions. Interpreted in this manner, the pupil and behavioral data would actually complement each other, rather than conflicting.

Pupil dilation (averaged over the entire interval from stimulus onset until response for each trial) was greater in the penalized condition compared to the neutral. There was no difference in pupil dilation between the consequential and neutral conditions. If we instead look at momentary pupil size at the time of response, dilation was greater in consequential than neutral, and greater in penalized than consequential. However, none of these observations was significant. Still, it is interesting to consider this as evidence that mental effort can be decomposed into separate dimensions of duration (captured by RT) and intensity (captured by pupil dilation). We predicted that because the duration dimension would be disincentivized in the “penalized” condition, participants should rely more on the intensity dimension. Our data support this prediction. We also predicted that pupil dilation should be higher when choices were consequential, due to the expected higher resource allocation. Finally, the fact that pupil

dilation was greatest in the penalized condition throughout the entire duration of the trials suggests that participants invested effort pro-actively in this condition—knowing that they should respond quickly on the current trial, participants might focus their attention more intensively, even before knowing what the choice options would be.



**Mean intra-trial pupil size across conditions**

### **Study 3: Implementation of Allocated Resources**

With our Study 3, summarized in the journal article titled, “Metacognitive Monitoring during Decision-Making,” we build on the theoretical foundation and empirical support that we provided in Studies 1 and 2. Specifically, we introduce the implementation phase of our MCD model, where the resources that were allocated in the first phase are actually invested. We explain how this model is essentially a hybrid of the basic MCD and basic DDM, and that it predicts several behavioral variables better than either of the basic models alone.

# **Metacognitive Monitoring during Decision-Making**

Douglas Lee<sup>1,2</sup>, Jean Daunizeau<sup>1,2,3</sup>

<sup>1</sup> Brain and Spine Institute (ICM), Paris, France

<sup>2</sup> Sorbonne University, Paris, France

<sup>3</sup> 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](mailto:jean.daunizeau@gmail.com)

## INTRODUCTION

There is an ever-growing interest in cognitive/behavioral neuroscience in the computational properties of value-based decision-making. This sort of decision permeates virtually every aspect of our modern lives, and thus the importance of understanding the processes involved cannot be understated. Furthermore, a thorough understanding of decision mechanisms in the healthy population can help to better understand aberrations in the clinical population. While there is already an abundance of process models that attempt to explain the cognitive dynamics of simple value-based decisions, most of these are variants of the drift-diffusion model (DDM). In this type of model, “evidence” for each option accumulates over time, where the relative “evidence” is monitored until it reaches a pre-determined threshold level and the option with the most “evidence” is thus chosen (Ratcliff & Rouder, 1998). Many renditions of the DDM include a threshold that collapses over time, representing a cost (e.g., opportunity cost of time, psychological cost of impatience) that grows with deliberation time and urges the system to make a choice (Ratcliff et al, 2016). The DDM performs well with respect to fitting empirical data, and it has been shown that it indeed implements a theoretically optimal algorithm such as the sequential likelihood ratio test, which maximizes choice accuracy for a given amount of deliberation time, or minimizes deliberation time for a given level of accuracy (Tajima, Drugowitsch & Pouget, 2016). Yet, this class of model remains ambiguous about choice confidence. The level of confidence that one has about one’s choices can be conceptualized as reflecting a “second-order” metacognitive evaluation, which can be useful for informing either the current or future choices (De Martino et al, 2013; Lee & Daunizeau, 2019). One known variant of the DDM does explicitly account for choice confidence, but even this model cannot account for other important decision features like value certainty (De Martino et al, 2013). This would reflect a different type of “second-order” metacognitive evaluation, one that should positively reinforce choice confidence, but this modified DDM predicts the opposite effect (Lee

& Daunizeau, 2019). Furthermore, no version of stochastic sampling model currently exists that can account for changes of mind. These models take the subjective values of the options as pre-determined inputs to the decision system, therefore classifying any choices that deviate from prior preferences as choice errors. However, the one's ability to change one's mind while deliberating about a decision is an inherent part of the process that must not be ignored. Indeed, if deliberation could never lead one to realize that one's initial "gut feelings" were misleading, there would be no adaptive advantage to it. From this perspective, not all choices that are inconsistent with prior value estimates should be considered as errors, yet this is something that is inherently outside the capacity of the DDM.

Recent work has proposed an entirely different type of model known as the Metacognitive Control (MCD) model of decision-making (Lee & Daunizeau, 2019). Under this model, the decision process starts with an automatic evaluation of decision difficulty based on an imprecise initial estimate of the value of the options. Depending on the importance of the decision, the system calculates an optimal amount of cognitive resources to invest. Here, optimality is defined as maximizing the expected value of the resource allocation (expected value = benefit – cost), where the benefit is choice confidence scaled by decision importance, and the cost is the subjective feeling of effort (e.g., biological resource depletion, opportunity cost of time, cognitive interference). The MCD model has been shown to account well for response time and subjective effort reports (proxies for resource allocation), choice accuracy (or rather, non-changes of mind), choice confidence, and choice-induced preference change (CIPC) (Lee & Daunizeau, 2019). While already two steps ahead of the competition, there is still room for improvement. Specifically, this model cannot account for empirical data that typically show a negative relationship between response time and confidence—the MCD model predicts a positive relationship, where greater effort effectively purchases greater confidence. Similarly,

the MCD model predicts a positive relationship between effort and both CIPC and changes of mind, although these relationships have not been tested.

In this study, we introduce a hybrid MCD-DDM model which can account for all of the primary variable relationships that the MCD model can (i.e., between choice difficulty and effort, confidence, preference change, and changes of mind), as well as the secondary relationships described above. In brief, the MCD-DDM is a two-phase model, where the first phase is identical to the MCD model, and the second phase is akin to the DDM. Here, the “evidence sampling” process will stop when either of the expected effort or confidence levels is reached. We first describe the model in detail. Next, we describe a behavioral experiment designed to validate the model and contrast it to the basic DDM and MCD models. Finally, we report the experimental results and compare them with model predictions. As will be clear, the MCD-DDM model qualitatively outperforms both the basic DDM and MCD models.

## **MODEL**

The first component of the MCD-DDM model can be labeled the “allocation phase”. It is during this phase that the system establishes a “budget” for how the amount of cognitive resources it is willing to invest in the current decision. This establishes a threshold for resource allocation (the optimal upper bound), along with another threshold for choice confidence (the minimal satisficing level). These thresholds will constrain the dynamics of the second component of the MCD-DDM model, which can be labeled the “implementation phase”. It is during this phase that the allocated resources are actually invested, initiating the cognitive processes that enable the reassessment of option values. Under the assumption that information processing is serial, one cannot simply expend a pre-allocated amount of resources all at once, thus we model resources as being invested *incrementally*. The system will continue to invest additional

resources until either of the two aforementioned thresholds is reached, at which point it will make a choice. The threshold for maximal resource allocation is merely  $z^*$ , and the threshold for satisficing choice confidence is merely  $P_c^* = E[P_c|z^*]$ :

$$z^* = \underset{z}{\operatorname{argmax}} (R * (P_c(z) - P_c(0)) - C(z))$$

$$P_c^* \approx \frac{1}{1 + e^{-\pi \sqrt{\frac{\phi + \beta z^*}{3}} E[|\delta\mu + \Delta| | z^*]}}$$

where  $z$  refers to the amount of resources invested,  $R$  is the incentive or decision importance,  $P_c$  is the confidence level expected to be achieved with resource investment  $z$ ,  $C$  is the subjective effort function of resource investment,  $\delta\mu$  is the difference in the initial value estimates of the options,  $\phi$  is the precision of the difference in the initial value estimates,  $\beta$  is an efficiency parameter for the impact of  $z$  on  $\phi$ , and  $\Delta$  is a random perturbation in the difference in value estimates brought about by  $z$ . As described in more detail in Lee & Daunizeau, 2019, the value of each option is represented as  $V_i \sim N(\mu_i, \sigma_i)$ ;  $i \in \{1, 2\}$ , and the value difference perturbation is represented as  $\Delta(z) \sim N(0, \gamma z)$ , where  $\gamma$  is an efficiency parameter of  $z$  on  $\delta\mu$ . The above formula for  $P_c$  is a simple analytical approximation to the normal cumulative density function (Daunizeau, 2017).

Computationally, the implementation phase of the BCD-DDM model can be represented as a drift-diffusion process that captures the evolution of choice confidence across resource investment. Across total effort investment ( $z$ ), the momentary confidence level will depend on the initial value estimates, the accumulation of incremental random perturbations in the value estimates, and the accumulation of incremental gains in the precision of the value estimates. The confidence monitor follows a random walk because of the stochasticity of the perturbation term (see Figure 1):

$$|d\mu_0| = d\mu(z = 0)$$

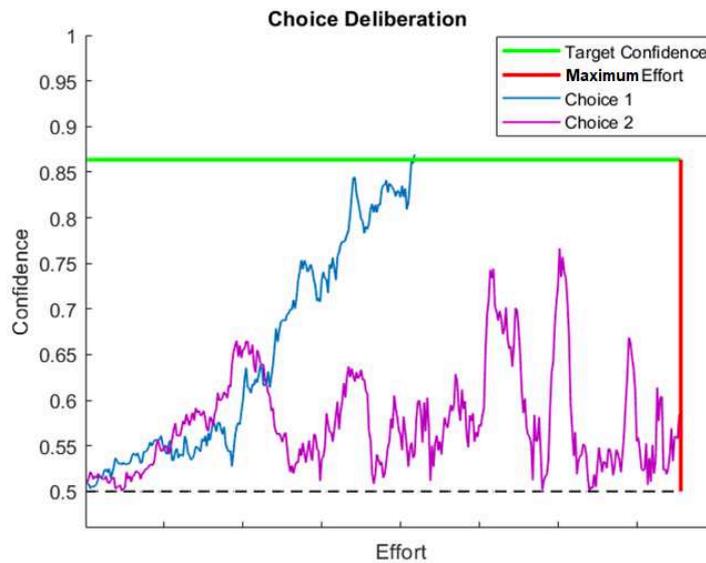
$$|d\mu_z| = |d\mu_{z-1} + \Delta|, \quad \Delta \sim N(0, \gamma), \quad z > 0$$

$$\phi_0 = \phi(z = 0)$$

$$\phi_z = \phi_{z-1} + \beta, \quad z > 0$$

$$P_z = \frac{1}{1 + e^{-\pi \sqrt{\frac{\phi_z}{3}} |d\mu_z|}}$$

In concurrence with the evolution of choice confidence, a resource monitor will tally the accumulation of invested resources. As stated above, the system will continue to incrementally invest additional resources until one of the thresholds is reached by its respective monitor ( $z = z^*$  or  $P \geq P^*$ ).



**Figure 1:** The implementation phase of cognitive resource allocation during a decision can be represented by a drift-diffusion process. Here, the two curves represent two individual decision trials where the decision maker (DM) initially had identical beliefs about the options. The starting point (i.e., with zero effort) shows the default level of choice confidence if the DM declines to deliberate, which is based solely on his uncertain prior beliefs. Through effort, the allocated resources allow for additional information to be processed, which in turn impacts how confident the DM feels about the imminent choice. The horizontal threshold represents the confidence level that the DM seeks to achieve prior to committing to a choice. The vertical threshold represents the maximum amount of resources that the DM is willing to allocate before terminating deliberation with a choice. Here, Choice 1 represents a situation where the information that the DM processed was helpful early on, allowing him to reach a high level of confidence and stop deliberating sooner than anticipated. Choice 2 represents a situation where the processed information was conflicting and unhelpful, and the DM ended up making his choice with a lower-than-desired level of confidence.

## RESULTS

First, we contrast the MCD-DDM against basic versions of its component models, namely: the drift-diffusion model (DDM) and the metacognitive control of decisions model (MCD). Each of these models makes quantitative predictions regarding the impact of pre-choice value and value certainty ratings on choice accuracy, response time, and choice confidence, as well as the relationships between each of these dependent variables. The MCD and MCD-DDM also make quantitative predictions about the impact of value ratings and certainty on choice-induced preference change (CIPC), and the relationship between CIPC on both response time and confidence. We compare the model predictions to data from two previous studies (Lee & Daunizeau, 2019), which we describe in the Supplementary Material.

- **Impact of value estimates and certainty on accuracy, response time, and confidence**

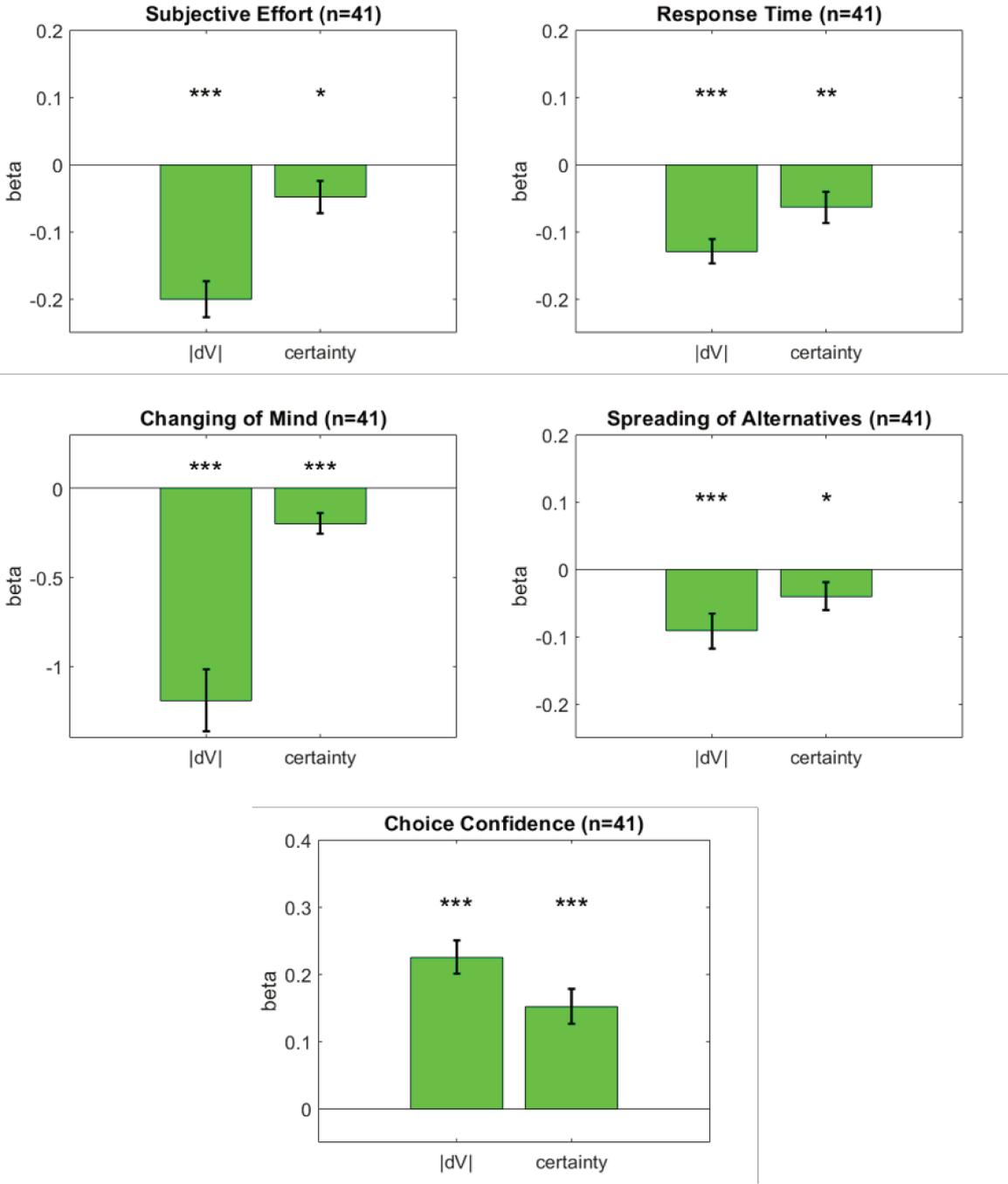
In each of the models we are comparing, a decision is said to be more difficult when the options are of similar value (i.e., the absolute distance between the value estimates are smaller). In the DDM, this difficulty manifests itself in the drift rate, where options whose values are close together have a lower drift rate than those whose values are far apart. In this way, choices between similarly-valued options (i.e., low drift rates) will be slower, more prone to errors, and less confident (note that in versions of the DDM with a constant threshold, trial-by-trial confidence levels cannot vary). The MCD makes similar predictions, but for different reasons. Recall that in the MCD, a difficult choice will call for a greater cognitive resource allocation. In this way, similarly-valued options will lead to greater mental effort (e.g., longer deliberation time). This effort, in turn, will lead to greater choice-induced preference change (CIPC) and more changes of mind (known in traditional models as choice “errors”). Preference change is measured by the so-called spreading of alternatives (SoA), which is simply the distance between the value estimates of the chosen and unchosen options post-choice versus pre-choice. The idea is that CIPC is not random, but rather in line with the pending choice. Choice confidence in the

MCD is simply the probability that the thought-to-be best option is indeed the best option, therefore it decreases with option value proximity. The predictions of the MCD-DDM follow those of the MCD, with respect to value proximity.

The second factor that directly impacts choices is the uncertainty in the comparison process. With the DDM, this uncertainty is typically described as noise in the decision system (e.g., neural firing stochasticity). Greater noise in the system will reduce choice accuracy, as one might expect. However, greater noise will counterintuitively lead to faster responses and higher choice confidence (Lee & Daunizeau, 2019). With the MCD, uncertainty refers to the lack of precision a decision maker feels about the options' value estimates. Value uncertainty is thus an additional component of choice difficulty under the MCD, and its impact is qualitatively the same as value proximity. Higher uncertainty will lead to higher mental effort (e.g., slower responses), greater choice-induced preference change, and more changes of mind. As with value proximity, choice confidence will decrease with value uncertainty. The predictions of the MCD-DDM follow those of the MCD, with respect to value uncertainty. Note that the DDM predicts the opposite direction as both the MCD and MCD-DDM, with respect to the impact of uncertainty on response time (RT) and confidence.

We now turn to the empirical data. To validate the model predictions, we ran, for each participant, separate multiple logistic regressions of RT, subjective effort, SoA, changing of mind (CoM) frequency (the complement of choice accuracy), and choice confidence against value distance (VD) and value certainty (VC). Random effect analyses shows that all regressors for all dependent variables have significant effects at the group level (see Figure 2). VD and VC negatively impact RT, effort, CoM, and SoA, and positively impact confidence (subjective effort: VD beta=-0.311, p<0.001; VC beta=-0.107, p=0.006; RT: VD beta=-0.179, p<0.001; VC beta=-0.108, p=0.003; changing of mind: VD beta=-1.190, p<0.001; VC beta=-0.199, p<0.001; SoA: VD beta=-0.226, p<0.001; VC beta=-0.050, p=0.059; choice confidence: VD

beta=0.306, p<0.001; VC beta=0.177, p<0.001). The data thus conform to the predictions of the MCD and MCD-DDM, but not those of the DDM.



**Figure 2: Impact of choice ease:** Here, we show the impact of choice ease (value distance and value certainty) on each of our dependent variables of interest. Ease has a negative impact on effort (both subjective effort and RT) and preference change (both changes of mind and SoA), and a positive impact on confidence.

- **Relationships between response time, confidence, and spreading of alternatives**

Beyond the impact of choice difficulty on RT, confidence, and SoA, the models that we consider can be further distinguished by examining the relationships they predict *between* RT, confidence, and SoA. We will start with the relationship between RT and confidence. With the DDM, this relationship is trivial, since confidence is merely proportional to the height of the evidence threshold at the time of choice, and the threshold collapses as a function of time. Thus, the DDM predicts a negative correlation between response time and confidence (again, this relationship disappears for versions of the DDM without collapsing bounds). The MCD predicts no relationship between RT (mental effort) and confidence. The reason for this is twofold. First, we have to consider the impact of choice difficulty on both effort and confidence: effort will increase and confidence will decrease with difficulty. Second, effort itself leads to higher confidence. So, the two effects offset each other and the net relationship is null. When controlling for choice difficulty (VD and VC), the MCD predicts no relationship between effort and confidence, because these are deterministic expected values under this model. The MCD-DDM follows the spirit of the MCD, but incorporates trial-by-trial variability like the DDM. Like the DDM, the MCD-DDM predicts a negative relationship between RT (mental effort) and confidence, but for different reasons. Recall that the MCD-DDM contains two stopping rules for incremental mental effort investment: either momentary confidence reaches the target threshold, or cumulative effort reaches the maximum limit. An algorithm of this sort automatically establishes a negative correlation between confidence and effort: deliberation will either stop earlier than expected with high confidence, or late with lower confidence than expected (see Figure 1 above).

We now turn back to the empirical data. To validate the model predictions, we ran, for each participant, separate multiple logistic regressions of choice confidence against value distance (VD), value certainty (VC), and either RT or subjective effort. Random effect analyses shows

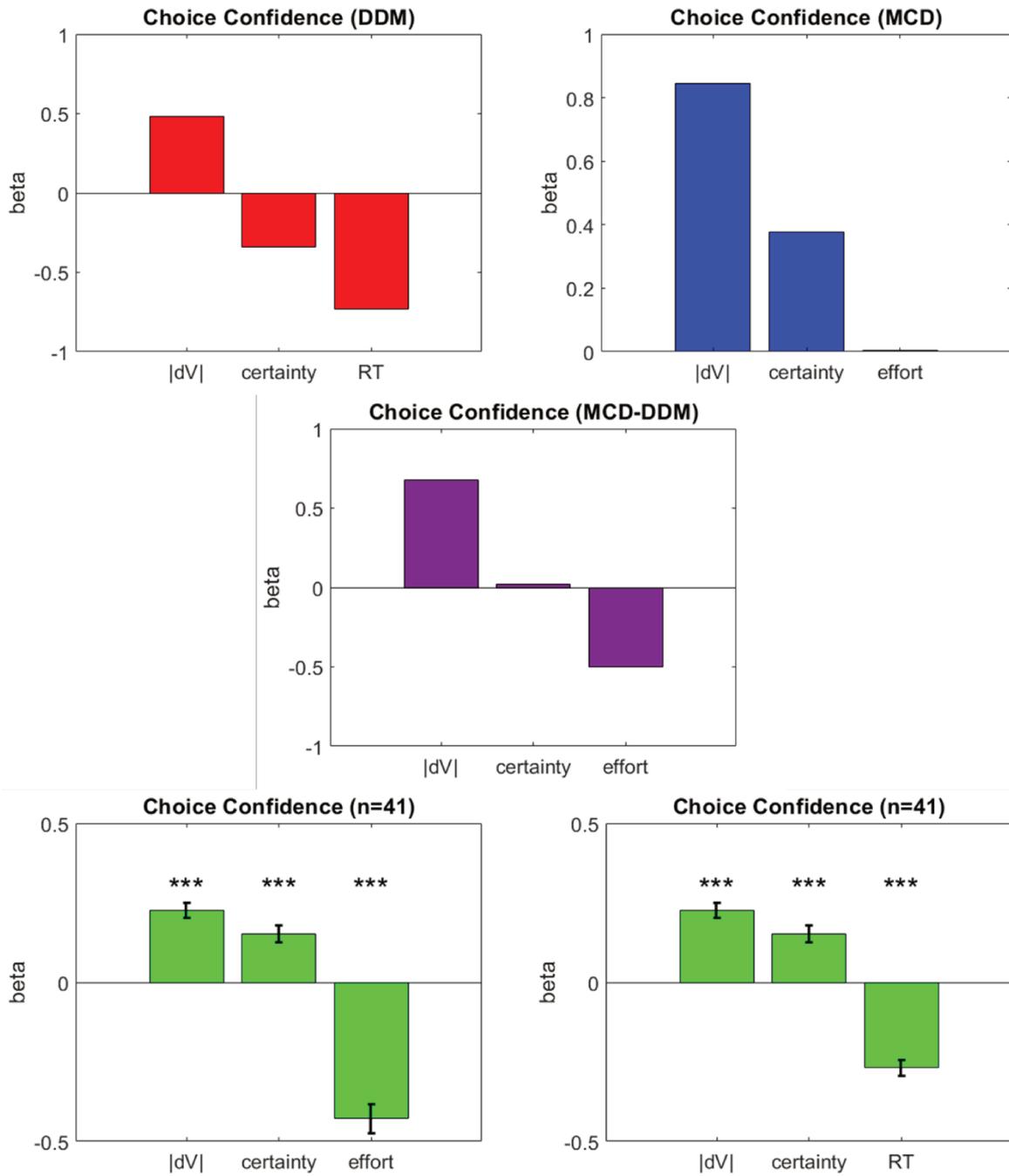
that all regressors for all dependent variables have significant effects at the group level (see Figure 3). VD and VC positively correlate with confidence, as before, and RT and effort each negatively correlate with confidence (RT-confidence:  $\beta = -0.345$ ,  $p < 0.001$ ; effort-confidence:  $\beta = -0.542$ ,  $p < 0.001$ ).

We now examine the relationship between SoA and both confidence and effort. Here, we will consider only the MCD and MCD-DDM, because the DDM does not allow for SoA. We will start with confidence. The MCD predicts a positive relationship between SoA and confidence. The reason for this is as follows. First, the impact of choice difficulty on both SoA and confidence leads to a negative relationship between them: SoA will increase and confidence will decrease with difficulty. Second, SoA itself leads to higher confidence (see below), creating a positive relationship. Here, the impact of SoA on confidence dominates the joint impact of choice difficulty on SoA and confidence. When controlling for choice difficulty (VD and VC), the MCD predicts no theoretical relationship between SoA and confidence, because these are deterministic expected values under this model. However, traces the positive relationship remain when examining the variables by GLM analysis, due to the complex nonlinear nature of the true relationship. With the MCD-DDM, SoA on a given trial can alternatively be labeled “increase in value distance”, so it can effectively be considered as a decrease in choice difficulty that leads directly to confidence. Hence, the MCD-DDM predicts a positive relationship between SoA and confidence.

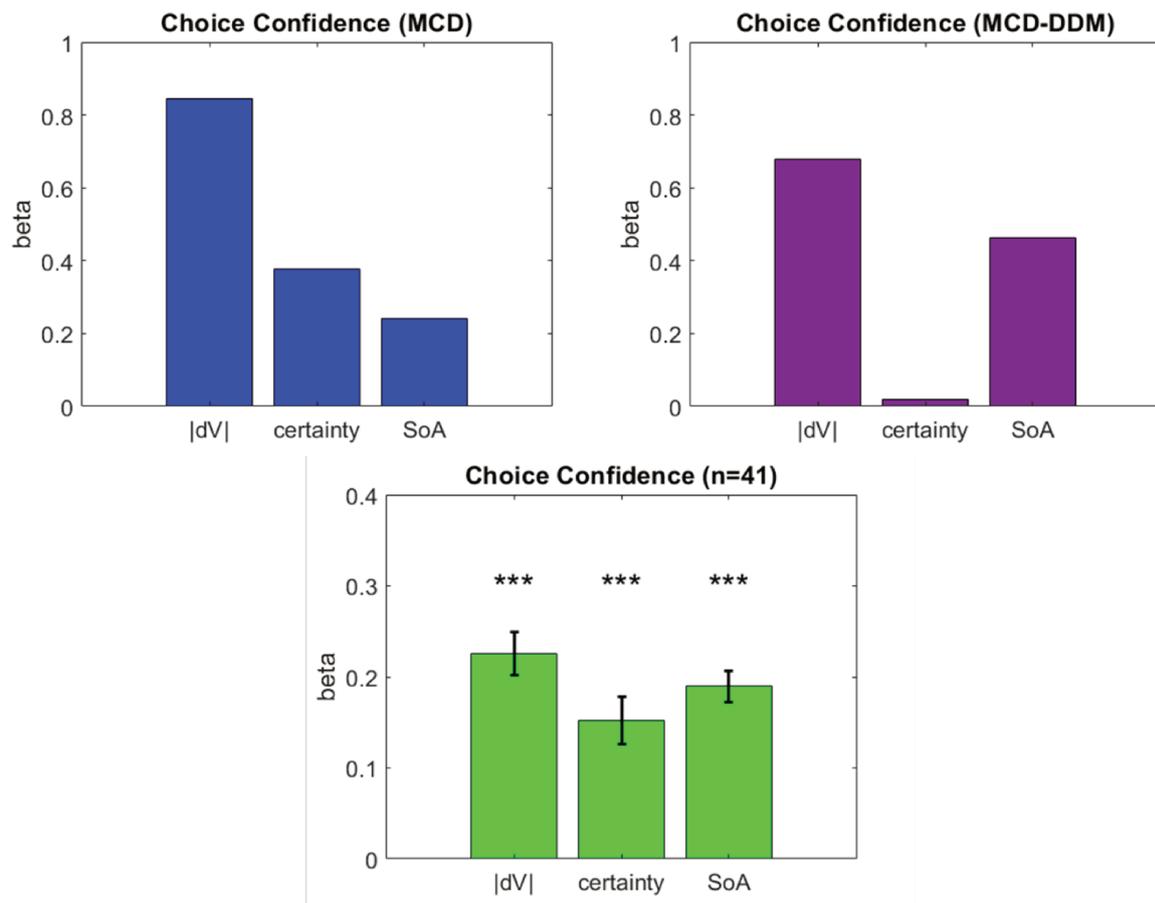
We now turn to effort. The MCD predicts a positive relationship between SoA and effort. Here, choice difficulty impacts positively on both variables, which leads to a positive relationship between them. Furthermore, effort itself leads to higher SoA, augmenting the positive relationship. Indeed, under the MCD, SoA is actually created by effort, making the positive relationship trivial. Evidence of this relationship remains even after controlling for choice difficulty (VD and VC) via GLM. With the MCD-DDM, trial-by-trial variability causes the

relationship to reverse. Recall that under the MCD-DDM, the expected effort, SoA, and confidence are all identical to what would be expected under the MCD. However, on a given trial, the decision maker will cease deliberation if momentary choice confidence reaches the target threshold, even if the expected mental effort allocation has not been fully implemented. Thus, higher SoA causes a higher likelihood of stopping the decision process early. Hence, the MCD-DDM predicts a negative relationship between SoA and effort.

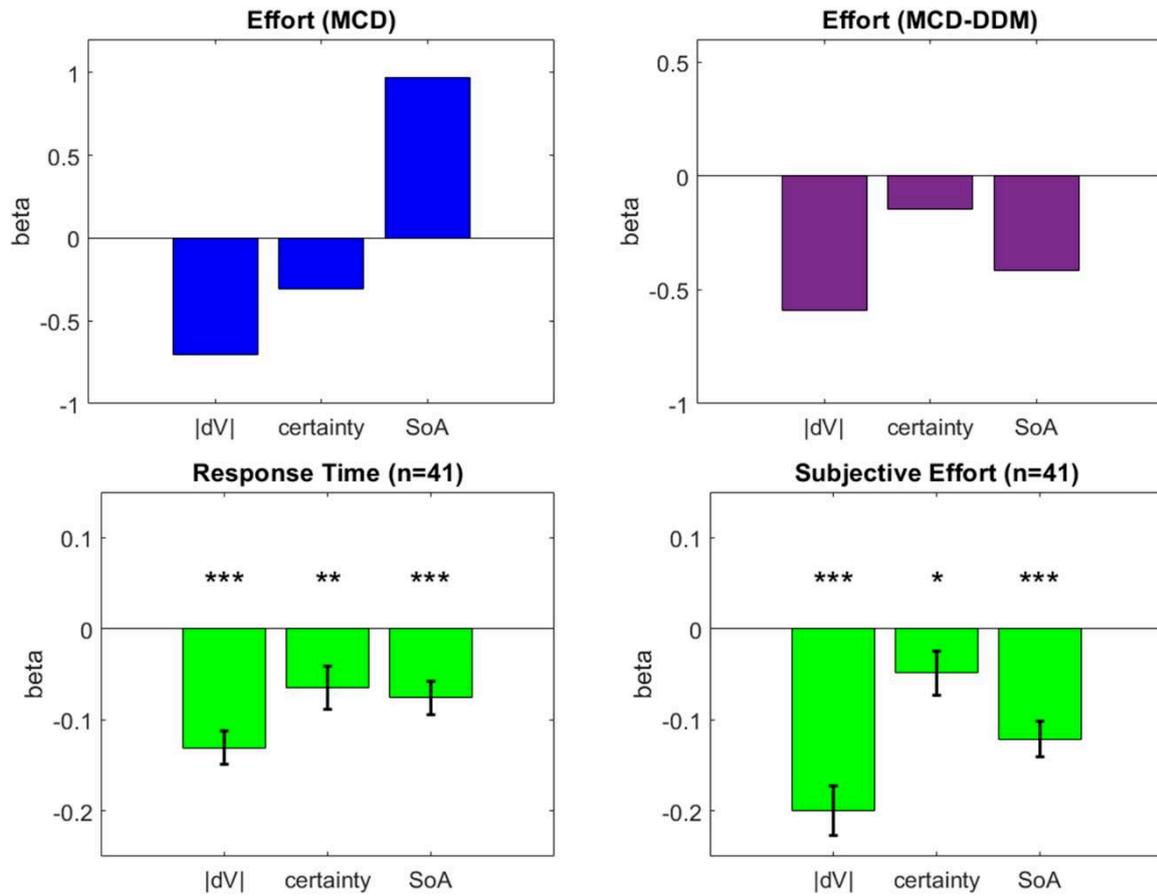
We now turn back to the empirical data. To validate the model predictions, we ran, for each participant, separate multiple logistic regressions of choice confidence and both RT and subjective effort against value distance (VD), value certainty (VC), and SoA. Random effect analyses shows that all regressors for all dependent variables have significant effects at the group level (see Figures 4 & 5). VD and VC correlate positively with confidence and negatively with both RT and effort, as before. SoA correlates positively with confidence (SoA-confidence:  $\beta=0.232$ ,  $p<0.001$ ) and negatively with both RT and effort (SoA-effort:  $\beta=-0.153$ ,  $p<0.001$ ; SoA-RT:  $\beta=-0.167$ ,  $p<0.001$ ).



**Figure 3: Impact of effort on confidence:** Intuitively, effort should invoke confidence. However, under the MCD-DDM, decision deliberation terminates with either high confidence and relatively low effort, or high effort and relatively low confidence. Here, we show the negative impact of effort (both subjective effort and RT) on confidence, after controlling for the impact of choice ease (absolute value difference and value certainty) (green plots). The data support the MCD-DDM (purple plot), but not the basic DDM (red plot) or MCD (blue plot).



**Figure 4: Impact of spreading of alternatives:** Here, we show the positive impact of SoA on confidence, after controlling for the impact of choice ease (absolute value difference and value certainty) (green plot). The data partially support the MCD-DDM (purple plot), but not the basic MCD (blue plot). The basic DDM ignores SoA.



**Figure 5: Impact of spreading of alternatives:** SoA can effectively be thought of as “gain in choice ease,” and should therefore have the same impact as choice ease itself. Here, we show the negative impact of SoA on effort (both subjective effort and RT), after controlling for the impact of choice ease (absolute value difference and value certainty) (green plots). The data support the MCD-DDM (purple plot), but not the basic MCD (blue plot). The basic DDM ignores SoA.

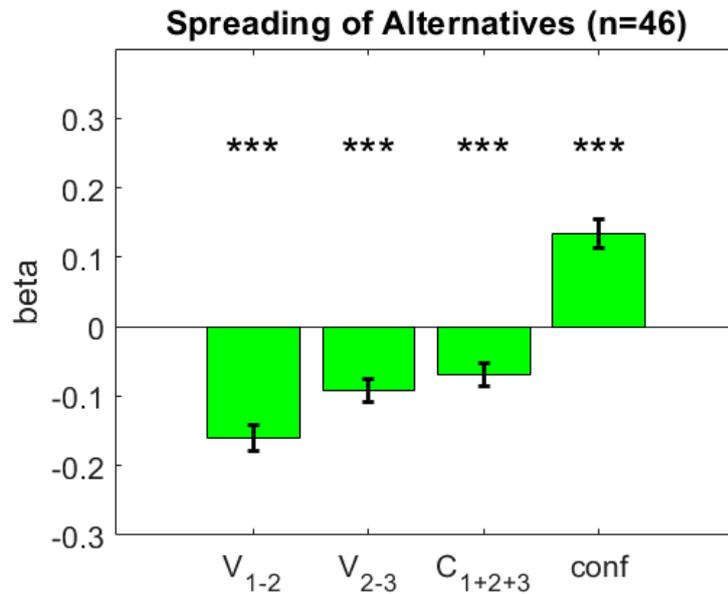
## **DISCUSSION**

In this study, we have introduced a new model of decision-making that is effectively a hybrid of two previous models, namely: the drift-diffusion model (broadly known and used across multiple fields) and the metacognitive control of decisions model (recently introduced in Lee & Daunizeau, 2019). Each of the component models offers its own strengths and weaknesses. The DDM has an impressive performance record in terms of predicting choice accuracy and response time distributions, but it cannot account well for metacognitive variables like choice confidence and rating certainty. Neither can it account for choice-induced preference change (CIPC), by design. The MCD, on the other hand, has not yet had the opportunity to prove itself, but initial studies have demonstrated that it can predict choice accuracy and RT, as well as choice confidence and CIPC. The MCD can also properly predict the impact of value certainty on accuracy, RT, confidence, and CIPC. However, the MCD cannot account for the inter-trial correlations between CIPC, RT, and confidence. The DDM, on the other hand, does predict the relationship between RT and confidence (but again, not CIPC). With our new MCD-DDM model, we can predict all of the aforementioned variables, as well as their inter-correlations. We demonstrated our model predictions with simulated data, and validated them with empirical data. In sum, we believe that the MCD-DDM model will soon find its place in the toolkit of researchers of decision-making, particularly those who have an interest in metacognitive factors in addition to overt choice behavior.

## REFERENCES

- Daunizeau, J. (2017). Semi-analytical approximations to statistical moments of sigmoid and softmax mappings of normal variables. arXiv preprint arXiv:1703.00091.
- De Martino, B., Fleming, S.M., Garrett, N., and Dolan, R. (2013). Confidence in value-based choice. *Nat Neurosci* 16, 105–110.
- Lee, D., & Daunizeau, J. (2019). Trading effort for confidence: the metacognitive control of decision-making. *bioRxiv*.
- Lee, D., & Daunizeau, J. (2019). Choosing what we like vs liking what we choose: How choice-induced preference change might actually be instrumental to decision-making. *bioRxiv*, 661116.
- Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in cognitive sciences*, 20(4), 260-281.
- Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological science*, 9(5), 347-356.
- Tajima, S., Drugowitsch, J., & Pouget, A. (2016). Optimal policy for value-based decision-making. *Nature communications*, 7, 12400.

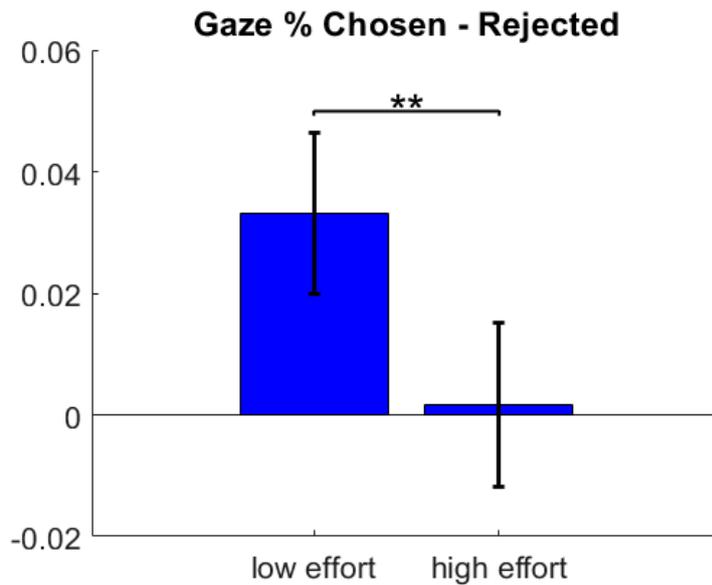
## Supplementary Material



**Figure S1:**

### **Eye-Tracking**

From the eye-tracking data, we were able 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, chosen items were gazed at more than rejected items, but we also found that this effect was limited to low effort choices. Figure S7 below shows the gaze bias for low and high effort trials, based upon a median-split of subjective effort (gaze bias mean difference=0.032, s.e.m.=0.016,  $p=0.055$ ). These results were previously reported in Lee & Daunizeau, 2019.



**Figure S2. 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.

During this current study, we then hypothesized that for difficult choices, those that eventually achieved higher spreading of alternatives (SoA), i.e., those that eventually became easier, would eventually have a higher gaze bias. Supporting this hypothesis, we found that SoA positively correlates with gaze bias (mean=0.054, s.e.m.=0.026,  $p=0.025$ ). The story we propose to explain these findings is as follows. During a simple choice like the ones in our experiment, people will gaze back and forth at each option while pondering which they prefer. If the choice is easy (i.e., one option clearly has a higher value than the other), people will quickly shift their gaze towards the preferred option. Otherwise, they will continue to shift their gaze until eventually making a choice. During the gazing time, one of two things will happen. Either the deliberation process will cause SoA, thus making the choice easier, thus creating a gaze bias; or, no SoA will occur, the choice will remain difficult, and no gaze bias will emerge. We believe

that the gaze patterns that we have observed are in line with and thus lend support to our claim that the MCD-DDM accurately represents the dynamics of the decision process.

## **Methods**

The experiment was developed using Matlab and PsychToolbox. The experiment was conducted entirely in French. 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.

### *Participants*

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 each participant was paid a flat rate of 20€ as compensation for his time plus an average of 4€ as a bonus. One group of 11 participants was excluded from the cross-condition analysis only (see below), due to technical problems. All participants were native French speakers. All participants were from the non-patient population with no reported history of psychiatric or neurological illness.

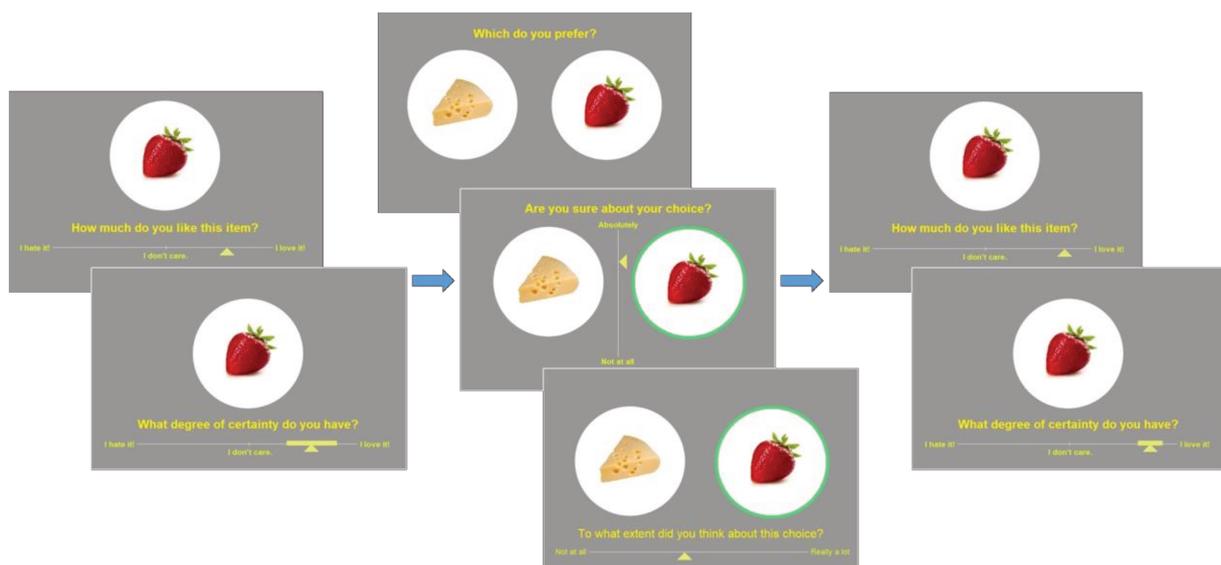
### *Materials*

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.

### *Procedure*

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).



**Figure S1. 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 this value rating (2nd-order 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).

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., Lee and Daunizeau, 2019): pre-choice item rating, choice, and post-choice item rating (see Figure S1 above). There was no time limit for the overall experiment, nor for the different sections, nor for the individual trials. Item rating and choice sessions are described below.

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 rating as the pre-choice first-order value rating. Participants then had to respond to the question, “How certain are you 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 second-order certainty rating. At that time, the next trial began.

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 decision 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.

#### **Study 4: Multiple Decision Options**

With our Study 4, summarized in the journal article titled, “Metacognitive control of multi-option decision-making,” we build on the theoretical foundation and empirical support that we provided in Studies 1-3. Specifically, we demonstrate that our MCD model is equally-well suited for value-based decision between multiple (i.e., more than two) options. We present the mathematical details of our multi-option MCD model, as well as experimental evidence to validate it and distinguish it from competitor models.

# Metacognitive Control of Multi-Option Decision-Making

## INTRODUCTION

Most of the decisions that we are faced with in our lives include many different options to choose from. This is true for consumer decisions (e.g., which groceries to buy, which apartment to rent), financial decisions (e.g., which portfolio to invest in, which insurance policy to enroll in), and even social decisions (e.g., which mate to pursue, which colleagues to mingle with). In the modern age, for most decisions, there is an increasing number and variety of options, as well as an increasing technical feasibility of realistically considering any of the potential options. While this obviously offers great advantages to a decision maker (DM), it simultaneously offers some important disadvantages. For one thing, the size of a set of options can overwhelm a DM with more information than he can efficiently process, potentially leading to suboptimal choice behavior. Published evidence of this sort of effect dates back to Hick [1952], who noted that people take longer to respond when they are presented with more options to choose from. Known as Hick's law, this phenomenon suggests that more options causes greater choice difficulty or uncertainty, which could lead to lower accuracy. Indeed, it is widely known from empirical evidence that choices become more stochastic when the choice set contains more elements. The literature on so-called "choice overload" shows that when the number of choice set options increases, the level of confidence or satisfaction that a DM has with his choice decreases, and the level of regret increases [see Chernev, Bockenholt & Goodman, 2015, for a review]. Choice overload has even been supported with neuroimaging, showing that neural activity in regions of the brain known to encode value (striatum and anterior cingulate cortex) initially increased with option set size, but then decreased as option set size increased further [Reutskaja et al, 2018]. This could be interpreted through the lens of the divisive normalization theory. Divisive normalization holds that neural activity representing the

value of each option is normalized by the sum of all option values [Louie, Khaw & Glimcher, 2013]. In this way, if the DM was presented with many options to consider, their neural representations would be sparse and therefore precision would be low. It would then immediately follow that choices between options represented in such a way would be more stochastic. Finally, the specific composition of a set of options can have unexpected effects on choice behavior, with inconsistencies arising according to decision context. One frequent observation is that when a third option is added to a choice set containing a pair of options, choice stochasticity increases. Specifically, choice stochasticity increases as a function of the value of the third option (even when it is always strictly lower in value than the first pair of options). This is explainable by divisive normalization [Louie, Khaw & Glimcher, 2013], but also by theories of attention. For example, as the value of the third option (often referred to as the “distractor”) increases, it captures more attention from the other options, thus causing interference in the comparison between the first and second options [Gluth, Spektor & Rieskamp, 2018].

Clearly, understanding situations where a DM is faced with more than just two options is important for the field of decision-making research. As computational modeling is used more extensively to search for mechanistic cognitive explanations for decision behavior, it is disappointing that there are so few computational models that can be applied for decisions with multiple (e.g., more than two) options. There do exist a few published models of the decision process that work with multiple options. For example, consider the race model, which consists of a separate “value evidence” accumulator variable for each option, where the first accumulator to reach a pre-determined threshold level is chosen as the winner and the process stops [Vickers, 1979]. The race model will work for an arbitrarily large set of options. However, it is known that the race model does not perform well when accounting for empirical data [Ratcliff & Rouder, 1998]. The most popular decision process model, the drift-diffusion model (DDM),

does perform very well when accounting for decision accuracy and response time [Ratcliff et al, 2016]. The DDM is similar to the race model, but it consists of only one accumulator variable, which records the difference in the “value evidence” between the two options. While not inherently obvious how this sort of model could be expanded to option sets greater than two, some authors have presented their approach to do just that [Krajbich & Rangel, 2011]. One group has actually formally demonstrated that the DDM can provide the optimal solution for choices between any number of options [Tajima et al, 2019]. Even still, these models cannot predict important metacognitive variables such as choice confidence. One recent model succeeded in accounting for choice confidence using a hybrid race-DDM [De Martino et al, 2013], but even that model fails when taking uncertainty about option values into account. Furthermore, there exists no model (to our knowledge) that attempts to account for a different sort of phenomenon that has been well-documented—choice-induced preference change (CIPC). This phenomenon demonstrates that decision-makers tend to change their value estimates for options after they are asked to choose between those options that they value similarly [e.g., Lee & Daunizeau, 2019; Voigt et al, 2019; Izuma et al, 2015]. While the underlying cause of this phenomenon is still debated, the phenomenon itself is robust. Nevertheless, decision models like the DDM, whose usage is widespread across many disciplines (e.g., neuroscience, psychology, economics), explicitly preclude the possibility of CIPC by design. With these models, the “evidence” that accumulates is a stochastic signal whose mean relates to the “true” value of the options, and this signal is stationary.

In this current work, we present a computational model that directly incorporates both choice confidence and CIPC. This model is a simple extension of our basic (two-option) model [Lee & Daunizeau, 2019]. Not only does our model perform well with respect to choice confidence and CIPC, it also outperforms any known version of race model or DDM with respect to response time (by taking into consideration certainty about option values). The performance of

our model is on par with the DDM with respect to error rate, although under our theory, so-called “errors” are instead classified as “changes of mind”. We first outline the mathematical details of our model, and then validate it using a behavioral experiment. Finally, we examine the differences in the predictions of our model, the race model, and the DDM.

## MODEL

The aim of this model is to describe the cognitive mechanisms that are engaged during a subjective value-based choice task between more than two options (although it would also work with two options). The agent will choose the option that it believes to be best, but only if its confidence reaches a target threshold. At any point during the deliberation process, if the agent’s instantaneous confidence level is sub-threshold, it will incrementally invest mental effort in order to process additional information in hopes of achieving greater confidence. Upon reaching either a satisficing level of confidence or exhausting the full amount of effort allocated to the task, the agent will finalize its choice.

The core of this multi-option computational model is similar to our basic two-option version (see above). Recall that our model is based on maximizing the expected value (EV) of mental resource allocation. Effort investment enhances performance, but carries a cost. The EV is thus comprised of a benefit term (B) and a cost term (C):

$$EV = E[B - C] \tag{28}$$

where  $E[.]$  refers to an expectation that accounts for predictable stochastic influences that ensue from allocating resources (details below). The benefit term can be broken down into the reward (R) for choosing the best option times the probability of choosing the best option ( $P_c$ ):

$$B = R * P_c \tag{29}$$

Recall that the reward scaling factor  $R$  captures the importance of making a "good" decision, whereas  $P_c$  really measures one's confidence in making a "good" decision. The true benefit of resource allocation, however, is the expected *gain* in confidence. The expected value of control (EVC), as well as each of its components, will be a function of the amount of cognitive resources invested in the task (which we label  $z$ ):

$$EVC(z) = R * (P_c(z) - P_c(0)) - C(z) \quad (30)$$

The only requirement for  $C$  is that it be monotonically increasing (in this paper, we assume it is linear for simplicity and without loss of generality). The EVC equation can thus be simplified as follows:

$$EVC(z) = R * (P_c(z) - P_c(0)) - \alpha z \quad (31)$$

where  $\alpha$  is simply the cost of effort per unit of resources.

To derive an expression for  $P_c$ , we will refer to a simple decision task where multiple options are available to choose from, with respect to a single dimension of comparison. For each option, we represent the DM's *pre-choice* subjective value ( $V$ ) as a normally distributed random variable with mean equal to the perceived value and variance equal to the uncertainty of that value:

$$V_i \sim N(\mu_i, \sigma_i); i \in \{1, \dots, n\}, n: \text{number of options} \quad (32)$$

In other words, the DM is endowed with a probabilistic representation of option values, whose mode  $\mu$  measures the strength of the subjective feeling of liking/wanting the option, and whose width  $\sigma$  measures how uncertain that feeling is. This probabilistic representation evolves along with the allocation of resources during the decision process.

When making a decision, however, it is not the individual option values that are important, but rather the relative value of each option with respect to the other options in the set. In particular,

the DM will search for the best option in the set of options. The best option is defined as the option whose value is higher than the value of each of the other options, and is therefore higher than the maximum of all of the other options. Selecting this best option is equivalent to making the “correct choice”:

$$\text{correct choice} \stackrel{\text{def}}{=} \text{option } i : V_i > V_j, \forall j \neq i \quad (33)$$

Because the value of each option is not precisely known, but rather represented as a probability distribution, the DM will choose according to the expected values of the options. The best option is simply the one with the highest mean. The other options will be summarized in terms of their expected maximum value:

$$\mu_{max} \stackrel{\text{def}}{=} E[\max(V_j)]; j \in \{1, \dots, n\} \setminus \{i\} \quad (34)$$

For example, for a choice set containing three options, with the first option having the highest expected value, the expected maximum value of the options other than the best is calculated as follows (Nadarajah & Kotz, 2008):

$$V_{max} = \max(V_2, V_3) \sim N(\mu_{max}, \sigma_{max}) \quad (35)$$

$$\mu_{max} = \int x f(x) dx \quad (36)$$

$$f(x): \text{pdf of } V_{max} \quad (37)$$

$$f(x) = \text{pdf}\left(\frac{x-\mu_2}{\sigma_2}\right) * \text{cdf}\left(\frac{x-\mu_3}{\sigma_3}\right) + \text{pdf}\left(\frac{x-\mu_3}{\sigma_3}\right) * \text{cdf}\left(\frac{x-\mu_2}{\sigma_2}\right) \quad (38)$$

For a choice set with four options, with the first option having the highest expected value, an iteration of this procedure will provide the expected value of the maximum of the options other than the best:

$$\mu_{max} = E[\max(\max(V_2, V_3), V_4)] = E[\max(V_{max_{2,3}}, V_4)] \quad (39)$$

The first and second moments of the maximum function can be calculated iteratively to obtain first the expected maximum value of two of the three non-best options, then the expected maximum value of this value and the third non-best option:

$$E[V_{max_{2,3}}] = \mu_2 * cdf\left(\frac{\mu_2 - \mu_3}{\theta}\right) + \mu_3 * cdf\left(\frac{\mu_3 - \mu_2}{\theta}\right) + \theta * pdf\left(\frac{\mu_2 - \mu_3}{\theta}\right) \quad (40)$$

$$\theta = \sqrt{\sigma_2^2 + \sigma_3^2} \quad (41)$$

$$E[V_{max_{2,3}}^2] = (\sigma_2^2 + \mu_2^2) * cdf\left(\frac{\mu_2 - \mu_3}{\theta}\right) + (\sigma_3^2 + \mu_3^2) * cdf\left(\frac{\mu_3 - \mu_2}{\theta}\right) + (\mu_2 + \mu_3) * \theta * pdf\left(\frac{\mu_2 - \mu_3}{\theta}\right) \quad (42)$$

$$\mu_{max_{2,3}} = E[V_{max_{2,3}}] \quad (43)$$

$$\sigma_{max_{2,3}}^2 = E[V_{max_{2,3}}^2] - \mu_{max_{2,3}}^2 \quad (44)$$

↓

$$E[V_{max}] = \mu_{max_{2,3}} * cdf\left(\frac{\mu_{max_{2,3}} - \mu_4}{\theta}\right) + \mu_4 * cdf\left(\frac{\mu_4 - \mu_{max_{2,3}}}{\theta}\right) + \theta * pdf\left(\frac{\mu_{max_{2,3}} - \mu_4}{\theta}\right) \quad (45)$$

$$\theta = \sqrt{\sigma_{max_{2,3}}^2 + \sigma_4^2} \quad (46)$$

$$E[V_{max}^2] = (\sigma_{max_{2,3}}^2 + \mu_{max_{2,3}}^2) * cdf\left(\frac{\mu_{max_{2,3}} - \mu_4}{\theta}\right) + (\sigma_4^2 + \mu_4^2) * cdf\left(\frac{\mu_4 - \mu_{max_{2,3}}}{\theta}\right) + (\mu_{max_{2,3}} + \mu_4) * \theta * pdf\left(\frac{\mu_{max_{2,3}} - \mu_4}{\theta}\right) \quad (47)$$

$$\mu_{max} = E[V_{max}] \quad (48)$$

$$\sigma_{max}^2 = E[V_{max}^2] - \mu_{max}^2 \quad (49)$$

Clearly, this iterative process could be repeated indefinitely, allowing for a choice option set of arbitrarily large size.

Once the expected maximum value of the non-best options is determined, the DM will compare this value with the expected value of the best option. Thus, the decision variable (DV) will be a normally distributed random variable with mean equal to the difference of the best and max-non-best means and variance equal to the sum of the best and max-non-best variances:

$$DV \sim N(\mu_{best} - \mu_{max}, \sqrt{\sigma_{best}^2 + \sigma_{max}^2}) \quad (50)$$

Subjective uncertainty regarding option values eventually translates into uncertainty regarding what the "correct choice" is. This is summarized in terms of the probability  $P_c$  of "committing to the correct choice":

$$P_c = 1 - cdf_{DV} \quad (51)$$

$$\approx \frac{1}{1 + e^{\frac{\mu_{best} - \mu_{max}}{\sqrt{\frac{3}{\pi^2}(\sigma_{best}^2 + \sigma_{max}^2)}}}} \quad (52)$$

where  $cdf$  is the cumulative density function of the corresponding normal distribution, evaluated at  $DV = 0$ , and the second line derives from a simple sigmoidal approximation to the normal cumulative density function [Daunizeau, 2017]. Critically,  $P_c$  can be viewed as measuring subjective *choice confidence*, in the sense that confidence stems from the belief that the currently preferred option is better than all of the other ones.

As discussed above, choice confidence (i.e.,  $P_c$ ) will be instrumental in the effort allocation process. Recall that, unless initial confidence is already high enough, the DM will invest mental effort in order to increase  $P_c$  towards a (subjective) threshold level. The multi-option version of our model works in the same way as the basic version: effort investment impacts  $P_c$  via an increase in the precision of DV, as well as via a change in the first-order moment of DV that allows for the possibility of revising the DM's initial preference.

The more information the DM takes into consideration, the greater is the precision of his value estimates. Thus, the precision of DV after effort is the initial precision plus an additive term that is a function of effort (herein linear, for simplicity and without loss of generality). The magnitude of the increase in precision is scaled by an efficiency parameter that we call  $\beta$ :

$$\phi \stackrel{\text{def}}{=} \frac{1}{\sigma_{best}^2 + \sigma_{max}^2} \quad (53)$$

$$\phi \xrightarrow{z} \phi + \beta z \quad (54)$$

The increase in precision causes an increase in  $P_c$ :

$$DV_z \sim N\left(\mu_{best} - \mu_{max}, \sqrt{\frac{1}{\phi + \beta z}}\right) \quad (55)$$

$$\sqrt{\frac{1}{\phi + \beta z}} < \sqrt{\frac{1}{\phi}} \quad (56)$$

$$\Rightarrow cdf_{DV_z} < cdf_{DV} \quad (57)$$

$$\Rightarrow P_{c_z} > P_c \quad (58)$$

Allocating resources also implies greater potential for changes in the DM's value estimates. These changes are equally likely to be positive or negative, with a range proportional to the amount of allocated resources. A mathematically convenient way to capture this phenomenon is via a stochastic perturbation term on the center of mass  $\mu$  of value representations. We thus include this in our model as a set of i.i.d. normally-distributed random variables each with mean zero and variance equal to effort scaled by an efficiency parameter that we call  $\gamma$ :

$$\mu_i \xrightarrow{z} \mu_i + \Delta_i \quad \forall i \quad (59)$$

$$\Delta_i(z) \sim N(0, \gamma z) \quad (60)$$

For simplicity, it would be convenient to summarize the cumulative effect of the perturbed individual value estimates. Because the perturbation is relevant insomuch as it represents an alteration of the DV, we choose to summarize it as the change in  $\mu_{best}$  minus the change in  $\mu_{max}$ . The change in  $\mu_{best}$  is straightforward. The change in  $\mu_{max}$  will depend on the values of the individual  $\mu$  terms as well as the individual  $\Delta$  terms:

$$\Delta_{best} = \Delta_i; i: \mu_i > \mu_j \forall j \neq i \quad (61)$$

$$\Delta_{max} = E \left[ \max(V_{new_j}) \right] - E[\max(V_j)], \forall j \neq i \quad (62)$$

$$V_{new_j} \sim N(\mu_j + \Delta_j, \sigma_j) \quad (63)$$

$$\Delta(z) = \Delta_{best}(z) - \Delta_{max}(z) \quad (64)$$

A congruent change (increase) in DV causes an increase in  $P_c$ , i.e.:

$$\frac{1}{1+e^{-\pi\sqrt{\frac{\phi}{3}}|DV+\Delta|}} > \frac{1}{1+e^{-\pi\sqrt{\frac{\phi}{3}}|DV|}} \text{ if } |DV + \Delta| > DV \quad (65)$$

but any incongruent change in DV (i.e.,  $|\mu_{best} - \mu_{max} + \Delta| < \mu_{best} - \mu_{max}$ ) can lead to a decrease in  $P_c$ . Because the perturbation is a random variable, the EVC derives from its *expected* impact on DV, which is given by:

$$E[|DV + \Delta|] = \sqrt{\frac{2\gamma z}{\pi}} e^{\left(\frac{-(DV)^2}{2\gamma z}\right)} + DV * erf\left(\frac{DV}{\sqrt{2\gamma z}}\right) \quad (66)$$

$$erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \quad (67)$$

where we have simply used the expression for the first-order moment of the so-called “folded normal distribution”. Recall that here, we would want to pass DV through the  $P_c$  function before taking its expectation, but for mathematical simplification, we instead rely on the approximation  $E[P_c(|DV + \Delta|)] \approx P_c(E[|DV + \Delta|])$ .

Technically, the expectation of DV after drift is not exactly as expressed above. The above formula is correct in the instance of only two choice options, because  $\mu_{max} = \mu_{worst}$  and therefore  $\mu_{best}$  and  $\mu_{max}$  merely swap roles if the drift causes them to reverse ordinal location. However, in the instance of more than two choice options,  $\mu_{max}$  does not represent a real option, it is merely a symbol that summarizes all options other than the best. Thus, even if the perturbation causes  $\mu_{best}$  to become smaller than  $\mu_{max}$ , it would not cause their roles to swap. The initial best option would remain the best option. The agent would still select the initial best option, yet it would believe that one of the other options would likely be better—it simply would not know which of the other options that would be. Only when the perturbation causes  $\mu_{best}$  to fall below  $\max(\mu_j), j \neq best$ , would the agent change its mind and now select the option that was associated with the new  $\max(\mu_j)$ . This means that the “folding” point of the folded Gaussian described above (where the to-be chosen option and the best of the to-be rejected options swap roles) would not be  $DV = 0$ . The DV could continue to decrease into the negative domain until the point at which  $\mu_{best} = \max(\mu_j), j \neq best$ , which occurs at  $DV = \mu_{max} - \max(\mu_j), j \neq best$ . Thus, the true expected value of DV, including the perturbation, would be strictly less than what is shown in Equation 39. The true expected value of  $P_c$ , in turn, would also be strictly less than it would be when derived using Equation 39.

Specifically, for a choice set containing n options with values  $V_i \sim N(\mu_i, \sigma_i); i \in \{1, \dots, n\}$ :

$$E[DV_z] = \sum_{i=1}^n \int_{-\infty}^{+\infty} \left\{ (\mu_i + \Delta_i - E[\max(\mu_j + \Delta_j)]) * \mathbb{1}_{\mu_i + \Delta_i \geq \max(\mu_j + \Delta_j)} \right\} * \prod_{k=1}^n pdf(\Delta_k) * \Delta_k$$

$$j \in \{1, \dots, n\} \setminus \{i\} \tag{68}$$

where  $\Delta_i$  is the perturbation in the expected value of option i, and  $pdf$  is the probability density function of the normal distribution corresponding to  $\Delta_i \sim N(0, \gamma z)$ . The above equation

calculates the expected post-perturbation value distance between the best and the maximum of the non-best options, for all possibilities of which option ends up being the best after the perturbation. As shown, this amounts to an n-dimensional summation of an n-dimensional integral.

As before, a crucial feature of the stochastic perturbation element of our multi-option model is the fact that it allows for the possibility that the DM will change his mind, eventually choosing an option that he was initially inclined to reject. The probability of changing one's mind is simply the probability that the perturbation causes the initial ordering of option values to change. The fact that this probability is non-zero is important because otherwise, the outcome of a choice should always be identical with or without effort, and therefore it would never be rational to allocate any resources to the task.

For any particular task at hand, the DM will select the optimal level of effort allocation ( $z^*$ ) by maximizing the EVC:

$$z^* = \underset{z}{\operatorname{argmax}} EVC \quad (69)$$

where the expressions for the expected impact of perturbation and precision gain on DV have been inserted into the mathematical definition of the expected choice confidence in the EVC. The optimal effort allocation ( $z^*$ ) will be the level at which the marginal cost equals the marginal benefit, beyond which the cost would outweigh the benefit. The EVC-optimal level of expected choice confidence ( $P_c^*$ ) is obtained by inserting  $z^*$  into the  $E[P_c]$  formula:

$$P_c^* \approx \frac{1}{1 + e^{-\pi \sqrt{\frac{(\phi + \beta z^*)}{3}} E[|DV + \Delta| | z^*]}} \quad (70)$$

The model thus predicts the amount of exerted mental effort as a function of three important features of the decision task (initial difficulty DV, initial precision  $\phi$ , incentive  $1/\alpha$ ) and two efficiency parameters ( $\beta$  and  $\gamma$ ). As before, the model also allows us to predict how the reported

choice confidence level, the probability of changing one's mind, the change in rating confidence, and the average spreading of alternatives should vary as initial difficulty, initial precision, and incentive vary.

During the “implementation phase”, the effort that was allocated is actually invested, in the form of deliberation. Computationally, the implementation phase of our model can be represented as a drift-diffusion process that captures the evolution of choice confidence with effort. Across total resource investment ( $z$ ), the location of the process will depend on incremental random perturbations in the valuations of the options, along with incremental gains in the precisions of said valuations. The perturbation element will directly impact the instantaneous DV and consequentially the instantaneous choice confidence level, whereas the precision gain element will directly impact the instantaneous confidence:

$$DV_z = |DV_{z-1} + \Delta| \quad (71)$$

$$\phi_z = \phi_{z-1} + \beta \quad (72)$$

$$P_z = \frac{1}{1 + e^{-\pi \sqrt{\frac{\phi_z}{3}} DV_z}} \quad (73)$$

In concurrence with the evolution of choice confidence, an effort monitor will tally the accumulation of invested effort. The DM will continue to invest additional increments of effort until one of the thresholds is reached by its respective accumulator process ( $z = z^*$  or  $P \geq P^*$ ).

## METHODS

To validate our model we conducted a behavioral experiment based on the traditional free-choice paradigm, where participants rate items, then choose between them, then rate them again. In addition to the standard measurements of choice and response time, we also asked

participants to report certainty about their ratings, confidence about their choices, and feelings of effort during their choices. For the choices in this study, triplets of items were used rather than the standard pairs, allowing us to focus on how our model performs for decisions between multiple (i.e., more than two) options.

### *Participants*

A total of 50 people (39 female; age: mean=27, stdev=5, min=19, max=40) participated in this study. The experiment lasted approximately 2 hours, and each participant was paid a flat rate of 24 euros as compensation for his time. One participant was excluded from our analysis because she was unable to complete the experiment. Three other participants were excluded from our analysis for not performing the tasks properly (i.e., responding randomly). All participants were native French speakers. All participants were from the non-patient population with no reported history of psychiatric or neurological illness.

### *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 the Matlab programming language (version R2018a) and the Psychophysics Toolbox (version 3.0.14). The experiment was conducted entirely in French.

The stimuli for this experiment were 300 digital images, each representing a distinct food item (fruits, vegetables, nuts, meats, cheeses, chips, crackers, cookies, cakes, pies, and a variety of other suitable snack items). 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 Choice sections of the experiment using The Eye Tribe eye tracking devices. Prior to each experimental section, the devices were re-calibrated using 16 points. 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.

### *Procedure*

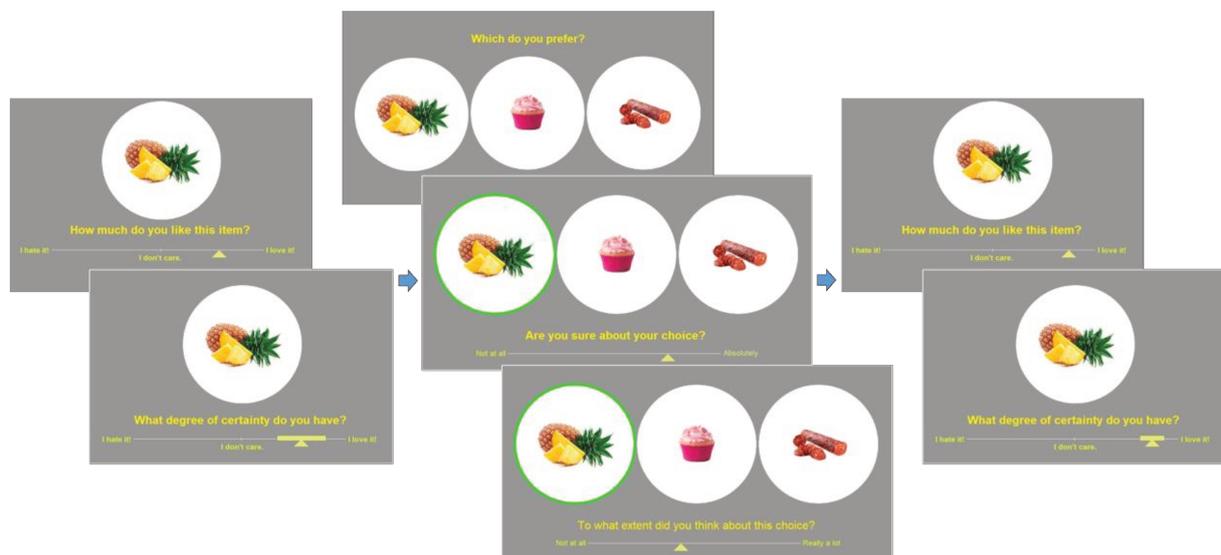
Prior to commencing the testing session of the experiment, participants underwent a brief training session. The training tasks were identical to the experimental tasks, although different stimuli were used (beverages).

The experiment itself began with an initial section where all individual items were displayed in a random sequence for 0.7 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: Rating, Choice, Rating. There was no time limit for the overall experiment, nor for the different sections, nor for the individual trials. Within-trial event sequences are described below (see Figure 1).

Rating: 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, "To what extent would you like to eat this item?" using a horizontal slider scale (from "Absolutely not!" to "Immensely!") to indicate their value rating for the item. Participants then had to respond to the question, "In what interval would you place your rating?" by expanding a solid bar symmetrically around the cursor of the slider

scale to indicate their certainty about the value rating that they had just provided. At that time, the next trial began.

Choice: Participants were asked to choose between triplets of items in terms of which item they preferred to eat. The entire set of items was presented one triplet at a time in a random sequence. Each item appeared in only one triplet. At the onset of each trial, a fixation cross appeared at the center of the screen for 750ms. Next, three images of snack items appeared on the screen: one towards the left, one in the center, and one towards the right. Participants had to respond to the question, “Which do you prefer to eat?” using the left, down, or right arrow key. Participants then had to respond to the question, “Are you sure about your choice?” using a horizontal slider scale (from “Not at all!” to “Absolutely!”). 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!”). At that time, the next trial began.

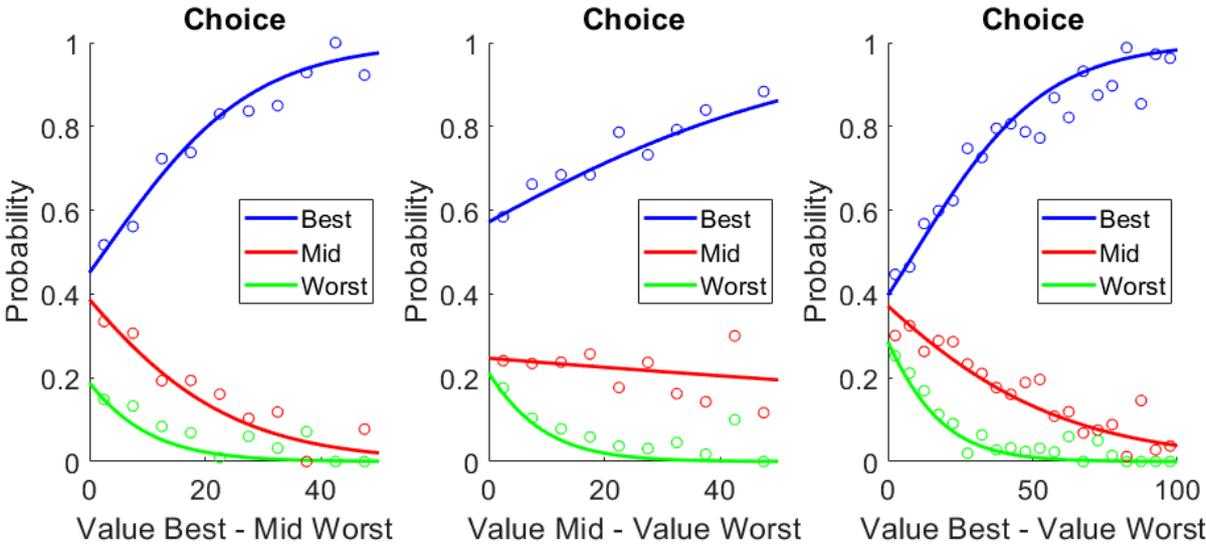


**Figure 1: Experiment Tasks:** participants individually estimated their subjective value of each option, along with their certainty about each estimation; next, they chose their preferred option from triplets of options, and provided their subjective confidence and effort estimates for each choice; finally, participants again estimated their subjective value and certainty for each option.

## RESULTS

Before attempting to validate our model, we performed a number of simple data quality checks. First, we assessed the test-retest reliability of both value judgments and their associated certainty reports. For each participant, we thus measured the correlation between ratings #1 and #2 (across items). We found that both ratings were significantly reproducible, although certainty estimates were less stable than value estimates (value ratings: correlation = mean 84%,  $r^2 = 0.98$ ,  $p < 0.001$ ; certainty ratings: mean correlation = 20%,  $r^2 = 0.55$ ,  $p < 0.001$ ). Second, we asked whether choices were consistent with value ratings #1. For each participant, we thus performed a logistic regression of paired choices against the difference in value ratings. We found that the balanced prediction accuracy was beyond chance level (group mean = 57%,  $r^2 = 0.98$ ,  $p < 0.001$ ) (see Figure 2 below for more details). Third, we checked that choice confidence increases both with the value difference between the best and middle items in each triplet, with the value difference between the middle and worst items in each triplet, and with the mean certainty reports about value judgments within each triplet. For each participant, we thus performed a multiple linear regression of choice confidence against value difference 1 (best-mid), value difference 2 (mid-worst), and mean judgment certainty (all using ratings #1). A random effect analysis shows that both value distance measures have a significant effect at the group level, but certainty does not ( $r^2 = 0.68$ ; value difference 1: GLM beta = 6.2055,  $p < 0.001$ ; value difference 2: GLM beta = 4.2103,  $p < 0.001$ ; mean judgment certainty: GLM beta = -0.2356,  $p = 0.666$ ). Fourth, we asked whether we could reproduce previous findings of choice-induced preference change (CIPC). For each participant, we thus measured the magnitude of CIPC in terms of the so-called "spreading of alternatives" (SoA), calculated as the mean difference in value rating gains between chosen and unchosen items ( $\text{SoA} = [\text{rating}\#2 - \text{rating}\#1]_{\text{chosen}} - [\text{rating}\#2 - \text{rating}\#1]_{\text{unchosen}}$ ). As expected, we found that SoA is significant (mean SoA = 5.615, SEM = 0.5213,  $p < 0.001$ ). Finally, we checked that our two measures of effort (i.e., subjective

effort ratings and response time) represented similar information. We observed that subjective effort was indeed positively correlated with RT (correlation = 51%,  $p < 0.001$ ). This suggests that RT is a fair, though not great, proxy for mental effort allocation, or that subjects have a decent, but not perfect, ability to gauge their effort exertion level.



**Figure 2: Choice Accuracy versus option values:** the probability of choosing the best (blue curves), 2<sup>nd</sup>-best (red curves), or worst (green curves) options as a function of the value difference between best and 2<sup>nd</sup>-best (left figure), 2<sup>nd</sup>-best and worst (middle figure), and best and worst (right figure) options.

Recall that, under our hypothesis, the deliberation that takes place during the decision process is expected to cause a refinement of internal value representations until a target level of choice confidence is met and the decider commits to a choice. To begin with, we thus asked whether certainty about value judgments improved after the choice had been made. For each participant, we estimated the mean difference between post-choice and pre-choice certainty reports (across all items). A random effect analysis shows that post-choice certainty reports are significantly higher than pre-choice certainty reports (mean certainty increase = 2.0623, SEM = 0.5217,  $p < 0.001$ ). This finding supports our claim. Under the same line of reasoning, we might expect

that post-choice ratings should be easier than pre-choice ratings, thanks to the contemplation that occurred during the choice itself. We thus asked whether response times (RT) for ratings were faster after the choice had been made. For each participant, we estimated the mean difference between post-choice and pre-choice rating RT (across all items). A random effect analysis shows that post-choice RT are significantly lower than pre-choice RT (mean RT change = -0.5710, SEM = 0.0663,  $p < 0.001$ ). This finding also supports our claim. We then asked whether post-choice ratings better predict choice (and choice confidence) than pre-choice ratings. First, we performed another logistic regression of paired choices, this time against the difference in post-choice value ratings (ratings #2). However, the ensuing choice prediction accuracy is not higher than with pre-choice value ratings (mean accuracy = 56%,  $r^2 = 0.98$ , accuracy gain:  $p = 0.79$ ). Second, we regressed choice confidence, this time against post-choice value difference 1, value difference 2, and mean judgment certainty. The ensuing amount of explained variance in choice confidence reports is not significantly different than with pre-choice ratings ( $p = 0.222$ ).

Next, we focus on testing the predictions of our model regarding a set of relevant dependent variables, and contrast our predictions with those of competing models. Before examining the experimental data, we first simulate data under our model and the competitor models. With the simulated data, we performed a multiple linear regression of our dependent variables of interest—choice confidence (probability of choosing correctly), RT/effort (allocated resources), SoA, and changing of mind frequency—on value difference 1, value difference 2, and value certainty. We first examined our EVC model, phase 1—the allocation of the effort budget. Here we find that choice confidence shows a strong positive effect of value difference 1, a weak positive effect of value difference 2, and a very weak positive effect of value certainty (GLM betas: 0.968, 0.071, 0.019;  $10^6$  simulated trials). Conversely, effort shows a strong negative effect of value difference 1, a weak negative effect of value difference 2, and a medium

negative effect of value certainty (GLM betas: -0.953, -0.091, -0.156;  $10^6$  simulated trials). Not surprisingly, there is a strong negative correlation between confidence and effort (correlation = -0.47,  $r^2 = 0.99$ ,  $10^6$  simulated trials). Spreading of alternatives shows a medium negative effect of value difference 1, a weak negative effect of value difference 2, and a negligible negative effect of value certainty (GLM betas: -0.217, -0.038, -0.003;  $10^6$  simulated trials). Changing of mind frequency shows a very strong negative effect of value difference 1, a very negative effect of value difference 2, and a weak negative effect of value certainty (GLM betas: -6.905, -0.060, -0.027;  $10^6$  simulated trials).

We then examined our EVC model, phase 2—the implementation of mental effort. Choice confidence shows a medium positive effect of value difference 1, a weak positive effect of value difference 2, and a weak negative effect of value certainty (GLM betas: 0.363, 0.081, -0.076;  $10^6$  simulated trials). Effort shows a strong negative effect of value difference 1, a weak negative effect of value difference 2, and a weak negative effect of value certainty (GLM betas: -0.828, -0.117, -0.122;  $10^6$  simulated trials). The correlation between confidence and effort is now much stronger (correlation = -0.75,  $r^2 = 1.00$ ,  $10^6$  simulated trials). Spreading of alternatives shows a strong negative effect of value difference 1, a weak negative effect of value difference 2, and a very weak negative effect of value certainty (GLM betas: -0.480, -0.125, -0.028;  $10^6$  simulated trials). Changing of mind frequency shows a very strong negative effect of value difference 1, a medium negative effect of value difference 2, and a weak negative effect of value certainty (GLM betas: -1.958, -0.179, -0.045;  $10^6$  simulated trials). Note that these results are all qualitatively similar as for the EVC phase 1 (though quantitatively different), except for the effect of value certainty on choice confidence.

The first competitor model that we examined is the so-called Race Model. This model consists of a separate accumulator variable for each choice option. Each accumulator collects a piece of “evidence” at each time step, in the form of a random sample from a normal distribution with

mean equal to the estimated value of the option and variance representing processing noise (equal for all options). Time advances until one of the accumulators reaches a pre-determined threshold, at which point the option represented by that accumulator is chosen. We performed a multiple linear regression of our dependent variables of interest on value difference 1, value difference 2, and value certainty. Choice confidence shows a strong positive effect of value difference 1, a weak negative effect of value difference 2, and a medium negative effect of value certainty (GLM betas: 0.636, -0.067, -0.409;  $10^6$  simulated trials). RT shows a medium negative effect of value difference 1, a medium negative effect of value difference 2, and a medium positive effect of value certainty (GLM betas: -0.283, -0.406, 0.171;  $10^6$  simulated trials). Changing of mind frequency shows a strong negative effect of value difference 1, a weak negative effect of value difference 2, and a medium negative effect of value certainty (GLM betas: -0.979, -0.106, -0.362;  $10^6$  simulated trials). Note that the concept of changing one's mind does not exist for the race model; here, these responses are referred to as errors. Relatedly, the concept of spreading of alternatives does not exist for the race model, as option values are assumed to be *a priori* known and stable.

In addition to the basic race model, we simulated a version where the variance of the distributions from which the evidence samples are drawn for each option is equal to the inverse of the certainty of its value estimate. Here, choice confidence shows a medium positive effect of value difference 1, a weak negative effect of value difference 2, and a strong negative effect of value certainty (GLM betas: 0.434, -0.047, -0.528;  $10^6$  simulated trials). RT shows a medium negative effect of value difference 1, a medium negative effect of value difference 2, and a medium positive effect of value certainty (GLM betas: -0.237, -0.378, 0.273;  $10^6$  simulated trials). Changing of mind frequency shows a strong negative effect of value difference 1, a weak negative effect of value difference 2, and a medium negative effect of value certainty (GLM

betas: -0.685, -0.102, -0.391;  $10^6$  simulated trials). Note that the results are qualitatively similar to those obtained with the basic race model.

The second competitor model that we examined is the so-called Drift-Diffusion Model (DDM). This model normally consists of a single accumulator variable that represents the comparison between two options. Implementing a DDM with multiple choice options is a bit less straightforward, but there are a few published examples of multi-option DDMs. We chose the model presented by Krajbich and Rangel [2011], as it is both simple and widely accepted in the field. Note that their model has an additional parameter that alters the evidence accumulation rate depending on attention, but we will remove this parameter for our purposes here (with no loss of model validity). Like the race model, this DDM consists of a separate accumulator variable for each choice option. Also like the race model, this DDM draws stochastic samples from normal distributions (mean = estimated value, variance = processing noise equal for all options) for each option at each time point, and the new evidence is added to each accumulator. Now, unlike the race model, the threshold that needs to be reached before a choice can be made is based on the relative values of the options. Here, a relative decision value is calculated at each time step, derived from the evidence accumulators as follows:

$$Value_t^1 = Evidence_t^1 - \max(Evidence_t^2, Evidence_t^3)$$

$$Value_t^2 = Evidence_t^2 - \max(Evidence_t^1, Evidence_t^3)$$

$$Value_t^3 = Evidence_t^3 - \max(Evidence_t^1, Evidence_t^2)$$

The process proceeds until one of the value signals reaches the threshold. We performed a multiple linear regression of our dependent variables of interest on value difference 1, value difference 2, and value certainty. Choice confidence shows a medium positive effect of value difference 1, a weak positive effect of value difference 2, and a strong negative effect of value

certainty (GLM betas: 0.193, 0.032, -0.863;  $10^6$  simulated trials). RT shows a strong negative effect of value difference 1, a medium negative effect of value difference 2, and a medium positive effect of value certainty (GLM betas: -0.744, -0.137, 0.435;  $10^6$  simulated trials). Changing of mind frequency shows a very strong negative effect of value difference 1, a small negative effect of value difference 2, and a large negative effect of value certainty (GLM betas: -1.773, -0.083, -0.706;  $10^6$  simulated trials). Note that the concept of changing one's mind does not exist for the DDM; here, these responses are referred to as errors. Relatedly, the concept of spreading of alternatives does not exist for the DDM, as option values are assumed to be *a priori* known and stable.

In addition to the basic DDM, we simulated a version where the variance of the distributions from which the evidence samples are drawn for each option is equal to the inverse of the certainty of its value estimate. Here, choice confidence shows a weak positive effect of value difference 1, a negligible positive effect of value difference 2, and a large negative effect of value certainty (GLM betas: 0.093, 0.005, -0.729;  $10^6$  simulated trials). RT shows a strong negative effect of value difference 1, a weak negative effect of value difference 2, and a strong positive effect of value certainty (GLM betas: -0.562, -0.109, 0.496;  $10^6$  simulated trials). Changing of mind frequency shows a strong negative effect of value difference 1, a weak negative effect of value difference 2, and a strong negative effect of value certainty (GLM betas: -0.944, -0.103, -0.667;  $10^6$  simulated trials). Note that the results are qualitatively similar to those obtained with the basic DDM.

We also examined our own novel version of a DDM with multiple choice options. As with the other DDM and race models, each option receives one stochastic sample of evidence at each time point, and the process terminates when one option accumulates enough evidence to reach a threshold. At each point in time, only the “winner” of the new evidence sample comparison moves closer to its threshold in a step the size of its value minus the next-best value. Also at

each point in time, the threshold for only the “loser” of the new evidence sample comparison moves farther away in a step the size of the second-to-worst value minus the worst value. The dynamics are thus as follows:

$$V_t^1 = V_{t-1}^1 + \max(\min(E_t^1 - E_t^2, E_t^1 - E_t^3), 0)$$

$$V_t^2 = V_{t-1}^2 + \max(\min(E_t^2 - E_t^1, E_t^2 - E_t^3), 0)$$

$$V_t^3 = V_{t-1}^3 + \max(\min(E_t^3 - E_t^1, E_t^3 - E_t^2), 0)$$

$$Th_t^1 = Th_{t-1}^1 + \min(\max(E_t^1 - E_t^2, E_t^1 - E_t^3), 0)$$

$$Th_t^2 = Th_{t-1}^2 + \min(\max(E_t^2 - E_t^1, E_t^2 - E_t^3), 0)$$

$$Th_t^3 = Th_{t-1}^3 + \min(\max(E_t^3 - E_t^1, E_t^3 - E_t^2), 0)$$

The process proceeds until one of the value signals reaches its respective threshold. We performed a multiple linear regression of our dependent variables of interest on value difference 1, value difference 2, and value certainty. Choice confidence shows a strong positive effect of value difference 1, a medium positive effect of value difference 2, and a weak positive effect of value certainty (GLM betas: 0.881, 0.223, 0.139;  $10^6$  simulated trials). RT shows a medium negative effect of value difference 1, a medium negative effect of value difference 2, and a strong positive effect of value certainty (GLM betas: -0.176, -0.176, 0.960;  $10^6$  simulated trials). Changing of mind frequency shows a strong negative effect of value difference 1, a medium negative effect of value difference 2, and a medium negative effect of value certainty (GLM betas: -0.806, -0.161, -0.439;  $10^6$  simulated trials).

In addition to the basic DDM, we simulated a version where the variance of the distributions from which the evidence samples are drawn for each option is equal to the inverse of the certainty of its value estimate. Here, choice confidence shows a medium positive effect of value difference 1, a weak positive effect of value difference 2, and a medium negative effect of value

certainty (GLM betas: 0.406, 0.086, -0.146;  $10^6$  simulated trials). RT shows a weak negative effect of value difference 1, a weak negative effect of value difference 2, and a strong positive effect of value certainty (GLM betas: -0.060, -0.061, 0.779;  $10^6$  simulated trials). Changing of mind frequency shows a medium negative effect of value difference 1, a weak negative effect of value difference 2, and a medium negative effect of value certainty (GLM betas: -0.468, -0.123, -0.389;  $10^6$  simulated trials). Note that the results are qualitatively similar to those obtained with the basic DDM, with the exception of the effect of value certainty on choice confidence.

Before returning to our experimental data, we briefly summarize the qualitative model comparison. We start by noting that all models make the same qualitative predictions for the effect of value difference 1, value difference 2, and value certainty on changing of mind frequency (referred to as error rate in most models). For response time (which is similar to resource allocation in our model), all models make the same qualitative predictions with respect to the effect of value difference 1 and 2, but they do not agree with respect to the effect of value certainty. Here, our EVC model (both phases) predicts that responses will be provided sooner when the DM is more certain about his value estimates, but the race model and DDM (all versions) predict the opposite. For choice confidence, all models make the same qualitative prediction with respect to the effect of value difference 1, but they disagree otherwise. With respect to the effect of value difference 2, our EVC model and the DDM both predict that the further apart the values of the middle and worst options, the more confident the DM will be about his choice; the race model predicts the opposite. Our EVC model is the only one we know of that can predict spreading of alternatives, so we cannot compare our predictions to those of any other model.

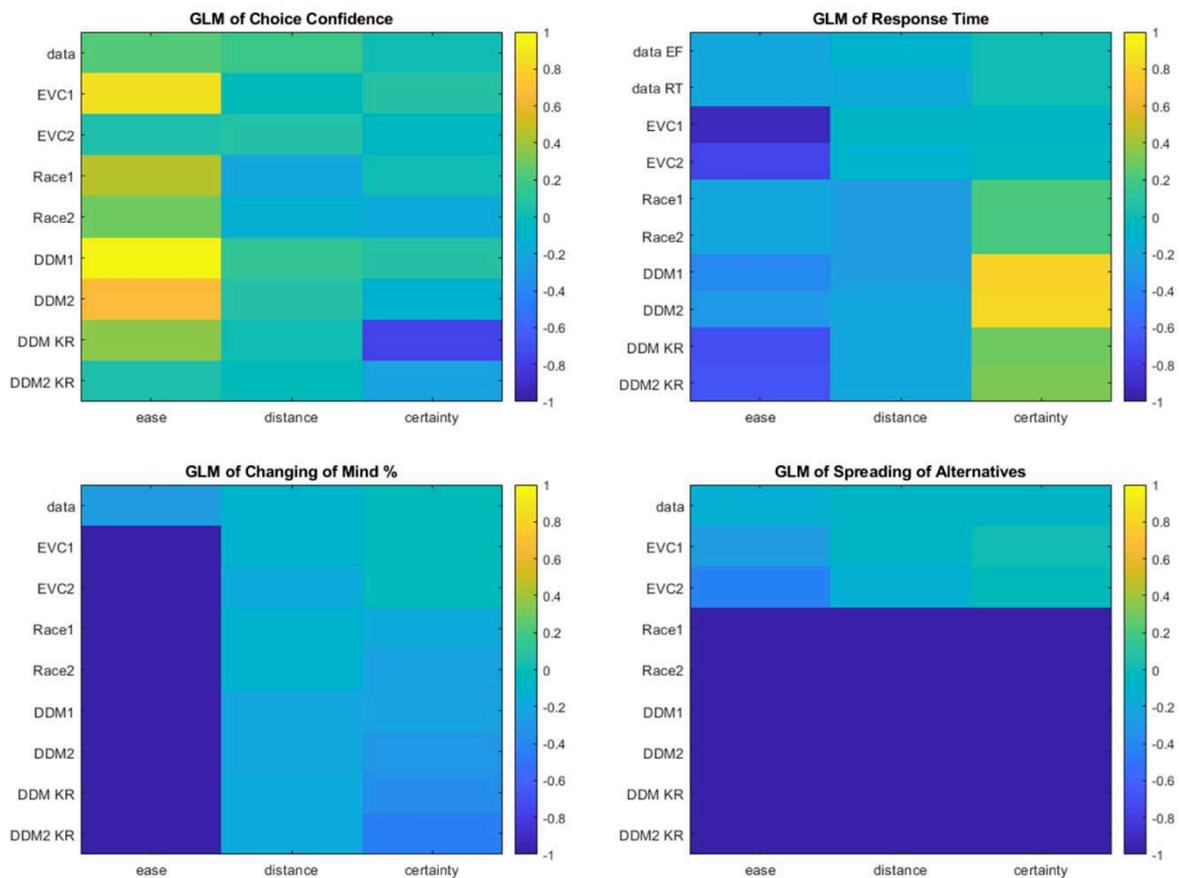
We then examined the possibility that there might be an interaction between value difference and value certainty that would manifest itself in different ways in the different models. We found no such manifestation with respect to value difference 2, but we did find it with respect to value difference 1. Indeed, adding this interaction term to our regressions made the models distinguishable across all independent and dependent variables. Of note, only our EVC model (phase 1) predicts a negative effect of value difference 1 on changing of mind frequency, with a positive effect of the interaction of value difference 1 and value certainty. The race model and DDM (all versions) predict the opposite. Our EVC model (phase 2) predicts that both effects will be negative. Interesting to note, all models except the Krajbich-Rangel DDM with unequal variance now predict no main effect of certainty on changing of mind frequency. For response time, our EVC model (both phases) predicts a negative effect of value difference 1, a negative effect of value certainty, and a positive effect of the interaction between value difference 1 and value certainty. The race model and DDM (all versions) predict the opposite. For choice confidence, there is a lot more variety in the model predictions. Our EVC model (phase 1) predicts a positive effect of value difference 1, no effect of value difference 2, a positive effect of value certainty, and a positive effect of the interaction between value difference 1 and value certainty. Our EVC model (phase 2) predicts a positive effect of value difference 1, a positive effect of value difference 2, no effect of value certainty, and a negative effect of the interaction between value difference 1 and value certainty. The race model (equal variance) predicts a positive effect of value difference 1, a negative effect of value difference 2, a positive effect of value certainty, and a negative effect of the interaction between value difference 1 and value certainty. The race model (unequal variance) predicts a positive effect of value difference 1, a negative effect of value difference 2, a negative effect of value certainty, and a negative effect of the interaction between value difference 1 and value certainty. The DDM (Krajbich-Rangel version with unequal variance) predicts a negative effect of value difference 1, a positive effect

of value difference 2, a negative effect of value certainty, and a positive effect of the interaction between value difference 1 and value certainty. The DDM (all other versions) predicts a positive effect of value difference 1, a positive effect of value difference 2, a negative effect of value certainty, and a positive effect of the interaction between value difference 1 and value certainty.

We now turn back to our experimental data, in order to assess the statistical relationships between choice difficulty (both value difference 1, value difference 2, and value certainty) and choice confidence, RT, subjective effort, SoA, and changing of mind frequency. For each participant, we performed a multiple linear regression of each dependent variable onto pre-choice rating difference 1, rating difference 2, and rating certainty. Random effect analyses on the ensuing parameter estimates provide the following results. Choice confidence shows a strong positive effect of value difference 1 (GLM beta = 0.2437,  $p < 0.001$ ), a medium positive effect of value difference 2 (GLM beta = 0.1645,  $p < 0.001$ ), and no effect of value certainty ( $p = 0.405$ ). Subjective effort and RT both show a strong negative effect of value difference 1 (effort: GLM beta = -0.2014,  $p < 0.001$ ; RT: GLM beta = -0.2056,  $p < 0.001$ ), a medium negative effect of value difference 2 (effort: GLM beta = -0.1088,  $p < 0.001$ ; RT: GLM beta = -0.1686,  $p < 0.001$ ), and no effect of value certainty (effort:  $p = 0.068$ ; RT:  $p = 0.146$ ). Spreading of alternatives shows a strong negative effect of value difference 1 (GLM beta = -0.1537,  $p < 0.001$ ), a medium negative effect of value difference 2 (GLM beta = -0.0924,  $p < 0.001$ ), and a medium negative effect of value certainty (GLM beta = -0.0696,  $p < 0.001$ ). Changing of mind frequency shows a strong negative effect of value difference 1 (GLM beta = -0.2787,  $p < 0.001$ ), a medium negative effect of value difference 2 (GLM beta = -0.1148,  $p < 0.001$ ), and no effect of value certainty ( $p = 0.171$ ). We re-ran the same series of regressions, this time including an interaction term between value difference 1 and value certainty, and an interaction term between value difference 2 and value certainty. The interaction term between value difference 1 and value certainty had a marginal effect on each of our dependent variables (choice confidence:

GLM beta = 0.0385,  $p=0.057$ ; subjective effort: GLM beta = -0.0320,  $p=0.055$ ; RT: GLM beta = -0.0311,  $p=0.089$ ; spreading of alternatives: GLM beta = 0.0559,  $p=0.026$ ; changing of mind frequency: GLM beta = 0.0327,  $p=0.068$ ).

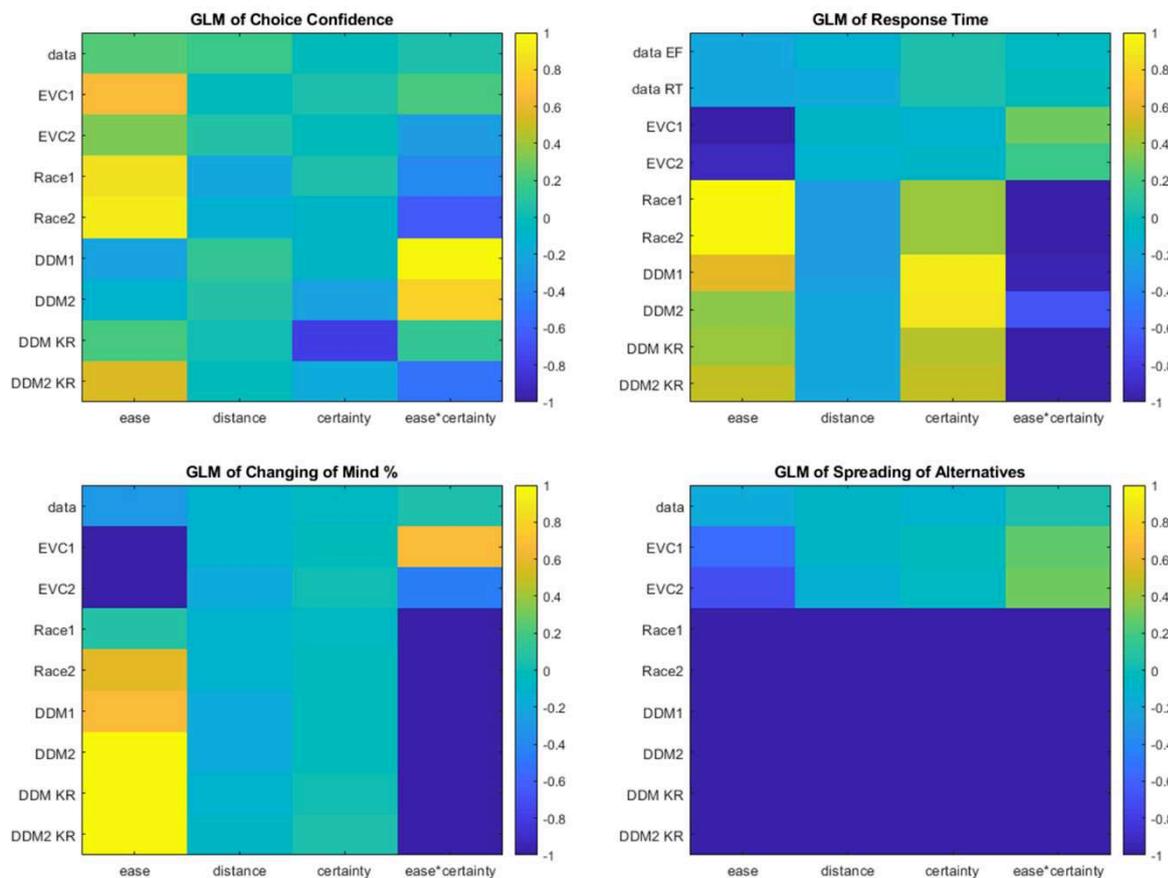
A comparison of GLM beta weights across experimental data and model predictions is summarized below (see Figure 3). (Note that the simulated data used in this model comparison differ from the data described above; here, the independent variables exactly match the experimental independent variables, whereas above, the entire range of possible input variables—defined according to our experimental paradigm—was uniformly covered.) In brief, regression analysis with no interaction term does not allow us to distinguish the performance of the models with respect to changing of mind frequency (or spreading of alternatives)—they all perform well. For response time / subjective effort, the race model and DDM (but not the EVC model) share the same sign as the data with respect to the effect of value certainty, but this effect is not significant. For choice confidence, both our EVC model (both phases) and the DDM model (equal variance) qualitatively match the experimental data; the other models do not. Our EVC model (both phases) match well the experimental data for spreading of alternatives.



**Figure 3: Model comparison:** GLM beta weights for choice confidence (upper left figure), RT (upper right figure), changing mind frequency (lower left figure), and spreading of alternatives (lower right figure). Each column in each figure corresponds to one of the regressors: ease (value best – value 2<sup>nd</sup>-best), distance (value 2<sup>nd</sup>-best – value worst), and certainty (averaged across options). Each row represents either the data (top rows) or the predictions of one of the examined models.

Regression analysis with an interaction term between value difference 1 and value certainty does allow us to distinguish model performance for all variables (see Figure 4). For changing of mind frequency, only our EVC model (phase 1) qualitatively matches the experimental data. For response time / subjective effort, our EVC model (both phases) does not match the experimental data with respect to value certainty or the interaction term, but the race model and DDM (all versions) do not match the experimental data with respect to value difference 1. For choice confidence, our EVC model (phase 1) well matches the experimental data; our EVC model (phase 2) does not match with respect to the interaction term; the race model (both

versions) does not match with respect to value difference 2 or the interaction term, and only the unequal variance version matches with respect to value certainty; and the DDM (basic versions) does not match with respect to value difference 1. It is worth pointing out that even though the DDM (all versions) matches the experimental data with respect to the sign of the effect of value certainty on choice confidence, this effect was not significant.



**Figure 4: Model comparison:** GLM beta weights for choice confidence (upper left figure), RT (upper right figure), changing mind frequency (lower left figure), and spreading of alternatives (lower right figure). Each column in each figure corresponds to one of the regressors: ease (value best – value 2<sup>nd</sup>-best), distance (value 2<sup>nd</sup>-best – value worst), certainty (averaged across options), and the interaction between ease and certainty. Each row represents either the data (top rows) or the predictions of one of the examined models.

Beyond the relationships between independent and dependent variables, our model also makes predictions about the relationships between the dependent variables. Specifically, controlling

for the effect of value difference and value certainty, the model predicts a negative correlation between spreading (SP) and effort (EF) (correlation = -22%,  $10^6$  simulated trials), a positive correlation between spreading and confidence (CC) (correlation = 59%,  $10^6$  simulated trials), and a (counter-intuitive) negative correlation between effort and confidence (correlation = -47%,  $10^6$  simulated trials). The explanation for these correlations relies on the fact that spreading of alternatives could alternatively be labeled as “increase in value difference”, which would be directly interpretable as “increase in choice ease”. Therefore, when deliberation leads to a larger “increase in choice ease” (spreading of alternatives), it is obvious that choice confidence will be higher. Similarly, effort will be lower, because the spreading of alternatives is a dynamic process that unfolds over the course of deliberation (i.e., it does not occur all at once at the time of choice). In this way, when spreading of alternatives occurs, and choice confidence is increased, the DM will terminate deliberation and make a choice (see Figure 5 below for an illustration of the process). Without spreading, the DM will continue to deliberate until eventually giving up and making a choice with lower confidence. We observe all of these predicted correlations in our experimental data (correlations: SP-EF = -6%,  $p=0.008$ ; SP-RT = -5%,  $p=0.022$ ; SP-CC = 11%,  $p<0.001$ ; EF-CC = -33%,  $p<0.001$ ; RT-CC = -35%,  $p<0.001$ ). To test these relationships further, we regressed effort, RT, and confidence on the value difference 1, value difference 2, and value certainty, plus spreading of alternatives. The data confirm all predicted relationships (SP-EF:  $\beta = -0.065$ ,  $p=0.004$ ; SP-RT:  $\beta = -0.071$ ,  $p=0.002$ ; SP-CC:  $\beta = 0.138$ ,  $p<0.001$ ).

Looking beyond the choice data, we also make some interesting observations with respect to the rating data. Specifically, we found that pre-choice rating certainty correlates negatively with rating change magnitude ( $\beta = -0.126$ ,  $p<0.001$ ). This is important, as it validates our measure of certainty. The idea is that the more certain the DM is about a value estimate, the less likely that estimate is to change during deliberation. At the same time, pre-choice rating RT correlates



## Study 4 – Appendix

### *Rating sum*

The data unexpectedly show clear patterns with respect to the sum of the ratings of options in a choice triplet. Although we say that this effect was unexpected theoretically, it was not surprising, as we had seen the same effect in each of our previous studies. There is a negative correlation with effort (subjective effort: correlation = -8%,  $p=0.003$ ; RT: correlation = -19%,  $p<0.001$ ) and spreading (correlation = -7%,  $p<0.001$ ). There is a positive correlation with confidence (correlation = 24%,  $p<0.001$ ). Note that neither our EVC model nor the DDM make any predictions with respect to rating sum. The race model does predict the negative correlation with effort, but it predicts a negative correlation with confidence.

### *Familiarity*

At the end of our experiment, participants provided familiarity ratings for each item, using a slider scale to answer the question, “How well do you know this item. Our familiarity measurement does not quite achieve significance for the expected positive correlation with value certainty (correlation = 3%,  $p=0.061$ ), but we do observe an unexpected positive correlation with value itself (correlation = 26%,  $p<0.001$ ). Familiarity does not show a correlation with subjective effort (correlation = -1%,  $p=0.200$ ), but it does show a negative correlation with RT (correlation = -8%,  $p<0.001$ ) and spreading (correlation = -6%,  $p<0.001$ ), and a positive correlation with confidence (correlation = 11%,  $p<0.001$ ). Our model does not currently include the concept of familiarity and thus makes no predictions with respect to familiarity.

After observing the correlations with familiarity, and the surprising correlations with certainty, we thought that perhaps the true variable of interest to represent precision in our model was the interaction of certainty and familiarity. The idea would be that in situations of low certainty but

high familiarity, effort could lead to higher precision; in situations of low certainty and low familiarity, effort would be futile. Checking the data with respect to the interaction of value certainty and familiarity, we find no correlation with subjective effort (correlation = 1%,  $p=0.269$ ), but a negative trend with RT (correlation = -3%,  $p=0.053$ ), a negative correlation with spreading (correlation = -7%,  $p<0.001$ ), and a positive correlation with confidence (correlation = 8%,  $p=0.002$ ).

## **Chapter 4**

### **General Discussion**

Through this project, we have presented our novel attempt to capture the mechanics of the mental effort allocation process for decision-making tasks. Our Metacognitive Control of Decision-making (MCD) model offers a simple yet robust explanation for when people will be inclined to invest more versus less mental effort, and what the effects of this investment on several key decision variables should look like. Our model has demonstrated its ability to accurately predict experimental data, and in particular to predict some non-trivial relationships between the dependent variables.

One of the more appealing aspects of our model is its generality. Even in its current underdeveloped form, it could already be adapted by clever researchers to put to use in a wide variety of choice domains or different types of cognitive tasks not typically referred to as decisions (e.g., categorization, judgment, analogical reasoning). Researchers could take the liberty to interpret the input parameters as they see fit, as there is no rigidity in our definition of the variables. For example, the specific nature of the cost associated with mental effort is currently a widely debated topic in the field, but in our model the nature of the cost is itself not relevant, only the mathematical form (and even that is only minimally constrained). The cost and/or benefit terms could be used to represent either internal or external factors, or combinations thereof. The definition of effort itself could be broken down into sub-classifications to represent different types of effort, effort divided across tasks, or any number of concepts. We believe that the MCD model could prove to be a useful tool for researchers in diverse fields such as neuroscience, cognitive psychology, behavioral economics, and computational psychiatry.

Our model is in the spirit of the literature on constructed preferences (Tversky and Thaler, 1990; Payne, Bettman and Johnson, 1993; Lichtenstein and Slovic, 2006). This literature demonstrates that decision makers often make choices that are inconsistent with their previously-stated preferences. This observation has been explained by suggesting that

preferences are not merely retrieved from a stable database in the decision maker's memory, but rather calculated during a decision task. Effectively, our model starts with this same motivation but goes further by explicitly proposing the mechanisms through which the process of the preference calculations takes place. Economics researchers could take advantage of our model to help structure their arguments about the nature of preferences and how the values that they are based on are constructed rather than retrieved. We believe that our work combines the theoretical frameworks established in the constructed preferences literature and recent work on rational inattention (Sims, 2003; Caplin & Martin, 2014; Caplin & Dean, 2015), and advances it. Our MCD model suggests that preferences are refined (constructed) at the time of choice, with information being selectively incorporated based on its associated expected cost and benefit. So, clearly our model is consistent with both of these previous frameworks. Moreover, we believe that our model brings these frameworks together cohesively and fills some of the gaps that exist when they are considered separately. For example, the work on rational inattention offers a formal description of why preferences might be constructed differently in different contexts, but it does not fully explain how psychological factors such as motivation and confidence will partially determine how much information a decision maker will consider. Our model was created to do just that. We therefore believe that our model will be of interest to economics researchers who wish to more fully explore the foundations of people's choices and preferences.

Another body of literature that demonstrates that preferences are not stable is based on what is known as choice-induced preference change (see Harmon-Jones and Harmon-Jones, 2007; Izuma and Murayama, 2013 for reviews). Here, the mere act of choosing one option over another tends to cause increased valuations for the chosen items and decreased valuations for the rejected items. We explained our belief that such valuation changes occur *before* the choice is finalized, not *after* as has traditionally been accepted (Festinger, 1962). In this work, we have

demonstrated how our MCD model can be used to describe in more detail how the preference changes might take place. For this reason, we believe that psychology researchers will also take interest in our model. It is surprising that a long-studied phenomenon like choice-induced preference change has not previously been formally described with a computational model. Most of the work describing this phenomenon has been descriptive and qualitative. Through our MCD model, we account for choice-induced preference change with a principled quantitative approach. We believe that this should provide psychology researchers with a more rigid structure upon which to base future studies. In addition, this more structured approach should allow the field to attract and sustain interest from researchers in related disciplines, who have been skeptical about choice-induced preference change in part due to the lack of a formal model to explain it.

Perhaps one of the most obvious ways in which our work could have an immediate impact is by providing an alternative to the growing number of studies that rely on the DDM to summarize the decision process. The DDM is now widely known, accepted, and promoted in psychology, neuroscience, and economics. Its presence across studies of decision neuroscience is pervasive, thanks to its simplicity and robust ability to predict choice accuracy and response time distributions. We acknowledge the importance of this model, and the role that it has played in advancing our understanding of decision-making. However, we believe that it is time to move forward, beyond the limits of the traditional DDM. Recall that the DDM does not perform well with relating metacognitive factors such as value certainty and choice confidence. Neither does it allow for constructed preferences, because the option values in this model are not allowed to change during deliberation. Our MCD model accounts for choice consistency (accuracy) and RT just as the DDM does, but it also accounts well for choice confidence and choice-induced preference change. It also allows for the levels of invested effort and confidence that the decision maker expects to differ as a function of motivation. The DDM and other related models

account for these sorts of differences only as they relate to specific factors such as the opportunity cost of time within series of sequential decisions. Our MCD model is more general, and does not require the assumption of any particular set of circumstances in order to be relevant.

Our model could be of benefit to the computational psychiatry community as well. Here, the goal would be to fit the model to individual choice data in order to recover a set of parameters for each individual. This could be of interest with respect to the general population, but it could be of particular interest to the clinical population if a certain patient cohort was found to have parameters that systematically differed from the healthy population. Recall that our beta parameter measures the impact of mental effort on certainty about value options. A low beta parameter might imply an inability to feel more certain about a belief even after thinking about it carefully, and could potentially lend insight into disorders such as anxiety or compulsivity. High beta could explain overconfidence or impulsivity. Recall that our gamma parameter measures the impact of mental effort on the reassessment of expected option values. A low gamma parameter might imply an inability to adjust one's prior beliefs even when presented with contradictory evidence, and could lend insight into disorders such as depression, obsession, or delusion. High gamma could explain attention deficit. Recall that our alpha parameter measures the subjective feeling of effort with investment of cognitive resources. A low alpha parameter might imply an inadequate metacognitive ability to ascertain the costs associated with mental effort, and could lend insight into disorders such as perfectionism. A high alpha could explain apathy. Even more interesting to consider than mere phenotyping, it is possible that neuroimaging could expose brain activity patterns that correlate with any of the various parameters of our model. While there is no way to know if it would work until we try it, we think that this presents an exciting potential opportunity for our model to be put to valuable practical use.

## CONCLUSION

Through this work, we have demonstrated our beliefs about what goes on in the mind of decision maker while deliberating. We have made clear our belief that the value of a particular option will be constructed at the time of decision, rather than simply retrieved from memory. We have explained our hypothesis that the extent to which a person will think about the options before choosing will depend on a variety of factors, including: how important he believes the decision to be; how similar he believes the options to be; how pleasant confidence makes him feel; and how unpleasant mental effort makes him feel. The computational model that we introduced serves to more formally demonstrate our ideas. The behavioral data that we summarized proves that our model can pass experimental validation. Moreover, we have described how our metacognitive control of decision-making model explains what other prominent models cannot, both theoretically and empirically. We hope that other scientists across disciplines will join us in future efforts to expand upon this model, and to put it to practical use in better understanding, and perhaps guiding, human decision-making behavior.

## BIBLIOGRAPHY

- Acharya, A., Blackwell, M., and Sen, M. (2015). Explaining Attitudes from Behavior: A Cognitive Dissonance Approach (Harvard University, John F. Kennedy School of Government).
- Akerlof, G. A., & Dickens, W. T. (1982). The economic consequences of cognitive dissonance. *The American economic review*, 72(3), 307-319.
- Albarracín, D., & Wyer Jr, R. S. (2001). Elaborative and nonelaborative processing of a behavior-related communication. *Personality and Social Psychology Bulletin*, 27(6), 691-705.
- Allahverdyan, A. E., & Galstyan, A. (2014). Opinion dynamics with confirmation bias. *PloS one*, 9(7), e99557.
- Arrow, K. J. (1950). A difficulty in the concept of social welfare. *Journal of political economy*, 58(4), 328-346.
- Bartra, O., McGuire, J. T., & Kable, J. W. (2013). The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage*, 76, 412-427.
- Beasley, Joslyn (2001). Cognitive Dissonance and Post-Decision Attitude Change in Six Presidential Elections. *Political Psychology*. 22 (3): 521–540.
- Becker, G. M., DeGroot, M. H., & Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral science*, 9(3), 226-232.
- Bem, D. J. (1967). Self-Perception: An Alternative Interpretation of Cognitive Dissonance Phenomena. *Psychological Review*, 74, 183-200.
- Benoit, R. G., Gilbert, S. J., Frith, C. D., & Burgess, P. W. (2011). Rostral prefrontal cortex and the focus of attention in prospective memory. *Cerebral Cortex*, 22(8), 1876-1886.
- Bettman, J. R., Luce, M. F., & Payne, J. W. (1998). Constructive consumer choice processes. *Journal of consumer research*, 25(3), 187-217.
- Bhui, R., & Gershman, S. J. (2018). Decision by sampling implements efficient coding of psychoeconomic functions. *Psychological review*, 125(6), 985.
- Bitzer, S., Park, H., Blankenburg, F., & Kiebel, S. J. (2014). Perceptual decision making: drift-diffusion model is equivalent to a Bayesian model. *Frontiers in human neuroscience*, 8, 102.
- Blain, B., Hollard, G., & Pessiglione, M. (2016). Neural mechanisms underlying the impact of daylong cognitive work on economic decisions. *Proceedings of the National Academy of Sciences*, 113(25), 6967-6972.
- Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological review*, 113(4), 700.
- Boldt, A., & Yeung, N. (2015). Shared neural markers of decision confidence and error detection. *Journal of Neuroscience*, 35(8), 3478-3484.

- Bostic, R., Herrnstein, R. J., & Luce, R. D. (1990). The effect on the preference-reversal phenomenon of using choice indifferences. *Journal of Economic Behavior & Organization*, 13(2), 193-212.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10, 433-436.
- Brehm, J. W. (1956). Postdecision changes in the desirability of alternatives. *The Journal of Abnormal and Social Psychology*, 52(3), 384.
- Caplin, A., & Dean, M. (2015). Revealed preference, rational inattention, and costly information acquisition. *American Economic Review*, 105(7), 2183-2203.
- Caplin, A., & Martin, D. (2014). A testable theory of imperfect perception. *The Economic Journal*, 125(582), 184-202.
- Caplin, A., & Dean, M. (2011). Search, choice, and revealed preference. *Theoretical Economics*, 6(1), 19-48.
- Caplin, A., Dean, M., & Martin, D. (2011). Search and satisficing. *American Economic Review*, 101(7), 2899-2922.
- Carland, M. A., Marcos, E., Thura, D., & Cisek, P. (2015). Evidence against perfect integration of sensory information during perceptual decision making. *Journal of neurophysiology*, 115(2), 915-930.
- Chaiken S. In: Zanna MP, Olson JM, Herman CP, eds. *Social Influence*. Hillsdale, NJ: Erlbaum; 1987, 3–39.
- Chammat, M., El Karoui, I., Allali, S., Hagège, J., Lehongre, K., Hasboun, D., ... & Navarro, V. (2017). Cognitive dissonance resolution depends on episodic memory. *Scientific reports*, 7, 41320.
- Chen, M. K., & Risen, J. L. (2010). How choice affects and reflects preferences: revisiting the free-choice paradigm. *Journal of personality and social psychology*, 99(4), 573.
- Chernev, A., Böckenholt, U., & Goodman, J. (2015). Choice overload: A conceptual review and meta-analysis. *Journal of Consumer Psychology*, 25(2), 333-358.
- Chib, V. S., Rangel, A., Shimojo, S., & O'Doherty, J. P. (2009). Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *Journal of Neuroscience*, 29(39), 12315-12320.
- Churchland, A. K., Kiani, R., & Shadlen, M. N. (2008). Decision-making with multiple alternatives. *Nature neuroscience*, 11(6), 693.
- Cisek, P., Puskas, G. A., & El-Murr, S. (2009). Decisions in changing conditions: the urgency-gating model. *Journal of Neuroscience*, 29(37), 11560-11571.
- Clithero, J. A., & Rangel, A. (2013). Informatic parcellation of the network involved in the computation of subjective value. *Social cognitive and affective neuroscience*, 9(9), 1289-1302.
- Colosio, M., Shestakova, A., Nikulin, V. V., Blagovechtchenski, E., & Klucharev, V. (2017). Neural mechanisms of cognitive dissonance (revised): An EEG study. *Journal of Neuroscience*, 37(20), 5074-5083.

- Coppin, G., Delplanque, S., Bernard, C., Cekic, S., Porcherot, C., Cayeux, I., & Sander, D. (2014). Choice both affects and reflects preferences. *The Quarterly Journal of Experimental Psychology*, *67*(7), 1415-1427.
- Coppin, G., Delplanque, S., Porcherot, C., Cayeux, I., & Sander, D. (2012). When flexibility is stable: implicit long-term shaping of olfactory preferences. *PloS one*, *7*(6), e37857.
- Coricelli, G., Dolan, R. J., & Sirigu, A. (2007). Brain, emotion and decision making: the paradigmatic example of regret. *Trends in cognitive sciences*, *11*(6), 258-265.
- Coricelli, G., Critchley, H. D., Joffily, M., O'Doherty, J. P., Sirigu, A., & Dolan, R. J. (2005). Regret and its avoidance: a neuroimaging study of choice behavior. *Nature neuroscience*, *8*(9), 1255.
- Daunizeau, J. (2017). Semi-analytical approximations to statistical moments of sigmoid and softmax mappings of normal variables. arXiv preprint arXiv:1703.00091.
- De Martino, B., Fleming, S.M., Garrett, N., and Dolan, R. (2013). Confidence in value-based choice. *Nat Neurosci* *16*, 105–110.
- De Martino, B., Kumaran, D., Holt, B., & Dolan, R. J. (2009). The neurobiology of reference-dependent value computation. *Journal of Neuroscience*, *29*(12), 3833-3842.
- Desrochers, T. M., Chatham, C. H., & Badre, D. (2015). The necessity of rostralateral prefrontal cortex for higher-level sequential behavior. *Neuron*, *87*(6), 1357-1368.
- Dhar, R., Nowlis, S. M., & Sherman, S. J. (2000). Trying hard or hardly trying: An analysis of context effects in choice. *Journal of Consumer Psychology*, *9*(4), 189-200.
- Dijksterhuis, A. (2004). Think different: the merits of unconscious thought in preference development and decision making. *Journal of personality and social psychology*, *87*(5), 586.
- Ditterich, J. (2006). Stochastic models of decisions about motion direction: behavior and physiology. *Neural networks*, *19*(8), 981-1012.
- Drugowitsch, J., Wyart, V., Devauchelle, A. D., & Koechlin, E. (2016). Computational precision of mental inference as critical source of human choice suboptimality. *Neuron*, *92*(6), 1398-1411.
- Drugowitsch, J., Moreno-Bote, R., Churchland, A. K., Shadlen, M. N., & Pouget, A. (2012). The cost of accumulating evidence in perceptual decision making. *Journal of Neuroscience*, *32*(11), 3612-3628.
- Dumontheil, I. (2014). Development of abstract thinking during childhood and adolescence: the role of rostralateral prefrontal cortex. *Developmental cognitive neuroscience*, *10*, 57-76.
- Egan, L. C., Bloom, P., & Santos, L. R. (2010). Choice-induced preferences in the absence of choice: Evidence from a blind two choice paradigm with young children and capuchin monkeys. *Journal of Experimental Social Psychology*, *46*(1), 204-207.
- Egan, L.C., Santos, L.R., and Bloom, P. (2007). The origins of cognitive dissonance evidence from children and monkeys. *Psychological Science* *18*, 978–983.
- Egner, T., Etkin, A., Gale, S., & Hirsch, J. (2007). Dissociable neural systems resolve conflict from emotional versus nonemotional distracters. *Cerebral cortex*, *18*(6), 1475-1484.

- Evans, N. J., Hawkins, G. E., Boehm, U., Wagenmakers, E. J., & Brown, S. D. (2017). The computations that support simple decision-making: A comparison between the diffusion and urgency-gating models. *Scientific reports*, 7(1), 16433.
- Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of personality and social psychology*, 50(2), 229.
- Feldman, J. M., & Lynch, J. G. (1988). Self-generated validity and other effects of measurement on belief, attitude, intention, and behavior. *Journal of applied Psychology*, 73(3), 421.
- Festinger, L. (1962). *A theory of cognitive dissonance* (Vol. 2). Stanford university press.
- Folke, T., Jacobsen, C., Fleming, S. M., & De Martino, B. (2017). Explicit representation of confidence informs future value-based decisions. *Nature Human Behaviour*, 1(1), 0002.
- Fox, C. R., & Weber, M. (2002). Ambiguity aversion, comparative ignorance, and decision context. *Organizational behavior and human decision processes*, 88(1), 476-498.
- Fox, C. R., & Levav, J. (2000). Familiarity bias and belief reversal in relative likelihood judgment. *Organizational Behavior and Human Decision Processes*, 82(2), 268-292.
- Fox, C. R., & Tversky, A. (1995). Ambiguity aversion and comparative ignorance. *The quarterly journal of economics*, 110(3), 585-603.
- Gabaix, X., Laibson, D., Moloche, G., & Weinberg, S. (2006). Costly information acquisition: Experimental analysis of a boundedly rational model. *American Economic Review*, 96(4), 1043-1068.
- Gehring, W. J., Goss, B., Coles, M. G., Meyer, D. E., & Donchin, E. (2018). The error-related negativity. *Perspectives on Psychological Science*, 13(2), 200-204.
- Gigerenzer, G., & Selten, R. (Eds.). (2002). *Bounded rationality: The adaptive toolbox*. MIT press.
- Gigerenzer, G., & Todd, P. M. (1999). Fast and frugal heuristics: The adaptive toolbox. In *Simple heuristics that make us smart* (pp. 3-34). Oxford University Press.
- Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: models of bounded rationality. *Psychological review*, 103(4), 650.
- Giguère, G., & Love, B. C. (2013). Limits in decision making arise from limits in memory retrieval. *Proceedings of the National Academy of Sciences*, 110(19), 7613-7618.
- Gilad, B., Kaish, S., & Loeb, P. D. (1987). Cognitive dissonance and utility maximization: A general framework. *Journal of Economic Behavior & Organization*, 8(1), 61-73.
- Gilovich, T., Griffin, D., & Kahneman, D. (Eds.). (2002). *Heuristics and biases: The psychology of intuitive judgment*. Cambridge university press.
- Gluth, S., Spektor, M. S., & Rieskamp, J. (2018). Value-based attentional capture affects multi-alternative decision making. *Elife*, 7, e39659.
- Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual review of neuroscience*, 30.
- Goldstein, W. M., & Einhorn, H. J. (1987). Expression theory and the preference reversal phenomena. *Psychological review*, 94(2), 236.

- Grether, D. M., & Plott, C. R. (1979). Economic theory of choice and the preference reversal phenomenon. *The American Economic Review*, 69(4), 623-638.
- Grueschow, M., Polania, R., Hare, T. A., & Ruff, C. C. (2015). Automatic versus choice-dependent value representations in the human brain. *Neuron*, 85(4), 874-885.
- Harlé, K. M., & Sanfey, A. G. (2007). Incidental sadness biases social economic decisions in the Ultimatum Game. *Emotion*, 7(4), 876.
- Harmon-Jones, E., Harmon-Jones, C., & Levy, N. (2015). An action-based model of cognitive-dissonance processes. *Current Directions in Psychological Science*, 24(3), 184-189.
- Harmon-Jones, E., Amodio, D. M., & Harmon-Jones, C. (2009). Action-based model of dissonance: A review, integration, and expansion of conceptions of cognitive conflict. *Advances in experimental social psychology*, 41, 119-166.
- Harmon-Jones, E., Harmon-Jones, C., Fearn, M., Sigelman, J. D., & Johnson, P. (2008). Left frontal cortical activation and spreading of alternatives: tests of the action-based model of dissonance. *Journal of personality and social psychology*, 94(1), 1.
- Harmon-Jones, E., & Harmon-Jones, C. (2007). Cognitive dissonance theory after 50 years of development. *Zeitschrift für Sozialpsychologie*, 38(1), 7-16.
- Hastie, R., & Park, B. (1986). The relationship between memory and judgment depends on whether the judgment task is memory-based or on-line. *Psychological review*, 93(3), 258.
- Hawkins, G. E., Forstmann, B. U., Wagenmakers, E. J., Ratcliff, R., & Brown, S. D. (2015). Revisiting the evidence for collapsing boundaries and urgency signals in perceptual decision-making. *Journal of Neuroscience*, 35(6), 2476-2484.
- Hawkins, G. E., Wagenmakers, E. J., Ratcliff, R., & Brown, S. D. (2015). Discriminating evidence accumulation from urgency signals in speeded decision making.
- Hebscher, M., & 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. *Frontiers in neuroscience*, 8, 150.
- Hick, W. E. (1952). On the rate of gain of information. *Quarterly Journal of experimental psychology*, 4(1), 11-26.
- Holroyd, C. B., & Coles, M. G. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychological review*, 109(4), 679.
- Holt, C. A. (1986). Preference reversals and the independence axiom. *American Economic Review*, 76(3), 508-515.
- Huber, J., Payne, J. W., & Puto, C. (1982). Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of consumer research*, 9(1), 90-98.

- Izuma, K., Akula, S., Murayama, K., Wu, D. A., Iacoboni, M., & Adolphs, R. (2015). A causal role for posterior medial frontal cortex in choice-induced preference change. *Journal of Neuroscience*, 35(8), 3598-3606.
- Izuma, K., & Murayama, K. (2013). Choice-induced preference change in the free-choice paradigm: a critical methodological review. *Frontiers in psychology*, 4, Art-No.
- Izuma, K., Matsumoto, M., Murayama, K., Samejima, K., Sadato, N., & Matsumoto, K. (2010). Neural correlates of cognitive dissonance and choice-induced preference change. *Proceedings of the National Academy of Sciences*, 107(51), 22014-22019.
- Jarcho, J. M., Berkman, E. T., & Lieberman, M. D. (2010). The neural basis of rationalization: cognitive dissonance reduction during decision-making. *Social cognitive and affective neuroscience*, 6(4), 460-467.
- Johnson, E. J., Payne, J. W., & Bettman, J. R. (1988). Information displays and preference reversals. *Organizational behavior and human decision processes*, 42(1), 1-21.
- Kaaronen, R.O. (2018). A Theory of Predictive Dissonance: Predictive Processing Presents a New Take on Cognitive Dissonance. *Front. Psychol.* 9.
- Kable, J. W., & Glimcher, P. W. (2007). The neural correlates of subjective value during intertemporal choice. *Nature neuroscience*, 10(12), 1625.
- Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.
- Kahneman, D., Slovic, S. P., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge university press.
- Kahneman, Daniel & Tversky, Amos, 1979. "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, *Econometric Society*, vol. 47(2), pages 263-291, March.
- Karni, E., & Safra, Z. (1987). "Preference reversal" and the observability of preferences by experimental methods. *Econometrica: Journal of the Econometric Society*, 675-685.
- Kepecs, A., Uchida, N., Zariwala, H. A., & Mainen, Z. F. (2008). Neural correlates, computation and behavioural impact of decision confidence. *Nature*, 455(7210), 227.
- Kitayama, S., Chua, H. F., Tompson, S., & Han, S. (2013). Neural mechanisms of dissonance: An fMRI investigation of choice justification. *Neuroimage*, 69, 206-212.
- Knetsch, J. L. (1989). The endowment effect and evidence of nonreversible indifference curves. *American economic review*, 79(5), 1277-1284.
- Kobayashi, S., de Carvalho, O. P., & Schultz, W. (2010). Adaptation of reward sensitivity in orbitofrontal neurons. *Journal of Neuroscience*, 30(2), 534-544.
- Krajbich, I., & Rangel, A. (2011). Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proceedings of the National Academy of Sciences*, 108(33), 13852-13857.
- Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature neuroscience*, 13(10), 1292.
- Kramer, T. (2007). The effect of measurement task transparency on preference construction and evaluations of personalized recommendations. *Journal of Marketing Research*, 44(2), 224-233.

- Kruglanski, A. W., & Webster, D. M. (1996). Motivated closing of the mind: "Seizing" and "freezing". *Psychological review*, 103(2), 263.
- Kvam, P. D. (2019). A geometric framework for modeling dynamic decisions among arbitrarily many alternatives. *Journal of Mathematical Psychology*, 91, 14-37.
- Lak, A., Costa, G. M., Romberg, E., Koulakov, A. A., Mainen, Z. F., & Kepecs, A. (2014). Orbitofrontal cortex is required for optimal waiting based on decision confidence. *Neuron*, 84(1), 190-201.
- Laughlin, S. B. (2001). Energy as a constraint on the coding and processing of sensory information. *Current opinion in neurobiology*, 11(4), 475-480.
- Lebreton, M., Abitbol, R., Daunizeau, J., & Pessiglione, M. (2015). Automatic integration of confidence in the brain valuation signal. *Nature neuroscience*, 18(8), 1159.
- Lebreton, M., Jorge, S., Michel, V., Thirion, B., & Pessiglione, M. (2009). An automatic valuation system in the human brain: evidence from functional neuroimaging. *Neuron*, 64(3), 431-439.
- Lee, D., & Daunizeau, J. (2019). Trading effort for confidence: the metacognitive control of decision-making. *bioRxiv*.
- Lee, D., & Daunizeau, J. (2019). Choosing what we like vs liking what we choose: How choice-induced preference change might actually be instrumental to decision-making. *bioRxiv*, 661116.
- Lennie, P. (2003). The cost of cortical computation. *Current biology*, 13(6), 493-497.
- Levy, D. J., & Glimcher, P. W. (2012). The root of all value: a neural common currency for choice. *Current opinion in neurobiology*, 22(6), 1027-1038.
- Levy, D. J., & Glimcher, P. W. (2011). Comparing apples and oranges: using reward-specific and reward-general subjective value representation in the brain. *Journal of Neuroscience*, 31(41), 14693-14707.
- Lichtenstein, S., & Slovic, P. (Eds.). (2006). *The construction of preference*. Cambridge University Press.
- Lichtenstein, S., & Slovic, P. (1973). Response-induced reversals of preference in gambling: An extended replication in Las Vegas. *Journal of Experimental Psychology*, 101(1), 16.
- Lichtenstein, S., & Slovic, P. (1971). Reversals of preference between bids and choices in gambling decisions. *Journal of experimental psychology*, 89(1), 46.
- Lieberman, M. D., Ochsner, K. N., Gilbert, D. T., & Schacter, D. L. (2001). Do amnesics exhibit cognitive dissonance reduction? The role of explicit memory and attention in attitude change. *Psychological Science*, 12(2), 135-140.
- Lim, S. L., O'Doherty, J. P., & Rangel, A. (2011). The decision value computations in the vmPFC and striatum use a relative value code that is guided by visual attention. *Journal of Neuroscience*, 31(37), 13214-13223.
- Lindman, H. R. (1971). Inconsistent preferences among gambles. *Journal of Experimental Psychology*, 89(2), 390.

- Locke, E. A., & Latham, G. P. (2006). New directions in goal-setting theory. *Current directions in psychological science*, 15(5), 265-268.
- Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American psychologist*, 57(9), 705.
- Loomes, G., Starmer, C., & Sugden, R. (1989). Preference reversal: information-processing effect or rational non-transitive choice?. *The Economic Journal*, 99(395), 140-151.
- Loomes, G., & Sugden, R. (1983). A rationale for preference reversal. *The American Economic Review*, 73(3), 428-432.
- Loomes, G., & Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *The economic journal*, 92(368), 805-824.
- Lopez-Persem, A., Domenech, P., & Pessiglione, M. (2016). How prior preferences determine decision-making frames and biases in the human brain. *Elife*, 5, e20317.
- Louie, K., Khaw, M. W., & Glimcher, P. W. (2013). Normalization is a general neural mechanism for context-dependent decision making. *Proceedings of the National Academy of Sciences*, 110(15), 6139-6144.
- Louie, K., & Glimcher, P. W. (2012). Efficient coding and the neural representation of value. *Annals of the New York Academy of Sciences*, 1251(1), 13-32.
- Louie, K., Gratton, L. E., & Glimcher, P. W. (2011). Reward value-based gain control: divisive normalization in parietal cortex. *Journal of Neuroscience*, 31(29), 10627-10639.
- Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization* (No. 8). Oxford University Press on Demand.
- Mansouri, F. A., Buckley, M. J., & Tanaka, K. (2007). Mnemonic function of the dorsolateral prefrontal cortex in conflict-induced behavioral adjustment. *Science*, 318(5852), 987-990.
- Mantel, S. P., & Kardes, F. R. (1999). The role of direction of comparison, attribute-based processing, and attitude-based processing in consumer preference. *Journal of Consumer Research*, 25(4), 335-352.
- March, J. G. (1978). Bounded rationality, ambiguity, and the engineering of choice. *The Bell Journal of Economics*, 587-608.
- Marcum, J. (1960). A statistical theory of target detection by pulsed radar. *IRE Transactions on Information Theory*, 6(2), 59-267.
- Marois, R., & Ivanoff, J. (2005). Capacity limits of information processing in the brain. *Trends in cognitive sciences*, 9(6), 296-305.
- Mengarelli, F., Spoglianti, S., Avenanti, A., & Di Pellegrino, G. (2013). Cathodal tDCS over the left prefrontal cortex diminishes choice-induced preference change. *Cerebral Cortex*, 25(5), 1219-1227.
- Milosavljevic, M., Malmaud, J., Huth, A., Koch, C., & Rangel, A. (2010). The drift diffusion model can account for the accuracy and reaction time of value-based choices under high and low time pressure. *Judgment and Decision Making*, 5(6), 437-449.

- Musslick, S., Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2015). A computational model of control allocation based on the Expected Value of Control. In The 2nd multidisciplinary conference on reinforcement learning and decision making.
- Nadarajah, S., & Kotz, S. (2008). Exact distribution of the max/min of two Gaussian random variables. *IEEE Transactions on very large scale integration (VLSI) systems*, 16(2), 210-212.
- Nakamura, K., & Kawabata, H. (2013). I choose, therefore I like: preference for faces induced by arbitrary choice. *PloS one*, 8(8), e72071.
- Navajas, J., Bahrami, B., & Latham, P. E. (2016). Post-decisional accounts of biases in confidence. *Current Opinion in Behavioral Sciences*, 11, 55-60.
- O'Connell, R. G., Shadlen, M. N., Wong-Lin, K., & Kelly, S. P. (2018). Bridging neural and computational viewpoints on perceptual decision-making. *Trends in neurosciences*, 41(11), 838-852.
- Oehr, C. R., Hanslmayr, S., Fell, J., Deuker, L., Kremers, N. A., Do Lam, A. T., ... & Axmacher, N. (2014). Neural communication patterns underlying conflict detection, resolution, and adaptation. *Journal of Neuroscience*, 34(31), 10438-10452.
- Palmer, J., Huk, A. C., & Shadlen, M. N. (2005). The effect of stimulus strength on the speed and accuracy of a perceptual decision. *Journal of vision*, 5(5), 1-1.
- Payne, J. W., Bettman, J. R., Schkade, D. A., Schwarz, N., & Gregory, R. (1999). Measuring constructed preferences: Towards a building code. In *Elicitation of preferences* (pp. 243-275). Springer, Dordrecht.
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge University Press.
- Payne, J. W., Bettman, J. R., Coupey, E., & Johnson, E. J. (1992). A constructive process view of decision making: Multiple strategies in judgment and choice. *Acta Psychologica*, 80(1-3), 107-141.
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1992). Behavioral decision research: A constructive processing perspective. *Annual review of psychology*, 43(1), 87-131.
- Payne, J. W., Bettman, J. R., Schkade, D. A., Schwarz, N., & Gregory, R. (1999). Measuring constructed preferences: Towards a building code. In *Elicitation of preferences* (pp. 243-275). Springer, Dordrecht.
- Perlovsky, L. (2013). A challenge to human evolution—cognitive dissonance. *Frontiers in Psychology*, 4, 179.
- Pike, R. (1973). Response latency models for signal detection. *Psychological Review*, 80(1), 53.
- Plassmann, H., O'Doherty, J., & Rangel, A. (2007). Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *Journal of neuroscience*, 27(37), 9984-9988.
- Pocheptsova, A., Amir, O., Dhar, R., & Baumeister, R. F. (2009). Deciding without resources: Resource depletion and choice in context. *Journal of Marketing Research*, 46(3), 344-355.

- Polania, R., Woodford, M., & Ruff, C. C. (2019). Efficient coding of subjective value. *Nature neuroscience*, 22(1), 134.
- Pommerehne, W. W., Schneider, F., & Zweifel, P. (1982). Economic theory of choice and the preference reversal phenomenon: A reexamination. *The American Economic Review*, 72(3), 569-574.
- Porcelli, A. J., & Delgado, M. R. (2009). Acute stress modulates risk taking in financial decision making. *Psychological Science*, 20(3), 278-283.
- Porcelli, A. J., Lewis, A. H., & Delgado, M. R. (2012). Acute stress influences neural circuits of reward processing. *Frontiers in neuroscience*, 6, 157.
- Quintana, J., Yajeya, J., & Fuster, J. M. (1988). Prefrontal representation of stimulus attributes during delay tasks. I. Unit activity in cross-temporal integration of sensory and sensory-motor information. *Brain research*, 474(2), 211-221.
- Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in cognitive sciences*, 20(4), 260-281.
- Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural computation*, 20(4), 873-922.
- Ratcliff, R., Thapar, A., Smith, P. L., & McKoon, G. (2005). Aging and response times: A comparison of sequential sampling models. *Measuring the mind: Speed, control, and age*, 3-32.
- Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological review*, 111(2), 333.
- Ratcliff, R., Thapar, A., & McKoon, G. (2004). A diffusion model analysis of the effects of aging on recognition memory. *Journal of Memory and Language*, 50(4), 408-424.
- Ratcliff, R., Thapar, A., & McKoon, G. (2003). A diffusion model analysis of the effects of aging on brightness discrimination. *Perception & psychophysics*, 65(4), 523-535.
- Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic bulletin & review*, 9(3), 438-481.
- Ratcliff, R., Thapar, A., & McKoon, G. (2001). The effects of aging on reaction time in a signal detection task. *Psychology and aging*, 16(2), 323.
- Ratcliff, R., Van Zandt, T., & McKoon, G. (1999). Connectionist and diffusion models of reaction time. *Psychological review*, 106(2), 261.
- Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological science*, 9(5), 347-356.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological review*, 85(2), 59.
- Ratneshwar, S., Shocker, A. D., & Stewart, D. W. (1987). Toward understanding the attraction effect: The implications of product stimulus meaningfulness and familiarity. *Journal of Consumer Research*, 13(4), 520-533.
- Reilly, R. J. (1982). Preference reversal: Further evidence and some suggested modifications in experimental design. *The American Economic Review*, 72(3), 576-584.

- Reutskaja, E., Lindner, A., Nagel, R., Andersen, R. A., & Camerer, C. F. (2018). Choice overload reduces neural signatures of choice set value in dorsal striatum and anterior cingulate cortex. *Nature human behaviour*, 2(12), 925.
- Risen, J. L., & Chen, M. K. (2010). How to Study Choice-Induced Attitude Change: Strategies for Fixing the Free-Choice Paradigm. *Social and Personality Psychology Compass*, 4(12), 1151-1164.
- Salti, M., El Karoui, I., Maillet, M., & Naccache, L. (2014). Cognitive dissonance resolution is related to episodic memory. *PloS one*, 9(9), e108579.
- Samuelson, P. A. (1938). A note on the pure theory of consumer's behaviour. *Economica*, 5(17), 61-71.
- Sanbonmatsu, D. M., & Fazio, R. H. (1990). The role of attitudes in memory-based decision making. *Journal of Personality and social Psychology*, 59(4), 614.
- Schkade, D. A., & Johnson, E. J. (1989). Cognitive processes in preference reversals. *Organizational Behavior and Human Decision Processes*, 44(2), 203-231.
- Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological review*, 84(1), 1.
- Segal, U. (1988). Does the preference reversal phenomenon necessarily contradict the independence axiom?. *The American Economic Review*, 78(1), 233-236.
- Shanteau, J. (1988). Psychological characteristics and strategies of expert decision makers. *Acta psychologica*, 68(1-3), 203-215.
- Sharot, T., Fleming, S. M., Yu, X., Koster, R., & Dolan, R. J. (2012). Is choice-induced preference change long lasting?. *Psychological science*, 23(10), 1123-1129.
- Sharot, T., Velasquez, C. M., & Dolan, R. J. (2010). Do decisions shape preference? Evidence from blind choice. *Psychological science*, 21(9), 1231-1235.
- Sharot, T., De Martino, B., & Dolan, R. J. (2009). How choice reveals and shapes expected hedonic outcome. *Journal of Neuroscience*, 29(12), 3760-3765.
- Shenhav, A., & Karmarkar, U. R. (2019). Dissociable components of the reward circuit are involved in appraisal versus choice. *Scientific reports*, 9(1), 1958.
- Shenhav, A., Musslick, S., Lieder, F., Kool, W., Griffiths, T. L., Cohen, J. D., & Botvinick, M. M. (2017). Toward a rational and mechanistic account of mental effort. *Annual review of neuroscience*, 40, 99-124.
- Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron*, 79(2), 217-240.
- Simon, D., Stenstrom, D. M., & Read, S. J. (2015). The coherence effect: Blending cold and hot cognitions. *Journal of Personality and Social Psychology*, 109(3), 369.
- Simon, D., Krawczyk, D. C., & Holyoak, K. J. (2004). Construction of preferences by constraint satisfaction. *Psychological Science*, 15(5), 331-336.
- Simon, D., & Holyoak, K. J. (2002). Structural dynamics of cognition: From consistency theories to constraint satisfaction. *Personality and social psychology review*, 6(4), 283-294.

- Simon, H. A. (1982). Models of bounded rationality. vol. 2, Behavioral economics and business organization. MIT Press.
- Simon, H. A. (1957). Models of man; social and rational.
- Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological review*, 63(2), 129.
- Simon, H. A. (1955). A behavioral model of rational choice. *The quarterly journal of economics*, 69(1), 99-118.
- Simonson, I. (1989). Choice based on reasons: The case of attraction and compromise effects. *Journal of consumer research*, 16(2), 158-174.
- Sims, C. A. (2003). Implications of rational inattention. *Journal of monetary Economics*, 50(3), 665-690.
- Slovic, P. (1995). The construction of preference. *American psychologist*, 50(5), 364.
- Slovic, P., & Lichtenstein, S. (1983). Preference reversals: A broader perspective. *The American Economic Review*, 73(4), 596-605.
- Slovic, P., Fischhoff, B., & Lichtenstein, S. (1980). Facts and fears: Understanding perceived risk. In *Societal risk assessment* (pp. 181-216). Springer, Boston, MA.
- Slovic, P., Fischhoff, B., & Lichtenstein, S. (1977). Behavioral decision theory. *Annual review of psychology*, 28(1), 1-39.
- Slovic, P., Griffin, D., & Tversky, A. (1990). Compatibility effects in judgment and choice. *Insights in decision making: A tribute to Hillel J. Einhorn*, 5-27.
- Slovic, P., & Lichtenstein, S. (1968). Relative importance of probabilities and payoffs in risk taking. *Journal of experimental psychology*, 78(3p2), 1.
- Sokol-Hessner, P., Camerer, C. F., & Phelps, E. A. (2012). Emotion regulation reduces loss aversion and decreases amygdala responses to losses. *Social cognitive and affective neuroscience*, 8(3), 341-350.
- Tajima, S., Drugowitsch, J., Patel, N., & Pouget, A. (2019). Optimal policy for multi-alternative decisions. *Nature neuroscience*, 22, 1503–1511 .
- Tajima, S., Drugowitsch, J., & Pouget, A. (2016). Optimal policy for value-based decision-making. *Nature communications*, 7, 12400.
- Thapar, A., Ratcliff, R., & McKoon, G. (2003). A diffusion model analysis of the effects of aging on letter discrimination. *Psychology and Aging*, 18(3), 415.
- Thorngate, W. (1980). Efficient decision heuristics. *Behavioral Science*, 25(3), 219-225.
- Thura, D., Beauregard-Racine, J., Fradet, C. W., & Cisek, P. (2012). Decision making by urgency gating: theory and experimental support. *Journal of neurophysiology*, 108(11), 2912-2930.
- Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *The quarterly journal of economics*, 106(4), 1039-1061.
- Tversky, A., Slovic, P., & Kahneman, D. (1990). The causes of preference reversal. *The American Economic Review*, 204-217.

- Tversky, A., & Thaler, R. H. (1990). Anomalies: preference reversals. *Journal of Economic Perspectives*, 4(2), 201-211.
- Tversky, A., & Kahneman, D. (1989). Rational choice and the framing of decisions. In *Multiple criteria decision making and risk analysis using microcomputers* (pp. 81-126). Springer, Berlin, Heidelberg.
- Tversky, A., Sattath, S., & Slovic, P. (1988). Contingent weighting in judgment and choice. *Psychological review*, 95(3), 371.
- Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211(4481), 453-458.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *science*, 185(4157), 1124-1131.
- Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological review*, 79(4), 281.
- Tversky, A. (1972). Choice by elimination. *Journal of mathematical psychology*, 9(4), 341-367.
- Tversky, A. (1969). Intransitivity of preferences. *Psychological review*, 76(1), 31.
- Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: the leaky, competing accumulator model. *Psychological review*, 108(3), 550.
- Van Veen, V., Krug, M. K., Schooler, J. W., & Carter, C. S. (2009). Neural activity predicts attitude change in cognitive dissonance. *Nature neuroscience*, 12(11), 1469.
- Vickers, D. (1979). *Decision processes in visual perception*. Academic Press.
- Voigt, K., Murawski, C., Speer, S., & Bode, S. (2018). Hard decisions shape the neural coding of preferences. *bioRxiv*, 298406.
- Von Neumann, J., & Morgenstern, O. (1947). *Theory of games and economic behavior*, 2nd rev.
- Wang, X. J. (2008). Decision making in recurrent neuronal circuits. *Neuron*, 60(2), 215-234.
- Wang, X. J. (2002). Probabilistic decision making by slow reverberation in cortical circuits. *Neuron*, 36(5), 955-968.
- Wang, Z., & Busemeyer, J. R. (2016). Interference effects of categorization on decision making. *Cognition*, 150, 133-149.
- Warren, C., McGraw, A. P., & Van Boven, L. (2011). Values and preferences: Defining preference construction. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2(2), 193-205.
- Weissman, D. H., Gopalakrishnan, A., Hazlett, C. J., & Woldorff, M. G. (2004). Dorsal anterior cingulate cortex resolves conflict from distracting stimuli by boosting attention toward relevant events. *Cerebral cortex*, 15(2), 229-237.
- Weissman, D. H., Giesbrecht, B., Song, A. W., Mangun, G. R., & Woldorff, M. G. (2003). Conflict monitoring in the human anterior cingulate cortex during selective attention to global and local object features. *Neuroimage*, 19(4), 1361-1368.

- West, S., Jett, S.E., Beckman, T., and Vonk, J. (2010). The phylogenetic roots of cognitive dissonance. *J Comp Psychol* 124, 425–432.
- Westphal, A. J., Chow, T. E., Ngoy, C., Zuo, X., Liao, V., Storozuk, L. A., ... & Rissman, J. (2019). Anodal Transcranial Direct Current Stimulation to the Left Rostrolateral Prefrontal Cortex Selectively Improves Source Memory Retrieval. *Journal of cognitive neuroscience*, 1-12.
- Wong, K. F., & Wang, X. J. (2006). A recurrent network mechanism of time integration in perceptual decisions. *Journal of Neuroscience*, 26(4), 1314-1328.
- Wyart, V., & Koechlin, E. (2016). Choice variability and suboptimality in uncertain environments. *Current Opinion in Behavioral Sciences*, 11, 109-115.
- Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004). The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychological review*, 111(4), 931.
- Zeelenberg, M. (1999). Anticipated regret, expected feedback and behavioral decision making. *Journal of behavioral decision making*, 12(2), 93-106.

## **APPENDIX - Unresolved Issues**

### *Taking a closer look at the variables*

In spite of its merits, our studies also highlighted some shortcomings of the original version of our model (what we now call the “allocation phase”). First among these is the fact that it was inflexible in the sense that a given set of inputs would always lead to the same set of outputs, which left it with no hope of ever perfectly fitting real data that is dynamic by nature (e.g., response time, pupil dilation). This version of the MCD model was a static and purely proactive model, which did not allow it to account for variability in observed data. Indeed, its “all-or-nothing” strategy for investing the optimal level of effort probably did not realistically describe actual human behavior. As we have detailed above, we now present our MCD model as being composed of two separate phases. A complete MCD model that includes both phases continues to correctly predict primary experimental output, and it additionally predicts the inter-variable intricacies that the original model (phase one only) could not (i.e., spreading of alternatives vs choice confidence or effort). Beyond behavioral data, our complete model could perhaps also explain certain inter-trial aspects of pupil dilation (e.g., mean or maximum dilation), but we do not yet know the best way of summarizing pupil dynamics on a trial-by-trial basis.

### *Reward and cost functions*

One of the fundamental components of our model is the reward term, which serves to incentivize higher effort allocation. We sought to demonstrate this by including both consequential and non-consequential conditions. For the consequential condition, participants were required to eat real food at the end of the experiment, as determined by their actual choices. In the non-consequential condition, all decisions were purely hypothetical. Our manipulation seems to have worked, since subjective effort, response time, and pupil dilation were greater in the consequential condition (vs neutral). This study has provided us with a partial answer to a

critical question for our work—what exactly is it that incentivizes a decision task such that the decision maker (DM) will be willing to invest more effort in hopes of choosing with greater confidence? With objective decisions, external incentive schemes would be easy to implement. A correct answer could simply be compensated with monetary or some other tangible reward. But, since we study subjective choices, where there is no concept of a knowable correct choice, we needed to look for another method. The bottom line is that we needed to make the task a true *decision*, rather than a mere statement about preferences or claim about what choice the DM might make if ever actually presented with such a true decision. We did this by *enforcing* the choices that participants made. In our interpretation, a consequential choice is more important than a non-consequential choice, which increases the reward term in our model, thus motivating a greater investment of mental effort for any given decision.

Just as the reward term in our model is not yet well-defined, neither is the cost term. For the moment, we have included a vague cost function that monotonically increases with effort, but the way in which mental effort allocation is costly could actually be much more complex than that. Perhaps there is a fixed cost to set up the necessary neural representations of the decision task. Maybe the variable cost that increases as a function of “effort” would actually be comprised of different types of costs for different types of efforts (e.g., required for different types of information processing). It seems reasonable to assume that there might be an additional cost of time itself, and yet another related to simultaneous demands (e.g., dual tasks). We might gain further insight into the cognitive mechanisms that we are attempting to describe if we include a more intricate cost function in future versions of our model. As a demonstration of one type of cost, we artificially inflated the cost of time on task in our experiment by introducing a financial penalty. Here, as predicted, effort as represented by both self-assessment and RT was lower with the penalty than without (although pupil dilation was greater). Observing the effects of different types of *external* costs could lend insight into the DM’s

*internal* cost function. It would be interesting to explore how the effects of different types of costs might manifest themselves differently. For example, preliminary analysis shows the non-intuitive prediction that adding a fixed cost of effort initiation would not change the EVC-optimal level of resource allocation (unless the magnitude of the EVC-optimal expected benefit is less than the magnitude of the fixed cost, in which case no effort would be allocated). A multi-attribute model (see below) would allow for a separate cost term for estimate refinements within each attribute dimension, as well as for the integration function across dimensions. This could enable us to predict effort/confidence levels for each attribute separately. It would likely allow us to explain default biases (e.g., choosing to eat a cupcake without deliberation, but choosing a bowl of oats with deliberation). The idea here is that certain types of information are more costly to process (e.g., factual knowledge) than others (e.g., experiential memories), in which case fewer resources might be allocated to the costly dimension. Furthermore, a cost of refinement of the integration function (e.g., weighted sum) could allow us to predict effort/confidence levels for items that are similar versus dissimilar, in terms of their attribute compositions.

#### *Relationships between spreading of alternatives, effort, and confidence*

Though we did not originally predict the observed relationships between spreading of alternatives and either response time or choice confidence, in retrospect, they are consistent with our model and rather easy to understand. When initially presented with a choice pair, an automatic value comparison will be made in the mind of the DM. This initial value comparison will give rise to an initial choice confidence level (the level of confidence with which the DM could make the choice immediately, without further contemplation). If this confidence level was already “high enough”, the DM would simply proceed to making the choice. Otherwise, he would start to invest mental effort by deliberating about the values of the options. Each incremental effort expenditure (i.e., the processing of each new piece of information), could

change the DM's relative valuations of the options. Sometimes the options would become more distinguishable, eventually reaching the point where choice confidence would reach the minimal threshold level and the choice would be made (resulting in a faster-than-expected choice). Sometimes the options would become less distinguishable, and the DM would eventually stop deliberating after the maximal effort threshold had been reached (resulting in a less-confident-than-expected choice). Notice that under this framework, spreading of alternatives is precisely the movement of the dynamic relative valuations. Spreading of alternatives can be rephrased as "making the choice easier", which directly implies higher confidence. Following this logic, it is apparent that higher spreading of alternatives will generally be matched with faster choices and higher choice confidence (when controlling for difficulty).

#### *Rating sum*

One unexpected finding that appeared in every one of our studies was the correlation between choice set rating sum and several dependent variables. Notably, rating sum correlated negatively with effort and positively with confidence. The explanation for this is not clear. It is possible that when both options are more disliked, the choice subjectively feels more difficult to the DM (as opposed to when both options are more liked, in which case, the DM might be happy to obtain either of them). This is reminiscent of the classic *loss aversion* effect, where minimizing the magnitude of a negative outcome might be more important than maximizing the magnitude of a positive outcome [Tversky & Kahneman, 1991]. If true, this could account for the observed pattern of higher RT and subjective effort (i.e., greater resource allocation motivated by greater choice importance). This would mean that the value of the options in a choice pair would be akin to the reward term in our model. Indeed, a larger reward term leads to a larger EVC-optimal resource allocation, all else equal. However, it also leads to a higher EVC-optimal expected confidence, which contradicts the empirical results. It could be possible

that the value of the options directly distorts choice confidence. The idea here would be that confidence did not merely represent the “probability of choosing the best option”, but also the “probability of being satisfied with the chosen option”. The combination of these two probabilities puts this idea in line with regret theory, wherein a DM will consider how much regret he would expect to feel after experiencing the outcome of his choice [Zeelenberg, 1999; Loomes & Sugden, 1982]. The DM will always be inclined to choose the best option, because otherwise he would regret his choice later. But, if both options were good, he would probably enjoy the outcome even if he chose the option that was not the best. This would support the prediction that higher option value leads to higher choice confidence. In our model, this would also lead to lower effort, because the beneficial effect of effort on confidence would be diminished. Neuroimaging studies have exposed activity in the orbitofrontal cortex (OFC) that reflects choice confidence [Lak et al, 2014; Kepecs et al, 2008] or anticipated regret [Coricelli, Dolan & Sirigu, 2007; Coricelli et al, 2005], but an intriguing next step would be to look for neural activity that simultaneously correlates with both.

### *Familiarity*

We added the familiarity assessment in our experiments in order to test for different results when items were more versus less known to the participants. Specifically, we wanted to contrast situations when value certainty was low due to lack of familiarity versus when it was low in spite of the item being well known. We expected that low familiarity items would always have low certainty, whereas higher familiarity items would have a wider range of certainty. However, we did not find a straightforward correlation between familiarity and certainty. Moreover, many items were scored with very low familiarity but very high value certainty. This implies that even for items that were unfamiliar, participants were often able to construct an adequate internal value representation. It is possible that such a value construction process was either different or altogether unnecessary for items that were familiar, whose values could instead be

derived from item-specific prior knowledge or perhaps simply retrieved from memory. If true, one could reasonably suspect that the amount of effort required to evaluate unfamiliar and familiar items would differ. Indeed, the correlations we observe between familiarity and almost all of our dependent variables support this idea. More curious is the clear positive correlation between familiarity and value, across the full range of measurements. This is reminiscent of the classic *familiarity bias*, which demonstrates that more familiar options are generally preferred to unfamiliar items (c.f., *ambiguity aversion*) [Fox & Weber, 2002; Fox & Levav, 2000; Fox & Tversky, 1995]. The idea here is that people would be more likely to choose options that they were more familiar with, even when the choice was supposed to be based solely on value. However, what is interesting here is the fact that we see evidence of this bias in the value ratings themselves, before any choice context was even introduced, adding an interesting new twist to the story.

There is another, more principled role that familiarity ratings could potentially play in our model. Recall that, in our model, mental resource allocation leads to information processing. But this assumes that there is necessarily information available to be processed. It is easy to envision a situation in which a decision maker is inclined to invest a high level of resources in order to achieve a high level of confidence, but is simply unable to find any relevant information no matter how hard he tries. For example, consider a decision maker presented with options that he does not even recognize, nor have any belief about which features might be relevant to the decision at hand. Clearly, this decision maker will not be able to process any information, nor gain any confidence, regardless of how much effort he invests. A familiarity rating such as ours could perhaps be used to represent the availability of information to be considered, a term which could be included in future versions of the MCD model.

### *Theoretical versus experimental variables: certainty*

In spite of our promising results, there still might have been an imperfect mapping between our experimental variables and their theoretical representations. One obvious contender would be self-reported rating certainty. It was our intention to use this variable as a proxy for the precision of the internal value representations of items. As such, deliberation should always serve to increase this certainty (or at least never decrease it). However, almost half of all items across all subjects had a lower certainty about the post-choice rating than about the pre-choice rating. This suggests that the subjects' stated certainty about an item's value might be more akin to the expected precision of the item's categorical value distribution (e.g., some red wines might be very much liked, while others might be very much disliked) rather than the certainty that the estimation of the mean of the distribution was accurate. In other words, rating certainty could be proportional to the range within which the DM believes the true value lies. Indeed, the way in which we solicited the certainty level might have encouraged this. Our observation that post-choice certainty was sometimes higher and sometimes lower than pre-choice certainty would be consistent with this idea. If choice deliberation is thought to be a time where the DM is processing multiple bits of information about each option, and if a conglomeration of information could be predominantly congruent or predominantly incongruent, the DM's representations of the value distributions could become either more or less narrow. We might thus claim that participants' interpretation of "certainty" about the ratings was not the same as the precision term in our model, which represents the degree of certainty that such ratings are accurate. One unexpected result that might further illustrate this point is the lack of correlation between pre-choice certainty and subjective effort. For the precision term as we intend it in our model, these should be negatively correlated, because an imprecise value representation would require effort to make it more precise in order to help distinguish the options. With the aforementioned distinction in mind, it could perhaps be interesting to adjust our model to

include “precision” rather than “certainty”, or perhaps both terms simultaneously. In this same spirit, we should perhaps also consider making the perturbation term in our model dependent on certainty/precision. For example, when the value estimate is certain/precise, it might make sense that it cannot be perturbed much; when it is uncertain/imprecise, it could have a lot more “wiggle room” for reassessment during deliberation.

### *Rating RT versus certainty*

One might claim that the time it takes to enter a rating for an item is a good measure of the level of certainty about that rating. Our data support this hypothesis, showing a strong negative correlation between RT and certainty (for the pre-choice ratings). The change in RT (post- vs pre-choice ratings) might then be a good proxy for change in certainty. Our data show this direct correlation, as well as some indirect evidence: RT change correlated nicely with (absolute) rating change. The idea here is that rating change implies additional information processing, which in turn implies higher certainty. This higher certainty, in turn, would be represented by a larger decrease in RT. Under this framework, we could perhaps simply take rating RT to represent rating certainty in our model. Indeed, the fact that most ratings were entered faster post-choice as compared to pre-choice would validate our model predictions that value certainty should always be greater (or equal) after choice deliberation than before.

### How to Measure Mental Effort?

#### *Response time*

Finally, if there is one element of our model that can be said to be its *raison d'être*, it would be what we call “z”—the amount of mental resources allocated to a decision task, or mental effort. But what is this mysterious “z”? Response time is frequently used as a proxy for mental effort,

but this is often met with skepticism. The typical argument is that although participants could take a long time to respond because they are focusing and thinking hard about the task, their response could also be delayed because they are *not* focused on the task and simply letting other thoughts distract them. In our experiment, we added an external cost of time in hopes of persuading participants not to waste their resources by letting their focus wander away from the task. We found qualitatively similar results in the penalty and non-penalty conditions, allowing us to conclude that time spent on task was indeed time spent investing mental effort. Furthermore, RT was highly correlated with the subjects' own perception of effort in all conditions. Therefore, we believe that RT is an acceptable proxy for mental effort expenditure in choice tasks. But, although mental effort can be represented by time on task, surely it is more complex than that. Time alone is therefore an unconvincing choice of proxy. Pupil dilation might be a better choice, as it is a physiological response that the DM cannot control (and is typically not even cognizant of). Furthermore, we believe that it is reasonable to claim that effort can be broken down into two dimensions: duration and intensity. We also assume that the total effort a participant is willing to invest in a task (duration x intensity) might remain constant even if the composition were to change. We analyzed our experimental data with this in mind—we thought that imposing a penalty on time should encourage participants to shift away from effort duration and towards effort intensity. As predicted, we observed much lower RT and much greater pupil dilation in the penalty versus non-penalty conditions. We therefore concluded that pupil dilation along with RT is a fine proxy for effort.

It could be interesting to separate the “effort” component of our model into the duration and intensity durations. It is unclear whether this would have any impact on the allocation phase of the model, but it would definitely play a part in the dynamics of the implementation phase. One obvious modeling choice would be to simply include a parameter that transforms response time into effort, which would be the intensity term. More complicated dynamics, such as an intensity

term that evolves across deliberation time (perhaps as a function of instantaneous confidence or distance to maximum effort), could be included as well, but we have not yet explored any such ideas.

### *Pupil dilation*

When examining the pupil response dynamics, we expected to find a positive correlation between choice difficulty and pupil size. We thought that this would represent a greater effort expenditure for more difficult choices. We also expected to find a positive correlation between subjective effort (as a more direct readout of effort expenditure) and pupil size. Indeed, we found both correlations. The effect was both stronger and more significant for subjective effort than for difficulty. The effect of effort remained even when controlling for difficulty. Furthermore, when using both difficulty and effort as regressors simultaneously, there was no longer any effect of difficulty. We believe that this should help settle the debate in the field as to whether physiological responses to mental effort and task difficulty can be disentangled and that purported measurements of effort might actually be reporting nothing more than automatic responses to difficulty. Clearly, difficulty and effort will not always be proportional, because the effort actually exerted might ultimately differ from the “required effort” (i.e., difficulty). In our data, the pupil response correlates better with effort than with difficulty.

Unfortunately, the link between effort and pupil dilation is still not very well understood. This and the fact that many other factors (known and unknown) can affect pupil dilation make it a dubious marker of mental effort allocation as well. Ideally, we would want to directly measure brain activity and use that as a more direct and pure indicator of the amount of effort expended during each trial. Inspired by this idea, future studies should include EEG/MEG recordings to demonstrate trial-by-trial effort levels. Beyond that, they should seek to trace the evolution of within-trial signals and to form links with the output of computational models such as ours.

### How are initial value estimates formed?

Our model takes as input the initial values of the mean and variance of the value distribution for each of the decision options. What is not clear, however, is how these values are formed in the first place. One idea is that from the moment the stimuli are first perceived, the brain automatically generates componential representations for each item (see Quintana, Yajeya & Fuster, 1988, for an exposition of neural populations encoding different stimuli attributes across decision time), along with a mapping from the components to a unitless value dimension within which the options can be compared (see Grueschow et al, 2015, and Lebreton et al, 2009, for evidence that value is automatically encoded in neural activity regardless of which attribute is relevant to the task at hand). This is tantamount to the construction of an imprecise set of relevant attribute dimensions (a temporary basis of comparison), imprecise estimates of each option's score within each dimension, and imprecise estimates of the contribution of each dimension towards the options' overall value ratings. This process is automatic, but nevertheless requires non-zero time. This is consistent with empirical findings (including in this current study), in which even the easiest of decisions require a certain minimal amount of time before a choice is made. Known in the literature as “non-decision time”, this is the time when a DM is thought to perceive the choice options before being able to actually assess their relative values or deliberate about them. Although most models of decision-making, of which there are many, fundamentally rely on subjective value, no models exist (to our knowledge) to demonstrate the construction of value estimates when no choice is involved (but, see Polania, Woodford & Ruff, 2019, for a model that simultaneously includes value estimates and choices). Perhaps, then, a necessary next step for computational models of choice is to first include a computational description of value estimation.

### Metacognitive control of multi-attribute decision-making

Continuing with the concept of decisions based on comparison across multiple attribute dimensions, we ask ourselves if our model could be refined to account for the different elements of this type of decision process. Presumably, this would amount to a separation of the costs and benefits associated with refinements of either the set, measurements, or weights of attributes. It would follow that an optimal DM would use the MCD model to select not only *how much* mental effort to allocate to the task, but also *how* to allocate it. It is obvious from the description of our model that too little or too much effort would be suboptimal. It would perhaps be more interesting, and possibly non-intuitive, to see how variations in the dispersal of total cognitive resource investment (e.g., to different brain regions or to support different calculations) might impact optimality. Modeling and testing multi-attribute decision-making was beyond the scope of this project, but this will be the focus of upcoming research.