Value Estimation and Comparison in Multi-Attribute Choice

Thesis by

Geoffrey Fisher

In Partial Fulfillment of the Requirements for the degree

of

Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2015

(Defended May 12, 2015)

ACKNOWLEDGEMENTS

Before applying to doctoral programs, I was lucky enough to attend two seminars on neuroeconomics given by Antonio Rangel. Simply put, these talks changed my life. The passion with which Antonio presented his work was infectious, and within an hour I was convinced I had to pursue similar topics as a graduate student. When I had the opportunity to go to Caltech and work with him, there really was no choice to be made. Certainly, the research that follows in this dissertation would not be possible without Antonio's generous support. From the rigorous training, to providing outstanding resources, to constantly working to make me a better scientist, Antonio has exceeded even the loftiest expectations one could have for any advisor or mentor.

Many others have been extremely influential throughout my time at Caltech. My committee members, Marina Agranov, Kota Saito, and Ben Gillen, have always offered extremely useful advice. Current and former members of the Rangel Lab have provided unlimited feedback and guidance throughout the years and include: John Clithero, Cary Frydman, Giovanni Gentile, Cendri Hutcherson, Peter Sokol-Hessner, Nikki Sullivan, Gabriela Tavares, and Anita Tusche. I'm also grateful to many of my fellow HSS graduate students, especially Matt Chao. Additionally, Jenny Niese and Laurel Auchampaugh have always helped make sure things run smoothly. Before Caltech, Dan Benjamin first introduced me to behavioral economics; who knows what I'd be doing now without his guidance.

I am grateful to my family for their constant support throughout graduate school. My parents, Alan and Wendy, who have been nothing but encouraging, and are always there to help out during rough patches. I'm grateful for my brother, Ethan, who I am extremely proud of and who has done far more hard work over the past several years than I have. Finally I'm thankful for my fiancée, Katie, who has experienced the many highs and lows of Caltech right along with me, but has always been a more supportive and loving partner than I could ask for.

ABSTRACT

The following work explores the processes individuals utilize when making multi-attribute choices. With the exception of extremely simple or familiar choices, most decisions we face can be classified as multi-attribute choices. In order to evaluate and make choices in such an environment, we must be able to estimate and weight the particular attributes of an option. Hence, better understanding the mechanisms involved in this process is an important step for economists and psychologists. For example, when choosing between two meals that differ in taste and nutrition, what are the mechanisms that allow us to estimate and then weight attributes when constructing value? Furthermore, how can these mechanisms be influenced by variables such as attention or common physiological states, like hunger?

In order to investigate these and similar questions, we use a combination of choice and attentional data, where the attentional data was collected by recording eye movements as individuals made decisions. Chapter 1 designs and tests a neuroeconomic model of multi-attribute choice that makes predictions about choices, response time, and how these variables are correlated with attention. Chapter 2 applies the ideas in this model to intertemporal decision-making, and finds that attention causally affects discount rates. Chapter 3 explores how hunger, a common physiological state, alters the mechanisms we utilize as we make simple decisions about foods.

TABLE OF CONTENTS

INTRODUCTION

What are the algorithms we employ when making simple multi-attribute decisions, such as choosing between foods that differ in taste and nutrition? How do we estimate and weight attributes as we make decisions? Chapter 1 proposes a new accumulator model, the multi-attribute attentional drift diffusion model (maDDM), that computationally describes the choice process and allows that process to be guided by visual attention. Using a laboratory experiment, we find the maDDM makes accurate quantitative predictions about several key variables including what we choose, how long it takes to make a choice, and how these variables are correlated with attention to different attributes. Furthermore, we estimate an attribute-based fixation bias that suggests attention to an attribute increases its subjective weight by 5%, while the unattended attribute's weight is decreased by 10%. Our findings imply we may use similar computational processes as we make multi-attribute choices.

Chapter 2 explores the consequences that such a model has for intertemporal choices. It is well known that discount rates vary widely, both across and within individuals. We propose that a sizable fraction of this variation results from differences in how relative attention is allocated to different features of the decision, such as immediate versus future rewards. We tested this hypothesis using an experiment in which subjects chose between receiving smaller-sooner versus larger-later monetary prizes, while we recorded their attention using eye tracking. We find that cross-subject variation in the allocation of attention explained between 20% and 35% of the individual differences in discounting, and that cross-trial variation explained about 30% of the subjects' ability to choose the delayed option. We carried out two additional experiments in which we exogenously manipulated the allocation of attention, and found that shifting attention to the attributes that are relatively more attractive in the larger-later option increased patience. Together, these results are consistent with the existence of a causal impact of relative attention on intertemporal choice.

Chapter 3 seeks to understand how certain physiological states can influence the attribute estimation and weighting process. Individuals commonly mispredict their future preferences when they make decisions in a visceral state different from their anticipated state at consumption. We used a bidding food task to test whether cold-to-hot and hot-to-cold errors are symmetric in size and driven by similar mechanisms, while we exogenously varied subjects' hunger levels at the time of decision and consumption. We found that the effect size is symmetric: hungry subjects overbid 20 cents for a snack to eat when they would be satiated, and satiated subjects underbid 19 cents for a snack to eat when they would be hungry. Furthermore, we found evidence that these gaps are being driven by symmetric mechanisms that operate on the evaluation of visceral features of food, like taste, as opposed to more cognitive dimensions, like health.

Chapter 1

THE MULTI-ATTRIBUTE ATTENTIONAL DRIFT DIFFUSION MODEL

Except for very simple and familiar choices, most decisions require the identification and weighting of multiple attributes. Examples include choosing between two meals that differ in their taste, nutrition, and costs, or choosing between slot machines that differ in the likelihood and size of the potential rewards. Given their prevalence, understanding the algorithms that we use to make multi-attribute choices, and how they are affected by contextual variables, is a central question in psychology, economics, and neuroscience (Busemeyer & Johnson, 2004; Mas-Colell, Whinston, & Green, 1995; Glimcher & Fehr, 2014; Fehr & Rangel, 2011).

While much evidence suggests we differentially weight attributes in decision-making, the additional impact of attending to particular attributes of a choice is currently unclear. For instance, suppose a restaurant menu contains a daily special of steak with a side of green beans. Furthermore, assume a decision-maker enjoys steak, but dislikes green beans. Although many theories suggest an optimal decision maker may properly weight these variables (according to some subjective weights) as they decide whether or not to order the meal, one open question asks whether we can quantitatively determine how differentially attending to the steak and green beans can impact the probability of ordering this meal. If most of the decision-maker's attention shifts to the steak, would they be more likely to order the meal and if so, by how much does the probability of ordering increase with additional attention? To address these questions we propose a computational model, which we call the multi-attribute attentional drift diffusion model (maDDM). Our model details the choice process by modeling how attention to attributes, at the level of random eye fixations, alters individual choices.

Our model builds on two main literatures. First, previous work has shown sequential integrator models of decision-making, such as the Drift-Diffusion model (Ratcliff, 1978; Ratcliff et al., 2003; Ratcliff & Smith, 2004), leaky-accumulator model (Usher & McClelland, 2001), and Decision Field Theory (DFT) (Busemeyer et al., 1993; Busemeyer & Townsend, 1992; Roe et al., 2001; Busemeyer & Diederich, 2002; Diederich, 1997) provide accurate quantitative accounts of how choice probabilities and reaction times vary with properties of the choice options. These models typically assume choices are made using a relative value signal that is dynamically computed by integrating an instantaneous noisy measure of the desirability of options, and that a choice is made when the accumulated relative value signal becomes sufficiently strong in favor of one of two options. Furthermore, a growing body of evidence from neuroscience has found that the implementation of certain sequential integrator models is biologically plausible (Britten et al., 1992; Gold and Shadlen, 2007; Heekeren et al., 2008; Rangel and Clithero, 2013; Hare et al., 2011).

Second, previous work suggests that this integration process exhibits an attentional bias: attended options and attributes are weighted more heavily. Examples include the attentional Drift-Diffusion Model (aDDM) (Krajbich et al., 2010; Krajbich & Rangel, 2011; Krajbich et al., 2012; Fehr & Rangel, 2011) and DFT (Busemeyer et al., 1993; Busemeyer &

Townsend, 1992; Roe et al., 2001; Busemeyer & Diederich, 2002; Diederich, 1997). While both the aDDM and DFT attempt to explain the underlying choice process, the two literatures have progressed slightly differently. For instance, the aDDM has investigated how the value computation process changes as a function of random eye fixations to different options (Krajbich et al., 2010) while DFT has modeled attention by appealing to a dynamic attention function that weights information over time (Roe et al., 2001; Diederich, 1997), but does not rely on fixation data to quantify how attention is distributed throughout a decision. On the other hand, DFT has certain benefits, as it has developed a deep understanding of both multi-attribute choice and choices over multiple options while the aDDM has only been extended to choice over a small number of options and has not modeled multi-attribute choice. Our work attempts to take a step in unifying these two literatures by extending the underlying principles of the aDDM to a case of simple multi-attribute choice.

To test our model, we conduct a laboratory experiment where participants make decisions over whether to consume multi-attribute bundles of food while we record their eye movements between two attributes. Critically, our results provide a quantitative estimate for how attending to particular attributes of a choice can alter the weights those attended features receive when computing value: we find subjects overweight the currently attended attribute and underweight the unattended attribute.

The theory and experiment here allows us to address two main questions about the computational process involved in making multi-attribute choices. First, are similar computations performed when making multi-attribute decisions as when making multi-alternative decisions? This question evaluates how well principles of the aDDM can be extended to the maDDM, an important test in understanding this new choice domain. Second, by how much do subjects overweight attended features and underweight unattended features, if at all? This second question is important because although there is previous evidence that fixations bias choice, there is little work estimating the quantitative impact of how attention to attributes alters the likelihood of choosing an option.

Results

As illustrated in Figure 1, every trial hungry subjects are shown a bundle of two foods, one appetitive and one aversive, and have to decide whether or not they want to eat both a minimum of three bites of both of the foods (see Methods for details). Subjects make such choices for 200 different bundles, and at the end of the experiment the decision that they made in a randomly selected trial is implemented. Subjects also completed two rating tasks in which they provided liking ratings for each food individually, and for each of the choice bundles.

We focus on this simple choice task because it is the simplest possible setting in which multi-attribute choice can be studied. Here, the choice objects are the bundles. Each bundle consisted of one appetitive food and one aversive food. Hence, the bundles contain two attributes: an appetitive and an aversive stimulus. The liking ratings provide a measure of the attribute values for each bundle.

Importantly, in order to study the role of relative attention to the attributes, we monitored fixations during the choice task with eye tracking. Although it is well known that fixations and attention can be dissociated (Posner et al., 1977; Egly et al., 1994), for the purposes of this experiment fixations appear to provide a reasonable measure of attention at any instant during the choice process.

Model. The experiment was designed to allow us to test the ability of the maDDM to account for the relationship between fixations, choices, and reaction times in a simple multi-attribute choice setting. To see why, we begin by describing the model and its properties.

The model assumes that the value of a bundle, denoted by V_B, is given by a linear combination of the values of the appetitive food (V_P) and the aversive foods (V_N); i.e.,

$$V_B = \beta_0 + \beta_P V_P + \beta_N V_N.$$

Note that the rating tasks provide a measure of each of these values, which allows us to test the general validity of this assumption. To do so, for every subject we estimated a linear regression of the bundle ratings on the ratings of the appetitive and the aversive foods, and found that the data approximates the assumption reasonably well (mean $\beta_0 = 0.59$, SD = 2.17; mean $\beta_P = 0.61$, SD = 0.40; mean $\beta_N = 0.99$, SD = 0.74; mean $R^2 = 0.32$, SD = 0.18). We further tested for an interaction effect by including an additional regressor, V_P*V_N, in the above linear combination. After estimating this regression for every subject, we found the mean coefficient on the interaction was 0.14 (SD = 0.31); however, the mean difference in R^2 before and after adding this term was only 0.007 (minimum = 0.00, maximum = 0.03). Furthermore, running a mixed-effects linear regression with the interaction term found the coefficient to be insignificant with an estimate of 0.077 ($p >$ 0.05). Hence, we find little evidence to suggest an interaction effect occurred and analyze the data without this term.

As depicted in Figure 2, the maDDM assumes that decisions are made by integrating a relative decision value (RDV) signal over time until enough evidence is accumulated in favor of one of the two options: choice = "yes" or choice = "no." In particular, the subjects choose "yes" if the barrier crossed is at $B = +1$, and choose "no" if the barrier crossed is at $B = -1$. The model also predicts reaction times, since choice time equals the time the barrier is crossed.

A key property of the model is that both the bundle properties and attention are allowed to influence the evolution of the RDV signal, and thus how choices are made. In particular, the model assumes that there is a fixation bias, so that attending to a particular attribute increases the weight that attribute is assigned in the integration process. Specifically, when looking at the appetitive attribute the RDV evolves according to

$$RDV_t = RDV_{t-1} + d(\beta_0 + \delta\beta_P V_P + \theta\beta_N V_N) + \varepsilon_t$$

and when looking at the aversive attribute, it evolves according to

$$RDV_t = RDV_{t-1} + d(\beta_0 + \theta\beta_P V_P + \delta\beta_N V_N) + \varepsilon_t.$$

Here, RDV_t indicates the value of the RDV signal at time t, d is a constant that controls the speed of integration (in units ms^{-1}), δ is a parameter that can take values greater than or equal to 1 and reflects the fixation bias towards the currently fixated attribute, θ is a constant between 0 and 1 that reflects a fixation bias to the currently non-fixated attribute, and ε_t is i.i.d. white Gaussian noise with variance σ^2 that reflects the stochastic nature of the process.

Importantly, the model assumes that the fixation process between the two attributes is independent of each individual attribute's value, or of the location of the positive and negative items. In particular, the first fixation is assumed to go to the left attribute with a constant probability p. Fixations then alternate between the two foods until a barrier is crossed. At the beginning of each fixation, a maximum fixation length is drawn from a distribution that depends on the type of attribute (appetitive or aversive), and whether the fixation is a first fixation or a later one. The fixation is then allowed to run its course unless a barrier is crossed before it terminates, which ends the choice process.

Several properties of the model are worth highlighting.

First, the model includes as a special case a multi-attribute DDM without attentional bias, which arises when $\delta = \theta = 1$. This model is almost identical to the standard DDM that has been widely used in the previous literature to study binary choices in a large number of domains, including simple choices (Milosavljevic et al., 2011; Ratcliff, 1978; Ratcliff & Smith, 2004).

Second, the model exhibits a fixation bias when $\delta > 1$ or $\theta < 1$. In that case, an exogenous relative increase in attention to the appetitive food biases choices towards consuming the bundle, while the opposite is true for an exogenous decrease. Figure 2 provides an intuition for why this is the case. Consider a case in which $\beta_P = \beta_N = 1$, $V_P = -V_N$, and $\beta_0 = 0$. Here, in the absence of an attentional bias (i.e., when $\delta = \theta = 1$), the slope of the RDV is always zero, and the choice is determined simply by the realization of noise. In contrast, when $\theta < 1 < \delta$ the slope of the RDV signal is positive when looking at the appetitive attribute, and negative otherwise. As a result, the probability of choosing "yes" depends on the relative allocation of attention.

Third, the model has four free parameters (d, δ, θ, σ) that can be fitted using the choice, fixation, and reaction time data. The model has a fifth parameter, given by the height of the barrier, which is assumed to be fixed at ± 1. This is without loss of generality because multiplying the barriers, slope, and noise by a fixed constant has no effect on the model's predictions.

Fourth, and somewhat more technical, the model allows for an asymmetric bias on the attended and unattended attributes (as opposed to requiring that $\delta - 1 = 1 - \theta$). This asymmetry can be identified from the data as long as β_0 is non-zero.

Model Fitting. We fitted the model using MLE on the pooled group data (see Methods for details). Importantly, parameters were fitted using only the even-numbered trials, and the odd trials were used to test the predictions of the best fitting model out of sample. The best fitting parameters where $d = 0.0013$, $\sigma = 0.02$, $\delta = 1.05$, and $\theta = 0.90$ (log-likelihood = -18,016).

In order to test for the presence of a fixation bias, we also fitted a model with the restriction $\delta = \theta = 1$. The best restricted model also had $d = 0.0013$ and $\sigma = 0.02$ (log-likelihood = -18053). A likelihood ratio test statistic provided support in favor of the unrestricted model with a small but significant fixation bias model ($p < 0.001$).

In order to test for the asymmetry of the fixation bias, we also fitted a model with the restriction $\delta - 1 = 1 - \theta$. The best restricted model had parameters $d = 0.0012$, $\sigma = 0.0225$, $\delta = 1.025$, and $\theta = 0.975$ (log-likelihood value = -18053). A likelihood ratio test provided support in favor of the asymmetric fixation bias ($p < 0.001$).

Together, these results are consistent with the existence of small and asymmetric fixation bias. It is worth emphasizing, however, that the size of these fixation biases are significantly smaller than those that have been found in previous studies of simple choice (Krajbich et al., 2010; Krajbich & Rangel, 2011; Krajbich et al., 2012).

Basic Psychometrics. Figure 3 compares the basic psychometric properties of the data with the predictions generated by the best fitting model. In this figure, and the following ones, black denotes data and red denotes out of sample predictions. Both data and predictions are shown only for odd trials, to insure that the comparison is out-of-sample. Predictions were made by simulation the best-fitting model 4,000 times for each bundle liking rating, and sampling fixation lengths from the empirical distribution of observed fixations, conditioning only on whether a fixation was to an appetitive or aversive attribute, and whether the fixation was a first a later one. See Methods for more details.

Figure 3A depicts the psychometric choice curve. It shows that the probability of choosing yes is a logistic function of the bundle value which matches well the predictions of the best fitting model (goodness of fit test: $p = 0.48$).

Figure 3B depicts the reaction time curve, which exhibits the typical inverted-U pattern of reaction time when plotted against the liking rating of the bundle, so that more difficult choices take longer. The data also matches the predictions of the best fitting model (goodness of fit test: $p = 0.27$).

Finally, Figure 3C depicts the fixation curve, which shows that the number of fixations that it takes to make a choice increases with the difficulty of the choice. Although both the data and predictions exhibit the same general pattern, the model over predicts the impact of choice difficult on the number of fixations, as well as the average number of fixations (data: coefficient on difficulty = -.24, mean = 2.79; model: coefficient on difficulty = -0.44, mean = 3.15). Part of the mismatch between actual and predicted fixations has to do with technical limitations of the fitting and prediction procedure, which are discussed in more detail in the final Discussion section.

Properties of the Fixation Process. As described above, the basic maDDM assumes that fixations are independent of the value of the foods. Here we test if the pattern of observed fixations is consistent with this assumption.

As shown in Table S1, subjects exhibited a left-first bias: they looked at the left attribute before the right 64% of the time ($p < 0.01$). However, the location of first fixation was not significantly different for positive and negative foods (Table S1, $p > 0.05$).

As shown in Figure 4, fixations to aversive foods were about 57 ms longer on average than fixations to appetitive items, both for first, middle, and last fixations ($p < 0.01$ in all cases). This is consistent with the assumptions of the model listed above, since fixations to different attribute types might follow a different process (e.g., they might have a different processing latency). The key assumption of the model, to which we turn attention to next, is that fixation duration is not dependent on the value of the fixated or unfixated items (controlling for their attribute type). We tested this assumption by examining how the duration of different types of fixations, either first or middle fixations, was affected by the value of the attended and unattended attribute. We ignore final fixations in this analysis since their duration is endogenous to the choice process.

The duration of the first fixation was not significantly related to the value of the attended item (mixed-effects regression of first fixation length on an indicator for whether the item is appetitive, the weighted value of the item and the weighted value of the unattended item: beta of indicator = -37.92, *t*-statistic: -1.84; beta for attended value = -1.90, *t*-statistic = -0.29; beta for unattended value = -5.38, *t*-statistic = 1.26). Furthermore, the duration of the first fixation was not related to the value of the bundle (mixed-effects regression of first fixation length on an indicator for whether the item is appetitive and the value of the bundle: beta of indicator = -64.44, *t*-statistic = -5.31; beta for value of bundle = -0.35, *t*-statistic = -0.11) or the difficulty of the choice (mixed-effects regression of first fixation length on an indicator for whether the item is appetitive and the absolute value of the bundle: beta of indicator = -64.44, *t*-statistic = -5.26; beta for absolute value of bundle = -1.49, *t*-statistic = -0.32). Clearly, the duration of the first fixation was not dependent on value.

For middle fixations, we found no significant relationship with the attended value and a significant but quantitatively small effect of the unattended value (analogous mixed-effects regression: beta of indicator = -24.85, *t*-statistic: -0.80; beta for attended value = 6.51, *t*-statistic = 0.92; beta for unattended value = 17.13, *t*-statistic = 2.32). Importantly, the size of this effect is relatively small as a change in value of the unattended attribute of 2.5, the maximum possible change, will only alter middle fixations by 43 ms, on average. Furthermore, the duration of the middle fixation was not related to the value of the bundle (analogous mixed-effects regression: beta for indicator = -62.04, *t*-statistic = -3.82; beta for value of bundle = -5.55, *t*-statistic = -0.91), but was slightly related to the difficulty of the choice (analogous mixed-effects regression: beta for indicator = -61.50, t-statistic = -3.78; beta for absolute value of bundle = -21.61, t-statistic = -2.25). Again, even though we found a significant effect here, the effect was quite small in size as a move from the most difficult choice, with bundle value 0, to the simplest, with an absolute bundle value of 3, only corresponds to a change in fixation duration of -65 ms.

Together, the results in this section suggest that the properties of the observed fixation process are largely consistent with the assumption that fixations are independent of changes in the value of specific attributes (e.g., example, changing a mildly appetitive item for a highly appetitive one). In addition, as shown in Figure S1, there is considerable variation in the duration of fixations across trials. This, together with the tests supporting the existence of a small but significant attentional bias, suggests that fluctuations in attention might be responsible for some of the observed differences in choices.

Model Predictions. The maDDM makes additional predictions about the pattern of the fixations, and their relationship to choices, which we test here.

First, the model predicts that final fixation durations should be shorter than middle fixation, since final fixations are terminated prematurely when a barrier is crossed. As shown in Figure 4, this also holds in our data (mean last = 376 ms; mean middle = 550 ms; $p < 0.01$).

Interestingly, we also found that first fixations were shorter than middle fixations (mean first= 309 ms; mean middle = 550 ms; $p < 0.01$). Note that although the model made no exante prediction about the relationship between these two types of fixations, this pattern was incorporated in our prediction exercise, since fixation durations were conditioned on whether a fixation was first or later.

Second, the model predicts a strong relationship between the fixation-averaged value at the start of the final fixation and the duration of the final fixation, conditional on the choice made. Specifically, the model predicts that conditional on a "no" choice, the duration of the final fixation should increase with the variable

$$F_P(\beta_0 + \delta\beta_P V_P + \theta\beta_N V_N) + (1 - F_P)(\beta_0 + \theta\beta_P V_P + \delta\beta_N V_N),$$

where F_P denotes the fraction of the trial spent attending to the appetitive item (as of the beginning of the last fixation). Essentially, this variable measures the average slope of the RDV signal during the initial phase of the choice, given the realization of fixations up to that point. The intuition for this relationship illustrates the key forces at work in the maDDM, and are best seen using a hypothetical case in which $\beta_P=\beta_N=1$, $V_P = -V_N$, and $\beta_0=0$. In this case, when $\theta < 1< \delta$ the slope of the RDV is positive when fixating to the appetitive attribute, and negative otherwise. As a result, the larger F_P, the farther the RDV signal is likely to be from the "choose no" barrier at the beginning of the last fixation. Thus, the process needs to cover more distance during the last fixation to reach the "no" barrier, leading to a longer last fixation.

To test this prediction, we estimated a mixed-effects regression of final fixation duration on the final fixation average value variable, for trials in which the subjects choose "no." Consistent with the prediction, we found a significant effect between the two (slope = 33.87, *t*-statistic = 4.22). The model makes an analogous prediction for trials in which the subject chooses yes, albeit with the opposite sign, which was also present on the data (slope = -24.17, *t*-statistic = -2.43).

These results demonstrate that several key predictions regarding the pattern of fixations and their relationship to choices hold in the data.

Choice Biases. When $\theta < 1 < \delta$, the maDDM predicts a number of attentional driven biases, that we test in this section. This provides an additional set of model tests, which we also carried out.

First, the model predicts that, controlling for bundle values, the probability of choosing "yes" increases with the relative attention to the appetitive attribute interacted with its subjective value, and decreases with additional time spent attending to the aversive attribute interacted with its subjective value. To test for this effect in our data, we run a mixed-effects logistic regression of choice on bundle rating, the weighted value of the appetitive food ($=\beta_P V_P$) interacted with its relative fixation time, and the weighted value of the aversive food ($=\beta_N V_N$) interacted with its relative fixation time. Consistent with the predictions, we found a negligible bias (constant = -0.20, p = .43), a significant increase in the probability of choosing "yes" with bundle value (slope = 1.14, $p < 0.01$) and the value of the appetitive food weighted by its share of relative attention (slope = 0.37, $p < 0.05$), and a significant decrease in the probability of saying yes with the value of the aversive item weighted by its relative attention (slope= -0.56, $p < 0.01$). Similar effects were found in the simulated data (constant = .11, $p < 0.01$; slope bundle rating = 0.62, $p < 0.01$, slope weighted value appetitive item interacted with relative fixation time = 0.07, $p < 0.01$; slope weighted value aversive item interacted with relative fixation time = -0.52, $p < 0.01$).

It is worth emphasizing that, despite the small effect of the fixation bias coefficients (θ and δ), the resulting choice biases need not be small. To see why, consider an example in which suppose the bundle liking rating is 0, the appetitive attribute has a rating of 2, the aversive attribute has a rating of -2, and the value weights take the mean value over all subjects. Our estimate predicts that when an individual spends 10% of the trial attending to the appetitive attribute, there is only a 23.9% chance of agreeing to consume the bundle; however, if that individual instead spends 90% of the trial attending to the appetitive attribute, the probability of responding "yes" increases to 52.4%. In contrast, in a model without a fixation bias ($\theta = \delta = 1$), this change in the fixation pattern has no effect on the choices.

Second, the model predicts that, controlling for bundle value, the probability of choosing "yes" depends on the relative amount of time spent attending to the appetitive item. As shown in Figure 5A, this was true in both the data and the model predictions, and the size of the two effects was remarkable similar. This test was computed as follows. For every trial, we computed a corrected choice measure by subtracting the observed choice (yes = 1, no = 0) from the average frequency with which the bundle was chosen for all trials with that bundle rating. We then estimated a linear regression of the corrected choice probabilities on the relative time advantage to the appetitive item and found the predicted effect in both the simulated data (slope = 0.04; $p < 0.01$) and the data (mixed-effects regression slope = 0.04, t-stat = 3.56).

Third, the model predicts that the longer the first fixation is to the appetitive item, interacted with its weighted value, the more likely the subject is to choose "yes." To see why, note that longer first fixations move the RDV signal towards the "yes" barrier, which all else equal biases choices towards "yes." To test this, we estimated a mixed-effects logistic regression of choice on the duration of the first fixation, conditioning on a first fixation to the positive attribute, and found the predicted effect in both the data and the simulations (simulated model: slope = 0.0009, $p < 0.01$; data: slope = 0.0017, $p < 0.01$). We also found the analogous effect for first fixations to the aversive item (simulated model: slope = -0.0010, $p < 0.01$; data: slope = -0.0017, $p < 0.01$).

In a related result, any initial biases in the first attended attribute should translate into choice biases. Figure 5B shows that this is the case: a linear regression of the probability that a subject looks first at the appetitive food first on the subjects' average probability of choosing shows a significant positive relationship (slope = 0.09; $p < 0.04$).

Fourth, the model predicts a relationship between the identity of the last fixation and choice. In particular, conditional on value of the bundle, it predicts that the probability of choosing "yes" is larger when the last fixation is to the chosen item. The intuition for this prediction stems from the fact that the slope of the RDV signal is more likely to be positive, and thus climbing towards the "yes" barrier, during fixations to the appetitive item.

We estimated a logistic regression of choice on a constant, the bundle liking rating, an indicator variable for when the last fixation is to the appetitive item, and the interaction of the bundle liking rating with the indicator variable, in the simulations and the actual data. In the simulations, we found a significant effect of the identity of the final fixation (beta = 0.71, $p < 0.01$), but not of the interaction term (beta = -0.00, $p = 0.87$). A similar pattern was found in the data (last fixation bias to positive indicator: beta = 0.21; $p = 0.06$; interaction term: beta = 0.03; $p = 0.60$). As shown in Figure 5C, although these biases are small, they follow a quantitatively similar pattern in the model predictions and data.

Discussion

The results described here suggest the maDDM is accurately able to describe the choice process in a simple multi-attribute environment. Specifically, the model quantitatively describes the relationship between choices, response time, and the correlation of these variables with attentional deployment, as measured by fixations. The data suggest that individuals increase the weight of an attended attribute by 5% and decrease the weight of an unattended attribute by 10%. Consistent with this estimation, a number of attention-based choice biases found in the data support the model's predictions.

Although our estimated model parameters suggest a fixation bias alters the decision process, we were surprised by its relatively small effect size. Notably, previous findings have estimated that only 30% of an unattended option's value is accounted for during the choice process (Krajbich & Rangel, 2011). There are several possibilities that can explain this finding. First, as our model is one of the first applications to model the fixation bias in attributes rather than options, it's possible the fixation bias over attributes is simply smaller than over options. Second, our task forces subjects to accept or reject an outcome. If they were instead making choices over two or more bundles, a larger bias may be prevalent. Understanding how the size of the bias changes with the task is an important step for future work.

A natural question about our model concerns the direction of causality between fixations and choice. Namely, while our model assumes that fixations bias the value estimation process, another possibility is that the value of the attributes directly affects the fixation process. Although we have shown some evidence suggesting such an explanation is not responsible for driving our results, the best way to address this question is through follow-up work that provides a causal test of this theory. Several related papers address this issue in the contexts of simple food choice and risky decision-making (Armel et al. (2008) and Kim et al. (2012)). Furthermore, Lim et al. (2011) find that the vmPFC encodes attention-modulated relative value signals, suggesting neurobiological evidence that fixations alter the choice process. While this literature speculates that there is a causal role from fixations to choice, we cannot rule out the possibility that causality also works in the other direction.

One limitation of our results is that we applied the maDDM to a scenario with a relatively small number of attributes. Many real-world decisions are likely more complex than our bundle choices. We chose this simple design for several reasons. First, we wanted to understand how the maDDM could be applied to the simplest possible multi-attribute choice scenario, an important feature when understanding whether and how these models may break down as the task becomes increasingly complex. Second, as the number of attributes and options on a screen grows, the model must be able to account for the fixation process between features. This is a complex task that may be a rich area for future work, but we currently lack a systematic understanding of how to model more than two fixations.

It is worth noting that in our results we find much evidence that has a qualitative flavor of loss aversion. Specifically, the duration of a fixation to an aversive attribute is on average longer than the duration to an appetitive one, individuals weight aversive attributes more heavily than appetitive attributes in choices, and aversive attributes are attended to for a longer period of time throughout the choice process. This differential attribute weighting is consistent with the literature on loss aversion (Kahneman & Tversky, 1984) while the differences in fixations to attributes are consistent with previous process tracking studies of loss aversion (Willemsen et al., 2011). This difference in fixation duration to the appetitive and aversive attributes can be further explained given that the amount of time one attends to a feature appears to influence its weight (Willemsen et al., 2011; Schkade & Johnson, 1989; Fiske, 1980; Wedell & Senter, 1997).

Our study builds on and contributes to several literatures. First, this paper differentiates itself from previous work in the cognitive modeling of choice in a number of ways. While the existing aDDM literature has focused on understanding the role of fixations in multi-alternative choice (Krajbich et al., 2010; Krajbich & Rangel, 2011), it has remained silent as to how fixations affect attribute based value estimation. Our model takes a step towards understanding this process and extends the aDDM to this choice environment. Furthermore, while DFT has previously explored decision-making in multi-attribute choice, it has not incorporated eye fixations in understanding how attention to attributes alters an attribute's perceived weights (Busemeyer & Townsend, 1992; Roe et al., 2001; Busemeyer & Diederich, 2002; Diederich, 1997). Our contribution is not to suggest that one model should be preferred to the other, but that we can accurately describe the choice process by utilizing tools from both of these literatures.

Second, our work adds to a large literature that uses process-tracing methods to understand the decision process (Russo & Rosen, 1975; Russo & Dosher, 1983; Willemsen et al., 2011; Johnson et al., 2008; Camerer & Johnson, 2004; Glockner & Herbold, 2011; Hortsmann et al., 2009; Towal et al., 2013). While much of this work makes use of eye tracking, others test process-based models by tracking mouse movements on a computer screen. Furthermore, our work adds to literature that explores how evidence accumulation may change throughout the choice process (Ratcliff, 1980; Ratcliff & McKoon, 1982). Our work complements these literatures by again demonstrating the usefulness and power of this approach in testing and employing choice process models.

Finally, our results have a number of implications for public policy and marketing that could aid policy designers in altering decisions. Specifically, it helps quantity the effect for how ‘nudging’ attention towards a particular attribute can impact behavior by quantifying the size of how that variable may be overweighted at the expense of others. For instance, by how much can changing a product’s design to highlight particular attributes increase sales? This is consistent with work that has found pushing attention towards the health attributes of a food, and away from its taste attributes, increases an individual’s ability to make healthy dietary choices (Hare et al., 2009). Investigating the size of this effect in other choice domains is a critical open question.

SI Methods

Subjects. Forty-six subjects recruited from the Caltech community participated in the experiment (63% male; mean age = 26.2). All subjects had normal or corrected-to-normal vision with the use of either contact lenses or glasses. Participants were paid a $5 show-up fee and received an additional $25 upon successful completion of the experiment. The study was approved by Caltech's IRB.

Task. Subjects were asked to fast for four hours prior to the start of the experiment. Compliance was verified through self-report upon subject arrival, and was required for participation.

The experiment consisted of three tasks. Subjects were informed of this at the outset, but the tasks were only described to them just before they took place.

In Rating Task 1, subjects performed a liking-rating task over individual snack food items shown in a computer screen, one at a time. The image size was 300 x 300, with a screen resolution of 1280 x 1024. Subjects were asked to enter a liking rating for each food using an integer scale (-3 to 3, framing: "How much would you enjoy that particular food at the end of today's experiment?"). The ratings were entered using the bottom row of the keyboard. Subjects could take as long as desired to enter each rating. Thirty unique foods were rated and each of them was shown twice to each subject in random order.

The foods were selected based on previous studies (Plassmann et al., 2007, Plassmann et al., 2010) and contained eighteen foods that were consistently rated as appetitive by previous subjects, and eighteen foods that were consistently rated as aversive. See Table S2 for details.

For every subject, we averaged the two ratings provided for each food in order to create subject-specific food classes. Snacks with a positive average rating were labeled as "appetitive," snacks with a negative average rating were classified as "aversive," and foods with a zero average rating were omitted from the remaining tasks. On average, subjects had 17 appetitive foods and 11 aversive foods.

In Rating Task 2, subjects saw bundles of two foods on the screen and had to provide liking ratings over the bundles, using the same integer scale (-3 to 3). In particular, subjects were asked to rate "How much would you enjoy taking at least three bites from both of the foods shown on the screen?" Every bundle contained one appetitive food and one aversive food from the previous round, and subjects could take as long as needed to enter their ratings. As shown in Figure 1, one of the items was shown in the left and the other on the right, with their location randomized every trial. The number of trials in this task varied across subjects, subjects were asked to rate every potential bundle made of one appetitive and one aversive food.

Finally, subjects participated in a Choice Task. Every trial they were shown one of the bundles from Task 2, and they had to decide (yes/no) if they wanted to take at least three bites from each of the foods at the end of the experiment. Choices were indicated using a keyboard button press using the subject's dominant hand with the index and middle fingers. Subjects could take as long as desired to make each choice. The choice task consisted of 200 trials, selected at random from the set of all possible bundles described above. Subjects were instructed that at the end of the experiment they would need to remain in the lab for an additional twenty minutes. During this time, one of the two hundred trials was randomly selected and their choice in that trial was implemented. This procedure encouraged subjects to give incentive compatible responses.

Importantly, we recorded eye movements throughout the task at 500 Hz using a desktop mounted SR research Eyelink 1000 eye tracker. The eye tracker was calibrated immediately after reading the instructions for this choice task.

Model fitting. As described in Results, the maDDM has four free parameters: the constant determining the speed of integration d, the positive fixation bias δ, the negative fixation bias θ, and the noise parameter σ. We fitted these parameters at the group level by pooling the data from all subjects into a single data set. The parameters were fitted to maximize the maximum likelihood of the observed choices and reaction times. Importantly, the model was fitted using only even trials, and the odd trials were reserved for out-of-sample comparisons, as described in the results section. We fitted the model at the group level to both choice and reaction time data for all 46 subjects by pooling all even numbered trials into a single data set. The model requires a large amount of data to estimate the parameters accurately, and fitting them at the individual level would result in highly noisy estimates.

The MLE procedure was conducted as follows.

First, we simulated the model 4000 times for each combination of the model parameters in the grid described below, and for each of the seven possible bundle ratings (ranging from -3 to 3). The simulations were carried out using 1 ms time steps.

In each simulation, individual liking ratings for both the appetitive and aversive foods were drawn from the empirical distribution of liking ratings conditional on the rating of the bundle. Furthermore, once a pair of liking ratings was drawn, we chose subject-estimated regression weights (β_0, β_P, and β_N) associated with the randomly selected simulated liking ratings. For instance, if the drawn liking rating for the appetitive item and the aversive item was drawn and belonged to subject *i*, then subject *i*'s regression weights were used throughout the simulated trial.

In each simulation, we randomly sampled fixations lengths from the empirical distribution for the group, conditional on whether it is a first or middle fixation in the trial, and whether the fixation was to either the positive or negative food. We also assumed that subjects looked first to the left item 68% of the time, which is the frequency observed in the data. As in the observed data, we assume fixations alternate between the foods. Although our model assumes that fixations between the two items on the screen occur instantaneously, in practice there are observed saccade length transitions in each trial. To take this into account, in every simulated trial we randomly sampled from the empirical distribution of transition times, and add that sampled transition time to the simulated total fixation time. To clarify, in every trial we have a response time (the length of time from stimulus onset to response) and total time spent fixating to both items. We say the transition time is the difference between these two variables, and the sum of the transition time and total simulated fixation time represents the simulated reaction time in a trial.

Second, we used the simulations to compute the likelihood of each observation, for each vector of parameters, as follows. Reaction time was discretized into bins of 100 ms, from 0 to 7500 ms, with an additional bin representing a trial that took longer than 7500 ms. Choice data is automatically discretized into yes/no bins. We then used the simulation results, conditional on the bundle rating, to compute the frequency with which responses followed into each time-choice bin.

Third, we used the data from the previous step to compute the log-likelihood of the data for each vector of parameters, and carried out a grid search to identify the vector of parameters with the largest maximum likelihood.

To reduce computational costs, this maximization was done in two steps. In step one we first did a coarse search over the following parameter space:

$$d \text{ in } \{0.0001, 0.0005, 0.001, 0.0015, 0.002, 0.0025\}$$

$$\sigma \text{ in } \{0.005, 0.01, 0.02, 0.025\}$$

$$\delta \text{ in } \{1, 1.05, 1.1, 1.15\}$$

$$\theta \text{ in } \{0.85, 0.9, .95, 1\}$$

which identified (d = 0.0015, σ = 0.02, δ = 1.05, θ = 0.95) as the parameters that maximized the log-likelihood. In step two we did a finer search around this vector using the grid:

$$d \text{ in } \{0.001, 0.0011, 0.0012, 0.0013, 0.0014, 0.0015, 0.0016, 0.0017\}$$

$$\sigma \text{ in } \{0.015, 0.0175, 0.02, 0.0225\}$$

$$\delta \text{ in } \{1, 1.025, 1.05, 1.075, 1.1, 1.125\ 1.15\}$$

$$\theta \text{ in } \{0.85, 0.875, 0.9, 0.925, 0.95, 0.975, 1\}.$$

References

Armel, C., Beaumel, A., & Rangel, A. (2008). Biasing Simple Choices by Manipulating Relative Visual Attention. *Judgment and Decision Making*, 3, 5, 396- 403.

Britten, K.H, Shadlen, M.N., Newsome, W.T., & Movshon, J.A. (1992). The Analysis of Visual Motion: A Comparison of Neuronal and Psychophysical Performance. *Journal of Neuroscience*, 12, 4745-4765.

Busemeyer, J.R. & Diederich, A. (2002). Survey of Decision Field Theory. Mathematical Social Science, 43, 345-370.

Busemeyer, J.R. & Johnson, J.G. in *Handbook of Judgment and Decision Making* (eds Koehler, D. & Narvey, N.) 133-154 Blackwell Publishing Co., New York, 2004.

Busemeyer, J.R., & Townsend, J. (1993). Decision Field Theory: A DynamicCognitive Approach to Decision Making in an Uncertain Environment. *Psychological Review*, 100, 432-459.

Camerer, C., and Johnson, E. J. (2004). "Thinking about attention in games: backward and forward induction," in The Psychology of Economic Decisions, eds J. Carillo and I. Brocas (New York: Oxford University Press), 111–129.

Egly, R., Driver, J., & Rafal, R.D. (1994). Shifting Visual Attention Between Objects and Locations: Evidence from Normal and Parietal Lesion Subjects. *Journal of Experimental Psychology: General*, 123, 161-177.

Fehr, E. & Rangel, A. (2011). Neuroeconomic Foundations of Economic Choice: Recent Advances. *Journal of Economic Perspectives*, 25(4).

Fiske, S.T. (1980). Attention and Weight in Person Perception: The Impact of Negative and Extreme Behavior. *Journal of Personality and Social Psychology*, 38, 6, 889-906.

Glimcher, P. & Fehr, E. *Neuroeconomics: Decision Making and the Brain*. Elsevier, London, 2014.

Glockner, A., and Herbold, A.-K. (2011). An eye-tracking study on information processing in risky decisions: evidence for compensatory strategies based on automatic processes. *J. Behav. Decis. Mak*. 24, 71–98.

Gold, J.I. & Shadlen, M.N. (2007). The Neural Basis of Decision Making. *Annual Review of Neuroscience*, 30, 535-574.

Hare, T., Camerer, C., & Rangel, A. Self-control in decision-making involves modulation of the vmPFC valuation system. *Science,* 2009, 324:646-648

Hare, T.A., Schultz, W., Camerer, C., O'Doherty, J., Rangel, A. (2011). Transformation of Stimulus Va *Proceedings of the National Academy of Sciences of the United States of America*, 108, 18120-18125.

Heekeren, H.R., Marrett, S., & Ungerleider, L.G. (2008). A General Mechanism for Perceptual Decision-Making in the Human Brain. *Nature*, 431, 859-862.

Horstmann, N., Ahlgrimm, A., and Glockner, A. (2009). How distinct are intuition and deliberation? An eye-tracking analysis of instruction induced decision modes. *Judgm. Decis. Mak*. 4, 335–354.

Johnson, E.J., Schulte-Mecklenbeck, M., & Willemsen, M.C. (2008). Process Models Deserve Process Data: Comment on Brandstätter, Gigerenzer, and Hertwig (2006). *Psychological Review*, 115, 1, 263-273.

Kahneman, D. & Tversky, A. (1984). Choices, Values and Frames. *American Psychologist*, 39, 341-350.

Kim, B.E., Seligman, D., & Kable, J.W. (2012). Preference Reversals in Decision Making Under Risk are Accompanied by Changes in Attention to Different Attributes. *Frontiers in Neuroscience*, 6, 109.

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual Fixations and Comparison of Value in Simple Choice. *Nature Neuroscience*, 13, 1292-1298.

Krajbich, I., Lu, D., Camerer, C., & Rangel, A. (2012). The Attentional Drift Diffusion Model Extends to Simple Purchasing Decisions. *Frontiers in Cognitive Science*, 3, 193.

Krajbich, I., & Rangel, A. (2011). A Multi-Alternative Drift Diffusion Model Predicts the Relationship Between Visual Fixations and Choice in Value-Based Decisions. *Proceedings of the National Academy of Sciences*, 108, 13853-13857.

Lim, S.L., O'Doherty, J., & 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, 13214-13223.

Mas-Colell, A., Whinston, M., & Green, J. *Microeconomic Theory* Cambridge University Press, Cambridge, 1995.

Plassmann, H., O'Doherty, J., & Rangel, A. Appetitive and aversive goal values are encoded in the medial orbitofrontal cortex at the time of decision making. *Journal of Neuroscience,* 2010, 30:10799-10808.

Plassmann, H., O'Doherty, J., & Rangel, A. Medial OFC encodes willingness-to-pay in simple economic transactions. *Journal of Neuroscience*, 2007, 27(37):9984-88.

Posner, M.I., Nissen, M.J., & Ogden, W.C. (1977). Attended and Unattended Processing Modes: The Role of Set for Spatial Location. Modes of Perceiving and Processing Information Hillsdale, N.J.: Erlbaum, 137-157.

Rangel, A. & Clithero, J.A. (2013). The Computation of Stimulus Values in Simple Choice. *Neuroeconomics: Decision-Making and the Brain, 2nd ed. (edited by Paul Glimcher and Ernst Fehr)*, 125-147.

Ratcliff, R. (1978). A Theory of Memory Retrieval. Psychological Review, 85, 59-108.

Ratcliff, R. (1980). A Note on Modeling Accumulation of Information When the Rate of Accumulation Changes over Time. *Journal of Mathematical Psychology*, 21, 2, 178-184.

Ratcliff, R., Cherian, A., & Segraves, M. (2003). A Comparison of Macaque Behavior and Superior Colliculus Neuronal Activity to Predictions From Models of Two-Choice Decisions. Jounral of Neurophysiology, 90, 1392-1407.

Ratcliff, R., & McKoon, G. (1982) Speed and Accuracy in the Processing of False Statements About Semantic Information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 8, 1, 16-36.

Ratcliff, R. & Smith, P. (2004). A Comparison of Sequential Sampling Models For Two-Choice Reaction Time. Psychological Review, 111, 333-367.

Roe, R.M., Busemeyer, J.R., & Townsend, J.T. (2001). Multialternative Decision Field Theory: A Dynamic Connectionist Model of Decision Making. *Psychological Review*, 108, 2, 370-392.

Russo, J.E. & Dosher, B.A. (1983). Strategies for Multiattribute Binary Choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9, 4, 676-696.

Russo, J.E. & Rosen, L.D. (1975). An Eye Fixation Analysis of Multialternative Choice. *Memory & Cognition*, 3, 267-276.

Schkade, D. & Johnson, E.J. (1989). Cognitive Processes in Preference Reversals. *Organizational Behavior and Human Decision Processes*, 44, 203-231.

Towal, R., Mormann, M., & Koch, C. (2013). Simultaneous Modeling of Visual Saliency and Value Computation Improves Predictions of Economic Choice. *Proceedings of the National Academy of Sciences*, 110, 3858-3867.

Usher, M. & McClelland, J.L. (2001). On the time course of perceptual choice: The leaky competing accumulator model. Psychological Review, 108, 550-592.

Wedell, D.H. & Senter, S.M. (1997). Looking and Weighting in Judgment and Choice. *Organizational Behavior and Human Decision Processes*, 70, 41-64.

Willemsen M.C., Böckenholt, U., & Johnson, E.J. (2011). Choice by Value Encoding and Value Construction: Processes of Loss Aversion. *Journal of Experimental Psychology: General*, 140, 3, 303-324.

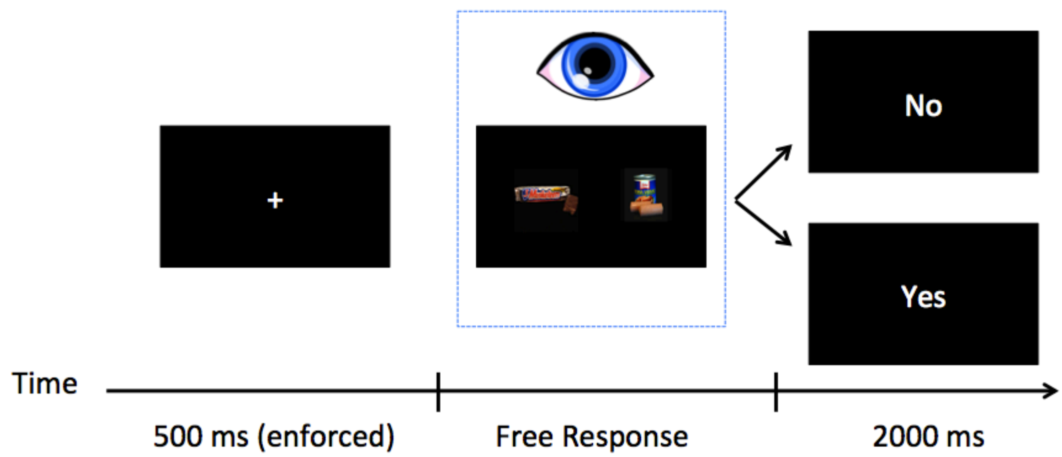
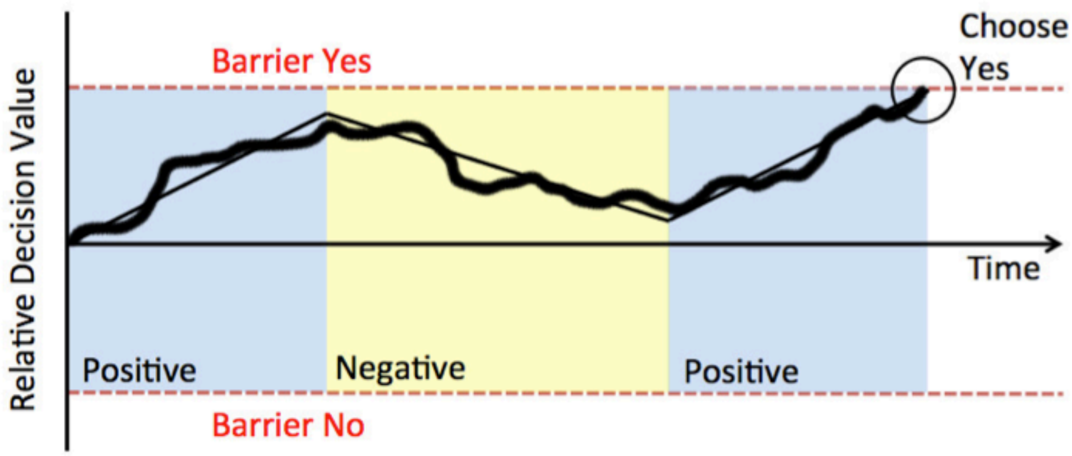


Figure 1

Experimental design. Subjects participated in a task where they made decisions over whether they were willing to take at least three bites from each of the two foods on the screen at the end of the experiment. One food was previously rated as appetitive while the other was previously rated as aversive. Participants' eye movements were recorded as they made these choices. The timing of each screen is depicted at the bottom of the figure. Each subject saw a fixation cross for 500ms, then had as long as they liked to enter a rating or make a choice. They then saw feedback for 2000 ms and moved to the next trial



Look at Positive Attribute: $RDV_t = RDV_{t-1} + d(\delta\beta_P R_P + \theta\beta_V R_V + \beta_0) + \varepsilon_t$

Look at Negative Attribute: $RDV_t = RDV_{t-1} + d(\theta\beta_P R_P + \delta\beta_V R_V + \beta_0) + \varepsilon_t$

Figure 2

Depiction of the maDDM. A relative decision value (RDV) signal evolves over time. Its slope is biased towards the fixated item, but random noise is added to the RDV at every millisecond. When the RDV hits a barrier, a decision is made. The shaded vertical regions represent what item is currently fixated. In this example, three fixations are made (positive, negative, positive) and the individual chosen "yes." The equations below the image describe how the RDV is integrated over time. The blue δ parameter describes an increase in weight that the attended item receives, while the red θ parameter describes a decrease in weight that the unattended item receives.

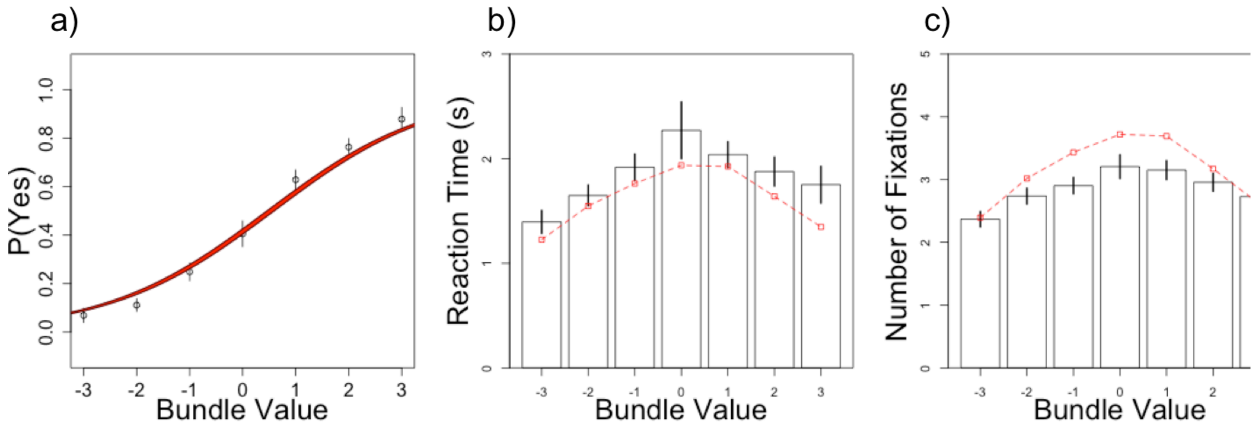


Figure 3

Basic psychometrics. (a) Psychometric choice curve as a function of the bundle liking rating. (b) Reaction times as a function of the bundle liking rating. (c) The number of fixations in a trial as a function of the bundle liking rating. The red lines indicate the model's predictions. The thickness of the red line in (a) represents the model's standard error, while bars in (b) and (c) represent the standard error. Subject data is shown for the odd-numbered trials, with standard errors clustered by subject.

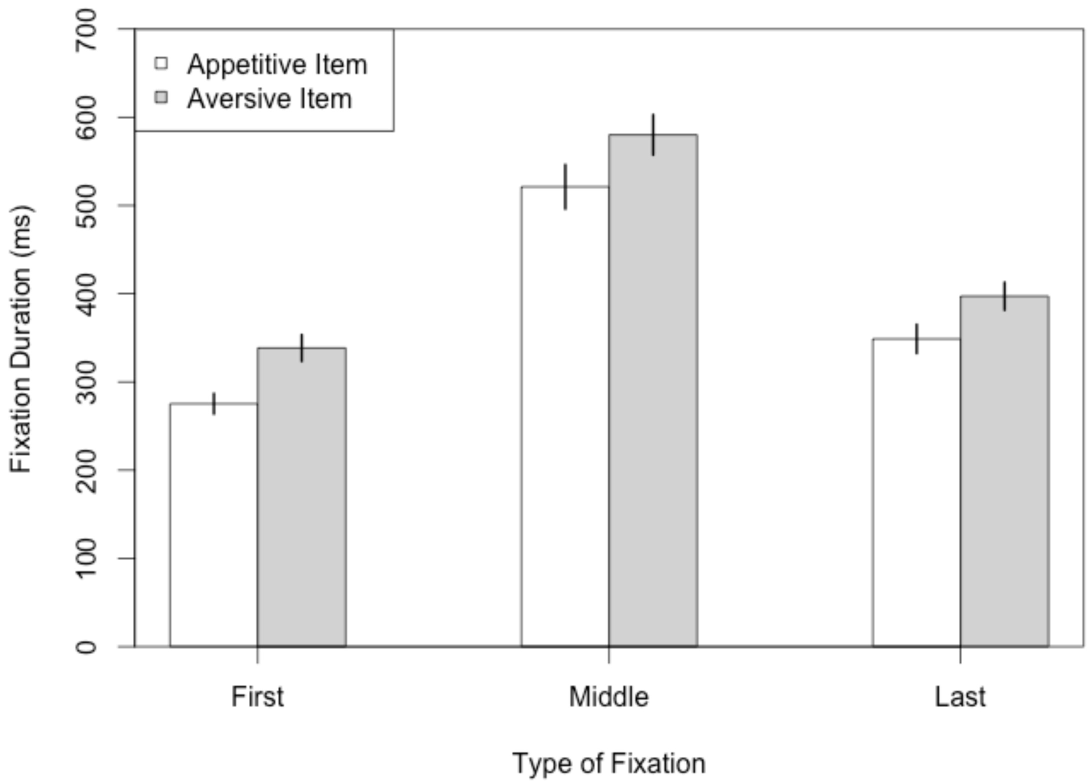


Figure 4

Fixation Durations. The mean fixation duration for first, middle, and last fixations split dependent on whether the fixation was to the appetitive or aversive item. Standard errors are clustered by subject.

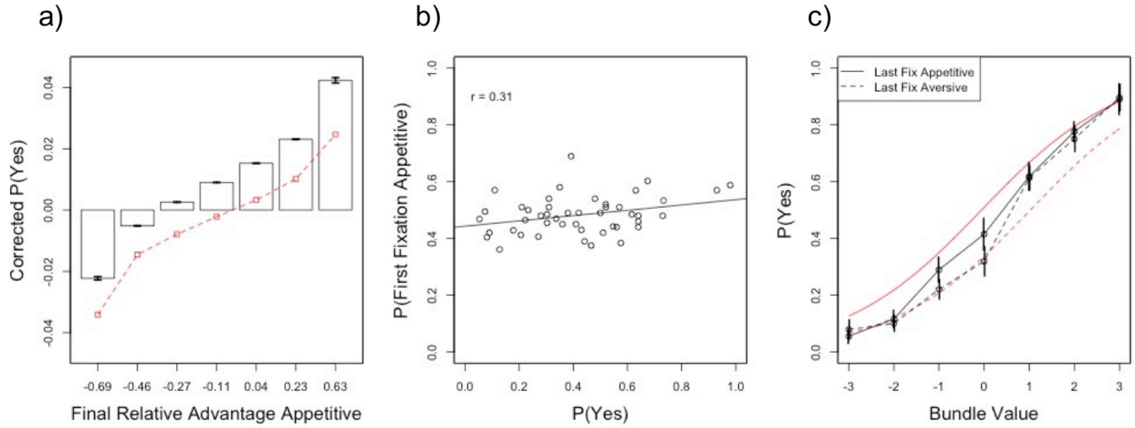


Figure 5

Choice Biases. (a) Corrected probability of agreeing to eat the bundle as a function of the relative time advantage looking at the appetitive item. Bins depict the odd-numbered trials, and the red dotted line is the model simulation. To compute the bins, the data was split into seven equal bins and the median of each is reported on the horizontal axis. (b) Probability of looking at the appetitive item as a function of agreeing to consume the bundle. Each circle represents a different subject. (c) The probability of agreeing to consume the bundle as a function of the bundle value and whether the last fixation was to the appetitive or aversive attribute. Red solid line indicates the model's prediction when the last fixation was to the appetitive item and the red dotted line indicates the model's prediction when the last fixation was to the aversive item. Black dots with a solid connecting line indicates the odd-numbered data when last fixation was to the appetitive item and black dots with a dotted connecting line indicates the data when the last fixation was to the aversive item. Standard errors clustered by subject.

Table S1

Percent of First Fixations to Each Item		
A. Spatial		
	Left	Right
Percentage	64.4 (22.1)	35.6 (22.1)
B. Attribute of Interest		
	Appetitive	Aversive
Percentage	48.2 (6.6)	51.8 (6.6)

Depicts the mean percent of first fixations to each item by spatial features (Panel A) and attribute features (Panel B). Standard deviation given below in parentheses.

Table S2

Food	Rating	Food	Rating
3 Musketeers Candy Bar	1.51 (1.52)	Nature Valley Granola Bar	1.26 (1.35)
Flamin' Hot Cheetos	0.72 (2.12)	Milano Cookies	2.00 (1.40)
Almond Joy Candy Bar	0.40 (2.27)	Peanut M&M's	1.88 (1.03)
MilkyWay Candy Bar	1.76 (1.22)	Canned Garbanzo Beans	-1.07 (1.78)
KitKat Candy Bar	2.21 (1.11)	Pureed Green Beans	-1.92 (1.35)
Crunch Bar	1.94 (1.19)	Canned White Meat Chicken	-0.80 (1.93)
Reese's Peanut Butter Cups	1.73 (1.70)	Soy Sauce	-1.71 (1.48)
Oreos	1.76 (1.25)	Canned Albacore Tuna	0.10 (2.16)
Tootsie Rolls	0.28 (1.73)	Canned Artichoke Hearts	-1.39 (1.86)
Doritos Cool Ranch Chips	1.61 (1.43)	Pureed Carrots	-1.85 (1.33)
Chocolate Pudding	0.65 (2.19)	Canned Vienna Sausage	-1.30 (1.92)
Twix Candy Bar	1.99 (1.26)	Canned Sweet Peas	-0.94 (1.84)
Snickers Candy Bar	1.69 (1.35)	Canned Deviled Ham Spread	-1.51 (1.67)
Butterfinger Candy	0.97 (1.70)	Canned Sardines	-1.33 (1.90)
Ghirardelli Milk Chocolate	2.24 (1.02)	Canned Spinach	-1.62 (1.62)

Food stimuli used in the experiment. Each stimulus contains the mean rating across subjects, with standard deviation below in parentheses.

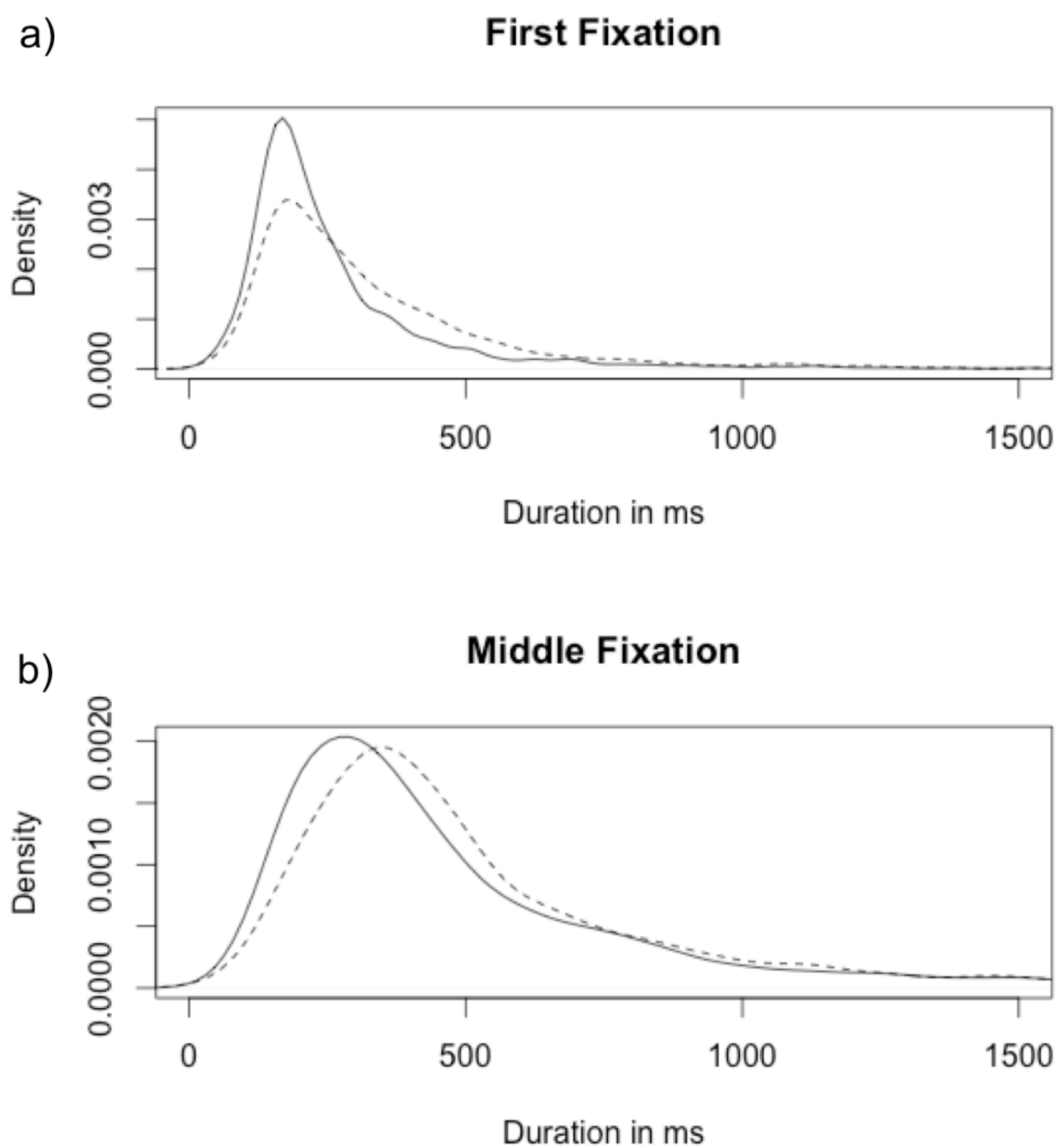


Figure S1

Distribution of fixation durations for (a) first fixations and (b) middle fixations. Solid line refers to when the last fixation was to the appetitive item, and the dotted line refers to when the last fixation was to the aversive item.

Chapter 2

INTERTEMPORAL CHOICES ARE CAUSALLY INFLUENCED BY RELATIVE ATTENTION

Many important choices involve tradeoffs between immediate and delayed rewards, and sound decision-making often requires delaying gratification. Examples include dietary choice (e.g., fruit or chocolate cake?), health behaviors (e.g., go to the gym or watch TV at home?), and saving (e.g., buy a new car or save for retirement?). Previous work has shown that we systematically struggle to delay gratification, that our ability to do so varies across decision contexts, and that there are sizable individual differences (1). Unfortunately, the mechanisms underlying the contextual and individual differences in discounting are not well understood.

Here we propose that variation in how attention is deployed during the choice process can account for a sizable portion of the behavioral differences across both contexts and individuals. In particular, we hypothesized that contextual variables that shift relative attention towards attributes favoring the patient option can induce a causal and sizable decrease on discount rates, and thus an increase in the ability to postpone gratification. Additionally, we hypothesized that individual differences in attentional patterns can explain a sizable portion of the individual differences in discount rates.

This hypothesis builds on several literatures. First, previous work has shown sequential integrator models of decision-making, such as the Drift-Diffusion Model (2-4), the leaky-accumulator model (5), and Decision Field Theory (DFT) (6-8), provide accurate quantitative accounts of how choice probabilities and reaction times vary with the properties of the choice options, including intertemporal monetary choice tasks (9-10). These models all assume choices are made using a relative value signal that is dynamically computed by integrating instantaneous noisy measures of the desirability of the attributes associated with the two options, and that a choice is made when the accumulated relative value signal becomes sufficiently strong in favor of one of the two options.

Second, previous studies suggest that this integration process exhibits an attentional bias: options and attributes are weighted more heavily while they are attended. Examples include the attentional Drift-Diffusion Model (aDDM) (11-13) and DFT (6-8). These models assume the allocation of attention to the attributes and options are independent of the state of the relative value signal or of the values of the attributes. Thus, any variable that shifts attention towards attributes that favor choosing the delayed option (e.g., a contextual manipulation that affects the saliency of the delayed reward, or a systematic individual trait) can increase the likelihood of making a patient decision. Evidence consistent this assumption comes from work showing that fixations are affected by relative visual saliency, independent of value (14-15), and that choices can be manipulated by exogenously changing attention (16-17).

The hypothesis that intertemporal choices are affected by relative attention is important for several reasons. First, most of the existing literature attributes individual differences to variation in fixed discount rates, which are preference parameters that appear hard to change. In contrast, we propose a significant fraction of the within and between subject variation is due to differences in attention. Second, an important goal of this literature is the design of policy interventions that increase people's ability to delay gratification. Knowing whether changes in attention can have a sizable impact on the likelihood of making a patient choice is useful because 'nudging' or 're-training' attention might be easier than changing more hard-wired preference parameters. Third, although there is a large literature demonstrating that context variables matter (18-21), we currently lack a systematic understanding for how these variables affect choices. Our proposal suggests a critical element of the problem is to understand how these contextual variables affect attention during the choice process.

We report the results of three laboratory experiments designed to test our hypothesis. The first experiment combines eye tracking with a common intertemporal choice task in which subjects choose between smaller-sooner and larger-later monetary rewards. The experiment tests the hypothesis that variation in relative attention, to different options and attributes, can explain a sizable fraction of the within and between subject variation in discount rates. The second and third experiments manipulate attention exogenously in order to test the causal impact of attention on intertemporal choice.

Experiment 1: Correlational Test

Experiment 1 (Figure 1 and SI Methods) combines a standard intertemporal monetary choice paradigm with eye tracking to address the following three questions. First, is the correlation between attention (as measured by fixations) and choices consistent with the hypothesis that relative attention affects intertemporal choice? Second, are cross-trial fluctuations in relative attention associated with sizable changes in the likelihood of making a patient choice? Third, what fraction of the individual differences in discount rates can be explained by attentional differences?

Choices. Subjects chose the patient option 54.4% of the time (SD = 27.0%). The average value of the estimated hyperbolic discount parameter, k, was 0.009 (SD = 0.013), which is comparable with previous results using this task (1, 22-23). Both indices provide a measure of the extent to which subjects discount future rewards. As shown in Figure S1, the two measures are significantly correlated (β = -0.20, $p < 0.001$). For robustness, below we report results using both choice measures. Furthermore, there was significant variation in discount rates between subjects (percent patient: max = 91.6, min = 1.4, k: max = 0.055, min = 0.0002), which we exploit in several of the analyses below.

Reaction times. Subjects took an average of 2.3 seconds to make a decision (SD = 0.9 seconds). Reaction times were not significantly different in trials in which a patient or an impatient choice was made (p = 0.48, paired t-test). We estimated a linear mixed-effects regression of reaction time on trial difficulty, which in this and all future similar analyses included random effects for all of the independent variables (including the constant), unless otherwise specified. Consistent with the predictions of sequential integration models of choice, we found that reaction times increased with difficulty (β = -0.057, $p < 0.01$).

Average fixation patterns. We used fixations to each of the four regions-of-interest (ROIs) as a measure of attention to the individual attributes. Although it is well known that it is possible to attend to information without fixating on it, as demonstrated in the literature on covert attention (24-25), it appears unlikely that there is a large dissociation between the two in this task.

Subjects made an average of 7.0 fixations per-trial (SD = 3.8), which implies that, on average, they fixated to the ROIs displaying the different attributes more than once. The number of fixations was not significantly different in trials in which a patient or an impatient choice was made ($p = 0.738$, paired t-test). We estimated a mixed-effects linear regression of the number of fixations on trial difficulty and found a significant relationship ($\beta = -0.177$, $p < 0.01$).

Table S1 summarizes the relative fixation time patterns across the four ROIs. Subjects spent more time looking at the upper fields in which the amounts were depicted than at the lower fields, which contained the delays (paired t-test, $p < 0.001$). Furthermore, there was not a significant difference between attending to the left and right locations (paired t-test, $p = 0.126$); however, subjects spent more time attending to the delayed option than the immediate option (paired t-test, $p < 0.001$). Finally, there was significant trial-to-trial variation for all ROIs, a fact that we exploit in the within subject analyses below.

Table S2 summarizes the pattern of first fixation locations across the four ROIs. The majority of first fixations were to the top-left location ($p < 0.01$, t-test of upper left ROI versus sum of all other 3 ROIs). It was rare for subjects to first fixate to one of the delays, as those always appeared in the bottom ROIs ($p < 0.01$, t-test of first looking at amounts versus first looking at delays). The vast majority of first fixations were to the amounts ($p < 0.001$, paired t-test). Given that amounts were always shown in the top location, we do not know if this is the result of a spatial bias or a top-down property of the attentional process in this class of tasks. There was a small bias towards first looking at the delayed option ($p < 0.001$, paired t-test), which suggests that subjects could identify the location of the delayed option through peripheral vision and use this information to influence the location of their first fixation before the information in the four ROIs had been sampled.

Within subject analysis. As shown in Table S1, relative fixations varied significantly from trial to trial. Here we investigate if this variation is associated with changes in the likelihood of making a patient choice, and if these changes are consistent with the predictions described above. Additionally, we quantify the size of these effects.

To do this, for each ROI we estimated a random-effects logistic regression of patient choices on the fraction of time spent fixating on a particular ROI. This was done in a separate regression for each ROI because the relative attention measures are not independent across ROIs. Thus, each regression should be interpreted as estimating the effect of shifting relative attention to the target ROI, while reducing relative attention on the other ROIs proportional to their average frequencies. Furthermore, for each ROI we computed the mean predicted impact of changing relative attention in that ROI from the 10^{th} to the 90^{th} percentile of the observed distribution of relative attention. To calculate this, we sampled random trial numbers 1,000 times for each subject, with replacement. We then extracted the fraction of time that was spent attending to the ROI in that trial, and used the individually estimated regression weights to calculate the probability that the subject chose the patient option in that trial. Next, we calculated the 10^{th} and 90^{th} percentile of the distribution of the probability of a patient choice, and report the mean change over subjects, which we denote as the mean effect size.

Table 1 summarizes these results. Consistent with the predictions, shifting attention to the immediate delay decreased the likelihood of choosing the delayed option, and shifting attention to the delayed amount had the opposite effect. Importantly, the predicted effect sizes were substantial. For example, a shift from the 10^{th} to the 90^{th} percentile of the observed distribution of relative attention to the immediate delay decreases the probability of making a patient choice by 35%, as given by the mean effect size.

The table also illustrates that shifting attention to the immediate amount decreases the likelihood of making a patient choice, and that shifting attention to the delayed date increases that likelihood. These results suggest that an option-based attentional bias, where attention to any attribute of an option gives the decision maker evidence in favor of choosing that option, might be dominant in this task. This stands in contrast to an attribute-based attentional bias, where attention to a class of attributes gives the decision maker evidence in favor of choosing a particular option.

These results suggest that cross-trial variation in attention can explain close to 30% of subjects' ability to choose the delayed outcome. The largest effects were found in the attentional differences to the immediate delay and the delayed amount, respectively.

Between subjects analysis. Next we tested if between-subject differences in relative attention could explain a substantial fraction of the individual variation in discount rates. To do this, we first computed the average relative attention that each subject paid to each of the four ROIs. For each ROI, we estimated a linear regression of our subject-level measure of patience (either fraction of patient choices or log(k)) on the subject-level measure of relative attention paid to the ROI.

Table 2 reports the results. We found a positive correlation between the propensity to shift attention to the delayed amount and the likelihood of making a patient choice, and a negative correlation between the propensity to shift attention to the immediate delay and the likelihood of making a patient choice. In contrast, shifting attention to the immediate amount or later delay was not correlated with individual differences in discounting. Since the subject-level measures of attending to the different ROIs are correlated (min = -0.84, max = 0.17), we also estimated a linear model in which the four attentional measures were included. As shown in the right-hand column of Table 2, only the propensity to shift attention to the immediate delay was significantly correlated with individual differences in discount rates.

Two aspects of these results are worth highlighting. First, the results suggest that between 20% and 35% of the individual differences in discount rates can be explained using differences in the average relative propensity to look at different ROIs. Second, the results across subjects are consistent with those found in the previous section; in both cases the attentional variable that has the largest impact in explaining variation in patience is the propensity to shift attention to the immediate delay, followed by the propensity to look at the delayed amount.

Changes in fixations across trials. An important assumption of the models motivating the hypothesis tested here is that the allocation of attention is largely exogenous to the state of the choice process, and to the value of the attributes. More concretely, these models assume that fixations can be modulated by visual features of the stimuli (e.g., text versus numbers or spatial location), but not by the state of the relative value signal that drives the choice, or by the absolute or relative value of the attributes. This assumption is important because it implies that fluctuations in attention have a causal impact in the choice process, instead of being driven by it.

Testing these assumptions about the orthogonality of attention directly is difficult as it requires having a measure of the relative value signal's state before a choice is made, which is quite challenging to obtain, and because attention terminates at the end of the choice process, which can produce spurious correlations between attentional measures and attribute values. One way to address these issues is to carry out external manipulations of attention, as we report in the final two experiments below. However, since these manipulations also have limitations, we finished the analysis of Experiment 1 by carrying out an indirect test of the orthogonality of attention.

The logic of the test is as follows. From the choice data, we know that the attribute values are correlated with the likelihood of choosing the patient option. Under the maintained hypothesis that choices are compatible with a sequential integration model, this implies that, on average, each of these variables should also be correlated with the state of the integrator across trials. Thus, if fixations were driven mostly by the state of the relative value signal, one would expect a strong association between the relative fixations and the attribute parameters across trials. Table S3 reports the results of this test. Importantly, in all cases the magnitude of the effects were quite small in size, contrary to what would be expected if attention were guided mostly by the state of the relative value signal, or by the relative value of the attributes. To quantify this, Table S3 also reports an estimate of the effect size, or the maximum percentage change in attention to each ROI that can be induced, which are also found to be relatively small.

Together, these analyses provide support for the hypothesis that some of the variation in attention is exogenous to the attribute values and to the comparison process used to make the choice. However, we emphasize that these tests cannot rule out the possibility that some of the attentional variation is endogenous, which highlights the importance of the last two experiments, where attention is manipulated exogenously.

Experiments 2 & 3: Causal Tests

The results of Experiment 1 demonstrate that variation in relative attention can account for a sizable fraction of the differences in discount rates, both within and across individuals. However, despite the evidence suggesting that a sizable fraction of the variation of attention is exogenous, the tests are purely correlational, and cannot rule out the possibility that the direction of causality runs in the opposite direction. The last two experiments were designed to address this issue.

Causal Test I. In Experiment 2 (see Figure 2 and SI Methods) we manipulated the relative attention paid to the different attributes and tested if this increased the likelihood that subjects made a patient choice. Here, subjects faced trials where they were forced to attend to the different attributes for a particular amount of time before they were allowed to enter their response. Relative attention was manipulated within subjects as each participant spent either more time fixating towards the amounts or more time fixating towards the delays depending on the trial. We manipulated exposure to amounts versus delays because the results of Experiment 1 suggest that this may affect patience, given the asymmetric impact of fixating on the immediate delay versus the other attributes.

We carried out two separate analyses on the choice data. First, we compared choices in trials where amounts were displayed for longer than delays to choices in trials where delays were displayed for longer than amounts. In particular, we computed the number of patient choices and estimated the k discounting parameter separately for each subject and group of trials, and compared them using two-sided paired t-tests. We found that subjects made 42.9 patient choices when amounts were displayed for longer than delays, and 41.3 patient choices in the other case. This amounts to a small, but significant, 4% increase in the number of patient choices as the result of the exposure manipulation ($p < 0.03$). A similar result was found when we examined at the estimated discount rates: the estimated $\log(k)$ is -5.66 when amounts are shown for longer than delays and -5.32 otherwise. Again, the effect is small, but significant ($p < 0.04$). The direction of these effects was consistent with the predictions made based on the findings of Experiment 1.

Second, we carried out a similar analysis comparing the trials in which amounts were shown first to those trials in which delays appeared first. We did not find a significant order effect using either the percentage of patient choices ($p = 0.40$) or the estimated $\log(k)$ ($p = 0.14$) indicating that total fixation time may play a larger role than order of fixations.

These results are consistent with the hypothesis that changes in relative attention to different attributes have a causal impact on the ability to make patient choices and suggests that some of the within and cross individual differences in discount rates, identified in Experiment 1, are due to attentional variation.

Causal Test II. While Experiment 2 found evidence that exogenously manipulating attention alters discounting, the effect was fairly small. One possible interpretation of the small effect size is that the relative deployment of attention is more endogenous than the previous literature and the previous discussion suggest. However, another possible interpretation is that carrying out meaningful manipulations of attention is hard, and that in the previous experiment attention might not have varied as much across conditions as intended, perhaps because subjects made up their minds before the exposure time terminated.

Experiment 3 was designed to address this issue (Figure 3 and SI Methods). Here, subjects were free to fixate between the immediate and delayed outcomes, but once a specified accumulation time had been reached for an option, where one option was always chosen to have a larger accumulation time than the other, it was removed from the computer screen. We still manipulated attention to certain features on the screen, but we sought to create an environment where fixation and behavioral patterns more closely matched those in Experiment 1. Furthermore, this design allows us to test for the presence of an option-based bias, rather than an attribute-based bias, as the within subject results from Experiment 1 suggest the option-based bias may be dominant.

We first verified that the experimental manipulation successfully biased fixations towards the target option. Subjects spent 1.14 seconds (SD = 0.08) fixating to the target option and 0.30 seconds (SD = 0.00) fixating to the non-target option, indicating the manipulation was successful in altering relative attention, as measured by fixations.

Next we compared behavior in this experiment to that in Experiment 1. Here, subjects made 2.7 fixations (SD = 0.6) between the options, meaning that they, on average, viewed each option more than once. Furthermore, subjects first looked left on 59% (SD = 29%) of the trials, but there was no relationship between whether they looked at the immediate or delayed option first ($p = 0.24$). The average time spent on the trial before entering a response was 2.9 seconds (SD = 0.7), and only 10.5% of the trials terminated at the five-second mark. These results indicate that certain behavioral outcomes, particularly response times, were similar to those in Experiment 1.

We then compared choices across the two attentional conditions and found a causal effect: when the immediate option was the target, subjects made 41.5 patient decisions, but when the delayed option was the target, subjects made 46.6 patient decisions. This difference of 5.1 patient decisions is significant (*t*-test of difference, $p < 0.01$). The effect size appears quite large and suggests that the number of patient choices increases by 12.3 percent as attention is exogenously shifted away from the impatient and towards the patient option. Note that, as in experiment 1, the range of how patiently subjects chose in this experiment greatly varied (pooled across both conditions: minimum = 2.0%, maximum = 97.5%, mean = 44.0%, SD = 25.0%). Ultimately, this may lead to an underestimate of the effect size as subjects who have a preference to almost always choose the impatient, or patient, option may only have room to slightly alter behavior until they hit a barrier.

Finally, we sought to quantify how the estimated k parameter changed as a result of the experimental manipulation. Since the choice pairs were designed so that subjects were close to indifferent between them, we do not have enough power to detect significant changes in the estimated k parameters in the data set ($p = 0.175$). Thus, in order to quantify how the estimated change in patience can translate to changes in k, we performed the following analysis. Given the mean estimated k from Experiment 1 we simulated choices for 1000 experimental sessions, where each session consisted of the exact questions from Experiment 1. For each session, we drew a random number from the distribution of estimated effect sizes from Experiment 3, with replacement, and scaled that number to reflect the predicted effect size in the simulated data set. When the drawn effect size, d, was positive, we randomly chose d simulated patient choices to switch to impatient; when d was negative, we randomly chose d simulated impatient choices to switch to patient. Thus, we estimated how each drawn effect size would impact decisions if the effect size could be generalized to the choice set from Experiment 1. Next, we estimated k both before and after making these effect size changes and found the estimated k increased from 0.0089 to 0.0115 ($p < 0.01$), which predicts a sizeable shift towards impatient behavior. To quantify, while subjects would have, on average, made 93.0 patient decisions, the estimated effect of the manipulation decreases this to 81.8 patient choices. Repeating this simulation exercise an additional 150 times shows these results are not outliers in the distribution of test statistics (Fig. S2).

Discussion

We have described the results of three experiments designed to test whether exogenous fluctuations in the relative attention paid to different features during intertemporal choices can influence the ability to delay gratification. Consistent with this hypothesis, we found shifting attention towards the delayed amount attribute was associated with a sizable increase in the likelihood of making a patient choice, whereas shifting attention towards the immediate delay attribute had a sizable effect in the opposite direction. Furthermore, cross-trial variation in attention explained about 30% of subjects' ability to choose the delayed option and between 20% and 35% of individual differences in discount rates could be explained by individual differences in the average relative propensity to look at different ROIs.

Since multiple alternative mechanisms are likely to affect the ability to delay gratification during intertemporal choice (26-28), we were surprised to find that between 20% and 35% of observed individual differences in self-control ability could be explained by differences in the deployment of relative attention. By comparison, previous studies have found that personality traits such as impulsivity or I.Q. can only explain around 13% and 5%, respectively, of individual differences in self-control (29-31).

A critical question underlying our hypothesis is the direction of causality between the fluctuations of attention and the decision-making processes. The models that motivate our hypothesis, such as DFT and the aDDM, assume observed fluctuations in attention are not driven by the state of the relative value signal that drives choice, or by the properties of the attributes. If this is correct, the results from Experiment 1 would suggest that a sizable fraction of the within and cross subject variation in discount rates is driven by fluctuations in attention. However, it is also possible to write simple variations of these models in which the state of the relative value signal has some influence on the deployment of attention, so that attributes that are "consistent" with the currently favored option are more likely to be fixated. Understanding how much of this variation can be causally attributed to attention is critical to evaluate the implications of our work. The analysis of the fixation data from Experiment 1 suggests there might be an influence from the choice process to the deployment of attention, but the effect may be small. The results of Experiments 2 and 3, in which attention was manipulated exogenously, suggest that there is a causal effect from attention to intertemporal choice.

We view the results here as demonstrating that there is a causal and sizable impact of variation in relative attention on choice. However, we emphasize that, in our view, it is likely that some of the differences in attention across trials and subjects is endogenous, especially later on in the course of a trial. Designing clear measures of the exogenous and endogenous channels of influence, and how they evolve over the course of a decision, is a critical open question for future studies.

Given the sizable differences in exposure time in Experiment 2, we were surprised by the small effect size. There are several potential explanations for this finding. First, it is possible that the direction of influence from fixations to choices runs in both directions, and that a sizable fraction of the correlations identified in Experiment 1 is due to an influence of perceived value on attention. Second, the experimental manipulation might have had a small impact on the actual relative processing of the attributes, as subjects may have made their decisions before the exposure was completed, which would lead to a reduced impact on choices. Disentangling these hypotheses is challenging, as it requires measuring the 'latent' state of the choice process before a response is made, and to measure processing time by the decision-making circuitry without using fixations. Third, the task parameters were not optimized to generate the maximum possible effect; it is possible a different manipulation could have led to a more sizable effect.

Reasuringly, Experiment 3 shows that a larger causal effect takes place under an alternative attentional manipulation. Given the differences between Experiments 2 and 3, it is difficult to precisely pinpoint why Experiment 3 has a larger effect size than Experiment 2. For instance, it may be related to testing for an option versus attribute bias, differences in exposure time, or differences in the choice response process. It's possible that altering any of these parameters in the experiment can lead to even larger effect sizes than those we observe.

Our study builds upon and contributes to several literatures. First, it is related to the work on sequential-integration models by multiple groups, especially on the DDM (2-4), and to versions of these models illustrating that the dynamic computation and comparison of value is influenced by attention, like DFT (6-8), and the aDDM (11-13), which lead to systematic attentional biases. Our results also build on a pioneering set of papers that have used attentional measures to test algorithmic models of how preferences are constructed and compared at the time of decision (32-35) as well as others that investigate how evidence accumulation can change throughout the choice process (36-37). The association between attention and contextual effects has also been used to explain preference reversals (38, 7-8), and context effects in risky choice (20, 35). Our paper's contribution to this existing work is to show that these ideas extend to the domain of intertemporal choice, and that attentional variation could potentially be a critical variable in explaining differences in the ability to delay gratification across individuals and contexts. An important direction for future research in this area is to carry out experiments that allow for more quantitative model fitting and testing, including a formal comparison of the various types of attentional biases that have been proposed by different types of models. This could not be done with the current design, as it requires decorrelating the different attributes and a larger number of trials. However, previous applications using the DDM, the aDDM, and other sequential integrator models have demonstrated the potential of this approach.

Second, our results build on a large body of work which has shown that individuals seem to exhibit hyperbolic discounting in intertemporal choice, which can lead to difficulties in delaying gratification when one of the options entail an immediate reward (1, 22-23, 39). One critical finding in this literature is that discount rates are not constant, and instead decrease with distance to the present. The results of Experiment 1 show that shifts in attention towards the immediate delay attribute had an especially strong impact on discount rates. This suggests that some of the 'hyperbolicity' of the discount function might be attributable to attentional effects, and might be more sensitive to training and context than 'deeper' preference parameters.

Third, our results have implications for how to design interventions that could increase people's ability to postpone gratification, as well as important implications for marketing. In particular, they suggest that any contextual variable, or 'nudge,' that directs attention towards the long-term benefits of self-control, and away from the immediate rewards, might improve self-control. This is consistent with previous work that has found directing attention towards the health attributes of foods, and away from their taste, increases an individual's ability to make healthy dietary choices (40). Furthermore, previous studies have found that changes in the low level visual features of stimuli can affect the relative attention they receive, and through it the likelihood that they are selected (15). A systematic investigation of these possibilities in the domain of intertemporal choice is an important open question for future research.

Methods

Participants. Twenty subjects participated in Experiment 1, thirty-three participated in Experiment 2, and twenty-one participated in Experiment 3. All subjects were Caltech students or Pasadena/Los Angeles community members and had normal or corrected-to-normal vision with the use of contact lenses. Caltech's Institutional Review Board approved all studies.

Tasks. Subjects completed various intertemporal decision making tasks where they chose between receiving some amount of money later that day, i.e. "today," and some amount of money on a future date as their eye movements were recorded (SI Methods). For their participation, they received a $5 show-up fee plus payment for one intertemporal choice question, chosen at random and implemented through PayPal. All experiments contained at least 200 trials.

See the SI for a more detailed description of the Methods.

References

1. Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time Discounting and Time Preference: A Critical Review. *Journal of Economic Literature*, 40, 351-401.
2. Ratcliff, R. (1978). A Theory of Memory Retrieval. *Psychological Review*, 85, 59-108.
3. Ratcliff, R., Cherian, A., & Segraves, M. (2003). A Comparison of Macaque Behavior and Superior Colliculus Neuronal Activity to Predictions From Models of Two-Choice Decisions. *Jounral of Neurophysiology*, 90, 1392-1407.
4. Ratcliff, R. & Smith, P. (2004). A Comparison of Sequential Sampling Models For Two-Choice Reaction Time. *Psychological Review*, 111, 333-367.
5. Usher, M. & McClelland, J.L. (2001). On the time course of perceptual choice: The leaky competing accumulator model. *Psychological Review*, *108*, 550-592
6. Busemeyer, J.R. & Diederich, A. (2002). Survey of Decision Field Theory. *Mathematical Social Science*, 43, 345-370.
7. Busemeyer, J.R., & Townsend, J. (1993). Decision Field Theory: A Dynamic-Cognitive Approach to Decision Making in an Uncertain Environment. *Psychological Review*, 100, 432-459.
8. Roe, R.M., Busemeyer, J.R., & Townsend, J.T. (2001). Multialternative Decision Field Theory: A Dynamic Connectionist Model of Decision Making. *Psychological Review*, 108, 2, 370-392.
9. Rodriguez, C.A., Turner, B.M., & McClure, S.M. (2014). Intertemporal Choice as Discounted Value Accumulation. *PLoS ONE*, 9, e90138.
10. Dai, J. & Busemeyer, J.R. (2014). A Probabilistic, Dynamic, and Attribute-Wise Model of Intertemporal Choice. *Journal of Experimental Psychology: General*, Advance online publication. http://dx.doi.org/10.1037/a0035976.
11. Krajbich, I., Armel, C., & Rangel, A. (2010). Visual Fixations and Comparison of Value in Simple Choice. *Nature Neuroscience*, 13, 1292-1298.
12. Krajbich, I., Lu, D., Camerer, C., & Rangel, A. (2012). The Attentional Drift-Diffusion Model Extends to Simple Purchasing Decisions. *Frontiers in Cognitive Science*, 3, 193.
13. Krajbich, I., & Rangel, A. (2011). A Multi-Alternative Drift Diffusion Model Predicts the Relationship Between Visual Fixations and Choice in Value-Based Decisions. *Proceedings of the National Academy of Sciences*, 108, 13853-13857.
14. Towal, R., Mormann, M., & Koch, C. (2013). Simultaneous Modeling of Visual Saliency and Value Computation Improves Predictions of Economic Choice. *Proceedings of the National Academy of Sciences*, 110, 3858-3867.
15. Milosavljevic, M., Navalpakkam, V., Koch, C., & Rangel, A. (2012). Relative Visual Saliency Differences Induce Sizeable Bias in Consumer Choice. *Journal of Consumer Psychology*, 22, 67-74.
16. Armel, C., Beaumel, A., & Rangel, A. (2008). Biasing Simple Choices by Manipulating Relative Visual Attention. *Judgment and Decision Making*, 3, 5, 396-403.
17. Shimojo, S., Simion, S., Shimojo, E., & Sheier, C. (2003). Gaze Bias Both Reflects and Influences Preference. *Nature Neuroscience*, 6, 1317-1322.

18. Ariely, D., Loewenstein, G., & Prelec, D. (2003). Coherent Arbitrariness: Stable Demand Curves without Stable Preferences. *Quarterly Journal of Economics*, 118, 73-105.
19. Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, 185, 1124-1131.
20. Johnson, E., & Schkade, D. (1989). Bias in Utility and Assessments: Further Evidence and Explanations. *Management Science*, 35, 406-424.
21. Simonson, I. (1989). Choice Based on Reasons: The Case of Attraction and Compromise Effects. *Journal of Consumer Research*, 16, 158-174.
22. Kable, J., & Glimcher, P. (2007). The Neural Correlates of Subjective Value During Intertemporal Choice. *Nature Neuroscience*, 10, 1625-1633.
23. Chabris, C., Laibson, D., Morris, C., Schuldt, J., & Taubinsky, D. (2008). Individual Laboratory-Measured Discount Rates Predict Field Behavior. *Journal of Risk and Uncertainty*, 37, 237-269.
24. Posner, M.I., Nissen, M.J., & Ogden, W.C. (1977). Attended and Unattended Processing Modes: The Role of Set for Spatial Location. Modes of Perceiving and Processing Information Hillsdale, N.J.: Erlbaum, 137-157.
25. Egly, R., Driver, J., & Rafal, R.D. (1994). Shifting Visual Attention Between Objects and Locations: Evidence from Normal and Parietal Lesion Subjects. *Journal of Experimental Psychology: General*, 123, 161-177.
26. Beaver, J.D., Lawrence, A.D., van Ditzhuijzen, J., Davis, M.H., Woods, A., & Calder, A.J. (2006). Individual Differences in Reward Drive Predict Neural Responses to Images of Food. *Journal of Neuroscience*, 26, 5160-5166.
27. Loewenstein, G. (1996). Out of Control: Visceral Influences on Behavior. *Organizational Behavior and Human Decision Processes*, 65, 272-292.
28. Muraven, M. & Baumeister, R.F. (2000). Self-Regulation and Depletion of Limited Resources: Does Self-Control Resemble a Muscle? *Psychological Bulletin*, 126, 247-259.
29. De Wit, H., Flory, J.D., Acheson, A., McCloskey, M., & Manuck, S.B. (2007). IQ and Nonplanning Impulsivity are Independently Associated with Delay Discounting in Middle-Aged Adults. *Personality and Individual Differences*, 42, 111-121.
30. Mitchell, S.H. (1999). Measures of Impulsivity in Cigarette Smokers and Non-Smokers. *Psychopharmacology*, 146, 455-464.
31. Shamosh, N.A. & Gray, J.R. (2008). Delay Discounting and Intelligence: A Meta-Analysis. *Intelligence*, 36, 289-305.
32. Russo, J.E. & Rosen, L.D. (1975). An Eye Fixation Analysis of Multialternative Choice. *Memory & Cognition*, 3, 267-276.
33. Russo, J.E. & Dosher, B.A. (1983). Strategies for Multiattribute Binary Choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9, 4, 676-696.
34. Willemsen M.C., Böckenholt, U., & Johnson, E.J. (2011). Choice by Value Encoding and Value Construction: Processes of Loss Aversion. *Journal of Experimental Psychology: General*, 140, 3, 303-324.

35. Johnson, E.J., Schulte-Mecklenbeck, M., & Willemsen, M.C. (2008). Process Models Deserve Process Data: Comment on Brandstätter, Gigerenzer, and Hertwig (2006). *Psychological Review*, 115, 1, 263-273.
36. Ratcliff, R. (1980). A Note on Modeling Accumulation of Information When the Rate of Accumulation Changes over Time. *Journal of Mathematical Psychology*, 21, 2, 178-184.
37. Ratcliff, R., & McKoon, G. (1982) Speed and Accuracy in the Processing of False Statements About Semantic Information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 8, 1, 16-36.
38. Kim, B.E., Seligman, D., & Kable, J.W. (2012). Preference Reversals in Decision Making Under Risk are Accompanied by Changes in Attention to Different Attributes. *Frontiers in Neuroscience*, 6, 109.
39. Madden, G.J., Petry, N.M., Badger, G.J., & Bickel, W.K. (1997). Impulsive and Self Control Choices in Opioid-Dependent Patients and Non-Drug-Using Control Participants: Drug and Monetary Rewards. *Experimental and Clinical Psychopharmacology*, 5, 256-262.
40. Hare, T.A., Camerer, C.F., & Rangel, A. (2009). Self-Control in Decision-Making Involves Modulation of the vmPFC Valuation System. *Science*, 326, 596-599.

ROI	Coefficient Estimates			
Constant	1.15* (0.34)	-1.96* (0.32)	1.12* (0.33)	-0.65 (0.38)
Immediate Amount	-3.53* (0.80)	- -	- -	- -
Delayed Amount	- -	5.74* (0.74)	- -	- -
Immediate Delay	- -	- -	-10.80* (1.16)	- -
Delayed Delay	- -	- -	- -	3.40* (0.56)
Mean Effect Size	0.16	0.27	0.35	0.15

Table 1.

Each column reports the results of a logistic mixed-model regression where an indicator variable for making a patient choice was regressed on a constant and the percentage of time spent fixating on the ROI of interest. Mean effect sizes denote the predicted effect of shifting relative attention towards each attribute, from the 10^{th} to the 90^{th} percentile.
* denotes significance at the 1% level

A. Fraction Patient					
Constant	-0.71	0.58	-0.05	0.99	-
	(0.56)	(0.47)	(0.59)	(0.14)	-
Delayed Amount	3.37*	-	-	-	0.03
	(1.51)	-	-	-	(1.44)
Delayed Delay	-	-0.16	-	-	2.04
	-	(2.05)	-	-	(1.53)
Immediate Amount	-	-	1.96	-	1.61
	-	-	(1.92)	-	(1.46)
Immediate Delay	-	-	-	-4.43**	-4.18**
	-	-	-	(1.28)	(1.56)
R^2	0.22	0.00	0.05	0.40	0.89
B. Estimated log(*k*)					
Constant	0.05	-5.41*	-2.91	-7.42	-
	(2.71)	(2.21)	(2.79)	(0.67)	-
Delayed Amount	-14.79$^+$	-	-	-	-2.70
	(7.26)	-	-	-	(7.03)
Delayed Delay	-	-0.14	-	-	-12.97
	-	(9.71)	-	-	(7.48)
Immediate Amount	-	-	-8.33	-	-10.29
	-	-	(9.14)	-	(7.10)
Immediate Delay	-	-	-	19.80**	16.12$^+$
	-	-	-	(6.26)	(7.63)
R^2	0.19	0.00	0.04	0.36	0.97

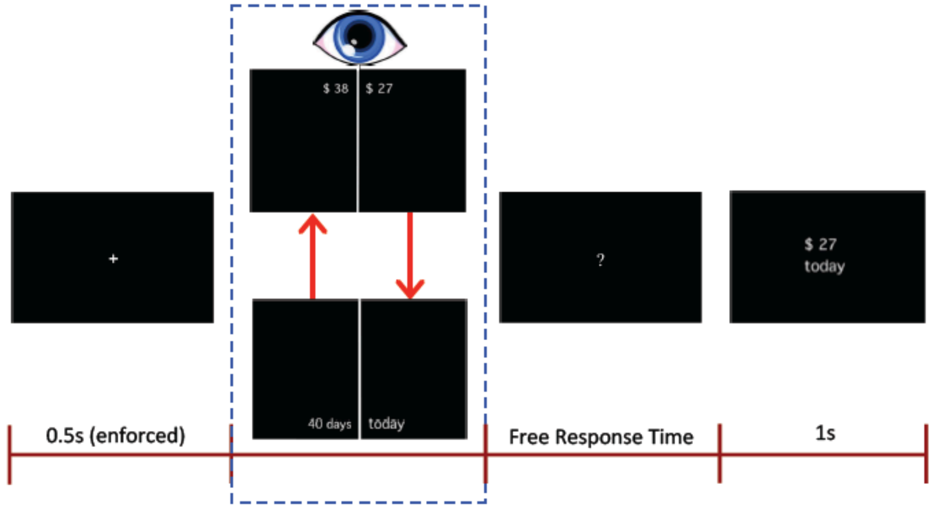
Table 2

Each column reports the results of a linear regression where subject-specific measures of patient (top = mean fraction of patient decisions, bottom = estimated log(k)) were regressed on the average fraction of time that each subject spent attending to particular ROIs. The slopes and constants from each regression are reported, with standard errors below in parentheses.
** Significant at the 1% level, * Significant at the 5% level, $^+$Significant at the 6% level.

Figure 1

Trial structure for Experiment 1. Every trial, subjects first stared at the central fixation cross for 500ms. Afterwards, the choice set was revealed and subjects had as long as they liked to make a choice between an option to be received today and an option to be received at some future date. Eye fixations were recorded at this point. After entering a response, subjects saw feedback for 1 second, and then moved to the next trial.

A)



B)

		Order	
		Amounts 1st	Delays 1st
Length on Screen	Amounts > Delays	2s **A** / 0.5s **D**	0.5s **D** / 2s **A**
	Delays > Amounts	0.5s **A** / 2s **D**	2s **D** / 0.5s **A**

Figure 2

Design for Experiment 2. (A) Trial structure for choice task in Experiment 2. Choices consisted of questions they should be close to indifferent between. First, subjects fixated on a central fixation cross for 500 ms. Next, they saw a pair of screens that alternated for a fixed length of time, depending on which of the four conditions was implemented. After switching between the two screens for a minimum of 5s, subjects were shown a question mark and had as long as they liked to enter a choice, but could only enter their response once they saw the question mark. Afterwards, feedback was shown for 1 second, and subjects continued to the next trial. (B) Exposure structure for the four different conditions. Both the length that each screen appeared as well as its order was varied. Each experimental cell repeated itself in each trial so that the total exposure time was not less than 5 seconds, and could last longer depending on whether the subject fixated to the options on the screen.

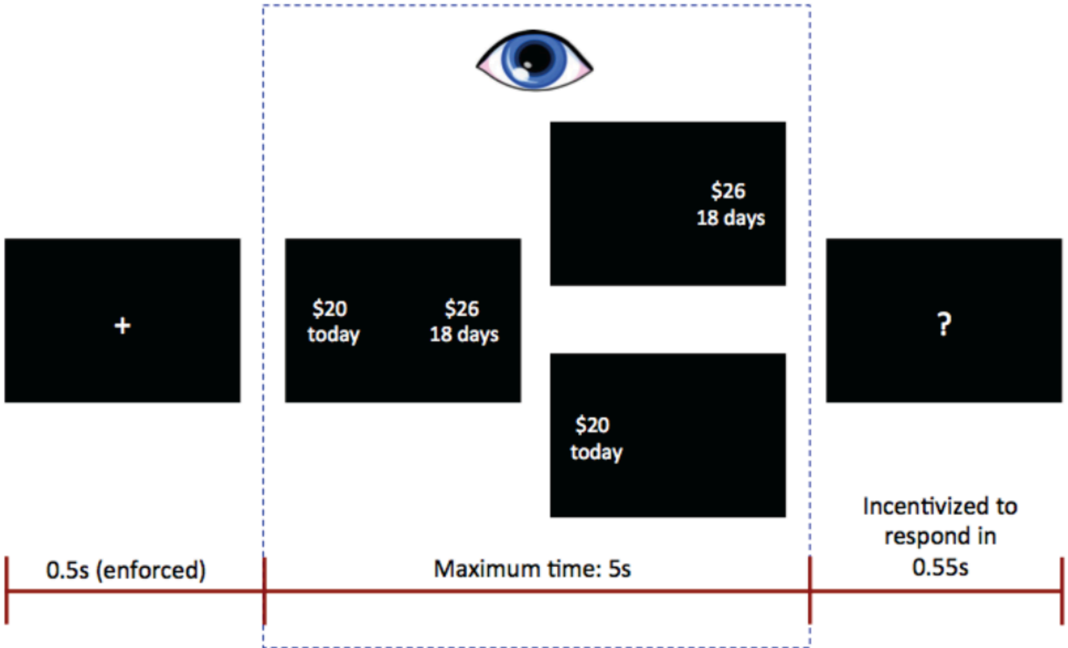


Figure 3

Trial structure for the choice task in Experiment 3. Choices consisted of questions they should be close to indifferent between. First, subjects fixated on a central fixation cross for 500 ms. Next, they saw both choice options presented and were free to fixate between them. One option was randomly designated the target, and the other was the non-target. Once the target (non-target) option was fixated at for 1.2 (0.3) seconds, it disappeared from the screen leaving only the non-target (target) visible. Then, once the non-target (target) option was attended to for a total of 0.3 (1.2) seconds, it also disappeared and a question mark appeared in the center of the screen. This was the subject's cue to enter their response and after doing so, feedback was shown and they continued to the next trial.

SI: Methods

Experiment 1

Task. Subjects completed 216 trials of an intertemporal monetary choice task. In each trial, subjects first viewed a fixation cross at the center of the screen and were asked to fixate on it for 500ms. Compliance was monitored with the eye tracker, so that the trial proceeded to the next stage only after 500 ms of continuous fixation. The duration of the initial fixation was enforced to ensure that subjects began every choice trial by fixating at the center of the screen. Next, subjects were shown a choice screen in which they had to make a decision between receiving a smaller, sooner monetary reward and a larger, later alternative. Subjects had as long as they needed to make a decision and indicated their choice by pressing either the left or right buttons on a keyboard with their dominant hand. Afterwards, subjects saw a 1 second feedback screen depicting their choice. Trials were separated by a 1 second black screen.

At the end of the experiment, one trial was selected at random and the subject's choice for that trial was implemented. This trial was determined by having the subject pick a random number out of an envelope. While determining the payments, subjects completed a short questionnaire. To minimize differences in transaction costs or credibility between immediate and delayed payments, choices were implemented via PayPal, with payments sent at the appropriate delay. In addition, regardless of the delay, subjects received an email at the time their PayPal account had been credited.

The choice screens were constructed as follows. Monetary amounts (range: $17-$60) always appeared at the top of the screen, while the delays (range: 0-200 days) always appeared at the bottom. Each trial included an immediate option with a delay of 0 days, which was labeled as “today.” The other delay was given as “X days,” where X ranged from 7-200. The size of the delayed amount was always constrained to be at least as large as the immediate amount. The order of the questions was randomized for each subject. The location (left or right) for the immediate and delayed options was randomized every trial.

Eye tracking. Eye movements were recorded at 50 Hz using a Tobii X50 desktop-mounted eye tracker. The eye tracker recorded throughout the choice task and produced a time series consisting of fixation locations for both right and left eyes at every time point. Subjects were required to keep their dominant hand on the response buttons throughout the task. This was done to eliminate eye movements related to the motor implementation of the choice, as opposed to the choice process, which is our object of interested.

We focus our analyses on the fixation locations during the choice screens. In particular, we measure the amount of time during choice that is spent looking at each of the four regions-of-interest (ROIs): immediate amount, immediate date, delayed amount, and delayed date. To do this, we define four boxes that add ten percent of the screen size, in pixels, around each of the ROIs, centered at the location of the associated text. The text for the ROIs was centered at locations (171, 120), (171, 598), (853, 120), and (853, 598) (coordinates in pixels based on a screen resolution of 1024x718). We then used the eye tracking data to measure which ROI, if any, was fixated at 20 milliseconds intervals during the choice.

Due to eye tracker and behavioral noise, some fixation data is missing. We deal with these missing fixations as follows. When a missing fixation was recorded between fixations to the same item, those missing fixations are changed to fixations to that item. For instance, the fixation pattern “upper right, missing, upper right” would become “upper right, upper right, upper right.” The assumption here is that the missing fixation corresponds to an eye blink or a temporary loss of eye location by the eye tracker. However, if the missing fixation occurred between fixations to different items, or at the beginning or end of the trial, those fixations are treated as missing data. So, the pattern “upper right, missing, upper left” would drop the missing fixations from the analysis. In this case, the assumption is that the subject made a saccade from one ROI into another, and some data was lost during this transition.

For every subject and trial, we computed the amount of time spent looking at one of the four ROIs. We then determined the relative time spent looking at each ROI by dividing the time spent looking at each ROI by the total time spent looking at all four ROIs for every trial.

Experiment 2

Task. The experiment consisted of two parts. In Part I, subjects performed a forty trial version of the choice task in Experiment 1, without any attentional manipulation. The goal of this part was to estimate a subject-specific discount rate that was then used to construct a choice set that the subject made decisions over in Part II. In particular, we constructed forty new choice questions for Part II such that the immediate amount varied between \$18 and \$27, the immediate delay was always "today," and the delayed date was in the set of {7, 14, 21, 30, 40, 60, 90, 180} days. The first two attributes were chosen randomly, and then the delayed amount was selected (using only integer monetary amounts) to generate questions in which the subject was as close as possible to being indifferent between choosing the immediate or delayed option, given the estimated k for the subject.

In Part II subjects encountered each choice problem four times, in random order, for a total of 160 trials. In each trial subjects first viewed a fixation cross at the center of the screen and were asked to fixate on it for 500ms. The trial would only proceed once they had done so. Next, we exogenously varied visual attention to the amounts and delays by alternating between two display screens: one screen depicted only the amounts and the second depicted only the delays. In all screens, the two attributes were displayed next to each other in order to facilitate processing them in parallel. Each of the forty questions appeared in all four conditions, which varied both the relative exposure to amounts and delays, as well as the order in which they appear. The length of the exposures was enforced by the eye tracker so that, for example, in trials in which amounts appeared for longer than delays, the screen would not advance until the subject had looked at amounts for a total of 2 seconds. Eye movements were monitored at either 250 or 500 Hz, using a desktop mounted SR Research Eyelink 1000.

Every trial involved 5 seconds of enforced exposure, which allowed for two showings of each of the attributes, one for a total of 4 seconds and the other for a total of 1 second.

Afterwards, a question mark appeared, which cued subjects to enter a response by key press, just as in Experiment 1. After seeing feedback for one second, the task advanced to the next trial.

Subjects were not informed that there were different types of trials. Instead, each subject took part in a short initial training round that consisted of 8 practice trials. They were informed that these practice questions would not count for payment, but were designed to ensure they understood the instructions for this part. The practice trials consisted of two questions from each of the four conditions, but with a different set of amounts and delays from those used in Part II.

Subjects were allowed to take a short break every 25 trials. At the end of the task, they picked a number from 1-200 out of an envelope. This number corresponded to the single trial, from either Part I or II, that would be implemented. All payments were implemented as in Experiment 1.

Remarks. Several features of the experiment are worth highlighting. First, given the sizable individual variation in discount rates, we used a within subjects design to increase statistical power. Second, we manipulated exposure to amounts versus delays because the results of Experiment 1 suggest that this may affect patience, given the asymmetric impact of fixating on the short delay versus the other attributes. Third, by manipulating the order of exposure, we control for the possibility of order effects. Finally, we emphasize that exposure need not be exactly equal to attention in this task. For example, subjects might have made a choice before the end of the exposure cycle and thus, might not process the stimuli throughout the last part of the trial. Subjects might also have been distracted by the need to detect screen changes and move their eyes in response. As a result, there is uncertainty about the size of the relative attention difference that is generated by the different conditions.

Experiment 3

Task As in Experiment 2, the experiment consisted of two parts. Part I was identical to Experiment 2's Part I. At the end of Part I, 100 new choice questions were constructed for Part II so that the subject would be close to indifferent in each choice.

In Part II, subjects encountered each choice problem twice, in random order, for a total of 200 trials. In each trial subjects first viewed a fixation cross at the center of the screen for 500ms, which was enforced by the eye tracker. Next, the two choice options were displayed on the screen and subjects were free to look between them. Unbeknownst to the subjects, in each trial the computer selected one of the options to be the "target" and the other to be the "non-target." Throughout the trial, the computer recorded the total duration that each option was attended to, and once an option reached its maximum fixation time, it disappeared from the screen. As one option must reach is maximum fixation time before the other, this resulted on having only one option was visible on the screen at this point. Once both options reached their maximum fixation time, or a total of 5 seconds since the start of the trial elapsed, a question mark appeared at the center of the screen and subjects were instructed to indicate their response as quickly as possible. The maximum fixation time for the target option was 1.2 seconds and the maximum fixation time for the non-target option was 0.3 seconds. In order to encourage subjects to respond immediately upon seeing the question mark, they were told that if in at least 190 out of the 200 trials they responded within 0.55 seconds of the question mark appearing they would receive an additional $5 at the end of the experiment. Eye movements were monitored at 500 Hz using a desktop mounted SR Research Eyelink 1000.

Importantly, subjects were not informed that there were different types of trials. Instead, each subject took part in a short initial training round that consisted of 8 practice trials. They were informed that these practice questions would not count for payment, but were designed to ensure they understood the instructions for this part, and how to enter their responses. The practice trials consisted of four questions from each of the two conditions, but with a different set of amounts and delays from those used in Part II. No information regarding their response speed was provided.

Subjects were allowed to take a short break every 25 trials. At the end of the task, they picked a number from 1-240 out of an envelope. This number corresponded to the single trial, from either Part I or II, that would be implemented. All payments were implemented as in the previous sections.

Remarks: Two features of the experiment are worth highlighting. First, whereas Experiment 2 found evidence that differential attention to amounts and delays could lead to choice biases, this experiment sought to test whether differential attention to the options could induce choice biases. Additionally, in this task the time from choice onset to decision more closely approximates the time it would take to make a choice without any experimenter attention manipulation. Subjects saw the options for a maximum total of 1.5 seconds, and were encouraged to enter their responses quickly after.

Estimating Discount Rates

We use two different measures of the intertemporal discount rate. First, for every subject we compute the fraction of time the delayed option was chosen, which we refer to as a patient choice.

Second, for each subject we estimated the discount rate that best explains the choice data using a hyperbolic model. We use the method proposed in Chabris et al. (2008), which is frequently used in the literature. The method assumes that subjects make choices by computing a value for each option and then comparing them. The value of receiving $\$Y$ in D days is assumed to be $\frac{Y}{1+kD}$, where k is a discount parameter controlling the subject's patience: a low k signifies patient decision-making, and a large k signifies impatient behavior. Subjects then choose the delayed option with probability

$$\frac{e^{\frac{\omega Y}{1+kD}}}{e^{\omega X}+e^{\frac{\omega Y}{1+kD}}},$$

where X is the monetary amount offered in the immediate option. The parameter ω controls the amount of the noise in the choice process: choices are fully random when $\omega = 0$, and their sensitivity to value differences increases with ω. For every subject, we used maximum likelihood estimation to estimate the $\hat{k}$ and $\hat{\omega}$ parameters that best explain the choice data.

Note that in many analyses entailing our estimates of k, we choose to analyze the logarithm of k, rather than k itself. We do this as a result of the functional form of the hyperbolic discounting model. When k is relatively low, small changes in k can produce large changes in decisions; yet when k is large, the same size changes will produce less noticeable decision alterations. Note that if the MLE estimate of k was 0, we used 0.0001 as the estimate when computing log(k) in the previous analysis.

In Experiment 1, we used the estimated k parameter for each subject to compute trial specific measures of the value of the sooner option (V_{sooner}), the value of the delayed option (V_{later}), the relative value of the patient option ($V_{diff} = V_{later} - V_{sooner}$), and the difficulty as measured by the absolute value of V_{diff}.

Total Fixation Time to Each ROI			
A. Spatial			
	Left	Right	
Up	34.4 (3.4)	33.1 (4.2)	67.4 5.4
Down	17.0 (3.7)	15.6 (2.3)	32.6 (5.4)
	51.4 (3.8)	48.6 (3.8)	
B. Feature of Interest			
	Immediate	Delayed	
Amount	30.3 (3.2)	37.1 (3.7)	67.4 (5.4)
Delay	10.0 (3.9)	22.6 (3.1)	32.6 (5.4)
	40.3 (4.0)	59.7 (4.0)	
C. Standard Deviation			
Amount	10.5 (2.0)	11.6 (2.3)	
Delay	8.4 (1.6)	10.0 (2.2)	

Table S1

Total fixation time to each ROI. (A) The mean percent of total fixation time that was spent attending to each region of interest split by the spatial orientation of the screen: columns are horizontal orientation and rows are vertical orientation. (B) The mean percent of total fixation time that was spent attending to each region of interest split by feature of interest: columns are the immediate and delayed options and rows are monetary amounts and delay dates. Means are taken over subject-specific means, and standard deviations reported below in parentheses. (C) Reports the mean standard deviation over subjects from (B). Means are taken over subject-specific standard deviations, and standard deviations appear below in parentheses.

Percent of First Fixations to Each ROI			
A. Spatial			
	Left	Right	
Up	68.8 (24.1)	24.5 (23.0)	93.3 (6.2)
Down	5.0 (5.0)	1.6 (1.8)	6.6 (6.2)
	73.8 (23.9)	26.2 (23.9)	
B. Feature of Interest			
	Immediate	Delayed	
Amount	45.3 (3.3)	48.0 (4.9)	93.3 (6.2)
Delay	2.0 (2.2)	4.7 (4.2)	6.6 (6.2)
	47.3 (2.5)	52.6 (2.5)	

Table S2

Percent of first fixations to each ROI. (A) The mean percent of first fixations that were made to each region of interest split by the spatial orientation of the screen: columns are horizontal orientation and rows are vertical orientation. (B) The mean percent of first fixations that were spent attending to each region of interest split by feature of interest: columns are the immediate and delayed options and rows are monetary amounts and delay dates. Means are taken over subject-specific means, and standard deviations reported below in parentheses.

Impact of Attribute Values of Attention				
	Immediate Amount	Delayed Amount	Immediate Delay	Delayed Delay
Constant	0.245	0.386	0.128	0.241
	19.15	26.52	8.96	17.74
Immediate Amount	0.004	-0.004	0.003	-0.003
	6.58	-5.98	5.82	-5.60
	4.1%	**-4.0%**	**3.3%**	**-3.3%**
Delayed Amount	-0.001	0.002	-0.003	0.002
	-4.52	7.53	-8.02	5.47
	-4.2%	**10.2%**	**-12.5%**	**6.5%**
Delayed Date	0.000	-0.000	-0.000	0.000
	3.83	-4.37	-0.77	1.62
	2.2%	**2.7%**	**0.3%**	**0.9%**

Table S3

Each column reports the results of a linear mixed-effects regression. In each case, the dependent variable was the fraction of time attending to the ROI, and the independent variables were a constant, the immediate monetary amount offered, the delayed monetary amount offered, and the delayed date. For each independent variable, coefficients and t-statistics from the regression are reported in the first two rows for each coefficient. For all independent variables, excluding the constant, a measure of effect size appears in bold in the third row. This effect size measures the average change in the fraction of time attending to each ROI, as the corresponding variable increases from the minimum amount shown to subjects to the maximum amount shown to subjects.

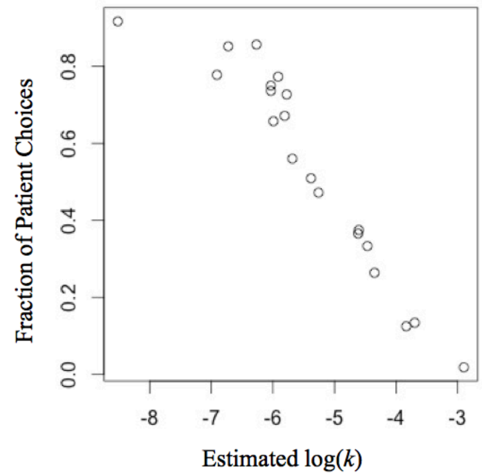


Figure S1

Relationship between log of the estimated k from the hyperbolic discounting model, plotted on the x-axis, and the fraction of patient choices, plotted on the y-axis, where each point on the graph corresponds to a different subject.

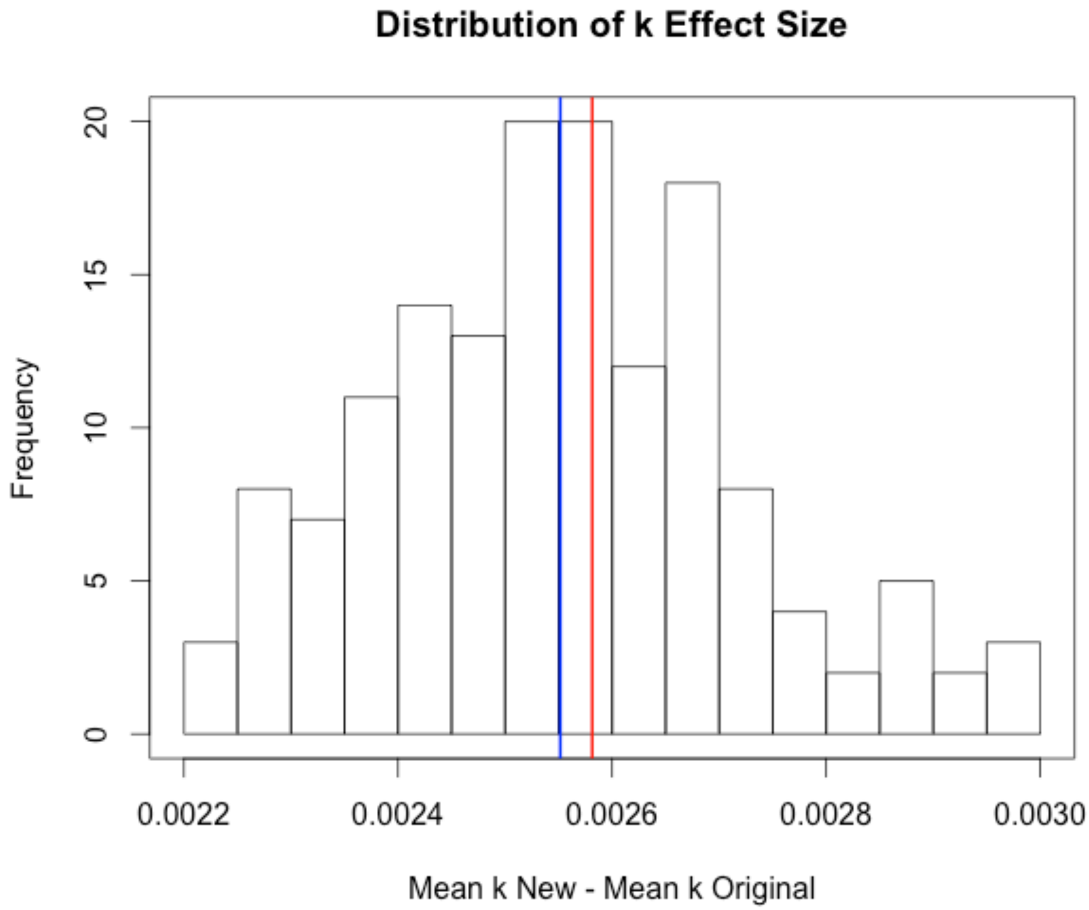


Figure S2

Histogram depicting effect size differences in the simulation exercise from Experiment 3. The x-axis depicts the difference between then mean estimated *k* after altering choices by randomly drawn effect sizes (k New) and the mean estimated k before such an exercise took place (k Original). The blue vertical line depicts the mean over all 150 times completing this simulation while the red vertical line depicts the statistics from the first time this was done, as described in the main text.

Chapter 3

SYMMETRY IN COLD-TO-HOT AND HOT-TO-COLD VALUATION GAPS

It is well known that the utility, or value, derived from consumption is modulated by emotional and physiological states at the time of consumption. For example, the pleasure of drinking water is larger when thirsty than when quenched. One basic question is whether individuals anticipate the effect of these “visceral” states in their utility when making decisions about future consumption. For instance, can a hungry grocery shopper buy the correct amount of food to consume throughout the week? A sizable body of evidence has shown that individuals in a “cold” state (e.g., satiated) systematically underestimate the increase in consumption value that they would experience in a “hot” state (e.g., hungry) (Badger et al., 2007; D. T. Gilbert, Gill, & Wilson, 2002; George Loewenstein, Nagin, & Paternoster, 1997; Nisbett & Kanouse, 1968; Sayette, Loewenstein, Griffin, & Black, 2008; Van Boven & Loewenstein, 2003). This phenomena is known as a cold-to-hot empathy gap in psychology (G. Loewenstein, 1996), and as a projection bias in behavioral economics (G. Loewenstein, O'Donoghue, & Rabin, 2003).

Empathy gaps can arise in two different scenarios. Cold-to-hot gaps refer to situations like forecasting the value of eating a hamburger in a hungry state, while being satiated at the time of decision. Hot-to-cold gaps refer to the opposite situation; like forecasting the value of eating dessert at the end of the meal in a satiated state, while being hungry at the time of decision. An important open question is whether both types of gaps are symmetric in the following ways. First, do individuals underestimate their change in preferences to the same degree when going from cold-to-hot as when going from hot-to-cold? Second, are symmetric mechanisms at work in generating both types of empathy gaps?

The answer to these questions matters for several reasons. First, they inform our beliefs about the likelihood that individuals make mistakes of similar magnitude in both types of situations, as well as the extent to which both type of mistakes can be addressed with similar policy instruments. Second, theories in behavioral economics and psychology have posited that cold-to-hot and hot-to-cold gaps are symmetric and driven by similar mechanisms, but this has not been previously tested.

Previous work has provided strong evidence for the existence of empathy gaps, but has not provided a definite answer to either of the two symmetry questions. In fact, the vast majority of experimental studies have focused on the cold-to-hot case (Badger et al., 2007; D. T. Gilbert et al., 2002; George Loewenstein et al., 1997; Nisbett & Kanouse, 1968; Sayette et al., 2008; Van Boven & Loewenstein, 2003). One important exception is (Read & van Leeuwen, 1998), which like us, compares hungry-satiated and satiated-hungry empathy gaps in real food choice. However, unlike us, their methodology does not permit a direct comparison of the extent to which changes in utility are underestimated in both cases. With respect to the second question, as far as we know, no previous experiments have investigated the mechanisms at work in projection bias, nor the extent to which they are symmetric.

Methods

Subjects. 101 Caltech students took part in two behavioral sessions. In order to encourage participants to return for the second session, subjects were paid \$10 after the first session and \$40 after the second. The experiment was approved by Caltech's Institutional Review Board.

Stimuli. Each subject saw two different food sets, each containing 50 different snacks. These snacks consisted of a variety of candies, fruits, chips, and energy bars. Subjects saw one set of foods in the first session, and a different set in the second. The identity of the sets was determined randomly for each subject. We used two different sets to avoid consistency biases in the tasks described below. All of the foods received a mean neutral-to-appetitive rating in previous experiments.

Task. The experiment consistend of two sessions that occurred at the same time of day, but were separated by 3 to 5 days. Fig. 1A provides a summary of the events in the experiment and Fig. 1B provides details on the timing of a typical trial for each of the tasks. In each session, subjects completed four tasks in the order specified below.

First, subjects performed a liking-rating task in which they rated how much they wanted to eat each of the 50 snacks at the end of the session (scale: integers from -2 to 2, "How much would you enjoy that particular food at the end of TODAY's experiment?"). The purpose of these ratings was to familiarize the subjects with the entire set of foods prior to the main bidding task.

Second, subjects entered bids for the right to eat each of the foods at the end of the second day of data collection, and were explicitly told whether they would be hungry or satiated at that time. Bids were entered by pressing a button and could take integer values from $0 to $4. Subjects bid on each food twice. At the beginning of the experiment they were informed that at the end of the second session they would need to remain in the lab for 20 minutes, and the only thing that they would be able to eat was whatever they purchased from us through their bids. At the end of day 2, one of the trials (from either date) was selected and implemented using the rules of a Becker-DeGroot-Marschak (BDM) auction (Becker, DeGroot, & Marschak, 1964).[1] Subjects were given $4 in bidding cash and kept whatever they did not spend. The bids provide a measure of the perceived value at the time of decision for eating the food at the end of session 2.

Third, subjects provided taste ratings (scale: integers from -2 to 2, "How tasty you believe that food to be, independent of any health considerations") and health ratings (scale: integers from -2 to 2, "How healthy you believe that food to be, independent of any taste considerations") for each of the foods. The ratings were collected in blocks, with the order randomized across subjects. These ratings provide a measure of the perceived attributes of each food at the time of decision.

---

[1] Briefly, the rules are as follows. Let b be the bid entered by the subject, and let n be a randomly selected number. If $b>=n$, the subject gets to eat the snack shown in that trial and only pays $\$n$ for it. If $b < n$, the subject gets nothing and pays nothing. We used this procedure because in is incentive compatible (i.e., the best strategy for the subjects is to bid their true value for the items), a fact that was emphasized during the training period.

As shown in Fig. 1C, the experiment had a 2x2 factorial design, with conditions varying across subjects. In each date we exogenously manipulated subjects' hunger by asking them to fast for 4 hours prior to the experiment (hungry), or to eat a large snack within half an hour prior to the start of the experiment (satiated). This led to four treatment groups: satiated-hungry (SH), satiated-satiated (SS), hungry-satiated (HS), and hungry-hungry (HH). The first dimension denotes whether subjects were hungry or satiated during the first session. The second dimension denotes whether subjects were hungry or satiated during the second session. The experiment consisted of 23 subjects in group SH, 27 in group SS, 27 in group HS, and 24 in group HH. Before entering the lab, subjects were verbally asked to report the last time they ate. If they gave an answer inconsistent with the instructions, they were excluded from further participation in the experiment (and not reported in the analyses). We use the following notation to simplify the description of the results. XYN denotes date $N = \{1,2\}$ for condition $XY = \{HH, SS, SH, HS\}$.

Results

Paradigm validation. Fig. 2 summarizes the bidding data for all of the conditions. Note that, in each date, the bids provide a measure of the perceived value of eating a snack at the end of the second session. The data illustrates several points. First, subjects bid consistently in both dates when there were no changes in the state (HH1 vs HH2: $p > 0.84$, SS1 vs SS2: $p > 0.57$; paired two-tailed t-tests).

Second, the bids at date 2 did not depend on the state in date 1 (HH2 vs SH2: $p > 0.80$, SS2 vs. HS2: $p > 0.90$; two-tailed t-tests). This implies that subjects bid the same amount in date

2 when they were in the same state, regardless whether their state at the time of the first bid was the same or different. In addition, we found no difference between the mean bid in SH2 and the average of HH1 and HH2 ($p > .82$, two-sided t-test), or between HS2 and the average of SS1 and SS2 ($p > .97$, two-sided t-test). For this reason, in some of the analyses below we pool the 8 bidding conditions into four cases: no-gap hungry (HH1, HH2, and SH2), no-gap satiated (SS1, SS2, HS2), hungry-satiated gap (HS1), and satiated-hungry gap (SH1).[2] Note that we pool the conditions HH1, HH2 and SH2 together, and call them the no-gap hungry case, because in all of those cases subjects are making decisions in a hungry state about what to consume in date 2 also in a hungry state. Analogously, we refer to the conditions SS1, SS2, and HS2 as the no-gap satiated case because subjects in all of those instances make decisions in a satiated state about consumption in date 2 in a satiated state.

Third, bids in the no-gap hungry case were on average 62 cents larger than in the no-gap satiated case ($p < 0.01$, two-tailed t-test), which demonstrates that our state manipulation affected subjects' food values.

Symmetric empathy gap. As shown in Fig. 2, we found a cold-to-hot (SH2 – SH1 = \$0.19 ± 0.08; $p < 0.018$, two sided t-test) and a hot-to-cold empathy gap (HS2 – HS1 = -\$0.20 ± 0.08; $p < 0.016$, two-sided t-test). The value difference is positive in the cold-hot case because subjects underestimate the value of eating when hungry when making decisions in

[2] To form the no-gap cases, we first averaged responses within conditions with the same subjects (e.g., average HH1 and HH2), and then averaged responses from the third group (e.g., SH2).

a satiated state. The opposite is true in the hot-cold case. A direct comparison revealed no differences between the magnitudes of both mistakes, which is consistent with a symmetric effect size for both types of value gaps (|SH2 - SH1| vs. |HS2 – HS1|: $p > 0.92$, two-sided t-test).

The next set of results is about the mechanisms at work in the empathy gaps, and the extent to which they work symmetrically in the two directions. We hypothesized that the mistakes in value forecasting could operate through three different mechanisms.

First, subjects might change their perception of the attributes of foods, such as how healthy or how tasty they are. For example, they might perceive junk foods to be healthier when making decisions in a hungry state. We refer to this channel as the *attribute perception mechanism*. We can test for this mechanism by comparing the distribution of taste and health ratings provided in the different conditions.

Second, hunger might increase the baseline value of all foods, regardless of their attributes. This would show up as a constant shift in the value of the foods. We refer to this channel as the *baseline value mechanism*. We can test for it by estimating a linear regression, for each subject and session, of the bids on the taste and health ratings, and then comparing the distribution of estimated constants.

Third, hunger might change how a food's attributes are weighted in computing its value. We refer to this channel as the *attribute weighting mechanism*. We can test for it using the same linear regression described above and comparing the distribution of estimated coefficients for the health and taste ratings.

Symmetric attribute perception mechanism. We tested for the role of this mechanism by comparing the mean taste and health ratings across the four cases: satiated no-gap, HS1, SH1, and hungry no-gap. We pooled the data this way to increase the statistical power of our tests. This is justified by the fact that, since there were no differences in the bids within each of the four cases, the mechanisms are also likely to be deployed in a similar way within each case. As shown in Table 1, we found that taste ratings were higher in HS1 than in the satiated no-gap case (HS1: 0.53 ± 0.07; satiated no-gap: 0.32 ± 0.05; $p < 0.02$, two-sided t-test). This is consistent with the idea that subjects in a hungry state overestimate the degree to which they will perceive the snacks as tasty when satiated. However, we did not find a difference between the SH1 and the hungry no-gap case (SH1: 0.37 ± 0.10; hungry no-gap: 0.50 ± 0.07; $p < 0.27$, two-sided t-test), although the sign of the difference is in the predicted direction, and the effect size is similar to the previous one. There were also no significant differences for health ratings. Together, this provides partial support for the hypothesis that the attribution perception mechanism is at work, and suggests that hunger affected the perception of the more "visceral" taste attributes, but not the perception of the more "cognitive" health attributes.

Symmetric baseline value and attribute weighting mechanisms. We tested for the role of these two mechanisms by estimating a linear mixed regression model. We regressed the amount bid on an indicator variable for each case (satiated no-gap, HS1, SH1, and hungry no-gap), as well as an interaction of each indicator with health and taste ratings. Random slopes were fit for each subject.[3]

The estimates are reported in Table 2. A comparison of the constants shows that that they exhibited a pattern similar to the bids, by underestimating the extent to which their value changes from the state at the time of bid to the state at the time of consumption. In particular, the constant in HS1 estimates a baseline value for consumption that is 11 cents higher than the satiated no-gap case, while the constant in SH1 estimates a baseline value 13 cents lower than the hungry no-gap case. This suggests that the baseline value effect was symmetric, which we tested by estimating a linear contrast of the distribution of estimated constants (with weights -1.5, -0.5, 0.5, 1.5 from left-to-right columns of Table 2; $p < 0.01$). A comparison of the taste coefficients reveals a similar pattern: a linear test of the taste coefficients suggests that the attribute weighting mechanism is symmetric in the case of taste (similar weights, $p < 0.01$). In contrast, a similar test found no significant differences for the health coefficients ($p > 0.28$). Together, these results suggest that both the baseline value and the attribute mechanism are at work in a symmetric fashion. Furthermore, it suggests the attribute weighting mechanism changes the valuation of the more "visceral" taste ratings, but not the valuation of the more "cognitive" health ratings.

Predicting cross-individual differences in empathy gaps. We carried out an additional post-hoc analysis to further test the validity of the mechanism results. We reasoned that if the

[3] Responses for subjects in the same experimental condition in the no-gap cases were equally weighted, since subjects saw different foods on each day.

identified mechanisms play a critical role in generating the empathy gaps, they should be correlated with cross-subject differences of the magnitude of the empathy gaps. To test this, we used the data from the HS and SH conditions to estimate a linear regression of an individual measure of the empathy gaps (given by mean bid in day 2 minus mean bid in day 1) on a measure of the individual taste perception effects (given by mean taste rating in day 2 minus taste rating in day 1), a measure of the individual baseline value effects (given by the estimated constant in day 2 minus the estimated constant in day 1, for each subject), and the taste attribute weighting effects (given by the estimated taste coefficient in day 2 minus the estimated taste coefficient in day 1, for each subject). The regression took into account the potential of measurement error on the independent variables, since they were estimated from linear regressions at the individual level. As shown in Table 3, we found that the size of the empathy gap was significantly correlated with the size of changes in our three relevant mechanisms. These results provide additional evidence in favor of the mechanism results described above.

Discussion

We carried out a modified version of the classic experiment by Read and van Leeuwen in order to address the following two basic open questions. First, do individuals incorrectly predict their change in preferences to the same degree when going from cold-to-hot as when going from hot-to-cold? Second, are symmetric mechanisms at work in generating both types of empathy gaps?

With respect to the first question, we found that the size of the empathy gap was symmetric: satiated individuals underestimated the value of foods to consume when hungry by a similar amount that hungry individuals overestimated the value of foods to consume when satiated. This provides support for the types of decision-making models proposed in the projection bias literature (Conlin, O'Donoghue, & Vogelsang, 2007; G. Loewenstein et al., 2003). In addition, this result suggests that both types of gaps lead to decision-making mistakes of similar magnitude, and thus ought to be of equal importance in public policy interventions.

With respect to the second question, we found evidence that three different mechanisms are at work in generating the empathy gaps, and appear to operate largely symmetrically. First, we found that subjects increase their perception of the tastiness of food when making decisions while hungry, regardless of their state at the time of consumption (attribute perception mechanism). Second, we found that they overestimate the value of the average food in hot-to-cold gaps, and underestimate it in cold-to-hot gaps (baseline valuation mechanism). Finally, we found that they overweight the anticipated tastiness of foods in hot-cold gaps, and underweight it in cold-hold gaps (attribute weighting mechanism).

Interestingly, the attribute perception and attribute weighting mechanisms seem to operate in the more “visceral” taste dimension, but not in the more “cognitive” health attribute. This suggests that changes in visceral states might lead to empathy gaps in part by changing how basic physiological attributes like taste are perceived and weighted, but that they do not affect how more abstract attributes like health are represented and weighted.

This is important because it suggests that one key to overcoming decision mistakes associated with empathy gaps may be to help individuals more accurately forecast these basic variables, instead of attempting to modulate representations of more abstract variables like health.

Our findings are also related to the important literature on mistakes in affective forecasting (Daniel T Gilbert, Pinel, Wilson, Blumberg, & Wheatly, 1998; D. T. Gilbert & Wilson, 2007; Riis, Loewenstein, Baron, & Jepson, 2005; Sackett & Torrance, 1978; Sieff, Dawes, & Loewenstein, 1999; Wilson, Wheatley, Meyers, Gilbert, & Axsom, 2000). A key difference between empathy gaps and the affective forecasting literature has to do with the types of values being forecasted and the mechanisms at work. In particular, much of the affective forecasting literature has focused on predicting the impact of current and future events on future well-being and mood, but not on decision-making per-se. In addition, a critical mechanism in many affective forecasting studies is an inability to forecast the speed at which visceral states change (e.g., how long will I be depressed after a divorce). In contrast, this mechanism is not part of the definition of empathy gaps, where subjects are assumed to know the future state, even if they cannot forecast their future utility properly.

We conclude by emphasizing two limitations of the study. First, the type of empathy gap studied here is likely to be relatively mild compared to those that arise in domains like addiction (Badger et al., 2007; Sayette et al., 2008) or sexual arousal (George Loewenstein et al., 1997). It is conceivable that the symmetry identified here breaks down in those cases, a possibility that should be investigated in future studies. Second, the list of attributes used in the study is far from comprehensive, and includes only taste and health, which lie in extreme positions of the visceral-cognitive spectrum. It is possible that there are attributes in the mid-part of the spectrum that also play a role in generating empathy gaps that we have not identified, even for the case of food choices.

References:

Badger, G. J., Bickel, W. K., Giordano, L. A., Jacobs, E. A., Loewenstein, G., & Marsch, L. (2007). Altered states: The impact of immediate craving on the valuation of current and future opioids. *Journal of Health Economics, 26*(5), 865-876. doi: Doi 10.1016/J.Jhealeco.2007.01.002

Becker, G., DeGroot, M., & Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral Science, 9*, 226-232.

Conlin, M., O'Donoghue, T., & Vogelsang, N. J. (2007). Projection bias in catalog orders. *American Economic Review, 97*(4), 1217-1249. doi: Doi 10.1257/Aer.97.4.1217

Gilbert, D. T., Gill, M. J., & Wilson, T. D. (2002). The future is now: Temporal correction in affective forecasting. *Organizational Behavior and Human Decision Processes, 88*(1), 430-444. doi: Doi 10.1006/Obhd.2001.2982

Gilbert, D. T., Pinel, E. G., Wilson, T. D., Blumberg, S. J., & Wheatly, T. P. (1998). Immune Neglect: A Source of Durability Bias in Affective Forecasting. *Journal of Personality and Social Psychology, LXXV*, 617-638.

Gilbert, D. T., & Wilson, T. D. (2007). Prospection: experiencing the future. *Science, 317*(5843), 1351-1354.

Loewenstein, G. (1996). Out of Control: Visceral Influences on Behavior. *Organizational Behavior and Human Decision Processes, 65*, 272-292.

Loewenstein, G., Nagin, D., & Paternoster, R. (1997). The effect of sexual arousal on predictions of sexual forcefulness. *Journal of Crime and Delinquency, XXXIV*, 443-473.

Loewenstein, G., O'Donoghue, T., & Rabin, M. (2003). Projection bias in predicting future utility. *Quarterly Journal of Economics, 118*(4), 1209-1248. doi: Doi 10.1162/003355303322552784

Nisbett, R. E., & Kanouse, D. E. (1968). Obesity, hunger, and supermarket shopping behavior. *Proceeding of the Annual Convention of the American Psychological Association, III*, 683-684.

Read, D., & van Leeuwen, B. (1998). Predicting Hunger: The Effects of Appetite and Delay on Choice. *Organizational Behavior and Human Decision Processes, LXXVI*, 189-205.

Riis, J., Loewenstein, G., Baron, J., & Jepson, C. (2005). Ignorance of hedonic adaptation to hemodialysis: A study using ecological momentary assessment. *Journal of Experimental Psychology-General, 134*(1), 3-9. doi: Doi 10.1037/0096-3445.134.1.3

Sackett, D. L., & Torrance, G. W. (1978). The utility of different health states as perceived by the general public. *Journal of Chronic Diseases, XXXI*, 697-704.

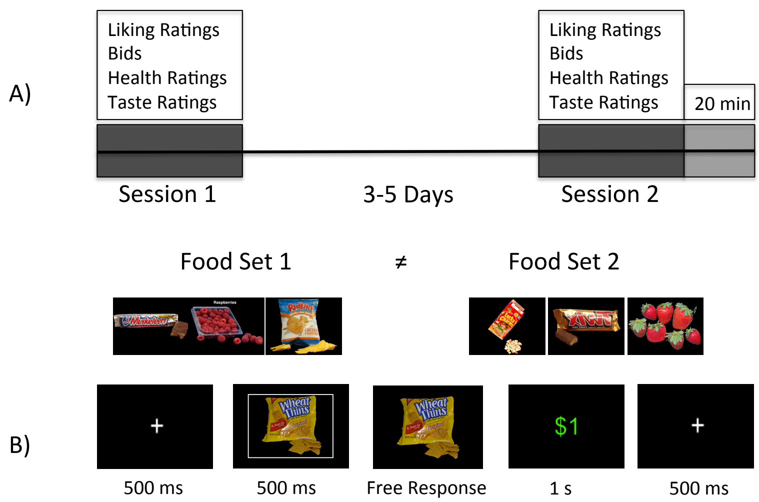
Sayette, M. A., Loewenstein, G., Griffin, K. M., & Black, J. J. (2008). Exploring the cold-to-hot empathy gap in smokers. *Psychological Science 19*(9), 926-932.

Sieff, E., Dawes, R., & Loewenstein, G. (1999). Anticipated versus actual responses to HIV test results. *American Journal of Psychology, CXII*, 297-311.

Van Boven, L., & Loewenstein, G. (2003). Social projection of transient drive states. *Personality and Social Psychology Bulletin, 29*(9), 1159-1168. doi: Doi 10.1177/0146167203254597

Wilson, T. D., Wheatley, T., Meyers, J. M., Gilbert, D. T., & Axsom, D. (2000). Focalism: A source of durability bias in affective forecasting. *Journal of Personality and Social Psychology, 78*(5), 821-836. doi: Doi 10.1037/0022-3514.78.5.821

Fig. 1. Experimental set-up. A) Subjects took part in 2 behavioral sessions, separated by 3-5 days. In each session they made liking, health, and taste ratings. They also made bids over foods to consume at the end of session 2. Regardless of whether they bought food, participants were required to remain in the lab for 20 minutes after the second session. On each day, subjects completed tasks over a different set of foods that was randomized across subjects. B) Timing of the task. On a computer monitor, subjects saw a fixation cross for 500ms before the trial food was revealed. After another 500ms, a white border disappeared and subjects had as long as they liked to give their rating/bid. Response feedback was provided for 1s before advancing to the next trial. C) The experiment consisted of four experimental conditions: hungry-hungry, hungry-satiated, satiated-hungry, and satiated-satiated.



		Session 2	
		Hungry	Satiated
Session 1	Hungry	HH	HS
	Satiated	SH	SS

Fig. 2. Mean bids by condition and experimental session. Standard error bars are shown.

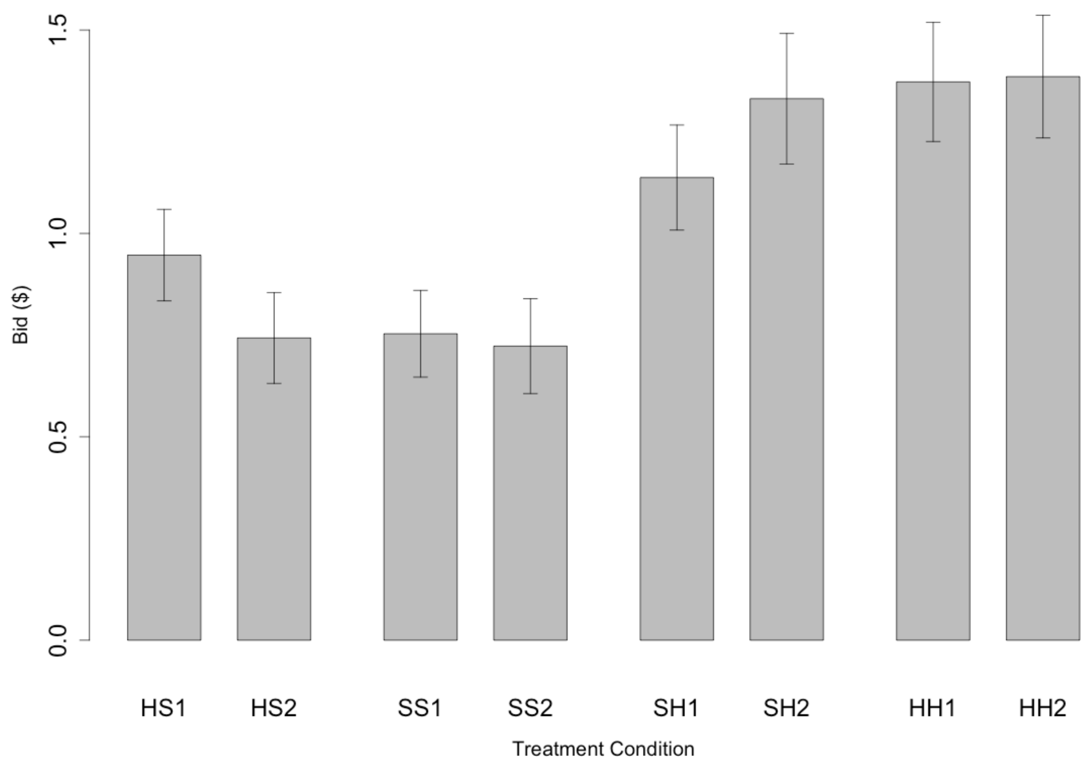


Table 1. Mean taste and health ratings by experimental condition. SDs reported in parentheses.

	Satiated No-Gap	HS-1	SH-1	Hungry No-Gap
Taste Rating	0.32 (0.36)	0.53 (0.34)	0.37 (0.49)	0.50 (0.45)
Health Rating	-0.53 (0.30)	-0.50 (0.31)	-0.41 (0.39)	-0.42 (0.34)

Table 2. Estimates of a linear mixed regression model where the bid for each food was regressed on an indicator variable for each case (satiated no-gap, HS1, SH1, and hungry no-gap), as well as an interaction of each indicator variable with health and taste ratings. Random slopes were fit for each subject. SEs are given in parentheses.

	Satiated No-Gap	HS-1	SH-1	Hungry No-Gap
Constant	0.62 (0.07)	0.73 (0.33)	0.93 (0.10)	1.07 (0.08)
Health	0.05 (0.02)	0.05 (0.03)	0.11 (0.06)	0.07 (0.03)
Taste	0.41 (0.04)	0.49 (0.09)	0.51 (0.06)	0.60 (0.04)

Table 3. Estimates of a linear regression of a measure of the empathy gaps on measures of the taste perception, baseline value, and taste weighting mechanisms. The regression accounts for measurement error on the independent variables. See text for details. Standard errors appear below in parentheses.
** significant at the 5% level, *** significant at the 1% level.

	Bid difference (all trials)	Bid difference (SH trials only)	Bid difference (HS trials only)
Baseline Value	0.92 *** (0.05)	0.92 *** (0.11)	0.87 *** (0.05)
Taste Perception	0.47 *** (0.07)	0.65 *** (0.11)	0.28*** (0.08)
Taste Weighting	0.39 *** (0.09)	0.26 ** (0.12)	0.65 *** (0.13)