COMPUTATIONAL BIASES
IN DECISION-MAKING

Thesis by

Vanessa Janowski

In Partial Fulfillment of the Requirements for the

Degree of

Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2012

(Defended May 9, 2012)

ACKNOWLEDGEMENTS

I would like to thank everyone who has helped and supported me throughout my Ph.D.

To start, I would like to thank my undergraduate mentors, Arturo Bris and Christos Cabolis, for introducing me to economic research and for their support and advice over the years. I would also like to thank my mother, Jadwiga Rogowska, for introducing me to neuroscience and functional magnetic resonance imaging (fMRI). They helped me to embark on the path to a Ph.D.

I would like to thank Ralph Adolphs for helping to convince me to come to Caltech and taking an interest in my development as a graduate student. He showed me what a successful scientist is like and inspired me to pursue my research in a thoughtful, determined way.

On a similar note, I would like to thank Colin Camerer for his mentorship in my first fMRI study and the always-stimulating discussions in lab meetings and in the classroom. He never fails to make me laugh and has inspired me to pursue better research.

I would also like to thank Peter Bossaerts, Pietro Ortoleva and Kota Saito for their help and advice along the way, as well as for serving on my committee. I would like to thank my co-authors Antonio Rangel, Colin Camerer, Eric Johnson, Martijn Willemsen and Erik Madsen for their contributions. Also, many thanks to Laurel Auchampaugh, Jenny Niese, Tiffany Kim, Barbara Estrada, Sheryl Cobb, and Victoria Mason for help with day-to-day logistics.

In addition, I gratefully acknowledge the National Science Foundation IGERT program as well as the HSS Department at Caltech for funding me through my Ph.D. Thanks also to the Kosciuszko Foundation and the Pulaski Scholarship for Advanced Studies for scholarships that helped support me during my Ph.D.

I owe the greatest debt for this thesis to my advisor, Antonio Rangel. His passion and drive for research motivated me to be a better scientist. He constantly encouraged me to be my best, and he taught me to always look at things from another angle. I truly admire his tireless dedication to the lab, and I feel fortunate to have been able to work with him. Without Antonio's feedback at all hours of the day and night, this thesis would not have been possible.

Thanks to all the members of the Rangel, Camerer and Adolphs labs who have not only provided mentorship and guidance, but also become my friends. I owe particular thanks to Cendri Hutcherson and Todd Hare, who helped me learn numerous fMRI techniques and never tired of my questions.

My time at Caltech was made enjoyable in large part thanks to many friends who became an important part of my life. In particular, I would like to thank Anush for putting up with me through both good and bad days and being a great partner in many adventures.

Finally, I would like to thank my parents and brother Radek for their love and support. Since my childhood, they have always encouraged my educational, intellectual and geographical pursuits. Most importantly, I would like to thank my Mom, who instilled in me a great deal of drive and determination and who has always encouraged me to pursue my passions.

ABSTRACT

Neuroeconomics has produced a number of insights into economics, psychology, and neuroscience in its relatively short existence. Here, I show how neuroeconomics can inform these fields through three studies in social decision making and decision making under risk. Specifically, I focus on computational biases inherent in our daily decisions.

First, using functional magnetic resonance imaging (fMRI), I examine how we make decisions for others compared to ourselves. I find that overlapping areas of the ventromedial prefrontal cortex (vmPFC) are involved in both types of decisions, though decisions for others are modulated by areas involved in social cognition. Specifically, activity in the inferior parietal lobule (IPL) encodes a variable measuring the distance between others' and our own preferences, suggesting that we may anchor our choices for others on our own preferences and attempt to modulate these preferences with what we know about others.

Second, I investigate how visual looking patterns can critically influence the computation and comparison of values. In a first study using eye-tracking, I investigate the relationship between loss aversion and attention and find a correlation between how loss averse subjects are and how long they look at losses vs. gains when evaluating mixed gambles. Importantly, I show that this effect is not due to subjects simply looking longer at items of higher value. In a second study using Mouselab, I show how attention influences multi-

attribute choice. I find that the display of different attributes has a significant effect on search among those attributes and, ultimately, choice.

TABLE OF CONTENTS

SUMMARY

In recent years, the field of neuroeconomics has shed insight on a number of traditional economics questions. While neuroscience has benefited from the collaboration with economics, there is still disagreement about the degree to which neuroscience can inform economics. In this thesis, I conduct three experiments to show how neuroeconomics can help address questions that remain unanswered with traditional economic methods.

In Chapter 1 I use fMRI to show how we make decisions for others compared to ourselves. Every day we make decisions on behalf of others, whether it is for a friend, a patient, a client, a coworker or a constituent. I investigate the neurobiological and computational basis of empathic choice using a human fMRI task in which subjects purchase items for themselves with their own money, or for others with the other's money. I find that empathic choices on behalf of others engage the same regions of the brain, the ventromedial prefrontal cortex (vmPFC), that are known to compute stimulus values when making choices for ourselves. During empathic choice, these value signals are modulated by activity in another brain region, the inferior parietal lobule (IPL), which has previously been found to play a role in social processes. These findings extend our understanding of social cognition and have broader implications for psychology and economics. The results suggest that the ability to make sound empathic decisions might depend on the ability to compute value signals in vmPFC that give sufficient weight to the differences between

others and ourselves. It follows that deficits in empathy and general social cognition might impair the ability to make sound empathic decisions, which could interfere with everyday social interaction.

In Chapter 2 I use eye-tracking to investigate the role of attention in loss aversion. While there is much evidence showing that individuals exhibit loss aversion in many domains, the mechanisms behind loss aversion remain unknown. Based on recent research showing that differences in attention affect the computation and comparison of values during simple choice, I hypothesize that differences in loss aversion across and within individuals could be modulated by differential attention to losses. I find that more loss averse subjects pay more relative attention to losses, and my results suggest a model in which attention must be included to make any inferences on choice. My results have important implications for the role of attentional processes in choice: systematic biases in fixations could lead to different choices. Moreover, as biases such as loss aversion can lead to deficits in decision-making, our findings raise the interesting possibility that we may be able to modulate our attention to make better choices.

In Chapter 3 I further examine the role of attention, and the computational biases it introduces, in choice. Using a two-item, two-attribute choice task, I examine how consumers might quickly parse and select among a variety of options. I find striking differences in search patterns depending on how attributes of each choice are visually arranged. These search patterns subsequently impact choice through the differential weighting and integration of the attributes. Higher-placed attributes for a given item will receive more attention and thus greater weighting in an overall value for the item. My results have important implications for a number of applications, including store display arrangement, product attribute emphasis and the organization of features on a website.

These three experiments show how neuroeconomics can help address questions that remain unanswered with traditional economics methods. These findings extend our understanding of social cognition and further our knowledge of how we assess potential outcomes in risky and complex choice situations. These results shed light on the role that computational biases play in decision-making and how we might be able to modulate such biases to improve our choices.

CHAPTER 1

Empathic Choice Involves vmPFC Value Signals that are Modulated by Social Processing Implemented in IPL

Abstract. Empathic decision-making involves making choices on behalf of others in order to maximize their well-being. Examples include the choices that parents make for their children, as well as the decisions of a politician trying to make good choices on behalf of his constituency. We investigated the neurobiological and computational basis of empathic choice using a human fMRI task in which subjects purchased DVDs for themselves with their own money, or DVDs for others with the other's money. We found that empathic choices engage the same regions of ventromedial prefrontal cortex that are known to compute stimulus values, and that these value signals were modulated by activity from a region of inferior parietal lobule (IPL) known to play a critical role in social processes such as empathy. We also found that the stimulus value signals used to make empathic choices were computed using a mixture of self-simulation and other-simulation processes, and that activity in IPL encoded a variable measuring the distance between the other's and own preferences, which provides a hint for how the mixture of self- and other-simulation might be implemented.

INTRODUCTION

Humans make different types of decisions. Self-oriented decisions mostly affect ourselves and are guided by the goal of maximizing our own well-being. Examples include what to have for lunch or which clothing to purchase. Pro-social decisions involve tradeoffs between our own well-being and the well-being of others. Examples include a donation to charity and purchasing a gift for a friend. Empathic decisions entail decisions made on behalf of other people, with the goal of choosing what is best for them, and without having to sacrifice our own resources. Examples include the myriad of choices that parents make for their children, the decisions of a politician trying to make good choices on behalf of his or her constituents, and economic agents (e.g., in real estate or entertainment) who strive to commit their clients' time and money to activities the clients prefer. Although a substantial amount of progress has been made in understanding self-oriented (Rangel, Camerer et al. 2008; Rushworth and Behrens 2008; Kable and Glimcher 2009; Rangel and Hare 2010) and pro-social decisions (Fehr and Camerer 2007), much less is known about the computational and neurobiological basis of empathic choice.

From a psychological and neurobiological perspective, empathic choice is particularly interesting because it is likely to involve the interaction of two different types of processes: those involved in basic decision-making, such as value computation and comparison, and those involved in social processing, such as empathy and mentalizing.

With respect to basic decision-making, a large body of work has begun to characterize in detail the computations involved in self-oriented decisions. For example, human neuroimaging studies have shown that activity in areas such as ventromedial prefrontal cortex (vmPFC) correlates with the value of stimuli at the time of choice (Kable and Glimcher 2007; Plassmann, O'Doherty et al. 2007; Tom, Fox et al. 2007; Valentin, Dickinson et al. 2007; Hare, O'Doherty et al. 2008; Plassmann, O'Doherty et al. 2008; Rolls, McCabe et al. 2008; Boorman, Behrens et al. 2009; FitzGerald, Seymour et al. 2009; Hare, Camerer et al. 2009; Hare, Camerer et al. 2009; Litt, Plassmann et al. 2009). Similar results have been found in non-human primate electrophysiology studies (Wallis and Miller 2003; Padoa-Schioppa and Assad 2006; Padoa-Schioppa and Assad 2008; Kennerley, Dahmubed et al. 2009; Kennerley and Wallis 2009; Padoa-Schioppa 2009). Activity in vmPFC has also been associated with the computation of stimulus values during pro-social choices (Moll, Krueger et al. 2006; Harbaugh, Mayr et al. 2007; Tankersley, Stowe et al. 2007; Hsu, Anen et al. 2008; Hare, Camerer et al. 2010; Tricomi, Rangel et al. 2010). Importantly, however, none of these previous studies include instances of empathic choice.

With respect to social processing, a separate body of work has begun to characterize the computations involved in social cognition. Empathy is normally defined as the ability to appreciate the emotions and feelings of others as separate from those of the self (Decety 2010; Shamay-Tsoory 2011). A significant number of studies, using a wide variety of paradigms, have shown that areas such as the inferior frontal gyrus (IFG), inferior parietal lobule (IPL), and dorsomedial prefrontal cortex (dmPFC) play a critical role in empathy computations (Mitchell 2009; Zaki, Weber et al. 2009; Decety 2010; Shamay-Tsoory 2011). Importantly, the previous literature on empathy has also not covered the case of empathic choice, since the tasks used involved the observation and evaluation of other's emotional states, but not decision-making on their behalf. A related literature has studied the neurobiological basis of mentalizing (often called theory of mind, ToM), and has found that areas such as medial prefrontal cortex (mPFC) and the temporo-parietal junction (TPJ) play a critical role in this process (Saxe and Kanswisher 2003; Mitchell, Banaji et al. 2005; Saxe and Wexler 2005; Mitchell, Macrae et al. 2006; Saxe 2006; Mitchell 2009).

Here we present the results of a simple human functional magnetic resonance imaging study (fMRI) in which subjects made otherwise identical decisions (purchasing DVDs) in either a self-oriented context, by buying them for themselves with their own funds, or in an empathic context, by buying them for an unknown third party, with this party's funds. This allowed us to investigate two basic questions regarding empathic decision-making.

First, is the same basic neural circuitry involved in making self-oriented and empathic decisions? And, if not, what are the critical differences? Based on the decision and social neuroscience literatures discussed above, we hypothesized that empathic decisions engage the basic elements of the decision making system, such as the computation of stimulus value signals in vmPFC, but that their computation during empathic choice requires the activation of regions, such as IPL and TPJ, that are known to play a critical role in empathy and mentalizing.

Second, what are the computational properties of the stimulus values used to make empathic choices? In particular, we were interested in disentangling the extent to which subjects computed the empathic stimulus value signals using self-simulation, other-simulation, or other-learning. Under self-simulation, subjects infer the other's DVD values by computing their own value for them. Under other-simulation, subjects use some model of the other individual to infer his value for the DVDs but make no use of their own preferences for them. Under other-learning, subjects learn to compute the other's DVD values by repeatedly observing their behavior. Conceptually there is an important difference between the last two approaches: other-simulation requires forming a social model of the other person (e.g., gender, nationality, age, etc.), whereas under other-learning, the other's preferences are learned simply by repeated observation and extrapolation. Thus, the other-simulation approach makes heavy use of social models and information, whereas other-learning involves much more basic forms of learning.

METHODS

Subjects. 32 normal-weight, American or Canadian, male subjects participated in the experiment (age: mean = 22.8, SD = 3.9). All subjects were right-handed, healthy, had normal or corrected-to-normal vision, had no history of neurological or metabolic illnesses, and were not taking any medication that interferes with the performance of fMRI. All subjects were informed about the experiment and gave written consent before participating.

Stimuli. Subjects viewed 100 high-resolution color images of DVD covers of popular films from the last 15 years. They included comedies (e.g., "Austin Powers"), action films (e.g., "Swordfish"), dramas (e.g., "Magnolia") and thrillers (e.g., "Panic Room").

Task. There were two types of subjects in the experiment: one passive subject and thirty-two active subjects. The role of the passive subject was to be the recipient of the active subjects' decisions.

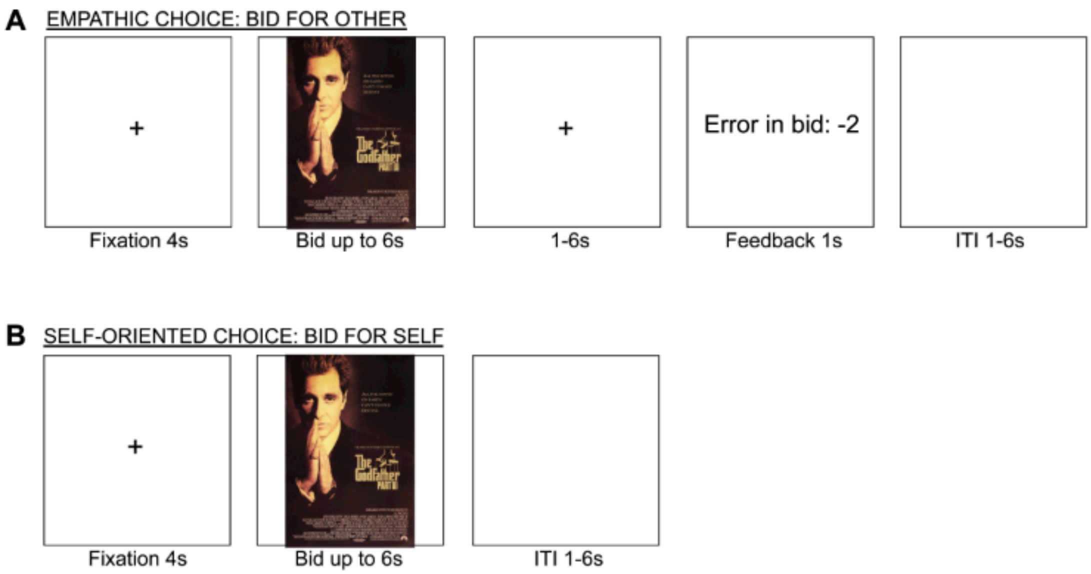
Active subjects made decisions inside the scanner in two types of trials performed on different days (average lag = 90 days). On the first visit they participated in an empathic choice task in which they made purchase decisions on behalf of the passive subject (Fig. 1A). They were given a budget of \$10 that belonged to the passive subject (any unspent funds were returned to him) and were given a summary sheet containing a photograph and some biographic information about the passive subject (see Appendix for detailed instructions). They were then shown images of 100 different DVDs and had to make a decision regarding how much to bid for each one of them on behalf of the subject. Bids were made using a six-point scale of \$0, 2, 4, 6, 8 and 10. After every bid subjects received feedback equal to the amount by which they had overbid or underbid relative to the passive subject's values (feedback = active subject's bid – passive subject's bid). Active subjects did not receive any form of compensation for making accurate bids. Instead, the instructions simply told them to try to maximize the passive subject's well-being. The mapping of bids to response buttons was counterbalanced across subjects.

At the conclusion of the experiment, one of the 100 trials was randomly selected and implemented using a Becker-DeGroot-Marschak (BDM) auction. The rules of the auction are as follows. Let b denote the bid made by the subject for a particular item. After the bid is made, a random number n is drawn from a known distribution (in our case, each integer dollar value from \$0 to \$10 was chosen with equal probability). If $b >= n$, the subject received the DVD and paid a price equal to n. Otherwise, if $b < n$, the subject did not get the DVD and did not pay anything. The optimal strategy in this type of auction is for the buyer to bid exactly how much he is willing to pay for the item being sold (Becker, DeGroot et al. 1964). The active subjects knew that the outcome of the auction would be implemented, and that the person for whom they were bidding would receive any DVD purchased plus any remaining cash from the \$10. Note that since only one trial was selected to count, the subjects did not have to worry about spreading the \$10 dollars across the different films, and could treat every decision as if it were the only one. No deception was used in the experiment. The passive subject actually received DVDs when the subject's decision led to a purchase of the DVD.

During the second day of scanning, active subjects participated in the self-oriented version of the task (Fig. 1B). In this case they performed a similar task, except that now they made purchase decisions for themselves out of a \$10 cash endowment that belonged to them. A randomly selected trial was again chosen, and the associated decision implemented, at the conclusion of the two sessions. At the end of the second session, subjects were asked to fill out a questionnaire detailing which DVDs they owned or had seen. In order to control for any potential order effects on bidding, the DVDs were shown in the same order as in the first experimental visit.

The passive subject played only the bid-for-self task outside the scanner. His responses were used to compute the feedback signals for the active subjects.

Fig. 1. Behavioral task for (A) empathic choice trials and (B) self-oriented choice trials.



About task order. Given the difficulty in guessing another's film preferences, we were concerned that subjects would exhibit an artificial tendency to use their own preferences to make the purchase decisions for the other. In order to minimize this concern, we decided not to counterbalance the order of the two tasks, and to introduce a long multi-month lag between them. The results described below suggest that we were successful in avoiding a full self-valuation bias during the empathic decisions. However, this raises the natural concern of order confounds. To address this concern we carried out a companion behavioral experiment (see SOMs for details) in which we directly compared the effect of order on bidding behavior. For each individual we carried out a linear regression of bid-for-other on bid-for-self and other-bid, separately for self-oriented and empathic choice trials. We found no significant differences across the two order conditions (min p=0.29, t-test), which rules out the order confound.

fMRI data acquisition and preprocessing. The fMRI data were acquired in a 3.0 Tesla Trio MRI Scanner (Siemens). We acquired gradient echo T2-weighted echoplanar (EPI) images with a BOLD contrast in an oblique orientation of 30 degrees to the anterior commissure-posterior commissure line. We also used an eight-channel phased array head coil. Each volume of images had 48 axial slices of 3 mm thickness and 3 mm in-plane resolution with a TR of 3 s. The imaging data were acquired in four separate sessions; the first two, in which subjects bid on behalf of the passive subject, lasted approximately 13 minutes each. The latter two, in which subjects bid for themselves, lasted approximately 9 minutes each. The first two sessions were performed on a separate date than the latter two sessions. Whole-brain high-resolution T1-weighted structural scans (1 x 1 x 1 mm) were acquired for each subject and coregistered to their mean functional EPI images. The structural scans were averaged across subjects to permit anatomical localization of the functional activations at the group level.

Image analysis was performed using Statistical Parametric Mapping software (SPM5; Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, UK). We preprocessed the data in the following way. First, slice-timing correction centered at the middle T2 scan was applied, followed by realignment to the first volume. We then applied spatial normalization to the standard Montreal Neurological Institute (MNI) EPI template with a resampled voxel size of 3 mm^2 and performed spatial smoothing using a Gaussian kernel with full width at half maximum of 8 mm. Intensity normalization and high-pass temporal filtering were also applied to the data.

fMRI data analysis. We estimated several models of the BOLD responses to test the various hypotheses.

GLM 1. This general linear model was designed to identify the similarities and differences between empathic and self-oriented choices. It was estimated in three different steps.

First, we estimated a GLM with AR(1) for each individual subject. The model contained the following regressors: R1) indicator function (equal to 1 when the event occurs and 0 otherwise) for bid-for-other screen; R2) indicator function for bid-for-other screen modulated by bid-for-other; R3) indicator function for bid-for-other screen modulated by the absolute value of the difference between bid-for-self and bid-for-other; R4) indicator function for feedback screen; R5) indicator function for feedback screen modulated by the negative absolute value of the feedback error; R6) indicator function for bid-for-self screen; R7) indicator function for bid-for-self screen modulated by bid-for-self. All regressors were modeled as box-car functions with a duration equal to the subject's reaction time for that trial, except for regressors 4 and 5, which had a duration of 1 second. These regressors were convolved with a canonical hemodynamic response. The model also included motion parameters and session constants as regressors of no interest. Trials with missed responses were not modeled. Second, we computed contrast statistics for all of the tests of interest for each individual subject. Third, we estimated second-level test statistics by computing one-sample t-tests on the single subject contrast coefficients for each contrast of interest.

For inference purposes, all results are reported at $p < 0.05$ whole brain corrected at the cluster level (using the corrected cluster size threshold algorithm by Thomas Nichols; http://www.sph.umich.edu/~nichols/JohnsGems5.html). The only exception is activity in the vmPFC for which, due to the strong prior hypotheses, we report activity at $p<0.05$ small volume cluster corrected (using an anatomical mask of the vmPFC area that included both sides of the medial orbitofrontal cortex and the rectus gyrus).

GLM 2. This model was very similar to GLM 1, except that activity at decision during empathic choices was modulated by two variables: bid-for-self and bid-for-other orthogonalized with respect to bid-for-self. All omitted details are as in GLM 1.

GLM 3. This model was very similar to GLM 1, except that activity at decision during empathic choices was modulated by two variables: bid-for-other and bid-for-self orthogonalized with respect to bid-for-other. All omitted details are as in GLM 1.

GLM 4. This model was very similar to GLM 1, except that activity at decision during empathic choices was modulated by two variables: bid-for-self and a difference signal (given by bid-for-other MINUS bid-for-self). All omitted details are as in GLM 1.

Psychophysiological interactions model (PPI). The goal of this analysis was to identify areas exhibiting differential connectivity with vmPFC during empathic and self-oriented choices. The model was estimated in the following steps.

First, we extracted individual average time-series of BOLD activity within an individually defined region of vmPFC, given by a 4mm sphere surrounding each individual's peak activation for the contrast 'R2 MINUS baseline' in GLM-1 within the anatomical mask of the vmPFC shown in Fig. 1C. We removed any variance from this time series associated with the motion regressors. The resulting time courses were deconvolved using standard procedures (Gitelman, Penny et al. 2003).

Second, we estimated a whole-brain GLM of BOLD responses with AR(1) and the following regressors: R1) interaction between the vmPFC deconvolved time series and an indicator function for bid-for-other screen; R2) interaction between the vmPFC deconvolved time series and an indicator function for bid-for-self screen; R3) indicator function for bid-for-other screen; R4) indicator function for bid-for-self screen; R5) the vmPFC deconvolved time series. These regressors were convolved with a canonical hemodynamic response. The model also included motion parameters and session constants as regressors of no interest. Note that Regressor 1 identifies areas exhibiting task-related functional connectivity with the vmPFC seed region during empathic choices. Regressor 2 does the same for self-oriented choices.

Third, we calculated the following single subject contrasts: C1) Regressor 1 vs. baseline, C2) Regressor 2 vs. baseline, and C3) Regressor 1 vs. Regressor 2.

Fourth, we conducted a second level analysis by calculating a one-sample t-test on the single subject contrast coefficients.

RESULTS

First we discuss tests designed to investigate if the same basic neural circuitry is involved in making self-oriented and empathic decisions, and to characterize the key differences.

Longer RTs in empathic choice. Mean reaction times when bidding for self were about 500 ms faster than when bidding for other (self: mean = 2.16 s, SD = 0.52; other: mean = 2.67

s, SD = 0.47; paired *t*-test $p < 0.05$). This is consistent with the hypothesis that empathic decisions involve the deployment of extra processes.

Common value coding in vmPFC. We hypothesized that a common area of vmPFC is involved in computing the stimulus values (SVs) assigned to DVDs at the time of decision in both the self-oriented and empathic trials. We focused our attention on vmPFC because a large number of studies have found SV signals in this area (see introduction). The bids-for-self provide a trial-by-trial measure of the SVs computed in self-oriented trials, whereas the bids-for-other provide a similar measure for empathic decisions.

We tested this hypothesis by estimating a general linear model of BOLD responses (GLM 1) that looked for correlations between the magnitude of the bids placed in each condition and BOLD activity (see Methods for details). Activity in vmPFC correlated with the bids-for-other during empathic choices (Fig. 2A, see Table 1 for a complete list of activations). Activity in the same area of vmPFC also correlated with bids-for-self during self-oriented choices (Fig. 2B, see Table 2 for a complete list of activations). A conjunction analysis showed that activity in a common area of vmPFC correlated with SVs in both conditions (Fig. 2C), as did activity in areas of the precuneus, middle frontal gyrus, and inferior parietal lobule (Fig. S3).

We also looked for differences in the strength of SV coding across the empathic and self-oriented conditions. We carried out this test in two ways. First, using a whole brain analysis and our omnibus threshold, we did not find any regions that exhibited stronger responsivity to bid-for-self during self-oriented choice than to bid-for-other during empathic choices at our omnibus threshold. Second, we carried an unbiased region-of-interest (ROI) analysis in the area of vmPFC that correlates with SVs in both conditions. A comparison of the average beta values within the ROI for the bid-for-self and bid-for-other regressors revealed no significant differences ($p=0.26$, paired t-test).

Fig. 2. A) Activity in vmPFC correlated with the bids-for-other during empathic choices ($p < 0.05$, SVC). B) Activity in vmPFC also correlated with the bids-for-self during self-oriented choices ($p < 0.05$, SVC). C) Conjunction analysis: activity in a common area of vmPFC correlated with the bids placed in both empathic and self-oriented choice trials.

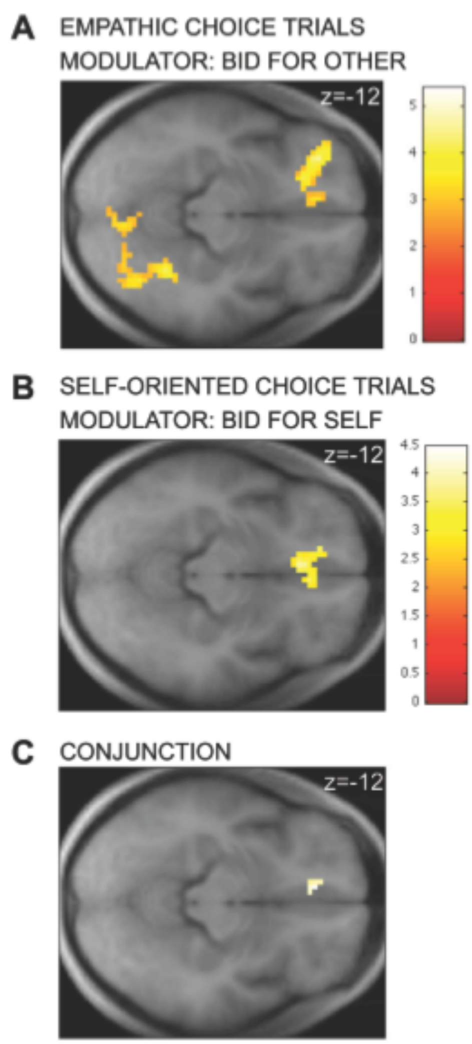


Table 1. Areas exhibiting a positive correlation with bid-for-other during empathic choice (GLM 1)

Region	**Side**	**k**	**T**	**MNI coordinates** x	y	z
Ventral striatum	L/R	153	5.39	-9	6	-6
Middle frontal gyrus	L	248	4.85	-27	33	-15
Precuneus / Inferior parietal lobule	L	255	4.78	-39	-57	42
Fusiform / Middle occipital gyrus	R	632	4.52	30	-66	0
Posterior cingulate	L	240	4.50	-6	-42	15
vmPFC*	L	21	3.57	-9	42	-15

Height threshold: $T = 2.74$, $p < 0.05$, whole brain cluster corrected.
Extent threshold: $k = 109$ voxels, $p < 0.005$.
*Survives small volume correction at $p < 0.05$.

Table 2. Areas exhibiting a positive correlation with bid-for-self during self-oriented choice (GLM 1)

Region	Side	k	T	MNI coordinates x	y	z
Inferior parietal lobule	L	295	4.47	-45	-36	39
Middle frontal gyrus	L	617	4.43	-39	36	12
Precuneus	L	135	4.07	-39	-72	30
vmPFC*	L/R	105	3.79	-6	27	-12

Height threshold: T = 2.74, p < 0.05, whole brain cluster corrected.
Extent threshold: k = 105 voxels, p < 0.005.
*Survives small volume correction at p < 0.05.

Differences in the network involved in empathic vs. self-oriented choices. We also hypothesized that empathic choice would require the activation of additional regions, such as IPL and TPJ, that are known to be involved in social cognition. We tested this hypothesis in two steps.

First, using GLM 1, we looked for regions that exhibit higher average activity during empathic choices, and areas that exhibit higher average activity during self-oriented choices. A large cluster of regions exhibited stronger activity during empathic choices, including bilateral inferior parietal lobule, bilateral middle frontal gyri, bilateral anterior insula (Fig. S4A, Table 3). We also found regions exhibiting stronger activity during self-oriented choices, including bilateral supramarginal gyri, middle temporal gyrus, right posterior insula and superior temporal gyrus (Fig. S4B, Table 3).

Second, we looked for differences in functional connectivity with the vmPFC valuation area between the empathic and self-oriented trials. We did this by estimating a psychophysiological interactions model that looks for areas that exhibit increases in functional connectivity at the time of decision separately in self-oriented and empathic trials. The model uses as a seed the area of vmPFC involved in SV coding in both conditions (see Methods for details). We found that activity in bilateral IPL exhibited stronger functional connectivity with vmPFC during empathic choices (Table 4, Fig. 3A).

In contrast, no regions exhibited stronger functional connectivity with vmPFC during self-oriented choices at our omnibus threshold. Interestingly, the regions of IPL that exhibit stronger functional connectivity with vmPFC overlap with those that exhibit stronger average activity during empathic trials (Fig. 3B).

Together, these results provide supporting evidence for the hypothesis that empathic choice engages the basic vmPFC valuation system, just as it does in self-oriented choice, but that the computation of these value signals in empathic choice involves the activation of regions of IPL that are known to play a critical role in social cognition.

Fig. 3. A) Areas of IPL exhibiting stronger connectivity with the vmPFC valuation region during empathic choices than during self-oriented decisions. B) Region of IPL exhibiting both stronger functional connectivity with vmPFC and higher average (unmodulated) activity during empathic choices. C) Region of IPL exhibiting both stronger functional connectivity with vmPFC during empathic choices and a correlation with the difference preference measure. The contrasts are thresholded at p<.05, WBC.

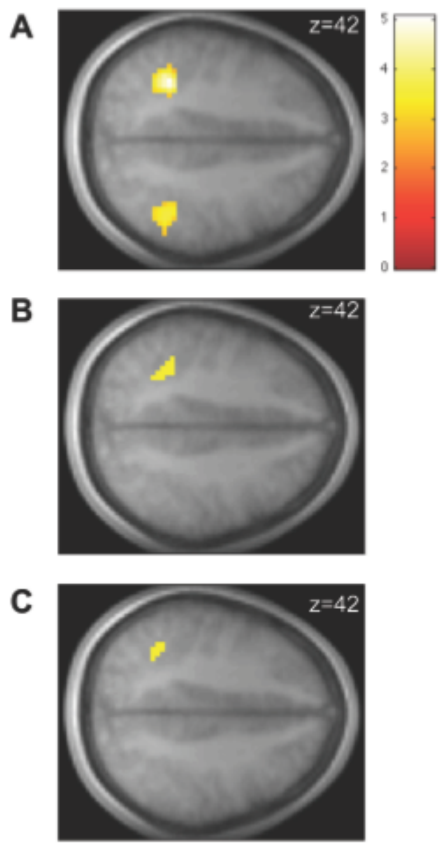


Table 3. Regions exhibiting stronger average (unmodulated) activation in self-oriented vs. empathic choice (GLM 1)

Region	Side	k	T	MNI coordinates x	y	z
Self-oriented > Empathic						
Inf parietal/Supramarginal gyrus	L	471	12.6	-51	-54	36
Inf parietal/Supramarginal gyrus	R	409	7.06	51	-57	42
Middle temporal gyrus	L	149	5.71	-63	-33	-9
Middle temporal gyrus	R	165	5.44	57	-27	-24
Cingulate gyrus	L	173	5.16	-9	-18	27
Middle frontal gyrus	R	218	4.71	39	12	57
Insula/Superior temporal gyrus	R	167	4.38	48	-6	0
Empathic > Self-oriented						
Middle occipital gyrus/cuneus	L	11460	-8.90	-24	-90	3
Putamen/caudate/thalamus	L	*	-8.63	-6	9	-3
Middle occipital gyrus/cuneus	R	*	-7.94	3	-93	9
Putamen/caudate/thalamus	R	*	-7.20	18	-27	3
Inf parietal lobe/Postcentral gyrus	R	*	-7.07	36	-27	69
Precentral/middle frontal gyrus	L	*	-7.04	-18	-75	51
Fusiform/middle temporal gyrus	L	*	-6.51	-39	-63	-18
Insula/Inf frontal gyrus	L	*	-6.23	-36	33	6
Midbrain	L	*	-6.16	-18	-24	-3
Precentral/middle frontal gyrus	R	*	-5.45	15	-72	54
Midbrain	R	*	-5.36	18	-24	3
Inf parietal lobe	L	*	-5.13	-27	-30	75
Insula/Inf frontal gyrus	R	180	-5.04	45	21	3

Height threshold: $T = 2.74$, $p < 0.05$, whole brain cluster corrected.
Extent threshold: $k = 112$ voxels, $p < 0.005$.
*Part of a larger cluster.

Table 4. Areas exhibiting positive task related functional connectivity with the vmPFC (PPI analysis)

Region	Side	k	T	MNI coordinates x y z
Self-oriented				
Inferior frontal gyrus	R	164	4.1	45 9 6
Supramarginal/sup temporal gyrus	R	169	4.3	66 -24 36
Inferior parietal lobe	L	142	3.9	-30 -57 60
Inferior parietal lobe	R	134	4.7	45 -39 66
Empathic				
Middle frontal gyrus	R	2383	6.0	45 45 15
Insula/Inferior frontal gyrus	L	354	4.5	-36 18 -3
Middle frontal gyrus	L	493	5.3	-39 33 39
Inferior parietal lobe	L	2727	7.1	-42 -48 42
Inferior parietal lobe	R	*	6.9	48 -45 54
Empathic > Self-oriented				
Inferior parietal lobe	L	145	5.1	-36 -45 42
Inferior parietal lobe	R	148	3.9	48 -48 57

Height threshold: T = 2.58, p < 0.05, whole brain cluster corrected.
Extent threshold: k = 102 voxels, p < 0.005.
*Part of a larger cluster.

Next, we investigated the extent to which stimulus value signals are computed using self-simulation, other-simulation, or other-learning, during empathic choices.

No behavioral evidence for other-learning. Under other-learning, the quality of bids-for-other should improve over time. A good measure of the quality of the individual's bids-for-other is given by

correlation(bid-for-other, other-bid) – correlation(bid-for-self, other-bid) (1).

The first term measures the extent to which the subject's bids-for-other correlates with the other's preferences. The second term corrects for the fact that the first term might be artificially large if both individuals tend to like the same movies. The mean quality statistic was 0.06 (S.E.= 0.017, $p<.0001$, t-test). Contrary to the other-learning model, we found no significant difference between the first and second half of trials ($p=0.72$, pairwise t-test), which provides evidence against other-learning.

Behavioral bids are consistent with a mixture of self- and other-simulation. A comparison of the differences between the bids that the subjects made for themselves (during self-oriented choice) and those that they made for the other (during empathic choice) provides a behavioral test of the extent to which the stimulus values were consistent with the self- versus the other-simulation models. The self-simulation model predicts a very high correlation between the bids-for-self and the bids-for-other. In contrast, the other-simulation model predicts a much lower correlation between the two types of bids.

One critical difficulty in carrying out this test is that, regardless of how the bids are computed, they may be correlated because individual preferences are not independent (for example, no one seems to like certain movies). This problem can be circumvented through the following two steps.

First, we estimated a mixed effects linear regression of bid-for-other on two regressors: other-bid and bid-for-self. Importantly, the bid-for-self regressor was orthogonalized with respect to the other-bid. This is important because, then, any variation on bid-for-other that is explained by the bid-for-self regressor cannot be attributed to common preferences. As a result, the relative magnitude of the bid-for-self regressor provides a lower bound on the contribution of self-simulation processes. Both coefficients were statistically significant and of approximately equal magnitude (other-bid: mean=.52, SE=.02, $p<.0001$; bid-for-self: mean=0.55, SE=0.03, $p<.0001$; t-tests).

Second, we estimated a related regression in which the independent variable was still bid-for-other, but the right-hand-side regressors were bid-for-self and other-bid orthogonalized with respect to bid-for-self. This alternative orthogonalization is useful because now the relative magnitude of the other-bid regression coefficient provides a lower bound on the contribution of other-simulation processes. Both coefficients were again statistically significant (other-bid: mean=.24, SE=.018, $p<.0001$; bid-for-self: mean=0.81, SE=0.03, $p<.0001$; t-tests).

Together with the previous result, the two regressions suggest that subjects computed stimulus values during empathic trials using a mixture of self-simulation and other-simulation processes. The relative magnitude of the regressors also suggests that the self-simulation component played a stronger role in our task.

Activity in vmPFC is also consistent with a mixture of self- and other-simulation. We also investigated the extent to which the stimulus value signals computed during empathic choices were consistent with self- or other-simulation. We did this by estimating two new GLMs of BOLD responses. The key difference with the previous models is that activity during empathic choices was now modulated by two variables: bid-for-self and bid-for-other. Importantly, to deal with the problem of preference correlation discussed above, in

GLM 2 the bid-for-other was orthogonalized with respect to the bid-for-self, and in GLM 3 the opposite orthogonalization was carried out.

We computed the average regression coefficients for bid-for-self and bid-for-other in both models within the vmPFC region that correlates with stimulus values in both empathic and self-oriented choice. We found that all regressors were significantly positive ($p<.0001$ in all cases, t-test). For completeness, we carried out similar ROI tests in all of the areas that correlated with stimulus values in either empathic or self-oriented choices and found similar results.

These results provide further neurobiological evidence that stimulus values during empathic choice are computed using a mixture of the self- and other-simulation processes.

We also carried out an additional post-hoc analysis designed to explore the computational role that IPL might play in empathic choice. Based on the results described above, as well as the literature discussed in the introduction, we speculated that IPL might contribute to the computation of stimulus values by measuring the extent to which the other's preferences differ from the subject's own preferences. In our task, this signal can be measured by

$$difference = bid\text{-}for\text{-}other - bid\text{-}for\text{-}self \quad (2).$$

This signal is computationally useful because it would allow subjects to compute their estimate of the value that the other places on the DVDs by computing their own value for it, and then carrying out the additive (and signed) adjustment given by the difference signal.

To test this hypothesis we estimated a new GLM 4 in which activity during empathic choices is modulated by bid-for-self and the difference signal. Consistent with our post-hoc hypothesis, activity in inferior parietal lobule and middle frontal gyrus was significantly correlated with the difference regressor (Table 5). Interestingly, the area of IPL identified in this model overlaps with those exhibiting increased functional connectivity with the vmPFC valuation areas during empathic choices (Fig. 3C).

Table 5. Areas exhibiting a positive correlation with the difference signal during empathic choice (GLM 4)

Region	**Side**	**k**	**T**	**MNI coordinates** x y z
Inferior parietal lobe/precuneus	L	242	5.22	-39 -54 42
Middle frontal gyrus	L	121	4.47	-39 45 -6

Height threshold: T = 2.74, p < 0.05, whole brain cluster corrected.
Extent threshold: k = 112 voxels, p < 0.005.

DISCUSSION

The results presented here provide the following insights about the computational and neurobiological basis of empathic choice. First, empathic choices engage the vmPFC valuation system used in self-oriented decisions, and these value signals seem to be modulated by activity in regions of IPL known to play a critical role in social processes such as empathy. Second, the stimulus values used to make empathic choices are computed using a mixture of self-simulation and other-simulation. Third, during empathic choices, activity in the IPL encodes a variable measuring the distance between the other's and own preferences. This variable could be used to compute the value of DVDs for other starting from the one's own value, which provides a hint for how the mixture of self- and other-simulation are implemented.

The results have implications for various areas of the neural and social sciences. The results extend our growing knowledge of how the brain makes decisions to the case of empathic choice, which had not been studied before. The results show that empathic decisions involve the combination of two types of processes: the basic valuation circuitry involved in self-oriented decisions and social processes such as empathy. In particular, in contrast to the case of self-oriented choice, in empathic choice, stimulus values in vmPFC seem to be modulated by a signal from IPL that reflects the difference in preferences between self and other. This result parallels a recent finding in an fMRI study of charitable donations (Hare, Camerer et al. 2010), which found that the value signals in vmPFC were modulated by an area of posterior superior temporal cortex commonly associated with mentalizing.

The results also extend our understanding of social cognition in several ways. First, they show that the same set of areas that have been shown to play a role in ‘passive’ empathy tasks are also at work during empathic choices. Second, they advance our understanding of the precise computations carried out by IPL (i.e., a measure of the difference between the other’s and the self’s preferences) as well as how they affect decision-making (i.e., by modulating activity in the vmPFC valuation circuitry). Third, they advance our understanding of the role of mPFC in social cognition, which has been previously implicated in person-related and not object-related knowledge (Mitchell, Heatherton et al. 2002; Mitchell, Macrae et al. 2006; Mitchell 2009). Our results show that, during empathic choice, mPFC is involved in the computation of stimulus values. Importantly, the area of mPFC characterized here is significantly more ventral than those identified in previous studies, none of which involved actual empathic choices. Fourth, the statistical influence of own-bids on bids-for-others replicates the “false consensus effect” from social psychology (Ross, 77, Marks and Miller 87). The evidence that IPL activity correlates with the difference in the two bids suggests a candidate region for explaining differences in the strength of false consensus across people or context.

The results also have potential practical applications in psychology and economics. They suggest that the ability to make sound empathic decisions might depend on the ability to compute value signals in vmPFC that give sufficient weight to the differences between others and ourselves. It follows that deficits in empathy and general social cognition might impair the ability to make sound empathic decisions, which could interfere with everyday social interaction. Additional evidence for the role of vmPFC in these types of processes comes from lesion studies (Krajbich, Adolphs et al. 2009).

We were surprised to find no other-learning process at work during empathic choice. It is possible that this is due to specific features of the current task that might not generalize to other settings. In particular, the DVD stimuli used here are highly multi-dimensional and complex, which makes it hard to generalize across very different types of films. Thus, we cannot rule out the possibility that other-learning processes might be at work in settings with a simpler stimulus set.

REFERENCES

Becker, Gordon, Morris H. DeGroot and Jacob Marschak. (1964). "Measuring utility by a single-response sequential method." *Behavioral Science* **9**: 226-232.

Boorman, Erie. D., Timothy E. Behrens, Mark W. Woolrich, et al. (2009). "How green is the grass on the other side? Frontopolar cortex and the evidence in favor of alternative courses of action." *Neuron* **62**(5): 733-743.

Decety, Jean. (2010). "To What Extent is the Experience of Empathy Mediated by Shared Neural Circuits?" *Emotion Review* **2**(3): 204-207.

Fehr, Ernst and Colin F. Camerer. (2007). "Social neuroeconomics: the neural circuitry of social preferences." *Trends Cogn Sci* **11**(10): 419-427.

FitzGerald, Thomas H. B., Ben Seymour and Raymond J. Dolan. (2009). "The role of human orbitofrontal cortex in value comparison for incommensurable objects." *J Neurosci* **29**(26): 8388-8395.

Gitelman, Darren R., William D. Penny, John Ashburner and Karl J. Friston. (2003). "Modeling regional and psychophysiologic interactions in fMRI: the importance of hemodynamic deconvolution." *Neuroimage* **19**(1): 200-207.

Harbaugh, William T., Ulrich Mayr and Daniel R. Burghart. (2007). "Neural responses to taxation and voluntary giving reveal motives for charitable donations." *Science* **316**(5831): 1622-1625.

Hare, Todd A., Colin F. Camerer, Dan T. Knoeple, et al. (2010). "Value computations in VMPFC during charitable decision-making incorporate input from regions involved in social cognition." *Journal of Neuroscience* 30: 583-590.

Hare, Todd A., Colin F. Camerer and Antonio Rangel. (2009). "Self-control in decision-making involves modulation of the vMPFC valuation system." *Science* **324**: 646-648.

Hare, Todd A., John O'Doherty, Colin F. Camerer, et al. (2008). "Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors." *J Neurosci* **28**(22): 5623-5630.

Hsu, Ming, Cedric Anen and Steven R. Quartz. (2008). "The right and the good: distributive justice and neural encoding of equity and efficiency." *Science* **320**(5879): 1092-1095.

Kable, Joseph W. and Paul W. Glimcher. (2007). "The neural correlates of subjective value during intertemporal choice." *Nat Neurosci* **10**(12): 1625-1633.

Kable, Joseph W. and Paul W. Glimcher. (2009). "The neurobiology of decision: consensus and controversy." *Neuron* **63**(6): 733-745.

Kennerley, Steven W., Aspandiar F. Dahmubed, Antonio H. Lara and Jonathan D. Wallis. (2009). "Neurons in the frontal lobe encode the value of multiple decision variables." *J Cogn Neurosci* **21**(6): 1162-1178.

Kennerley, Steven W. and Jonathan D. Wallis. (2009). "Evaluating choices by single neurons in the frontal lobe: outcome value encoded across multiple decision variables." *Eur J Neurosci* **29**(10): 2061-2073.

Krajbich, Ian, Ralph Adolphs, Daniel Tranel, et al. (2009). "Economic games quantify diminished sense of guilt in patients with damage to the prefrontal cortex." *J Neurosci* **29**(7): 2188-2192.

Litt, Ab, Hilke Plassmann, Baba Shiv and Antonio Rangel. (2011). "Dissociating valuation and saliency signals during decision making." *Cereb Cortex* **21:** 95-102.

Marks, Gary and Norman Miller. (1987). "Ten years of research on the false-consensus effect: An empirical and theoretical review." *Psyc Bull* **102**: 72-90.

Mitchell, Jason P. (2009). "Inferences about mental states." *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **364**(1521): 1309-1316.

Mitchell, Jason P., Mahzarin R. Banaji and C. Neil Macrae. (2005). "The link between social cognition and self-referential thought in the medial prefrontal cortex." *Journal of Cognitive Neuroscience* **17**(8): 1306-1315.

Mitchell, Jason P., Todd F. Heatherton and C. Neil Macrae. (2002). "Distinct neural systems subserve person and object knowledge." *Proc Natl Acad Sci U S A* **99**(23): 15238-15243.

Mitchell, Jason P., C. Neil Macrae and Mahzarin R. Banaji. (2006). "Dissociable medial prefrontal contributions to judgments of similar and dissimilar others." *Neuron* **50**(4): 655-663.

Moll, Jorge, Frank Krueger, Roland Zahn, et al. (2006). "Human fronto-mesolimibic networks guide decisions about charitable donations." *PNAS* **103**(42): 15623-15628.

Padoa-Schioppa, Camillo. (2009). "Range-adapting representation of economic value in the orbitofrontal cortex." *J Neurosci* **29**(44): 14004-14014.

Padoa-Schioppa, Camillo and John A. Assad. (2006). "Neurons in the orbitofrontal cortex encode economic value." *Nature* **441**: 223-226.

Padoa-Schioppa, Camillo and John A. Assad. (2008). "The representation of economic value in the orbitofrontal cortex is invariant for changes of menu." *Nat Neurosci* **11**(1): 95-102.

Plassmann, Hilke, John O'Doherty and Antonio Rangel. (2007). "Orbitofrontal cortex encodes willingness to pay in everyday economic transactions." *J Neurosci* **27**(37): 9984-9988.

Plassmann, Hilke, John O'Doherty and Antonio Rangel. (2010). Aversive goal values are negatively encoded in the medial orbitofrontal cortex at the time of decision making. *Journal of Neuroscience*, 30: 10799-10808.

Rangel, Antonio, Colin F. Camerer and P. Read Montague. (2008). "A framework for studying the neurobiology of value-based decision making." *Nat Rev Neurosci* **9**(7): 545-556.

Rangel, Antonio and Todd A. Hare (2010). "Neural computations associated with goal-directed choice." *Curr Opin Neurobiol* **20**(2): 262-270.

Rolls, Edmund R., Ciara McCabe and Jerome Redoute. (2008). "Expected Value, Reward Outcome, and Temporal Difference Error Representations in a Probabilistic Decision Task." *Cereb Cortex* **18**: 652-663.

Ross, Lee, David Greene and Pamela House. (1977). "The false consensus effect: An egocentric bias in social perception and attribution processes." *J Exp Soc Psy* **13**(3): 279-301.

Rushworth, Matthew F. and Timothy E. Behrens. (2008). "Choice, uncertainty and value in prefrontal and cingulate cortex." *Nat Neurosci* **11**(4): 389-397.

Saxe, Rebecca. (2006). "Uniquely human social cognition." *Curr Opin Neurobiol* **16**(2): 235-239.

Saxe, Rebecca and Nancy Kanswisher. (2003). "People thinking about thinking people: The role of the temporo-parietal junction in “theory of mind"." *Neuroimage* **19**(4): 1835-1842.

Saxe, Rebecca and Anna Wexler (2005). "Making sense of another mind: The role of the right temporo-parietal junction." *Neuropsychologia* **43**(10): 1391-1399.

Shamay-Tsoory, Simone G. (2011). "The neural bases for empathy." *The Neuroscientist : a review journal bringing neurobiology, neurology and psychiatry* **17**(1): 18-24.

Tankersley, Dharol, C. Jill Stowe and Scott A Huettel. (2007). "Altruism is associate with an increased neural response to agency." *Nature Neuroscience* **10**(2): 150-151.

Tom, Sabrina M., Craig R. Fox, Christopher Trepel and Russell A. Poldrack. (2007). "The Neural Basis of Loss Aversion in Decision-Making Under Risk." *Science* **315**: 515-518.

Tricomi, Elizabeth, Antonio Rangel, Colin F. Camerer and John P. O’Doherty. (2010). "Neural evidence for inequality-averse social preferences." *Nature* **463**(7284): 1089-1091.

Valentin, Vivian V., Anthony Dickinson, and John P. O'Doherty. (2007). "Determining the neural substrates of goal-directed learning in the human brain." *J Neurosci* **27**(15): 4019-4026.

Wallis, Jonathan D. and Earl K. Miller. (2003). "Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task." *Eur J Neurosci* **18**(7): 2069-2081.

Zaki, Jamil, Jochen Weber, Niall Bolger and Kevin Ochsner. (2009). "The neural bases of empathic accuracy." *Proceedings of the National Academy of Sciences of the United States of America* **106**(27): 11382-11387.

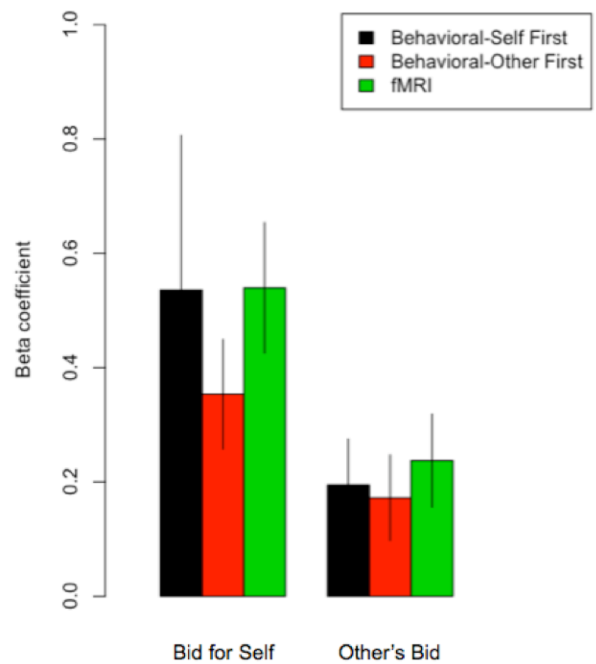
APPENDIX

CONTROL BEHAVIORAL EXPERIMENT

Twenty subjects (mean age: 21.85, SE: 0.76), with the same demographic characteristics required for the main fMRI experiment, participated in a control behavioral study. None of the subjects had previously participated in the fMRI experiment. The experiment was almost identical to the fMRI task, except for the following differences. First, half of the subjects completed the empathic choice task first and the self-oriented choice task second; the other half of the subjects completed the tasks in the opposite order. Second, the inter-trial interval was reduced to 1s, as was the interval between bids and feedback in empathic choice trials. Third, subjects completed both tasks on the same day.

For each individual we carried out a linear regression of bid-for-other on bid-for-self and other-bid, separately for self-oriented and empathic choice trials. This regression measures the extent to which the bids-for-other were driven by their own preferences or by the actual preferences of the other subject. As shown in Fig. S1, we found no significant differences across the two order conditions (min p=0.29, t-test).

Fig. S1. Results of the companion behavioral experiment



TIME COURSE ANALYSIS FOR vmPFC

We carried out a post-hoc ROI analysis to estimate time courses of BOLD activity in the vmPFC valuation areas. This was done as follows. First, we extracted a time course of average BOLD activity at each TR within the region of vmPFC that was found to correlate with stimulus values in both self-oriented and empathic choice trials (Fig. 1C depicts the ROI). Second, we removed from this time course the variance associated with the six motion regressors estimated during the pre-processing process. Third, we resampled the time series into ten time bins per TR, and smoothed it using cubic spline interpolation. Fourth, we then estimated a finite impulse response model that included separate regressors for the following conditions: 1) empathic choice trials for which the bid-for-other was above average for those trials, 2) empathic choice trials for which the bid-for-other was below average, 3) self-oriented choice trials for which the bid-for-self was above average, 4) self-oriented choice trials for which the bid-for-self was below average. The model was estimated at each time bin within a 20 second window starting at the onset of the DVD cue. Finally, the parameter estimates for each condition were averaged across participants at each time point. Fig. S2 depicts the results of the analysis.

Fig. S2. Time course of BOLD responses in the vmPFC valuation area.

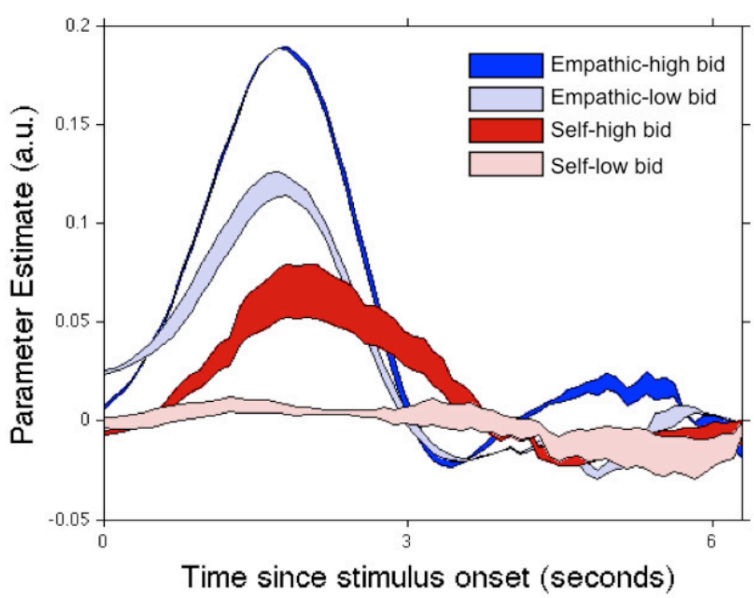


Fig. S3. Additional regions exhibiting a positive correlation with stimulus values in both empathic and self-oriented choice trials (GLM 1).

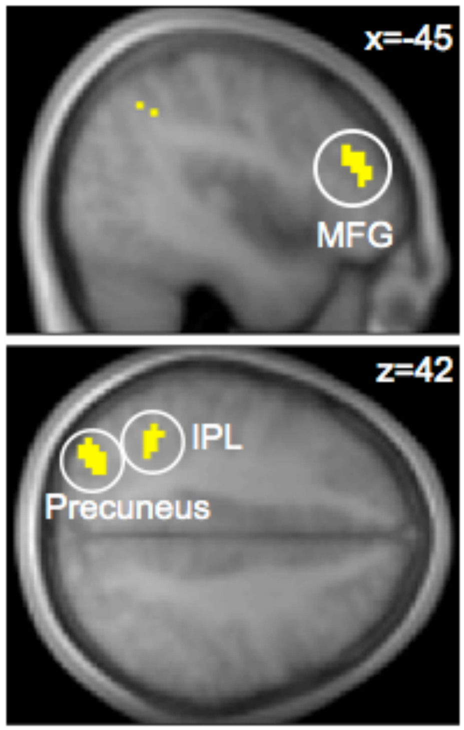
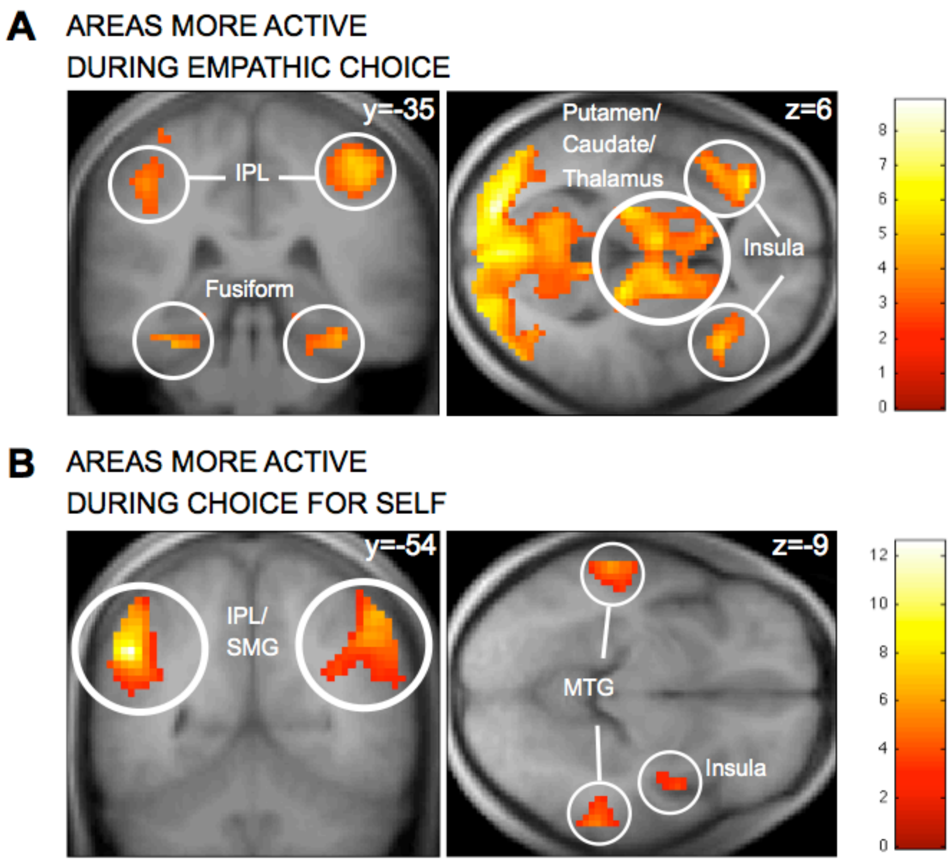


Fig. S4. A) Areas exhibiting higher average (unmodulated) activity during empathic choices. B) Areas exhibiting higher average activity during self-oriented choices (GLM 1).



EXPERIMENTAL INSTRUCTIONS FOR SELF-ORIENTED CHOICE TASK

In this experiment you will be bidding on a number of DVDs. To enable this, please note that $10 has been placed on the table in front of you. This money is yours to use for this experiment. Whatever money you do not use will be yours to keep at the conclusion of the experiment.

In each round of the experiment, you will see the cover of a DVD displayed on the screen. Please look carefully at the DVD and decide how much you would be willing to pay for the DVD on a scale from $0 to $10, in increments of $2. So the possible bids are $0, $2, $4, $6, $8, and $10. If you already own the DVD, bid as if you do not own it. When you are ready, enter your bid for that item using the relevant key on the keypad. If you take longer than 6 seconds to enter a bid, one will randomly be selected for you. **You should bid in the following way: press "z" for $0, "2" for $2, "4" for $4, "6" for $6, "8" for $8, and "0" for $10.**

After you enter your bid, you will see a blank screen for several seconds, after which the next DVD will appear. You will bid in this way for 100 DVDs. After you have completed all rounds of the experiment, ONE of the rounds will be selected at random. The computer will generate a random number from 0 to 10. If this number is larger than or equal to your bid for the DVD in the selected round, you do not purchase the DVD and keep the entire $10. If the number is less than your bid, you pay the amount of the random number for the DVD, obtain the DVD, and keep the rest of the $10. Note that if a DVD you already own is selected, you will simply keep the $10 and will not have to buy the DVD.

For example, let's say trial number 22 was selected, and that in this trial you were shown a picture of the DVD "The Godfather" and entered a bid of $6. Let's say the random number drawn was 8. In this case, you would not be able to buy the DVD and would keep the entire $10. However, if the number drawn were 4, you would purchase the DVD for $4 and keep the remaining $6.

Note that only ONE of the rounds will be executed in this way. Note also that you have every incentive to be truthful in your bid for each DVD, since any single one of the bidding rounds can be implemented at the conclusion of the experiment. You must therefore treat every bid as if it will be implemented, though in the end only one of the bids will count. You should also note, then, that you do not have to divide your budget of $10 over all the rounds – only one of the rounds will actually be implemented, and hence you can treat each round as if you have all $10 available for that item. Thus bidding truthfully is the best strategy you can use in this experiment.

You will now go through five practice rounds to ensure you have understood the instructions. These practice rounds will not have any consequences for the actual experiment and they will not be used in selecting a random round at the end. **Please observe the DVD carefully and enter a bid when you are ready, but within 6 seconds of the DVD appearing.**

EXPERIMENTAL INSTRUCTIONS FOR EMPATHIC CHOICE TASK

In this experiment you will be asked to think about the preferences of another person (whom you will learn about shortly) and how much that person would be willing to pay for a number of DVD titles that will be shown to you.

On the next page, you will learn about the person whose preferences and choices you will be trying to guess. This person is a former Caltech student. Please read this carefully, as it will contain information that may be valuable to you during the experiment.

Meet Todd. He’s from Omaha, NE and studied neuroeconomics at Caltech. He enjoys rock climbing, hiking, and yoga. He also enjoys watching movies, and his favorite film is Braveheart. He likes sports, pizza, and has a fondness for whisky, but dislikes traffic and having to wait for anything.

Todd came to our lab last year and went through 100 rounds of the following experiment. In each round, he was shown a DVD cover and was asked how much he would be willing to pay for that DVD from $0 to $10 in increments of $2 (e.g., either $0, $2, $4, $6, $8, or $10). We used a type of auction that ensured that Todd was telling the truth.

This experiment will also consist of 100 rounds. In each round of the experiment, you will see the cover of a DVD displayed on the screen for 6 seconds. This DVD cover is EXACTLY the same one seen by Todd. So for each DVD that you will see, we know how much Todd was willing to pay for it.

Your task in this experiment will be to try to guess the amount that Todd was willing to pay for each DVD shown.

In each round of this experiment, you are to look carefully at the DVD cover shown, and decide how much you think *Todd* was willing to pay for the DVD. This value should be between $0 and $10 and in increments of $2. Once you have decided on a value, you will indicate your bid of $0, $2, and so on by entering the number corresponding to your bid on the keypad. You have a maximum of 6 seconds to enter a value, after which a value will be entered for you randomly. **You should bid in the following way: press “z” for $0, “2” for $2, “4” for $4, “6” for $6, “8” for $8, and “0” for $10.**

After you have entered your choice, you will see a fixation screen with a “+” sign for 1 second. After this fixation screen, in order to help you to get to know Todd better, you will be told how far off your guess was from Todd’s actual bid. For example, if your guess was $6 while Todd’s guess was $4, you will be told that the error in your guess was 2. If your guess was $6 while Todd’s guess was $10, you will be told that the error in your guess was -4. Your goal is to minimize the absolute value of the error – you want to obtain errors as close to 0 as possible for your guesses.

Following the feedback, after several seconds you will be shown another DVD cover, and you will complete the same task again for each round of the experiment.

Remember, this experiment is not about your own preferences, but what you think Todd's preferences are based on what you know about him. To ensure you are trying your best to guess his preferences, $10 has been placed next to you. This $10 is Todd's money, and you will use it to bid on each DVD.

After you have completed all rounds of the experiment, ONE of the rounds will be selected at random. The computer will generate a random number from 0 to 10. If this number is larger than your bid for the DVD in the selected round, you do not purchase the DVD with Todd's money, and we give the entire $10 to Todd. If the number is less than or equal to your bid, you pay out from Todd's money the amount of the random number for the DVD, and Todd obtains the DVD and keeps the rest of the $10.

For example, let's say trial number 22 was selected, and that in this trial you were shown a picture of the DVD "The Godfather" and thought that Todd would be willing to pay $6 for the DVD. Let's say the random number drawn was 8. In this case, Todd would not be able to buy the DVD and would keep the entire $10. However, if the number drawn were 4, Todd would have to purchase the DVD for $4 and keep the remaining $6.

Note that only ONE of the rounds will be executed in this way. So you do not have to divide Todd's budget of $10 over all the rounds - only one of the rounds will actually be implemented, and hence you can treat each round as if Todd has all $10 available for that DVD. So you want to bid what you really think Todd would be willing to pay for the DVD shown in each round.

You will now do five practice rounds to ensure you have understood the instructions. These rounds are not based on real data from Todd's decisions; the 'feedback' at the end of each round is randomly generated for the practice rounds. These rounds will not have any consequences for the experiment; they are simply for you to become familiar with the task and to give you an opportunity to ask any questions you may have. Remember to first observe the DVD and then enter your decision of what you think Todd was willing to pay by entering the corresponding number on the keypad. **Remember: press "z" for $0, "2" for $2, "4" for $4, "6" for $6, "8" for $8, and "0" for $10.**

CHAPTER 2

Variation in Loss Aversion is Associated with Differential Attention to Losses

Abstract. Risk aversion is widespread in our daily decisions. A popular explanation is loss aversion, or the tendency to prefer avoiding a loss over obtaining a gain by overweighting losses relative to gains. A large body of behavioral evidence has shown that individuals exhibit loss aversion in many domains; however, the mechanisms behind loss aversion remain unknown. Based on recent research that has shown that value-independent differences in attention affect the computation and comparison of values during simple choice, we hypothesized that differences in loss aversion could be modulated by differential attention to losses. In particular, we hypothesized that paying greater attention to losses would result in greater loss aversion both across and within subjects. We tested this hypothesis using a simple eye-tracking choice task in which subjects made binary choices between a mixed-valence lottery and a constant sure outcome. We found that more loss averse subjects paid more relative attention to losses: value-independent differences in attention account for 72.5% of differences in loss aversion across individuals.

INTRODUCTION

Many of our daily decisions involve risk, from financial investments to social interactions. Human behavior in such scenarios is generally risk averse: people require a much larger potential upside to compensate for any potential downside. For example, most people would reject a gamble offering them a 50-50 chance of winning or losing $50. They would require nearly twice as much in gains, or $100, to compensate for the potential loss of $50. This firmly established feature of risk preferences is called loss aversion: losses are weighed more heavily, and thus have more impact on choice, than gains of equivalent magnitude and likelihood (Kahneman & Tversky, 1979). Preferences incorporating loss aversion can reconcile modest-scale risk aversion where other theories, such as expected utility theory, fail (Rabin, 2000). Loss aversion has been conceptualized as a multiplicative overweighting of losses relative to gains and has been well established both in the laboratory and in real world data. Loss aversion has been found to be responsible for a wide range of phenomena, including the endowment effect (Knetsch, Tang, & Thaler, 2001), labor market decisions (Camerer, Babcock, Loewenstein, & Thaler, 1997), the pricing and purchasing of consumer goods (Hardie, Johnson, & Fader, 1993; Putler, 1992) and behavior in financial markets (Barberis & Huang, 2001; Benartzi & Thaler, 1995; Odean, 1998). Studies with primates have shown that they also exhibit loss aversion (Chen, Lakshminarayanan, & Santos, 2006). These studies suggest that loss aversion may be a fundamental feature of how we assess potential outcomes in risky choice.

Little, however, is known about the mechanisms responsible for individual variations in loss aversion. A number of studies have shown that loss aversion can be affected by framing (Gneezy & Potters, 1997; Thaler, Tversky, Kahneman, & Schwartz, 1997). Some studies have proposed that loss aversion may be due to some basic hedonic property of our reaction to losses, or to an error in judgment caused by an exaggeration of losses' actual proportion (Camerer, 2005; Kermer, Driver-Linn, Wilson, & Gilbert, 2006; Novemsky & Kahneman, 2005). It is difficult to distinguish the drivers of loss aversion from purely behavioral data alone because different cognitive processes might result in the same outward behavior. For example, it may be that some people are more loss averse than others because they spend more time evaluating the potential downside of their decisions (say, a day of their company's revenue) compared to the upside (a boost to the company's brand), or because they may be more fearful to face a loss in one domain (a new mate) than in another (a new job). Behavioral data can be integrated with psychophysiological methods to shed light on the mechanisms behind differences in loss aversion across individuals, as well as variation within individuals.

A more recent approach thus uses process tracking to understand choice. A number of studies have examined visual fixation patterns during simple choice to show that attention affects the computation and comparison of values. Johnson et al. (2007) develop a theoretical model that includes attention, in the form of visual fixation, as a decision weight over possible outcomes. Willemsen et al. (2011) use computer mouse tracking data to study a simple choice task using the Asian disease question (Tversky & Kahneman, 1981) and an employment choice paradigm (Tversky & Kahneman, 1991) and suggest that framing may lead to directional comparisons that distort attribute valuations and thus choice. Krajbich et al. (2010) found that visual fixations drive value computation and integration in a simple binary choice task: the amount of time subjects spent looking at their options had a critical effect on choice. Armel and Rangel (2008) found that willingness to pay for appetitive items increases significantly with computation time, while the opposite is true for aversive items. Similarly, changing the relative amount of time that subjects fixate on an item while making a choice can change the probability the item is chosen (Armel, Beaumel, & Rangel, 2008). Busemeyer et al. (1993) develop a cognitive, dynamic model of decision making called decision field theory (DFT), which describes how preferences might evolve over time before a choice is made. The approach encompasses a range of information accumulation models and has been shown to account for a wide range of phenomena, including the relation between choice and decision time as well as preference reversals (Busemeyer & Diederich, 2002). Glöckner et al. (2011) use an eye-tracking task in which participants select between two non-negative outcome gambles to test several models of information search. They find that choice proportions are in line with the predictions of cumulative prospect theory, and their process data indicate support for decision field theory models.

In this paper, we propose to better understand the underlying choice process to uncover what drives loss aversion. We investigate attentional processes in the context of financial decision making under risk. Specifically, we examine visual fixation patterns to see how values for losses compared to gains might be constructed and integrated differently both across and within individuals. Given the effects of attention on the computation and comparison of values during simple choices in these and other studies, we hypothesized that differences in loss aversion across and within individuals might be driven by differential attention to losses compared to gains. We tested this hypothesis by using an eye-tracking decision making experiment in which subjects made binary choices between risky options and a constant sure outcome.

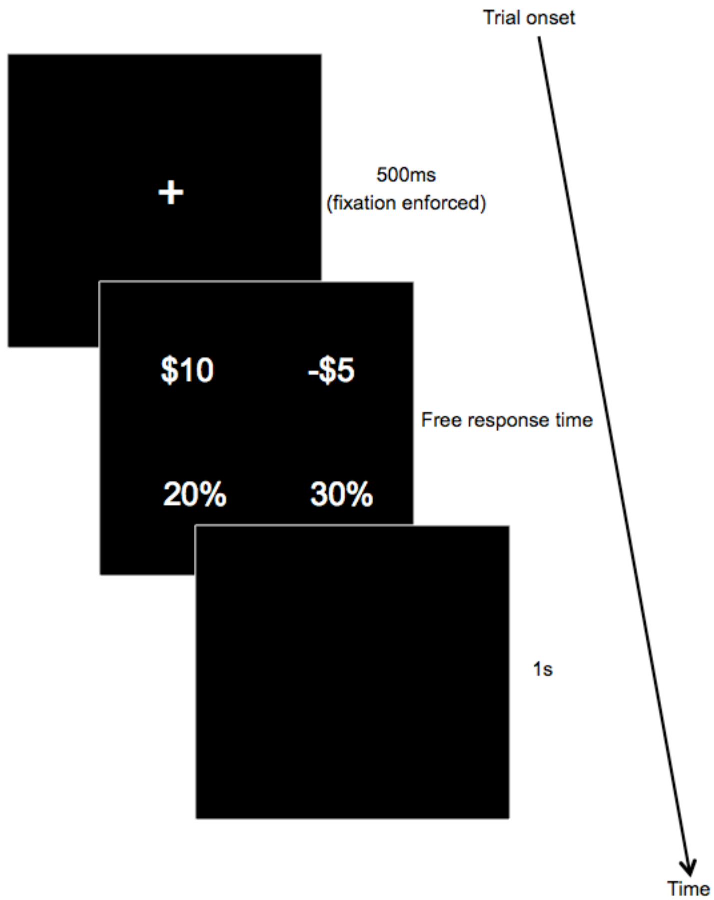
METHODS

Subjects. Twenty-two California Institute of Technology students participated in the experiment (age: mean = 24.3, SD = 4.7; 9 female). Two subjects were excluded because the eye-tracker had difficulty in capturing their gaze. All subjects had normal or corrected-to-normal vision. All subjects were informed about the experiment and gave written consent before participating.

Task. Subjects received written instructions for the task and underwent five practice rounds to ensure their understanding of the task. They were informed that these trials would have no effect on their earnings in the actual experiment. In each trial, subjects first viewed a fixation cross in the center of the screen and were asked to fixate on it for 500ms. The trial would not commence until they had done so. Subjects then viewed the choice screen. Each lottery was comprised of a gain and loss component, as well as percentages indicating the likelihood of receiving each, all of which varied across trials (Fig. 1A). Gain and loss values always appeared on top, while the percentages appeared on bottom, though the locations of gain and loss were randomized. In each trial, subjects made a choice to accept or reject the gamble in favor of a constant sure outcome of $0. Subjects were instructed to press “1” if they strongly accepted the gamble, “2” if they weakly accepted the gamble, “3” if they weakly rejected the gamble and “4” if they strongly rejected the gamble. Subjects completed 384 trials of the task, with a break every 100 trials. The gain outcomes for the lotteries were drawn from the set {$2, 4, 6, 8, 10, 12} and corresponding losses were obtained by multiplying the gain outcomes by a factor ranging from [-¼,-2] in increments of ¼ in a factorial design pairing each gain with each multiplier, yielding a total of 48 gain-loss combinations. These parameters were chosen based on a parameter recovery exercise to find lottery values that were efficient for measuring changes in loss aversion (see Sokol-Hessner et al., 2009). Eight percentage pairings (in which the combined percentages were less than 100%, to increase task difficulty) for each of these combinations resulted in a total of 48*8 = 384 trials. Subjects were paid a show-up fee and experiment completion fee. In addition, five randomly selected trials were implemented for real money at the conclusion of the experiment.

Eye Tracking. Eye movements were recorded at 50 Hz using a Tobii desktop-mounted eye tracker. Before each trial, subjects were required to maintain fixation at a cross at the center of the screen for 500ms before the gamble would appear, ensuring that subjects began each trial fixating on the same location.

Fig. 1. (a) The time course of a sample trial. Subjects are forced to fixate at the center of the screen for 500ms. They are then presented with the lottery, divided into its gain and loss components and the relative probability of obtaining each underneath, and are given as much time as they want to make their choice. After selection, subjects see a blank screen for 1s before the next trial begins.



Data Analysis. We defined four regions of interest (ROIs), or square boxes surrounding each of the four numbers appearing on the screen during each trial. The ROIs were located in the upper left, lower left, upper right and lower right quadrants of the screen. The eye tracker recorded whether the subjects' fixations fell into one of the ROIs or was not recorded (a missing fixation). On average, the latency period (time elapsed between stimulus appearance and first recorded fixation) was 247.50 ms (SD = 45.74 ms). The latency period was assumed to be due to peripheral attentional processes involved in first fixation selection and not part of the decision time. Subjects spent 11.94% (SD = 5.12%) of each trial looking at a point other than one of the four ROIs. Missing fixations during the trial were treated as follows:

1) If the missing fixations were recorded between fixations to the same item, then those missing fixations were changed to that item and assumed to be response time. For example, a fixation pattern of "upper left, missing, upper left" would become "upper left, upper left, upper left."
2) If the missing fixations were recorded between fixations to different items, then those missing fixations were discarded and not counted in response time.

These missing fixations were likely due to momentary transitions between items or momentary loss of fixation from the eye tracker. The mean number of trials dropped per subject was 0.70 (SD = 1.42). The mean response time was 3952 ms (SD = 1231 ms).

Prospect Theory Model. We estimated the parameters of a prospect theory model for each subject (Tversky & Kahneman, 1992). The subjective utility of a lottery L is defined by four parameters: the gain or loss amount x, the percentage chance of receiving that amount *p(x)*; the loss aversion coefficient λ, and the curvature of the utility function α (representing risk aversion due to the presence of diminishing sensitivity to changes in value as the absolute value increases). The subjective utility of each lottery was estimated with Equation (1), while Equation (2) translates the difference between the subjective value of the lottery and the subjective value of the certain amount (0) into a probability of gamble acceptance using the logit sensitivity μ:

$$u\,(L) = \begin{cases} p(x) \cdot x^{\alpha}, & x \geq 0 \\ -\lambda \cdot p(x) \cdot |x|^{\alpha}, & x < 0 \end{cases} \tag{1}$$

$$p(accept) = \frac{1}{1 + e^{-\mu(u(gamble) - u(certain))}} \tag{2}$$

The lottery values themselves were originally chosen based on a parameter recovery exercise to find lottery values that were efficient for measuring changes in loss aversion, similar to that employed by Sokol-Hessner et al. (2009). In essence, a hypothetical participant was created by selecting a range of psychologically plausible values for the three model parameters based on results from earlier studies. Stochastic choices were simulated, using those parameter values and Eq. 2, over the initial monetary amounts. Given these simulated choices, we then used the maximum-likelihood procedure to estimate parameters by maximizing the following likelihood function:

$$l(\alpha, \lambda, \mu | y) = \sum_{i=1}^{384} y_i \log(p(accept)) + (1 - y_i) \log\big(1 - p(accept)\big) \qquad (3)$$

where α, λ and μ are the parameters to be estimated, y is the subject response, i is the trial number, and p(accept) is as defined in Eq. 2. The Nelder-Mead Simplex Method as implemented in Matlab 2007a was used to obtain estimates for each parameter. If the estimated parameters were close to the actual ones used to create the simulated data, then we could say that the modeling procedure could ''recover'' parameter values accurately. We used this method of creating our stimuli to improve our ability to accurately recover a range of parameter values from actual participants given the choices made and therefore increase the power of statistical tests to detect differences across and within subjects.

RESULTS

Basic psychometrics. The mean parameter values estimated across all subjects were as follows: μ = 6.98 (SE = 1.70), α = 0.79 (SE = 0.07), λ = 1.53 (SE = 0.12) (See Table S1 for details). Note that there were five subjects with unusually low values of α, α < 0.50. To ensure the robustness of our main results, we repeated several keys analyses with these five subjects removed (see Appendix, Fig. S1). The results did not differ from our main findings.

The model fit the choice curve well. The choice data indicate that choices were a logistic function of the subjective lottery value lottery value (pseudo-R^2=0.55; Fig. 1B). Reaction times and number of fixations both correlated with difficulty (mixed effects regression estimates: -16.23, p=0.05, and -0.09, p=0.0000, respectively; Fig. 1C-D).

Fig. 1. (b) Psychometric choice curve.

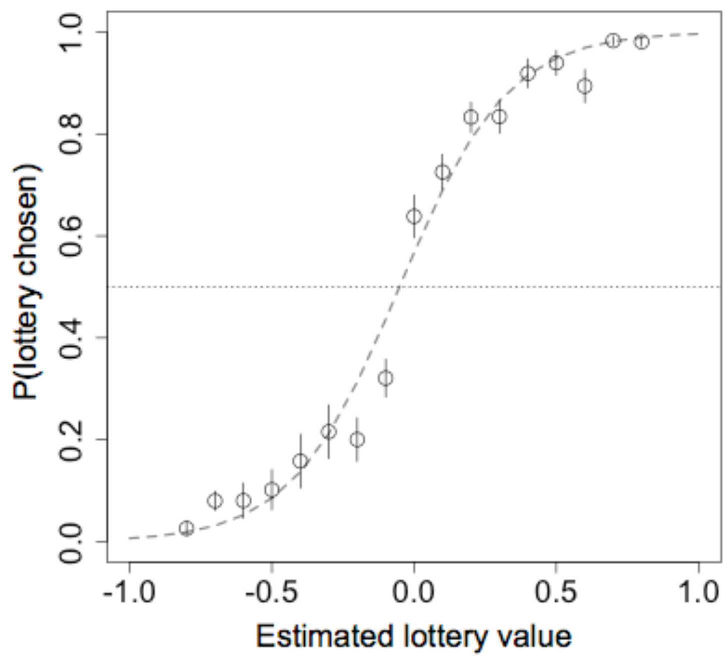


Fig. 1. (c) Reaction time as a function of difficulty (the absolute value of the subjective value of the lottery).

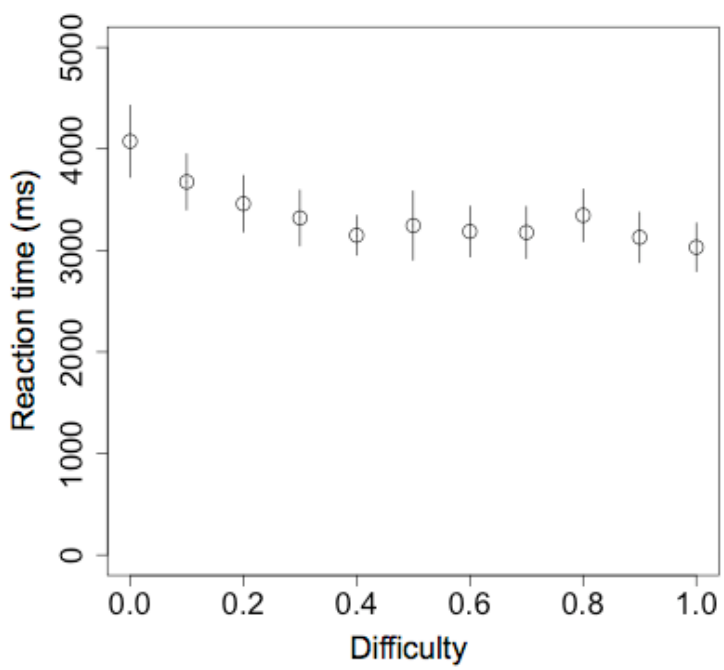
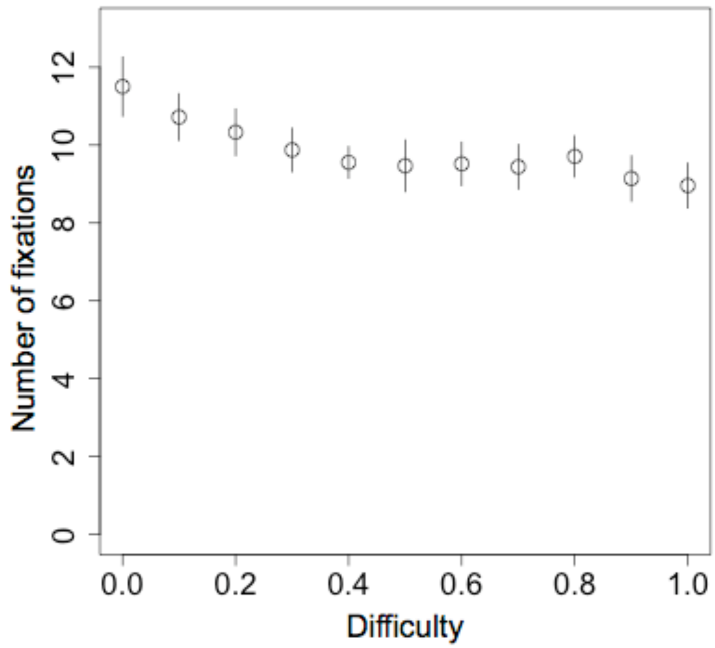


Fig. 1. (d) Number of fixations as a function of difficulty (the absolute value of the subjective value of the lottery).



Attentional biases across subjects. Consistent with the first hypothesis, we found that more loss averse subjects paid more relative attention to losses. The loss aversion coefficient λ is positively correlated with the relative time spent looking at losses compared to gains (mixed effects regression estimate: 0.08, p=0.02; Fig. 2A).

To calculate the magnitude of this effect across subjects, we performed the following analysis. We took the 5% and 95% individual loss aversion coefficients across subjects, $\lambda_{5\%}$ and $\lambda_{95\%}$, respectively, according to the distribution of the relative time spent looking at losses compared to gains. We divided the difference $\lambda_{5\%} - \lambda_{95\%}$ by the difference between the maximum and minimum loss aversion coefficients, $\lambda_{min} - \lambda_{max}$, to obtain a statistic indicating the percentage of the differences in loss aversion across individuals that is accounted for by value-independent differences in attention: 72.5%.

Across subjects, the correlation between λ and the total time spent looking at the loss amount was 0.49 (p=0.03), while the correlation between λ and the total time spent looking at the gain amount was 0.17 (p=0.49). The correlation between λ and the difference in the percent of time spent looking at the loss amount vs. the gain amount was 0.51 (p=0.02).

Attentional biases within subjects. We also conducted another analysis in which we examined whether the magnitude of differences in loss aversion are correlated with the magnitude of attentional fluctuations within subjects. For each subject, we divided the trials into two halves based on whether excess total fixation to gains over losses is above or below the median. We then estimated the model parameters for each subject in each of the samples independently. We found that the mean difference in λ between the below and above median samples is 0.12 (SD = 0.026, p=0.02). Fig. 2B shows a scatter plot of the difference in lambda vs. the difference in percent time looking at losses in the two samples (mixed effects regression estimate: 0.06, p=0.05).

To calculate the magnitude of this effect within subjects, for each subject, we looked at the 5% and 95% probability of accepting the lottery, $p_{5\%}$ and $p_{95\%}$, as a function of the difference in time spent looking at gains compared to losses. The percent of the variation in λ that is explained by their fixation is then given by dividing the difference $p_{5\%} - p_{95\%}$ by the difference in the time spent looking at gains compared to losses for each individual. On average, we found that the percentage of variation in the probability of accepting the lottery explained by the amount of time spent looking at the loss amount vs. the gain amount is 6.25% (SE = 2.09%). Fig. 2C shows a histogram of the individual percentage variations.

In addition, we found a strong relationship between the last fixation and choice. Specifically, for the last fixation only, subjects spent more time looking at gains compared to losses as expected value increased (mixed effects regression estimate: 28.64, p=0.0001; Fig. 3A). There is thus a bias toward the chosen item.

Properties of the general search process. However, this choice bias does not extend to the general nature of the fixation process, which is independent of underlying value. To rule out that our main effect is simply due to subjects paying more attention to large gains or losses compared to smaller ones, we examine whether the time spent looking at gains compared to losses is a function of the expected value of the lottery. If subjects were paying more attention to more attractive options (e.g., large gains), we would expect this relationship to be positive, resulting in an upward sloping curve. However, we find that there is a nearly flat relationship between expected value and time spent looking at gains compared to losses (mixed effects regression estimate: 2.42, p=0.05; Fig. 3B). While the relationship is significant, it is extremely small and cannot account for the effect. Note that as the last fixation displays a choice bias, the last fixation has been discarded here.

We also examine several additional fixation properties. First, the probability that the first fixation was to the upper left was much higher than for any of the other areas (Fig. S2A). This is likely a cultural artefact from reading left to right and top to bottom. As a result, the first fixation was more likely to be to either the gain or loss amount, while later fixations were more likely to be to the probabilities (Fig. S2B). The last five fixations did not show any bias towards area (Fig. S2C) or type (Fig. S2D). Fixation duration was relatively constant regardless of the item location (Fig. S3A) or type (Fig. S3B).

An analysis of subjects' transitions among the gain and loss amounts and probabilities indicated a common pattern (Fig. S4). Subjects were much more likely to exhibit horizontal and vertical rather than diagonal fixations.

Fig. 2. (a) Lambda coefficients estimated for each subject as a function of the percent of time spent looking at the dollar value of the loss – the dollar value of the gain.

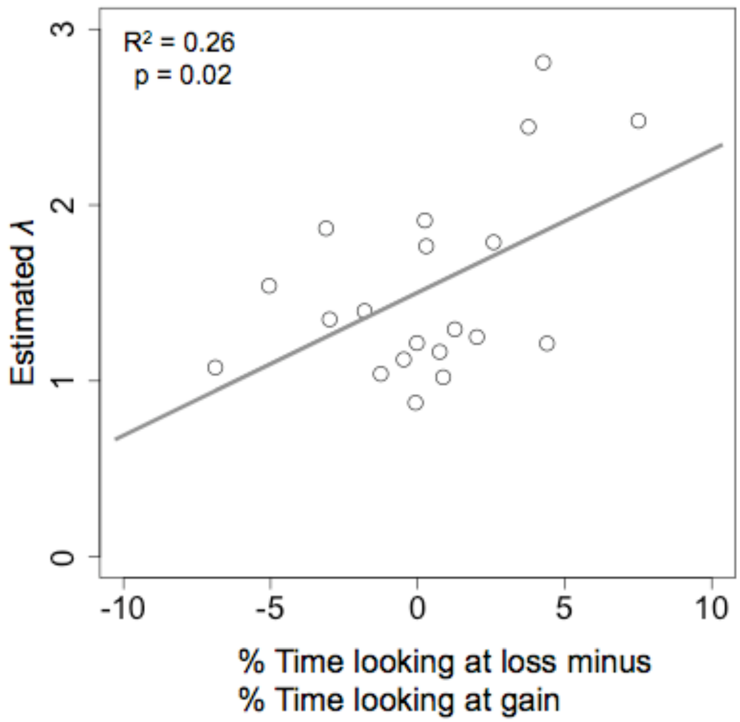


Fig. 2. (b) Scatter plot of the difference in lambda vs. the difference in percent time looking at losses after dividing individual trials into two halves based on whether excess total fixation to gains over losses is above or below the median.

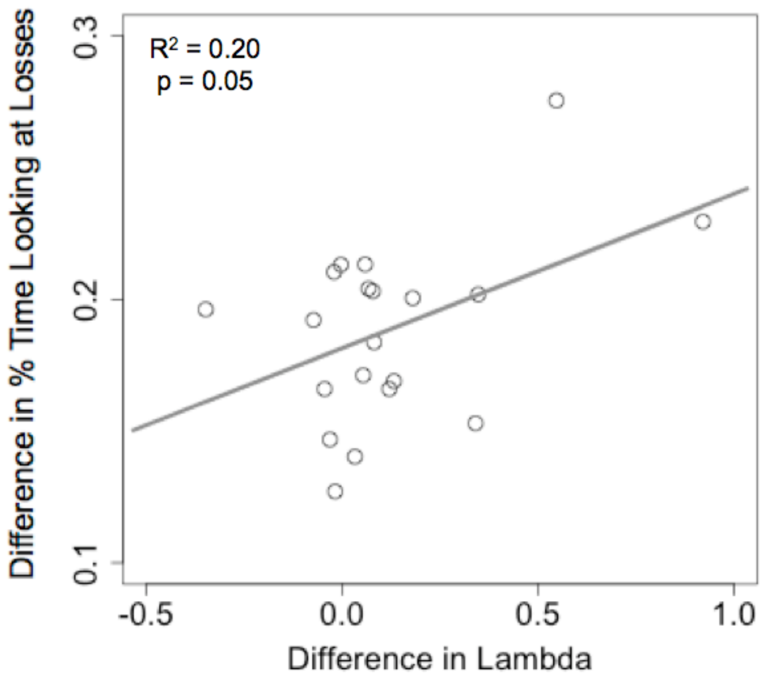


Fig. 2. (c) Histogram of individual percentage variations in the probability of accepting the lottery explained by the amount of time spent looking at the loss amount vs. the gain amount.

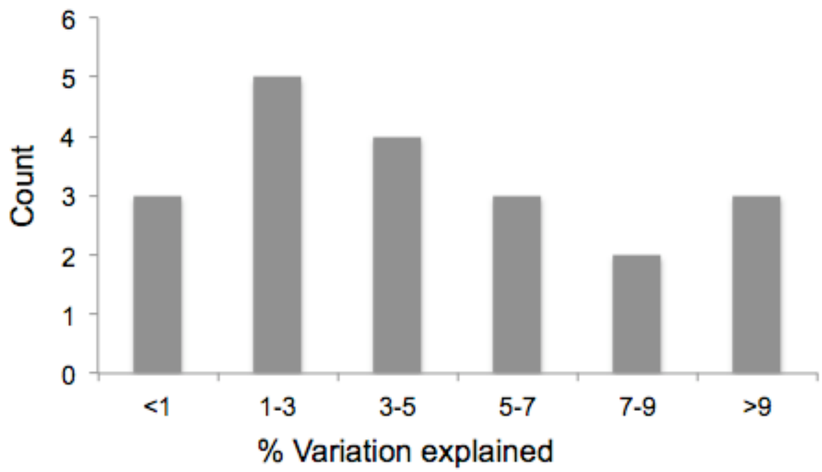
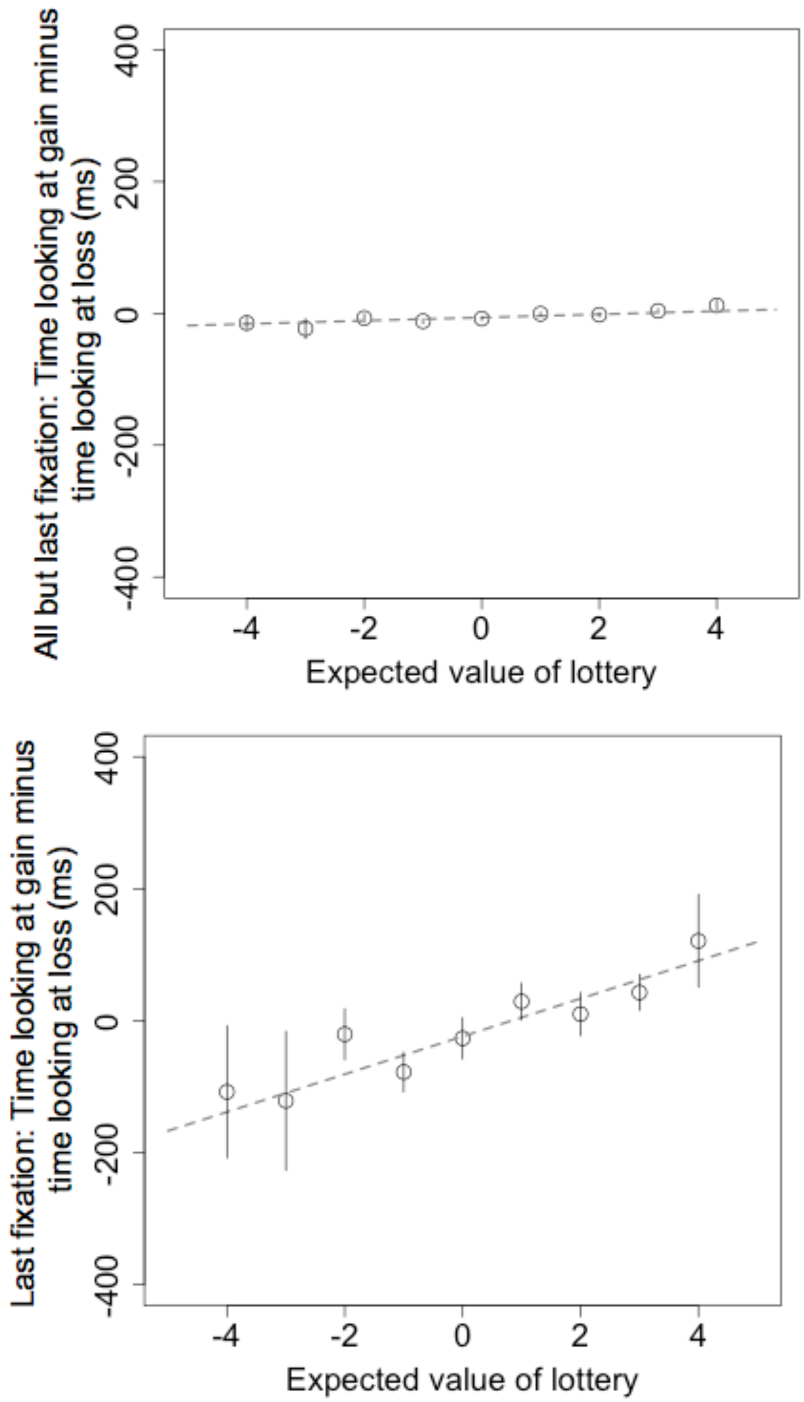


Fig. 3. (a) For the last fixation only, time looking at the dollar value of the gain minus time looking at the dollar value of the loss as a function of the expected value of the lottery. **(b)** Across all but the last fixation, time looking at the dollar value of the gain minus time looking at the dollar value of the loss as a function of the expected value of the lottery.



DISCUSSION

Our study provides insight into how we process risk. The results describe the nature of the attentional processes that guide choice in a simple lottery task. Specifically, we show that attention plays a key role in differences in loss aversion both across and within subjects. Our results suggest a model in which attention must be included to make any inferences on choice. In addition, our data also demonstrate that prospect theory can account for a number of correlations between choice and fixation patterns.

Our results contribute to the literature on the cognitive processes behind information search and integration. While subjects in our study were not under any time constraints, we find evidence to support some of the key predictions of DFT, including preference variability and a strong relationship between preference and reaction time. Similarly to Krajbich et al.'s (2010) attentional drift diffusion model (DDM), we found that the first and middle fixations were independent of the value of the fixated item. Our data also display a last-fixation bias, which in the DDM framework, is a direct implication of the fact that the value of the non-fixated item is discounted. In our experiment, subjects discount the lottery when they spend more time looking at losses during the last fixation. Other results, however, were inconsistent with the predictions of the DDM. This is likely due to the fact that lotteries are much more complex stimuli than snack foods, and attentional processes are deployed differently.

An important question raised by our results is the directionality of the relationship between attention and loss aversion. Is it that more loss averse individuals pay more attention to losses, or that paying more attention to losses causes one to become more loss averse? Several related studies have shown that it is possible to bias choices by exogenously manipulating relative fixation durations and that the fixation process may have a causal effect on the value comparison process (Armel et al., 2008; Shimojo, Simion, Shimojo, & Scheier, 2003). Willemsen et al. (2011) use a path model relating gain vs. loss frames and reference options to information acquisition and choice. They find that the framing of gains vs. losses as well as reference points affect attentional differences, and these changes significantly affect both search and choice. However, the evidence in this study is not sufficient to establish a causal relationship between attention and loss aversion. An experiment manipulating subjects' fixations to establish causality is thus a promising idea for future research. Another topic for future study is how changes in emotional states, such as stress, may change loss aversion and thus risk aversion. Understanding how loss aversion changes within individuals in different contexts can further uncover the mechanisms driving loss aversion.

In addition, our results do not rule out the possibility that subjects' values may have an effect on fixation patterns. While expected value did not have an effect on fixation patterns, reaction time did increase with choice difficulty. Thus random variation in fixation duration might affect the search process and thus choice. An open question is therefore to test whether the relationship between values and fixations is actually exogenous and how the fixation process takes value into account.

An additional question that was not tackled in this paper is how such choices might be implemented in the brain. One brain region that is likely to be involved is the medial orbitofrontal cortex (mOFC), which has been found to encode value at the time of choice in a number of different studies and contexts. Tom et al. (2007) found that activity in mOFC during a monetary risky choice task correlated with stimulus values consistent with the predictions of prospect theory. Furthermore, the authors found that the same area of mOFC correlates with both positive and negative potential outcomes. Similarly, Plassmann et al. (2010) found that mOFC activity correlates with the appetitiveness and aversiveness of foods. Levy et al. (2010) further showed that the mOFC encodes subjective values in choice under risk as well as ambiguity. Basten et al. (2010) found evidence that the brain weighs costs against benefits by combining neural benefit and cost signals into a single representation of value that is accumulated over time, in accordance with perceptual diffusion models. Most recently, Lim et al. (2011) showed that value computations in the mOFC and the ventral striatum are fixation-dependent: activity in the these areas correlates with the difference in value between attended and unattended items. The question of how attention influences these computations and whether these signals are driven by attentional processes elsewhere in the brain remains an open question.

Our results have important implications for decision making and the role of attentional processes in choice. Our data show that systematic biases in fixations could lead to different choices. Moreover, as biases such as loss aversion can lead to deficits in decision-making, our findings raise the interesting possibility that we may be able to modulate our attention to make better choices. Further uncovering the relationship between attention and loss aversion is of direct interest to a number of fields, including psychology, economics and neuroscience.

REFERENCES

Armel, K. C., Beaumel, A. and Rangel, A. (2008). Biasing simple choices by manipulating relative visual attention. *Judgment and Decision Making, 3*(5), 396-403.

Armel, K. C. and Rangel, A. (2008). The impact of computation time and experience on decision values. *American Economic Review, 98*(2), 163-168. doi: Doi 10.1257/Aer.98.2.163

Barberis, N. and Huang, M. (2001). Mental accounting, loss aversion, and individual stock returns. *The Journal of Finance, 56*(4), 1247-1292.

Basten, U., Biele, G., Heekeren, H. R. and Fiebach, C. J. (2010). How the brain integrates costs and benefits during decision making. *Proceedings of the National Academy of Sciences, 107*, 21767-21772.

Benartzi, S. and Thaler, R. H. (1995). Myopic loss aversion and the equity premium puzzle. *Quarterly Journal of Economics, 110*, 73-92.

Busemeyer, J. R. and Diederich, A. (2002). Survey of decision field theory. *Mathematical Social Sciences, 43*, 345-370.

Busemeyer, J. R. and Townsend, J. T. (1993). Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review, 100*(3), 432-459.

Camerer, C. (2005). Three cheers - psychological, theoretical, empirical - for loss aversion. *Journal of Marketing Research, 42*(2), 129-133.

Camerer, C., Babcock, L., Loewenstein, G. and Thaler, R. (1997). Labor supply of New York City cabdrivers: One day at a time. *Quarterly Journal of Economics, 112*(2), 407-441.

Chen, M. K., Lakshminarayanan, V. and Santos, L. R. (2006). How basic are behavioral biases? Evidence from capuchin monkey trading behavior. *Journal of Political Economy, 114*(3), 517-537.

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. *Journal of Behavioral Decision Making, 24*(1), 71-98. doi: Doi 10.1002/Bdm.684

Gneezy, U. and Potters, J. (1997). An experiment on risk taking and evaluation periods. *Quarterly Journal of Economics, 112*(2), 631-645.

Hardie, B. G. S., Johnson, E. J. and Fader, P. S. (1993). Modeling Loss Aversion and Reference Dependence Effects on Brand Choice. *Marketing Science, 12*(4), 378-394.

Johnson, J. G. and Busemeyer, J. R. (2007). A computational model of the attentional processes used to generate decision weights in risky decision making. *Under revision for Cognition.*

Kahneman, D. and Tversky, A. (1979). Prospect Theory - Analysis of Decision under Risk. *Econometrica, 47*(2), 263-291.

Kermer, D. A., Driver-Linn, E., Wilson, T. D. and Gilbert, D. T. (2006). Loss aversion is an affective forecasting error. *Psychological Science, 17*(8), 649-653.

Knetsch, J. L., Tang, F. and Thaler, R. H. (2001). The endowment effect and repeated market trials: is the Vickrey auction demand revealing? *Experimental Economics, 4*(2), 257-269.

Krajbich, I., Armel, C. and Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience, 13*(10), 1292-1298. doi: Doi 10.1038/Nn.2635

Levy, I., Snell, J., Nelson, A. J., Rustichini, A. and Glimcher, P. W. (2010). Neural Representation of Subjective Value Under Risk and Ambiguity. *Journal of Neurophysiology, 103*(2), 1036-1047. doi: Doi 10.1152/Jn.00853.2009

Lim, S. L., O'Doherty, J. P. and Rangel, A. (2011). The Decision Value Computations in the vmPFC and Striatum Use a Relative Value Code That is Guided by Visual Attention. *Journal of Neuroscience, 31*(37), 13214-13223. doi: Doi 10.1523/Jneurosci.1246-11.2011

Novemsky, N. and Kahneman, D. (2005). How do intentions affect loss aversion? *Journal of Marketing Research, 42*(2), 139-140.

Odean, T. (1998). Are investors reluctant to realize their losses? *Journal of Finance, 53*(5), 1775-1798.

Plassmann, H., O'Doherty, J. P. and Rangel, A. (2010). Appetitive and Aversive Goal Values Are Encoded in the Medial Orbitofrontal Cortex at the Time of Decision Making. *Journal of Neuroscience, 30*(32), 10799-10808. doi: Doi 10.1523/Jneurosci.0788-10.2010

Putler, D. S. (1992). Incorporating Reference Price Effects into a Theory of Consumer Choice. *Marketing Science, 11*(3), 287-309.

Rabin, M. (2000). Risk aversion and expected-utility theory: A calibration theorem. *Econometrica, 68*(5), 1281-1292.

Shimojo, S., Simion, C., Shimojo, E. and Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nature Neuroscience, 6*(12), 1317-1322. doi: Doi 10.1038/Nn1150

Sokol-Hessner, P., Hsu, M., Curley, N. G., Delgado, M. R., Camerer, C. F. and Phelps, E. A. (2009). Thinking like a trader selectively reduces individuals' loss aversion. *Proceedings of the National Academy of Sciences of the United States of America, 106*(13), 5035-5040. doi: Doi 10.1073/Pnas.0806761106

Thaler, R. H., Tversky, A., Kahneman, D. and Schwartz, A. (1997). The effect of myopia and loss aversion on risk taking: An experimental test. *Quarterly Journal of Economics, 112*(2), 647-661.

Tom, S. M., Fox, C. R., Trepel, C. and Poldrack, R. A. (2007). The neural basis of loss aversion in decision-making under risk. *Science, 315*(5811), 515-518. doi: Doi 10.1126/Science.1134239

Tversky, A. and Kahneman, D. (1981). The Framing of Decisions and the Psychology of Choice. *Science, 211*(4481), 453-458.

Tversky, A. and Kahneman, D. (1991). Loss Aversion in Riskless Choice - a Reference-Dependent Model. *Quarterly Journal of Economics, 106*(4), 1039-1061.

Tversky, A. and Kahneman, D. (1992). Advances in Prospect-Theory - Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty, 5*(4), 297-323.

Willemsen, M. C., Bockenholt, U. and Johnson, E. J. (2011). Choice by Value Encoding and Value Construction: Processes of Loss Aversion. *Journal of Experimental Psychology-General, 140*(3), 303-324. doi: Doi 10.1037/A0023493

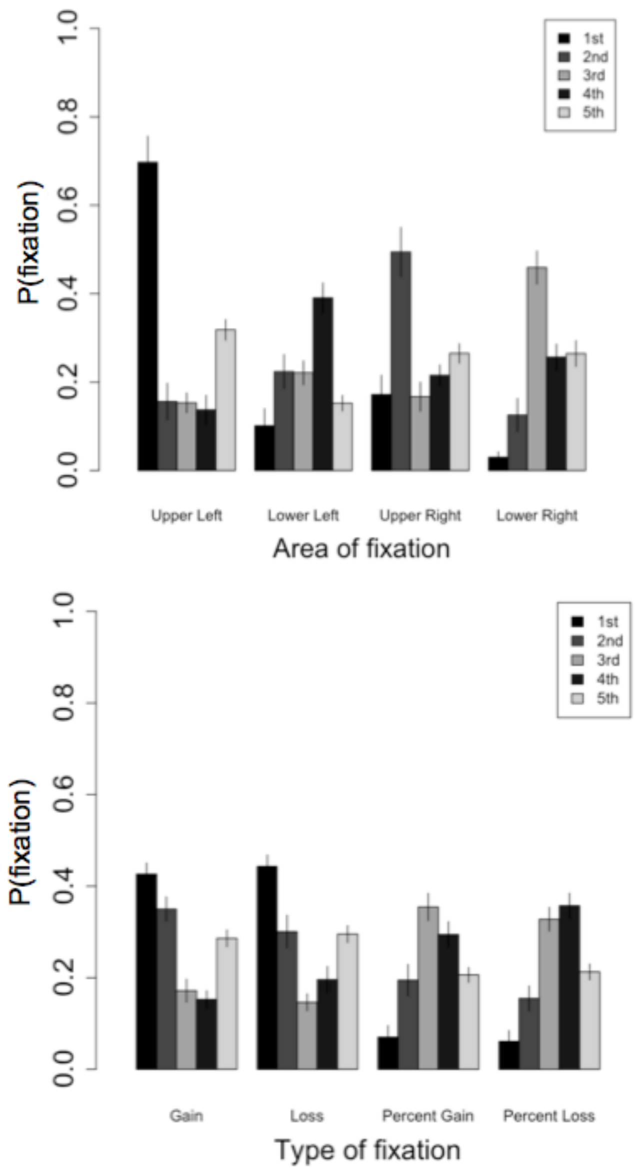
APPENDIX

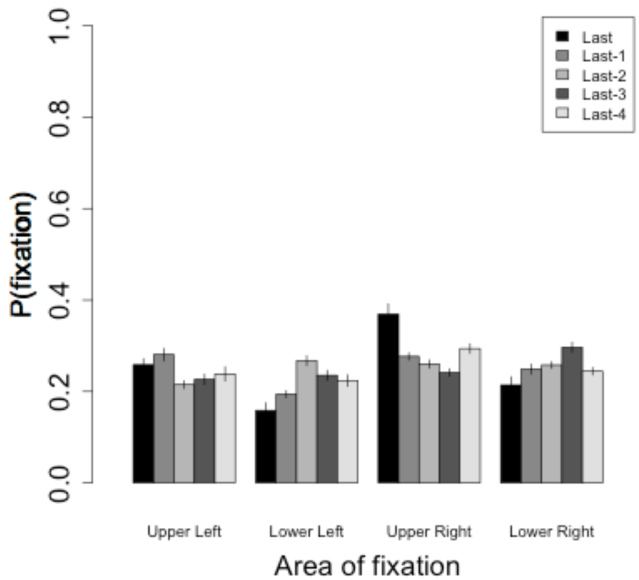
SUPPLEMENTARY TABLES AND FIGURES

Table S1. Parameter estimates for all subjects. BIC = Bayesian information criterion.

Subject	μ	α	λ	Log likelihood	BIC
1	8.73	0.95	1.12	128.47	-239.09
2	10.38	0.43	1.08	236.56	-455.27
3	2.60	0.95	1.40	144.36	-270.87
4	33.66	0.13	1.02	300.2	-582.55
5	13.21	0.49	1.21	177.29	-336.73
6	1.15	0.82	1.77	236.29	-454.73
7	4.30	0.92	1.04	171.16	-324.47
8	17.68	0.22	1.35	243.56	-469.27
9	2.19	0.79	2.44	213.19	-408.53
10	8.41	0.42	1.54	183.21	-348.57
11	5.75	0.74	2.81	232.86	-447.87
12	3.16	1.12	0.88	194.36	-370.87
13	5.52	1.01	1.22	120.22	-222.59
14	2.45	1.17	1.91	138.45	-259.05
15	0.38	1.50	2.48	178.37	-338.89
16	4.21	0.84	1.25	144.65	-271.45
17	5.07	0.91	1.17	137.94	-258.03
18	2.16	0.87	1.87	172.78	-327.71
19	5.11	0.65	1.79	163.94	-310.03
20	3.41	0.97	1.29	132.46	-247.07
Mean	**6.98**	**0.79**	**1.53**	**182.52**	**-347.18**
SE	*1.70*	*0.07*	*0.12*	*10.73*	*21.46*

Fig. S1. Probability of fixation **(a)** to a given area and **(b)** for a given type for the first five fixations, and probability of fixation **(c)** to a given area and **(d)** for a given type for the last five fixations.




Last
Last-1
Last-2
Last-3
Last-4
P(fixation)
0.0
0.2
0.4
0.6
0.8
1.0
Upper Left
Lower Left
Upper Right
Lower Right
Area of fixation

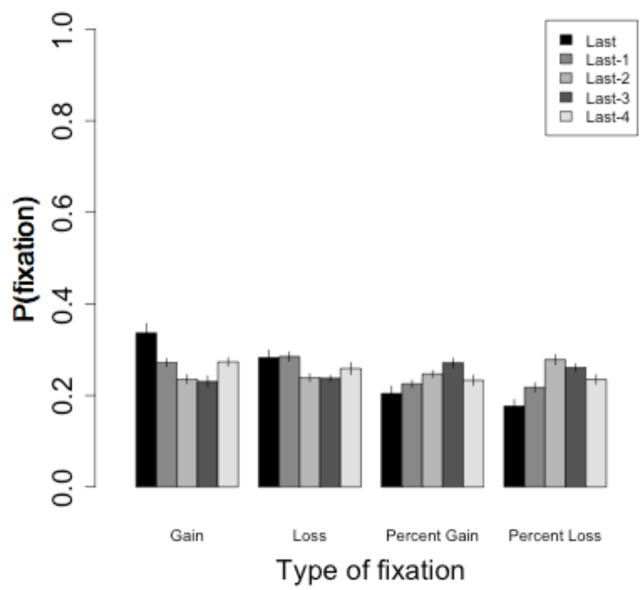
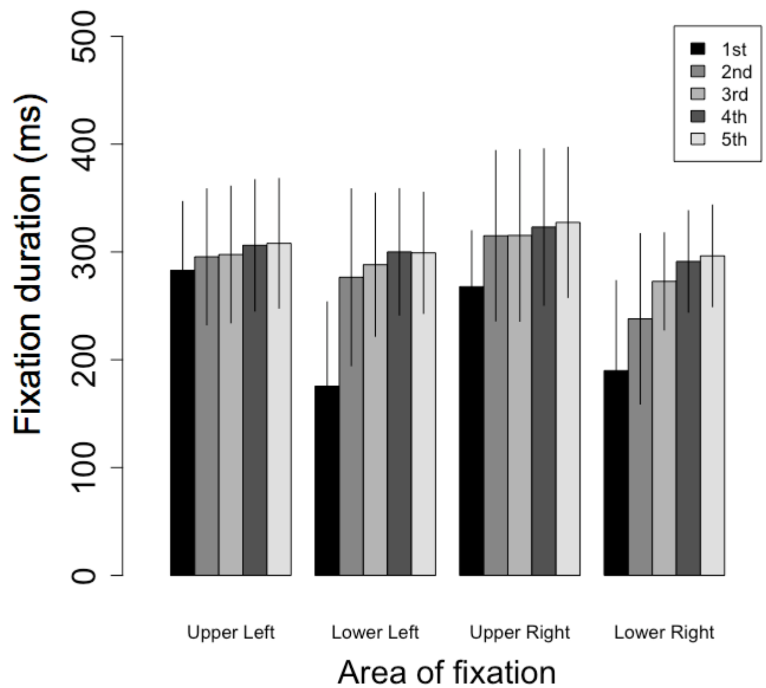

Last
Last-1
Last-2
Last-3
Last-4
P(fixation)
0.0
0.2
0.4
0.6
0.8
1.0
Gain
Loss
Percent Gain
Percent Loss
Type of fixation

Fig. S2. Fixation duration as a function of **(a)** area and **(b)** type.



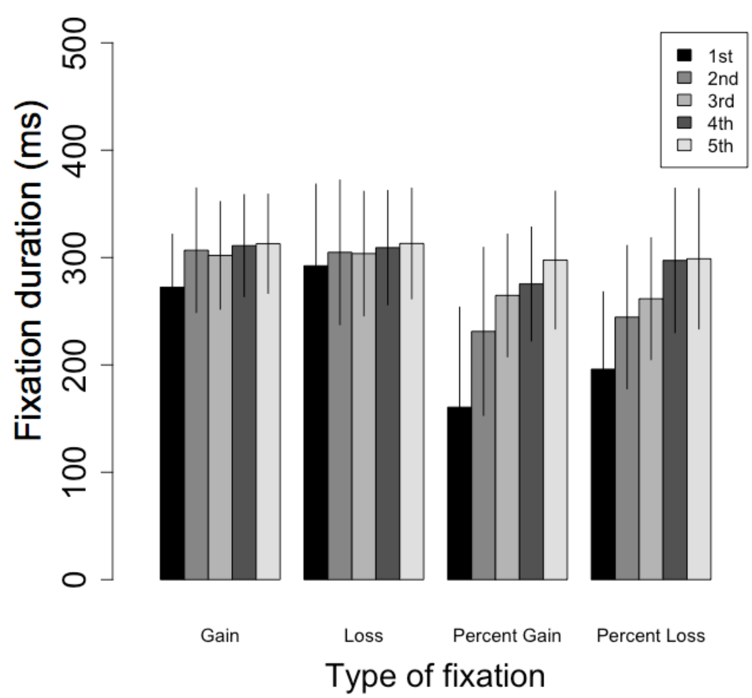


Fig. S3. Probability of transition to an area given current area of fixation.

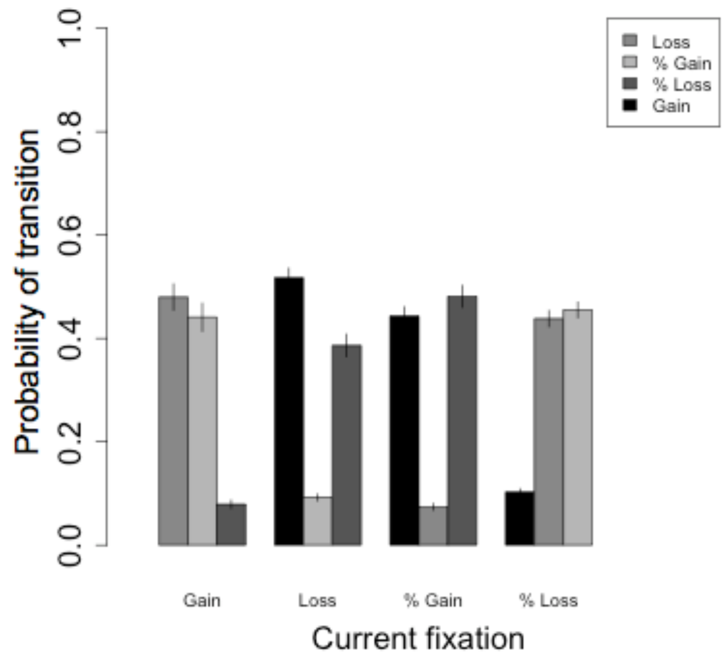
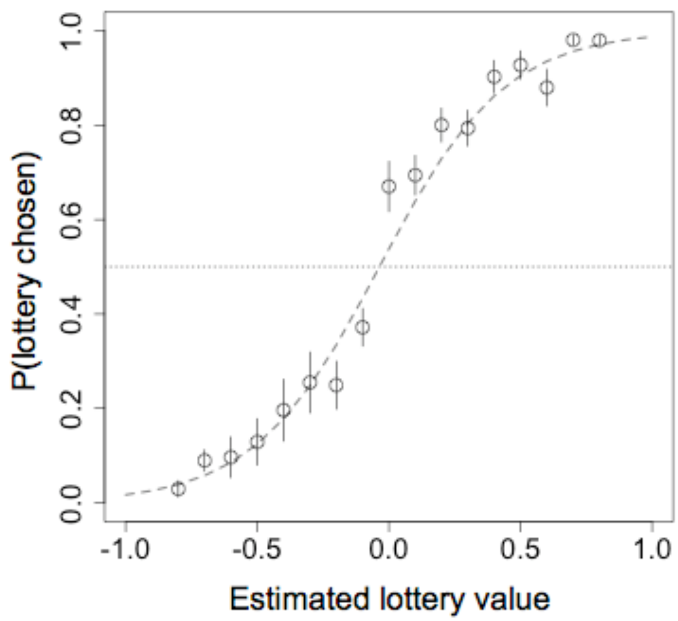
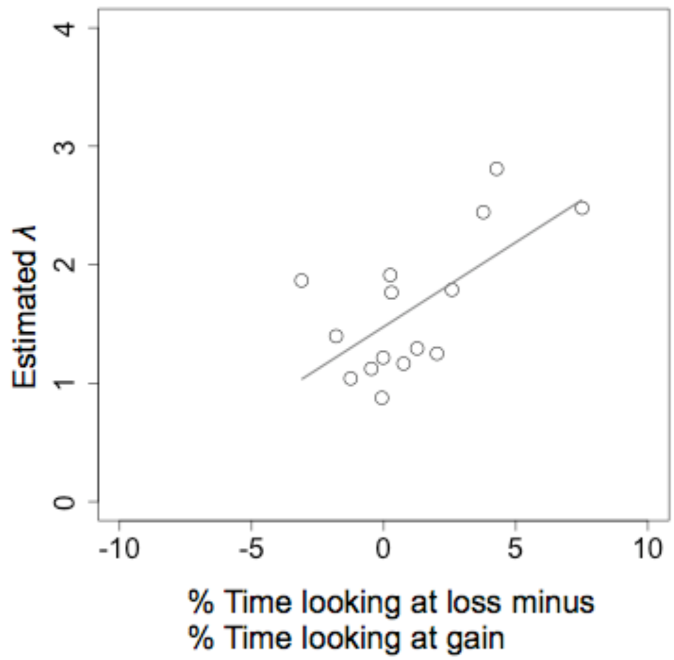
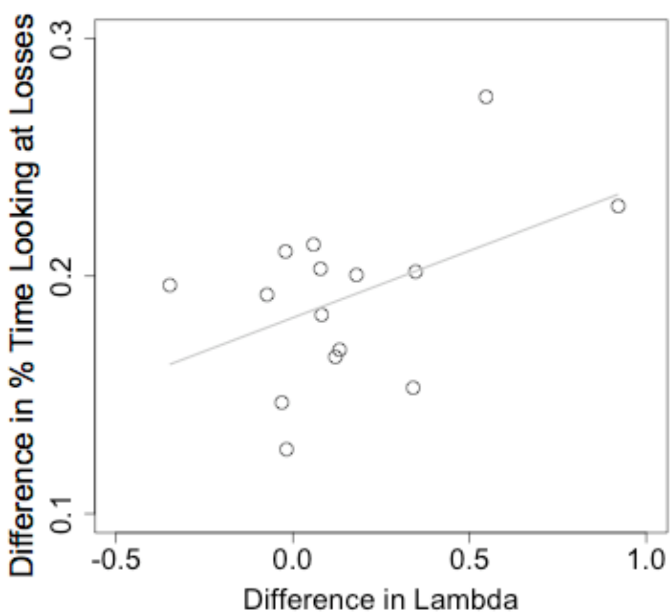


Fig. S4. Key figures with five subjects with α < 0.50 excluded. **(a)** Fig. 1b, **(b)** Fig. 2a, **(c)** Fig. 2b.




Estimated λ
0
1
2
3
4
-10
-5
0
5
10
% Time looking at loss minus
% Time looking at gain


Difference in % Time Looking at Losses
0.1
0.2
0.3
-0.5
0.0
0.5
1.0
Difference in Lambda

CHAPTER 3

Display and Search Dynamics in Multi-Attribute Choice

Abstract. Consumer choices are a crucial component of everyday decisions. When entering a store to make a purchase, consumers must dynamically search and select among a variety of options. Understanding the details of this decision-making process is therefore essential for developing models of choice behavior. This raises two questions that are studied in this paper: (1) What attentional processes are involved in the choice between multi-attribute consumer goods? (2) Do simple display changes affect these attentional processes and thus choice? We hypothesized that (1) value-independent differences in attention impact multi-attribute choice, and (2) simple changes to the display impact the attentional fixation process and thus choice. We tested these hypotheses using a two-item, two-attribute choice task to determine how decision-making processes are modified by the introduction of multiple discrete attributes for each item. Subjects were presented with pairs of posters, each of a different design and size, and asked to choose the more desirable one. Attributes of each choice were arranged in two visual conditions: designs on top, and sizes on top. MouseLab was used to gather detailed search-process data. We found that attentional processes play a key role in integrating across multi-attribute goods in consumer choice. Moreover, we found that simple display changes resulted in striking differences in visual search patterns, which subsequently impacted choice through differential weighting and integration of the attributes. Our results demonstrate the value of understanding

computational processes in consumer choice and have important applications in economics and marketing.

INTRODUCTION

Consider the hundreds, if not thousands, of choices made every day between items that differ in many dimensions, such as color, taste, and size. Consumers need to dynamically search over an often complex choice set. Understanding the details of this decision-making process is essential for developing models of choice behavior. This raises two basic questions that we study in this paper: (1) How do attentional fixation processes impact choice for multi-attribute consumer goods? (2) Do simple display changes affect these attentional processes and thus choice?

This study builds on previous literature from economics and marketing that has used visual attention to provide a window into the computational processes involved in consumer decisions. Such studies have typically used process data, such as eye-tracking or MouseLab (Payne et al., 1993; Camerer et al., 1993), to examine choice among several alternatives because they allow experimenters to track where subjects look in a display. Process data can provide important insight to help us understand and influence consumer choice, an endeavor on which companies spend billions of dollars every year. An understanding of the computational processes behind consumers' decisions can yield a better model of consumers' well-being and help make predictions about how changes in the display of multi-attribute consumer goods can affect attention and thus choice.

In economics, a number of studies have examined the computational process used to make strategic decisions. For example, Camerer et al. (1993) and Johnson et al. (2002) showed that subjects often did not look ahead or look at their partners' payoffs in determining offer patterns in bargaining settings. Gabaix et al. (2006) used MouseLab to investigate information acquisition under both time and financial constraints. They found evidence to support a directed cognition model, which assumes that subjects use partially myopic option-value calculations to determine how to search among several options. Knoepfle et al. (2009) and Wang et al. (2010) used eye-tracking to study information transmission and found that visual fixations as well as pupil dilation could be used to predict an otherwise unobservable, private information state. Similarly, Caplin et al. (2011) collected interim choice data to better understand the dynamic search algorithm used by subjects when faced with complex choices. However, none of these studies investigate consumer choice.

Several marketing studies have looked into consumer choice using different types of displays. However, most such studies have been limited to hypothetical choice. For example, Van der Lans et al. (2008) studied how subjects found brand information within a display; however, no choices were made in the experiment. Russo et al. (1975) found a critical role for fixation patterns in multi-alternative, multi-attribute choice with car descriptions, though only hypothetical choices were made. Chandon et al. (2008) analyzed commercial eye-tracking data collected for hypothetical magazine product ads and found that packaging played a more dominant role than pricing in attentional fixations, but the authors did not control for visual saliency or display location. Lohse (1997) used a business phone directory search task and found that display ads were more significantly noticed than plain listings, and viewed color ads more quickly and for longer duration than non-color ads. However, such effects might simply be due to the greater saliency and complexity of larger and more colorful items.

More recently, Krajbich et al. (2010) used eye-tracking to study simple binary choice between snack foods and found strong evidence in favor of a drift-diffusion model (DDM) in which integration is driven by visual attention. In particular, while the fixation process is random with respect to the value of the items, fixations have a considerable effect on value integration in item comparison, through which attentional fixations subsequently bias choice. Krajbich et al. (2011) extended the model to the case of multi-alternative choice. Reutskaja et al. (2011) used a similar eye-tracking paradigm to examine multi-alternative choice under time pressure and found that subjects appear to use a stopping rule to terminate the search process, which leads them to choose items they looked at first as well as more often.

A related class of models to the DDM is decision field theory (DFT), a cognitive, dynamic model of decision making developed by Busemeyer et al. (1993). DFT describes how preferences might evolve over time before a choice is made; in such models, fixations matter for sequential integration across multidimensional items. In DFT, fixations focus the integration of value to a subset of dimensions, unlike in DDMs, in which fixations bias the integration of value in favor of one of the items. The approach encompasses a range of information accumulation models and has been shown to account for a wide range of phenomena, including the relation between choice and decision time as well as preference reversals (Roe et al., 2001; Busemeyer & Diederich, 2002). However, there is little direct testing of the impact of fixations on choice in DFT using eye-tracking.

The above studies have yielded a deeper investigation of the underlying computational processes, which is central to understanding how preferences are built, and, as a result, how choices are made. However, these studies have not addressed how multiple attributes might be integrated when making choices between multiple items, and there is no evidence with real choice demonstrating the effects of display on choice through the attentional fixation process.

In this paper, we disentangle how different attributes of a product, such as visual design or size, are evaluated and integrated when choosing between two similar products. We build on the attentional DDM model put forth by Krajbich et al. (2010), which makes stark predictions about the relationship between attentional fixation patterns and choice. We extend their framework on the attentional DDM model to the domain of multi-attribute choice by examining the computational processes used to search and identify multiple attributes and integrate across these attributes to make a choice. Our experimental design also allows us to test if the visual fixation process depends on display effects. Specifically, we hypothesize that (1) value-independent differences in attention impact multi-attribute choice, and (2) simple changes to the display impact the attentional fixation process and thus choice.

We tested these hypotheses using a two-item, two-attribute choice task to determine how decision-making processes are modified by the introduction of multiple discrete attributes for each item. We first elicited subjects' valuations over poster designs and sizes. Subjects were then presented with pairs of posters, each of a different design and size, and asked to choose the more desirable one. Attributes of each choice were arranged in two visual conditions: designs on top, and sizes on top. MouseLab was used to gather detailed search-process data. Our experimental design allowed us to test if the choice process is influenced by display effects—in other words, whether the location of an item in the display affects its probability of selection. First, we found evidence to support the hypothesis that value-independent attentional fixation processes play a key role in integrating across multi-attribute goods in consumer choice. Second, we found that simple display changes resulted in striking differences in visual search patterns, which subsequently impacted choice through differential weighting and integration of the attributes.

METHODS

Participants. Seventy-four participants completed the experiment (age: mean = 33.03, SD = 8.28; 71% female), which was administered over the internet as part of a larger online survey. Participants were recruited from an online pool registered with the Center for Decision Sciences at Columbia University. Each participant completed the entire survey and received compensation for their participation.

Stimuli and recording. Participants viewed images of seven artistic posters (sourced from zazzle.com) displayed on their computer screens. Stimulus presentation and response recording were controlled using the MouseLab Web process tracing tool (Payne et al., 1993).

Task. Participants received extensive training in the use of MouseLab at the beginning of the session. Participants were then instructed at the start of the experiment that they would make a series of choices between pairs of artistic posters of different designs and sizes. To incentivize participants to take their choices seriously, they were told that they would be mailed a poster based upon the choice made during one randomly-selected trial. A copy of the instructions is included in the Appendix.

Each participant then performed two rating tasks. In the first task they viewed and rated seven poster designs. On each trial one poster design was displayed on-screen, and participants rated the design on a scale from 1 (do not like it at all) to 7 (like it very much) by clicking an appropriate radio button below the design. Once participants had entered a rating and pressed the “continue” button, the trial ended. Participants had unlimited time to make each rating. Each of the seven poster designs was presented once, with order of presentation randomized across participants.

In the second rating task participants viewed and rated three poster sizes. Before the task began participants were shown a relative comparison of the sizes. Three solid gray boxes displayed side-by-side (size increasing left-to-right) depicted the relative sizes of the poster types. Above each box a label indicated the actual size of each poster type: "portfolio (circa 15'' x 11'')"; "small (circa 20'' x 15'')"; and "large (circa 30'' x 23'')." Participants were given unlimited time to examine the screen before beginning the rating task. On each trial of the task one of the solid gray boxes from the pre-task screen was displayed along with its corresponding size label. Participants rated the size on a scale from 1 (do not like it at all) to 7 (like it very much) by clicking an appropriate radio button below the rectangle. The trial ended and choices were recorded when the participant pressed the "continue" button. Participants had unlimited time to make each rating. Each of the three sizes was presented once, with order of presentation randomized across participants.

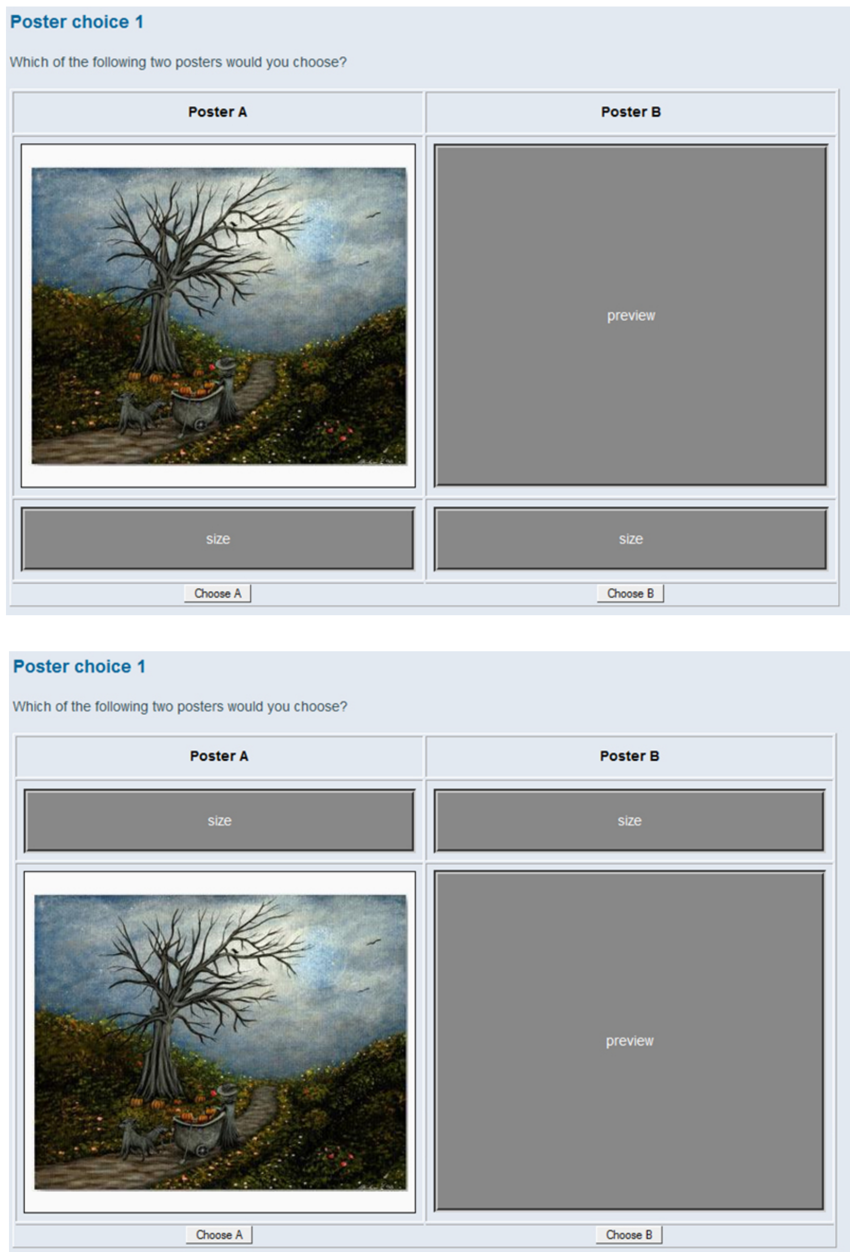
Participants then performed the main choice task, consisting of 30 trials. On each trial the participant chose between a pair of posters. Every poster was composed of two attributes: a design and a size. Attributes of one poster were displayed top to bottom on the left size of the screen, while attributes of the other were displayed in the same order on the right side. Display order was counterbalanced across participants: half viewed designs on top on all trials (C1; Figure 1a), while the other half viewed sizes on top (C2; Figure 1b).

Poster pairs were constructed by combining every unordered pair of the participant's five highest-rated designs (ties broken by random selection) with every unordered pair of sizes, resulting in 10 * 3 = 30 distinct unordered poster pairs. Order of presentation of the pairs was then randomized to construct 30 trials. On each trial the display order (left-to-right) of the designs was randomized. The pairing of each design and size was chosen to ensure non-dominated choices (e.g., so that the higher-rated design was not also paired with the higher-rated size)—the higher-rated design was paired with the lower-rated size, and vice versa. In cases of equally-rated designs or sizes, the presentation of sizes was determined by random selection.

In order to track search processes, each attribute was hidden behind a solid gray box that revealed its contents only while the participant held the mouse cursor over the box. Boxes were labeled with the attribute revealed by opening them. To prevent selection bias due to initial positioning of the mouse cursor, participants began each trial by clicking on a cross positioned between the four attribute boxes at the center of the display. Participants had unlimited time to complete each trial. The trial ended and the choice was recorded when the participant pressed the "choose" button beneath a poster. The locations, timings, and durations of each "fixation" (defined by the opening and closing of a box) and the ultimate choice were recorded by MouseLab Web.

After the choice task was completed, one trial was randomly selected and the poster selected on that trial was mailed to the participant.

Fig. 1. (a) Sample choice screen for condition 1 (C1), with poster designs on top. **(b)** Sample choice screen for condition 2 (C2), with poster sizes on top.



Data analysis. As the study was administered over the internet, we could not directly monitor participants' activity to ensure compliance with instructions or ensure that participants remained attentive during every trial. We used a data-cleaning procedure screen for improperly completed trials and inattentive participants.

To screen for accidental fixations caused by moving the mouse briefly across one box to reach another, we discarded any fixations with a duration less than 100 ms. We chose a low threshold to ensure that such fixations were truly accidental and not part of the participant's search pattern and that the subject gained no information about the stimuli even through the short-duration fixation.

To ensure that participants remained attentive throughout each trial, we employed several tests to screen for anomalous trials. First, we discarded all trials with reaction times more than 4 SD above or below the mean across all participants (mean = 6.53 s, SD = 4.12 s). Second, we discarded all trials in which the average fixation duration was more than 4 SD above the mean across all participants (mean = 0.735 s, SD = 0.242 s). Third, we discarded all trials in which the average transition time between fixations was more than 4 SD above the mean of across all participants (mean = 0.105 s, SD = 0.103 s). These tests were collectively designed to screen for trials in which participants were not attentive to the task, as signaled by an excessively fast choice (suggesting that the choice was not fully considered and was made haphazardly) or excessively slow performance during the trial (suggesting that the participants became distracted during performance of the trial or walked away from the computer). As an additional safeguard, we discarded any subject who successfully completed fewer than 15 trials (half of the total) as determined by the data-cleaning procedures.

As a result of the cleaning, we discarded an average of 1.4 trials per participant (M = 1.43, SD = 2.39). One participant was eliminated entirely.

Finally, sometimes subjects returned to the same area of fixation immediately following a fixation; such fixations were treated as one fixation with a duration equal to the sum of the two separate fixations.

Theoretical model. The theoretical model developed here extends the attentional DDM put forth by Krajbich et al. (2010) to the domain of multi-attribute choice. Their model takes as a framework DDMs of binary response selection (Stone, 1960; Ratcliff, 1978) that have proved to be quite accurate in describing a range of perceptual decision making data (Gold et al., 2001; Gold et al., 2002; Mazurek et al., 2003). The key idea behind such models as applied to binary choice is that the stochastic evidence for one response (the relative decision value, or RDV) is accumulated over time until the integrated evidence passes a decision threshold, and a choice is made. There is thus a tradeoff between the benefit of accumulating more information and the cost of taking more time to make a choice. Krajbich et al. (2010) apply this type of model to the domain of binary choice for simple snack items and find that it explains many relationships between choice and fixation patterns.

Following Krajbich et al. (2010), our model assumes that an RDV, which evolves over time as a Markov Gaussian process, is computed until a choice is made. In our case, we have a choice between two posters with two attributes each, design and size. The RDV starts each trial at 0 and continually evolves over time at one of two possible rates, μ_d, μ_s (both in units of ms^{-1}), depending on which attribute is fixated, design or size, respectively. A choice is made when a threshold is reached at either +1 or -1. If the RDV reaches the +1 threshold, the left item is chosen; if the RDV reaches the -1 threshold, the right item is chosen. We assume that consumers have independent valuations over designs and sizes that can be elicited through prior ratings, just as in the single-attribute case. These valuations are termed $r_{d1}, r_{d2}, r_{s1}, r_{s2}$, representing the two design ratings and the two size ratings for the items on the screen for any trial. Importantly, the slope with which the RDV evolves depends on the fixation location at that instant. The slope is proportional to the weighted difference between the values of the fixated and non-fixated items and is given by the following expression:

$$f(r_{d1}, r_{s1}, r_{d2}, r_{s2}, E_t) = \alpha_{d1t} r_{d1} + \alpha_{s1t} r_{s1} - \alpha_{d2t} r_{d2} - \alpha_{s2t} r_{s2} \qquad (1)$$

The coefficients of this expression are given by:

$$\alpha_{ijt} = \begin{cases} \mu_j, & f_{ij} = 1 \\ \mu_j \theta, & f_{i,3-j} = 1 \\ 0, & \text{otherwise} \end{cases} \qquad (2)$$

where f_{ij} is an indicator variable for fixation on the *i*th attribute of the *j*th item, with $i \in \{d, s\}$ and $j \in \{1,2\}$ and θ between 0 and 1 is the bias term toward the fixated option.

In words, this model gives full weight to the rating of the fixated attribute and item, a reduced weight to the fixated attribute of the non-fixated item (with the bias term fixed across attributes for simplicity), and no weight to the non-fixated attribute. When the consumer is looking at a given item, the RDV changes according to:

$$V_{t+1} = V_t + \begin{cases} \mu_d(r_{d1} - \theta\, r_{d2}), \text{current fix} = d_1 \\ -\mu_d(r_{d2} - \theta\, r_{d1}), \text{current fix} = d_2 \\ \mu_s(r_{s1} - \theta\, r_{s2}), \text{current fix} = s_1 \\ -\mu_s(r_{s2} - \theta\, r_{s1}), \text{current fix} = s_2 \end{cases} + N(0, \sigma^2) \tag{3}$$

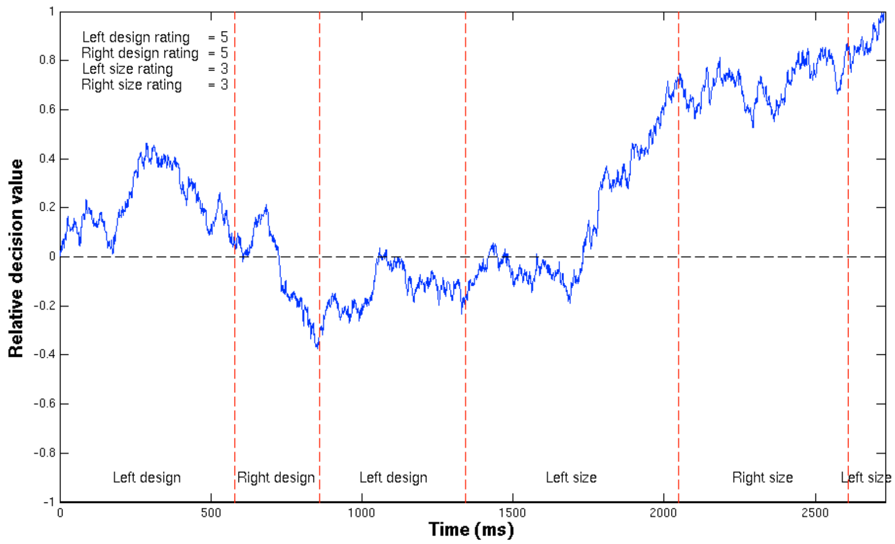
where V_t is the value of the RDV at time t and σ^2 is the variance. Note that there may be other possible functional forms for the value of the RDV. Here we use the above model as a starting point to derive qualitative predictions, with the idea that potential refinements can be made in the future.

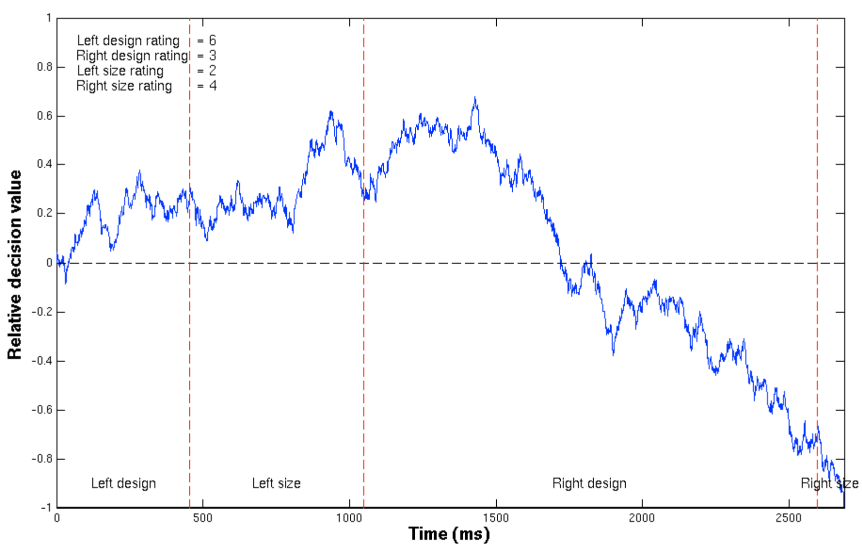
With respect to the fixation process, we adopt the same treatment of fixation lengths as in the basic model in Krajbich et al. (2010), sampling from an empirical distribution of first and middle fixations instead of attempting to model fixation length explicitly. We also condition the fixation distributions on the fixated attribute, as it is natural to suppose that fixation lengths may be different for different attributes. The order of fixations, however, is more complex in the multi-attribute case. We have four distinct fixation locations and three possible transitions from one fixation to the next (with four possible transitions for the first fixation). For the first fixation, as in the basic model, we simply use an empirical distribution of first fixation locations. For subsequent fixation patterns, we could assume that the transition process is an exogenous Markov chain with given probabilities for a next transition conditional on the current fixation location. However, empirical data shows that transitions occur in more complex ways, so we take the transition process to be exogenous but conditional on the entire past history of fixations for that trial. For example, the probability of transitioning from d_l to s_l changes based on the pattern of fixations up to that point. In order to avoid sparse search patterns in the data, after the sixth fixation, we conditioned the search tree only on the location of the immediately previous fixation.

We thus have a four-parameter model with $(\theta, \mu_d, \mu_s, \sigma^2)$ that we can simulate using the Bayesian updating procedure described above. Note that as we have a low number of trials per subject and thus relatively high variation across subjects, our data is not particularly well suited to fitting this type of model. Instead, we simulate the model based on plausible parameter values of $(\theta, \mu_d, \mu_s, \sigma^2) = (0.6, 0.0001, 0.00001, 0.015)$. In simulated runs, the RDV generally moved toward the fixated item, but the slope depended on the values of the designs and sizes (Fig. 1c, d). For example, the RDV signal integrated toward the right item barrier when the subject fixated on the right design, even though it had a lower value than the left design (Fig. 1d). This suggests that, as in simple binary choice, visual fixations are also important for the integration process in multi-attribute binary choice.

Fig. 1. (c, d) Simulated runs of the model using

$(\theta, \mu_d, \mu_s, \sigma^2) = (0.6, 0.0001, 0.00001, 0.015)$.





RESULTS

Basic psychometrics. As predicted by the model, the choice data indicated that items that were rated more highly were more likely to be chosen across both display conditions (Fig. 2a). Note that the plot shows only design, and not size, ratings. Due to the design ratings' high correlation with size ratings (see Supplementary Materials), size ratings were excluded in the relevant figures to avoid dominance.

A logit regression of differences in poster and size ratings on choice shows that choices are a logistic function of the ratings and size in both conditions (design coefficient = 1.29, $p = 0.0002$ for C1; 2.02, $p = 0.0000$ for C2; size coefficient = 0.27, $p = 0.052$ for C1, 0.54, $p = 0.0002$ for C2; Fig. 2b). There were no significant differences between C1 and C2 in poster design ($p=0.053$) or size ratings ($p=0.14$). Within conditions, however, the difference between poster and size rating effect was highly significant ($p = 0.0026$ for C1 and $p = 0.0000$ for C2), suggesting subjects place more weight on poster design compared to size.

We also examined response time and number of fixations per trial. Response time was correlated with difficulty, or the difference in poster design ratings (linear regression coefficient = -537.9, $p = 0.0019$ for C1, and -475.7, $p = 0.0001$ for C2, respectively; Fig. 2c), and there were no significant differences between the two conditions ($p = 0.75$). The same was true of the number of fixations per trial (linear regression coefficient = -0.46, $p = 0.0047$ for C1, and -0.47, $p = 0.0001$ for C2, respectively; p-value between C1 and C2 =

0.95; Fig. 2d). This correlation between both response time and number of fixations and difficulty is a common property of drift-diffusion models.

Fig. 2. (a) Choice curve as a function of the difference in poster design ratings.

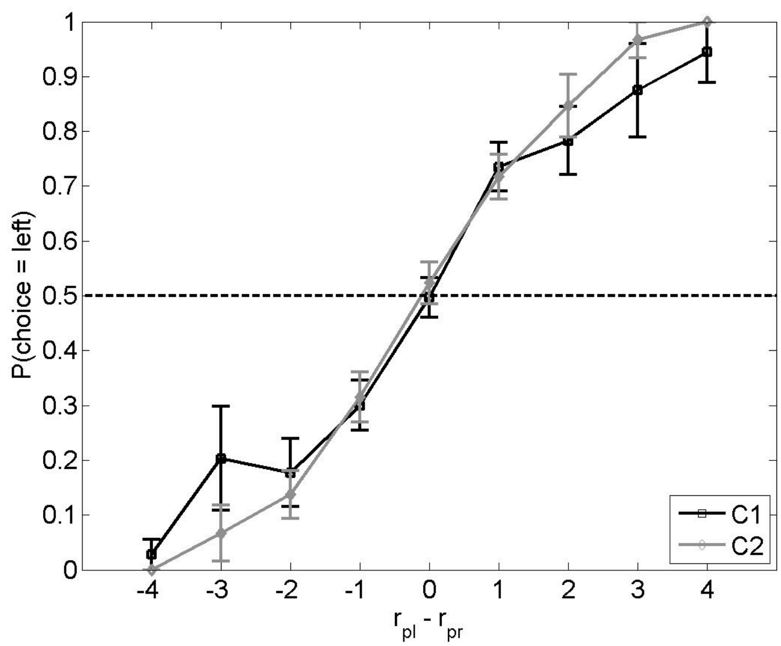


Fig. 2. (b) Logit regression coefficients of choice vs. constant and differences in poster design and size ratings. **(c)** Response time as a function of the difference in poster design ratings.

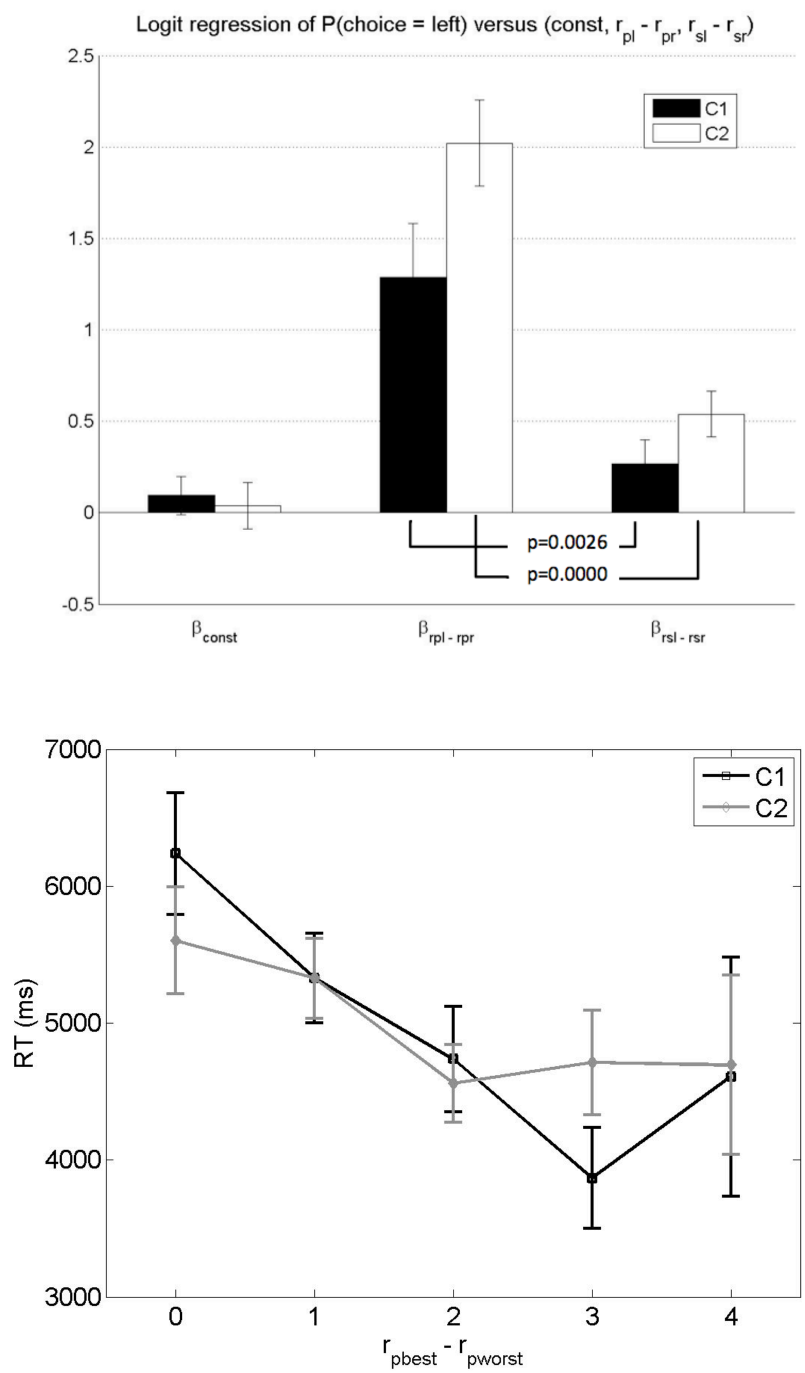
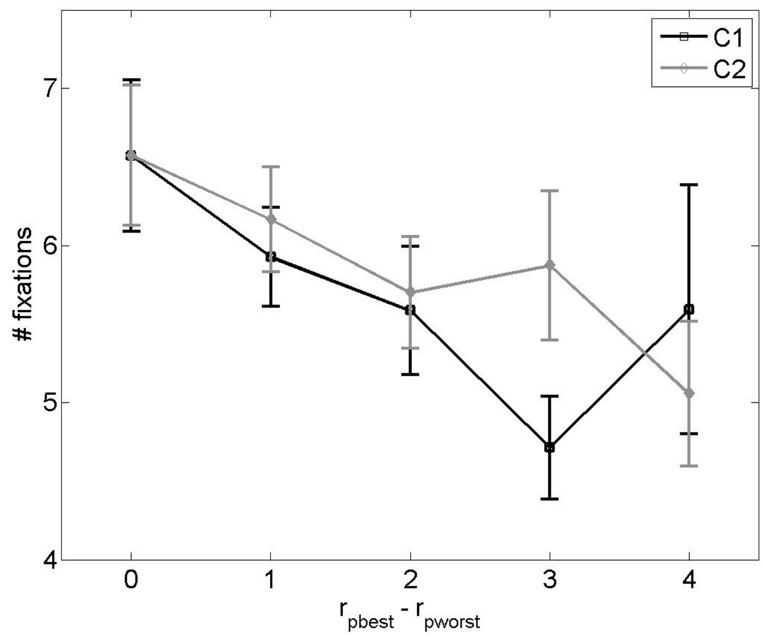


Fig. 2. (d) Number of fixations as a function of the difference in poster design ratings.



Choice biases. The differential impact of display on search seen in the data is important because search impacts value integration and thus choice. The attentional DDM also predicts that the decision process should show several choice biases. First, the model predicts a last-fixation bias: subject should be more likely to choose an item if their last fixation is to that item, a result of the fact that the value of the non-fixated item is discounted. As predicted, we find that the last fixation exhibits a choice bias towards the last fixation location (logit regression slope = 0.58, $p = 0.0064$ for C1, 1.26, $p = 0.0001$ for C2; Fig. 3a). There is a significant difference in the slopes of C1 and C2 ($p = 0.037$).

Second, the model predicts that final fixations should be shorter than middle fixations since fixations are interrupted when a decision threshold is reached. We find that, consistent with the model predictions, last fixations are shorter compared to middle fixations for both design and size in C1, though not significantly for C2 (for design: $p = 0.0016$ for C1, $p = 0.29$ for C2; for size: $p = 0.0000$ for C1, $p = 0.069$ for C2; Fig. 3b). In addition, there is a significant difference between the two conditions for last fixation duration to poster design ($p = 0.022$; Fig. 3c), but not size ($p = 0.91$). This result provides some evidence that the integration aspect only occurs during the viewing of complex stimuli, like poster designs.

Finally, the model predicts that there should be a choice bias due to relative increased looking time for one item over the other. As predicted, Fig. 3d shows the bias in choosing the relatively more viewed item as a function of the rating difference (logistic regression slope = 0.25, $p = 0.084$ for C1, 0.17, $p = 0.013$ for C2).

In addition, the results show a large overall differential in attention to design vs. size—poster designs are looked at longer than size in both conditions (p=0.0000 for C1, p=0.0000 for C2; Fig. 3e). There are no significant differences in latencies between the two conditions when looking at poster design (p=0.1801) and size (p = 0.2887). This differential attention to designs is consistent with the larger weights that design ratings have in the logit regression described in Fig. 2b.

Fig. 3. (a) Choice bias for the last fixation as a function of the difference in poster design ratings.

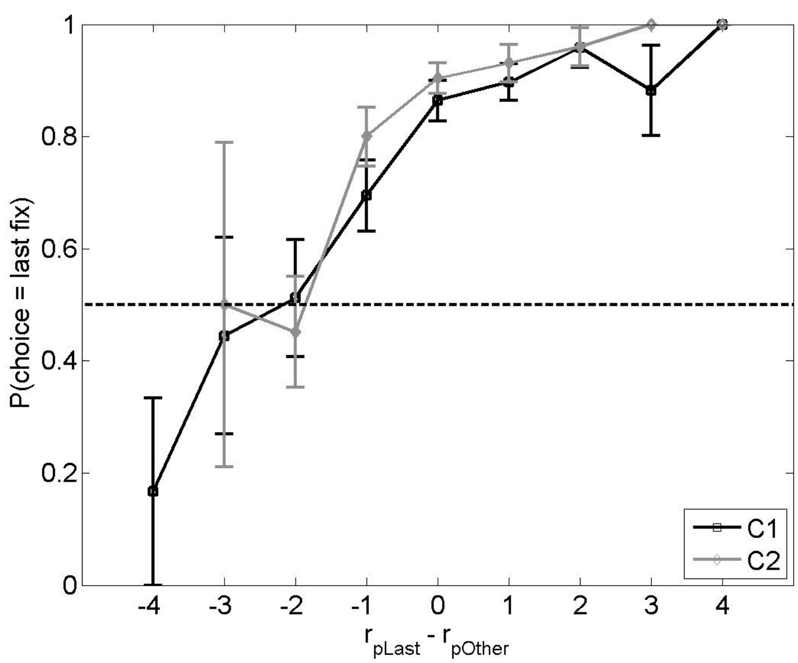
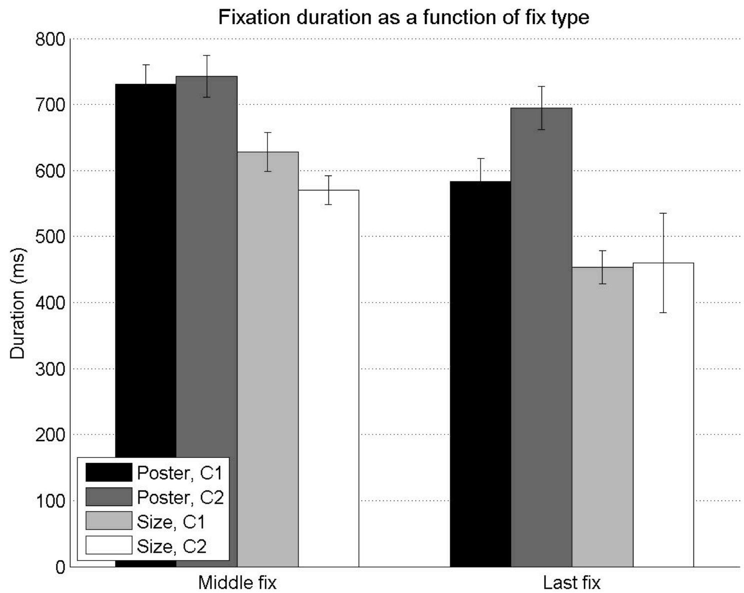


Fig. 3. (b) Fixation duration as a function of fixation type. **(c)** Choice bias for the last fixation to a poster as a function of the difference in poster design ratings.



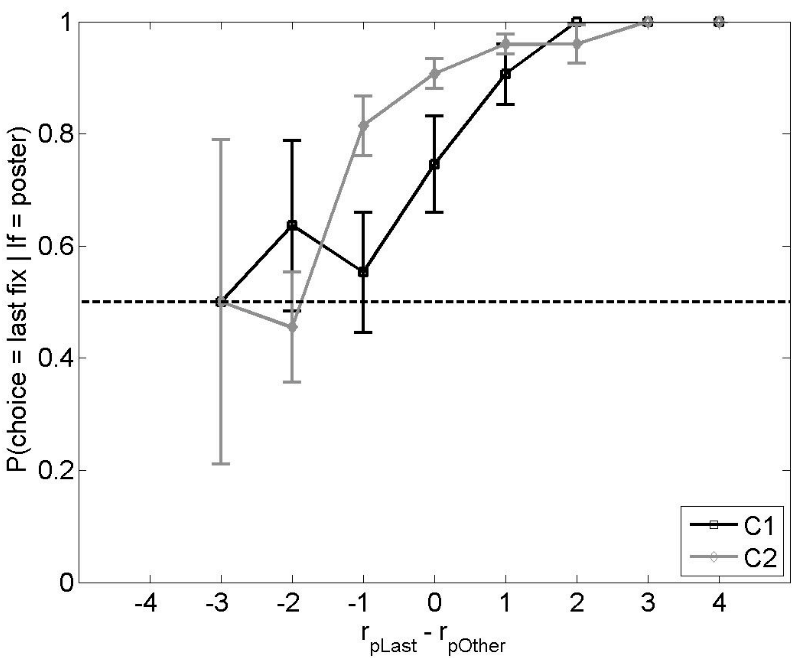
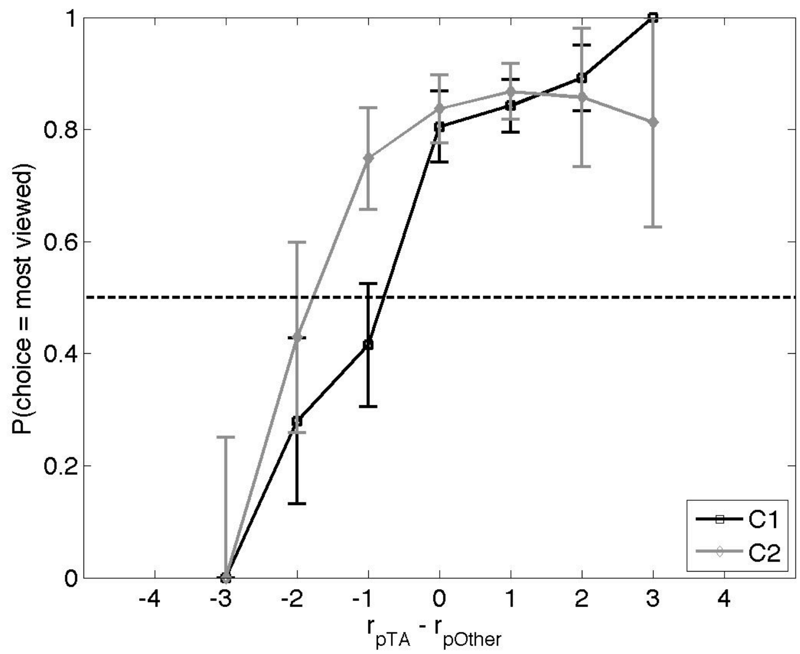
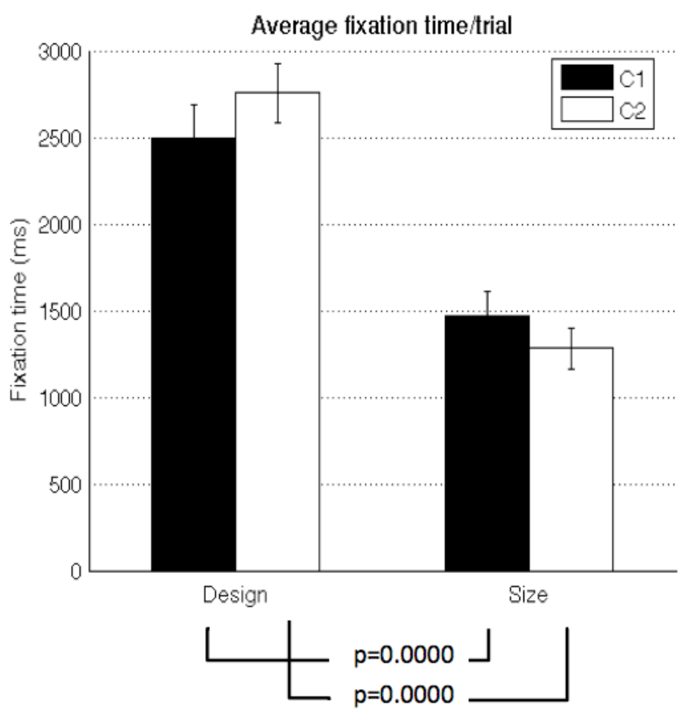


Fig. 3. (d) Choice bias for the more viewed item as a function of the difference in poster design ratings. **(e)** Average fixation time by attribute and condition.





Properties of the general search process. However, the choice biases do not extend to the general nature of the fixation process, which is independent of underlying value. To rule out that our effect is simply due to subjects paying more attention to preferred items, we examine whether the time spent looking at items is a function of their value across all but the last fixation. Fig. 4a-b shows that fixation durations are independent of the value of the fixated design for C1 (Fig. 4a; mixed effects regression estimate: -9.8, p=0.47), though not for C2 (Fig. 4a; mixed effects regression estimate: 40.8, p=0.0028) and size for C1 (Fig. 4b, mixed effects regression estimate: -1.4, p=0.93), though not for C2 (Fig. 4b, mixed effects regression estimate: 26.4, p=0.019). However, note that while the relationships for C2 are significant, they are extremely small and cannot account for the effect.

The fixation duration does depend slightly on the difference in value between the fixated and non-fixated design for C2 (Fig. 4c, mixed effects regression estimate: 20.4, p=0.010), but not for C1 (Fig. 4c, mixed effects regression estimate: -16.8, p=0.067) and size for C2 (Fig. 4d, mixed effects regression estimate: 19.6, p=0.053), but not for C1 (Fig. 4d, mixed effects regression estimate: 4.9, p=0.55), though, again, these effects are extremely small. Fixation duration depended more strongly on the difficulty of the decision for poster design in both conditions (Fig. 4e, mixed effects regression estimate: -45.5, p=0.0053 for C1; -41.5, p=0.0096 for C2) but not for size (Fig. 4f, mixed effects regression estimate: -1.7, p=0.90 for C1; -11.4, p=0.52 for C2). Note that as the last fixation displays a choice bias, the last fixation has been discarded here.

Fig. 4. For middle fixations, fixation duration as a function of **(a)** poster rating, **(b)** size rating.

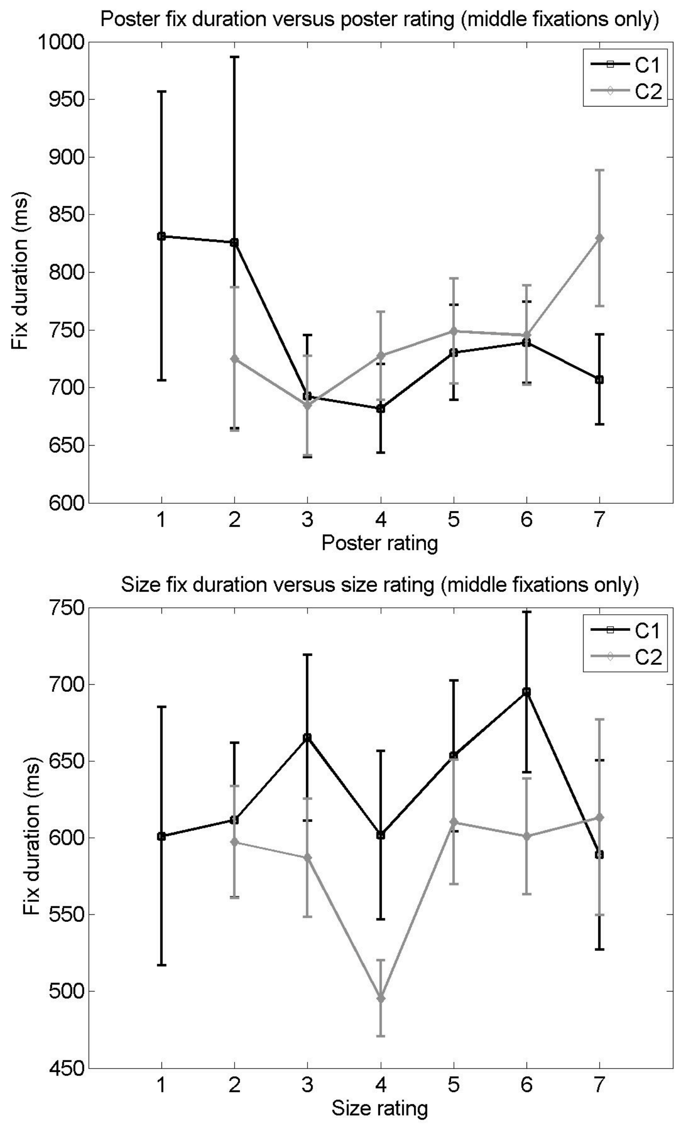


Fig. 4. For middle fixations, fixation duration as a function of **(c)** difference in design rating between fixated and non-fixated item, **(d)** difference in size rating between fixated and non-fixated item.

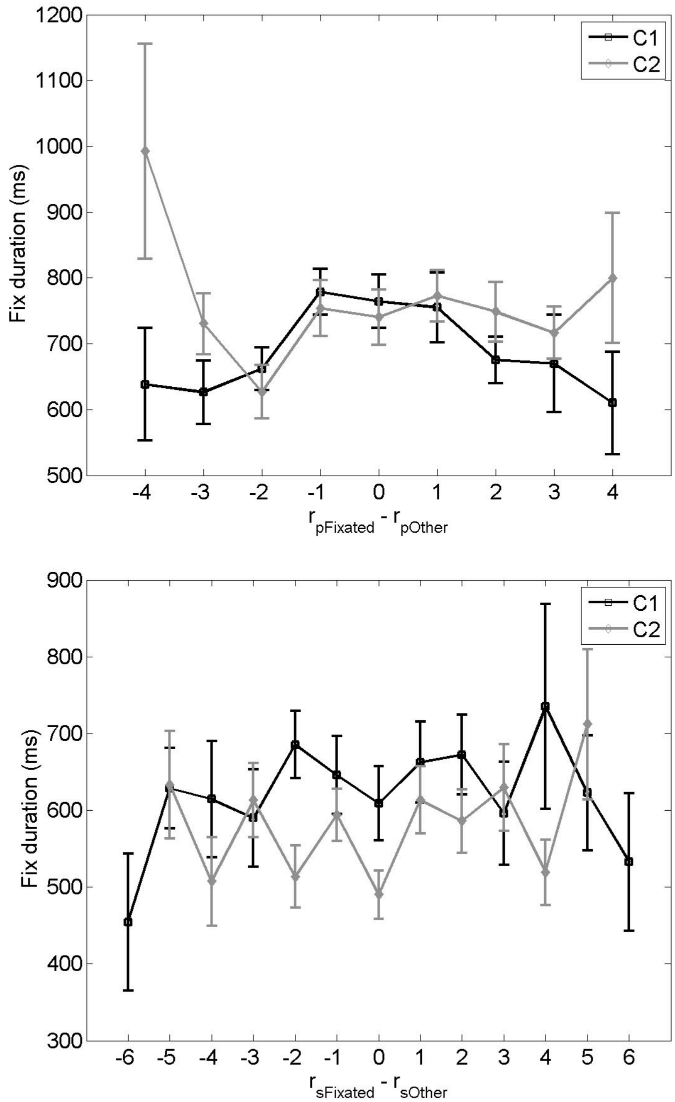
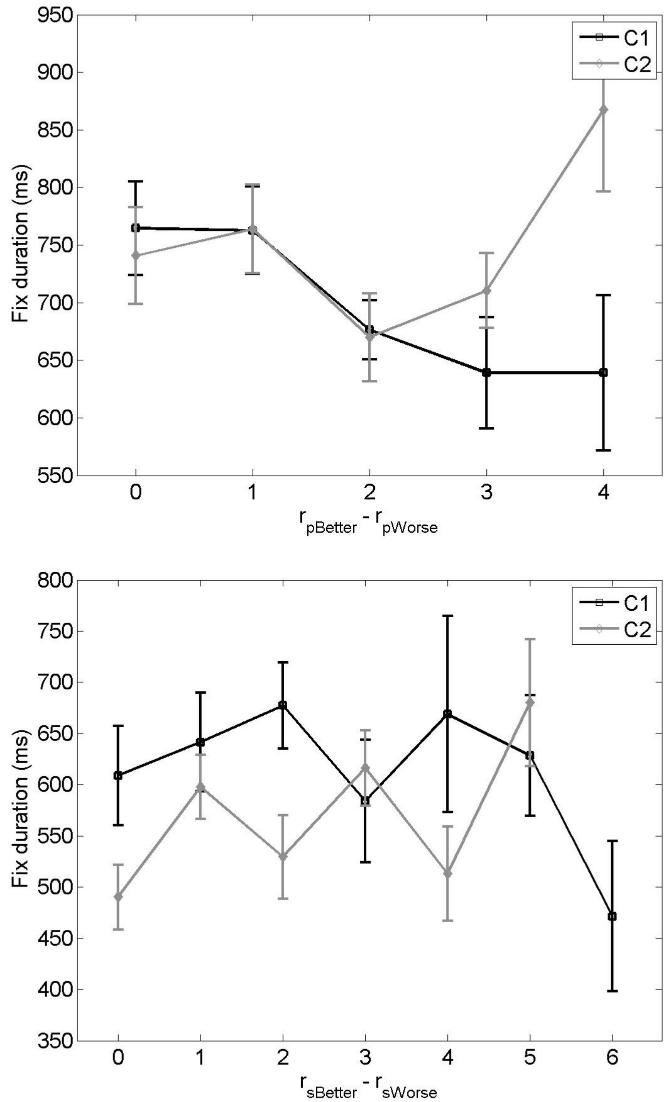


Fig. 4. For middle fixations, fixation duration as a function of **(e)** difference in design rating between better and worse item, **(f)** difference in size rating between better and worse item.



Display configuration strongly impacts search. More striking differences between the two display conditions were found upon examining the search data. Fig. 5a shows the difference in fixation location between the two conditions as a function of fixation number. Poster designs are fixated earlier on in C1, and the same is true for poster sizes in C2. Similarly, a plot of the transitions between design and size in the two conditions suggests that subjects in C1 are more likely to search within the two attributes of one item before looking at the other item, while there is no such clear pattern for C2 (Fig. 5b).

Figures 6a and 6b report fixation transition probabilities or "search trees" with branches trimmed to reflect only the most-traversed paths. Search trees showing the modal search patterns in C1 (Fig. 6a) and C2 (Fig. 6b) show evidence for path-dependent search behavior that varies dramatically between conditions. In C1, over 97% of first fixations are to the poster design, while this is true of only 57% in C2. In C1, there is a very clear search pattern of design, design, size, size. In C2, however, the modal search path is much less clear: there appears to be a multiplicity of patterns. Figures 6c-d show the same data in a different format: the darker squares indicate more likely fixation patterns. While Fig. 6c demonstrates a clear search pattern, Fig. 6d is much more varied. The results show that attention appears to be driven by the latent values of stimuli and by top-down processes: search is not random.

We see a similar, striking pattern when looking at the distribution of last fixations between the two display conditions (Fig. 7a). Subjects in C1 are much more likely to look at size, while subjects in C2 are much more likely to look at design. Moreover, this trend is constant regardless of the total number of fixations in the trial (Fig. S2a-d).

Fig. 5. (a) Probability of fixation to poster design as a function of fixation number. **(b)** Probability of transition to an area given current area of fixation.

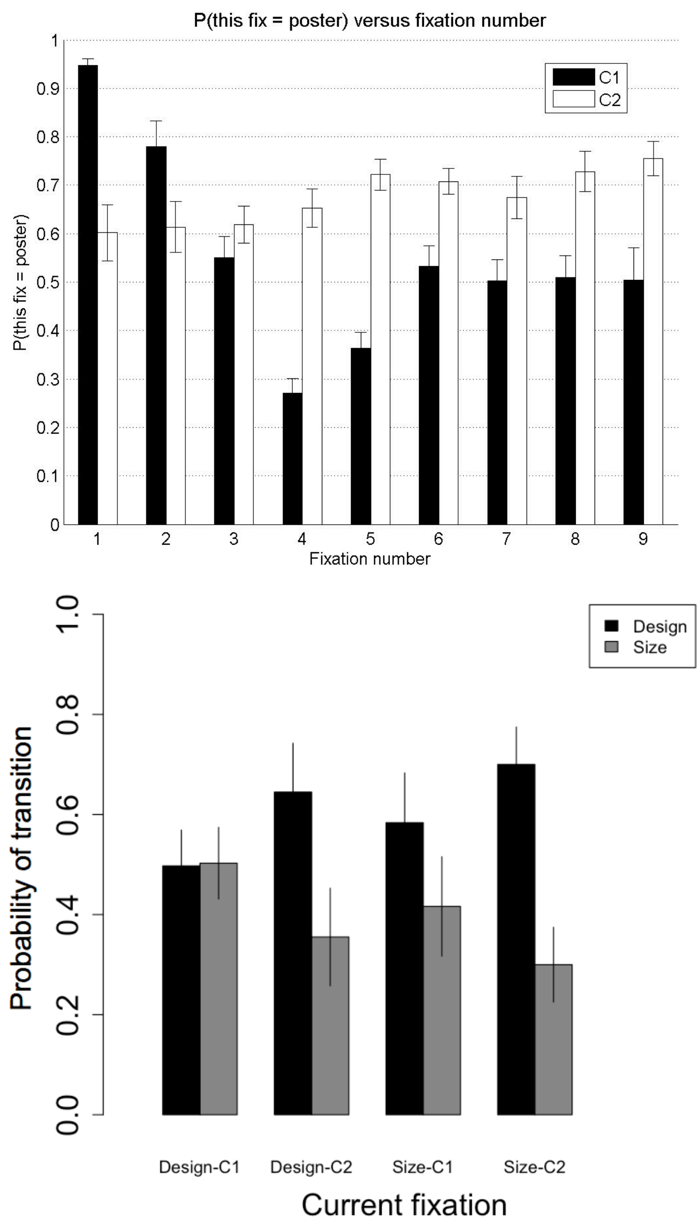


Fig. 6. (a) Search tree displaying modal search patterns for C1. **(b)** Search tree displaying modal search patterns for C2.

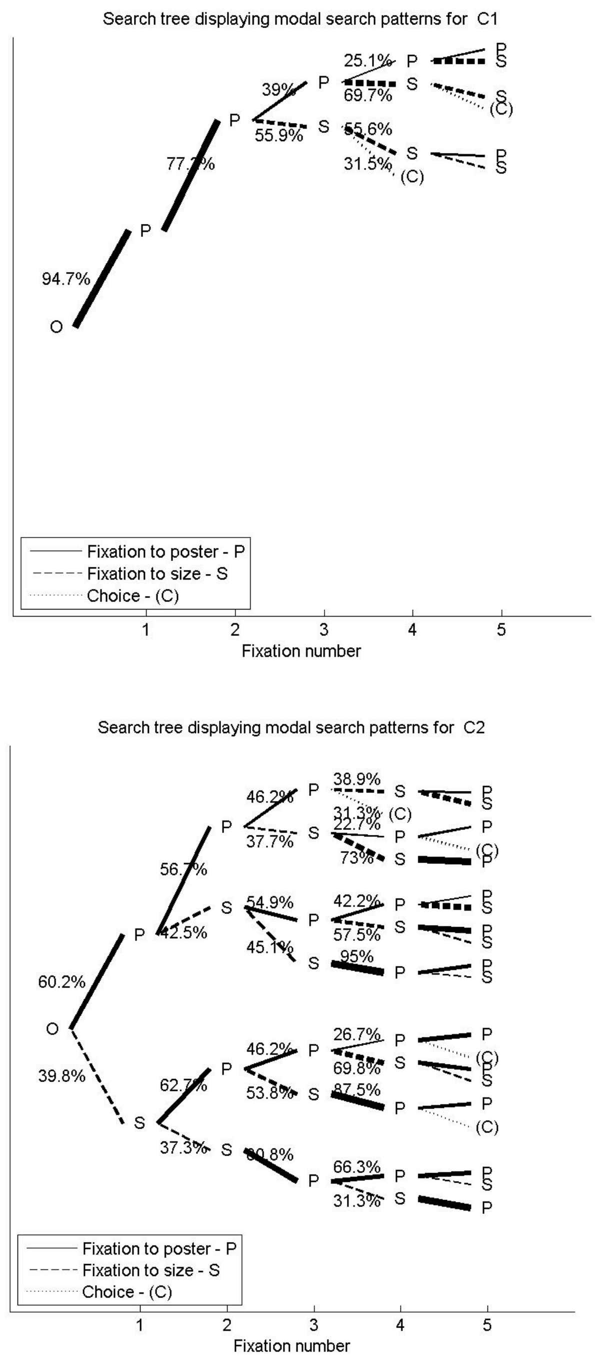
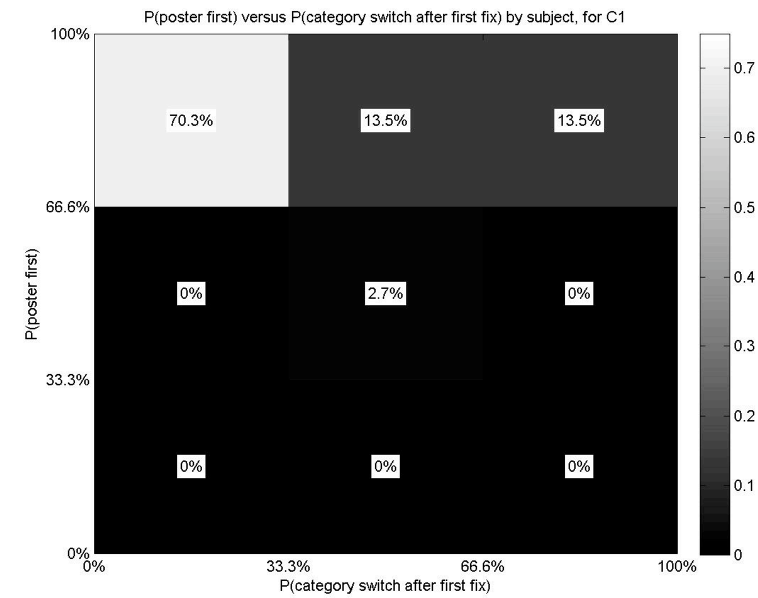


Fig. 6. (c) Heat map of the probability of looking at the poster design first as a function of the probability of switching attribute category after the first fixation for C1. **(d)** Heat map of the probability of looking at the poster design first as a function of the probability of switching attribute category after the first fixation for C2.



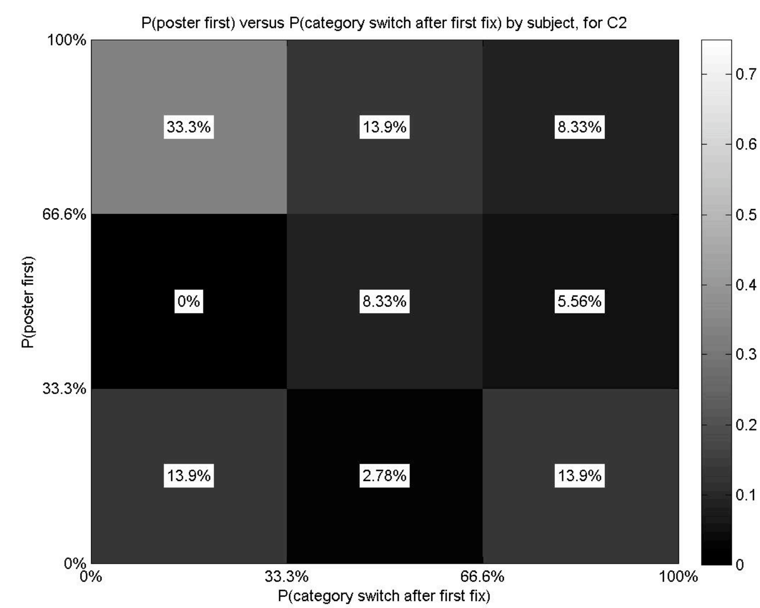
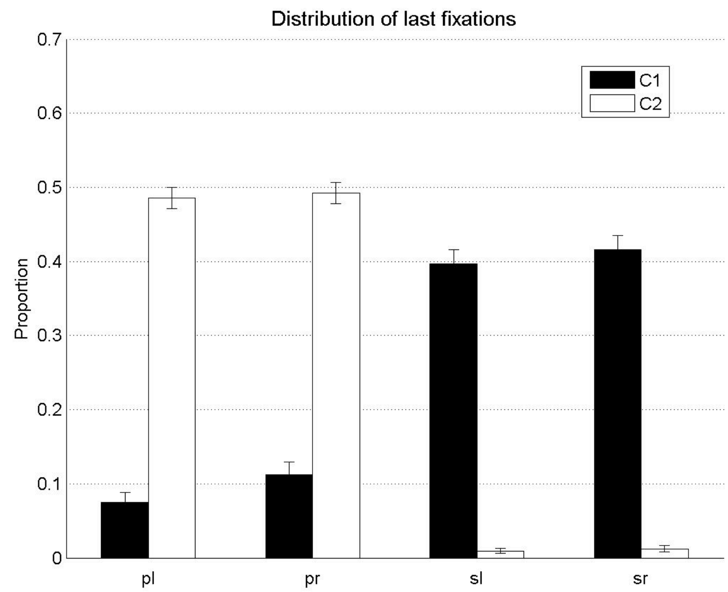


Fig. 7. (a) Distribution of the last fixation by attribute and condition.



DISCUSSION

Our results provide insight into the nature of the computational processes that guide multi-attribute, binary choice. We found evidence to support an extension of the attentional DDM to multi-attribute choice, in which visual fixations play a key role in the value integration process. In particular, we found that value-independent differences in attention impact multi-attribute choice. Moreover, we found that simple changes to the display impact the attention fixation process and thus choice, which suggests that the visual fixation process may play a causal role in the value integration process. This supports the results of related studies that have shown that manipulating attentional fixations can bias choice (Shimojo et al., 2003; Armel et al., 2008).

It is important to emphasize that values do have an effect on the pattern of fixations. For example, as shown in Fig. 2c-d, the response time and number of fixations increase with choice difficulty. However, our display manipulation results also suggest that small variations in attentional fixations can affect the choice process in a significant way. C1 might be interpreted as a more "natural" display condition, as the poster designs were found to be a far more significant determinant of choice than the sizes. Subjects may therefore have habitually expected the designs to be displayed at the top of the screen, with the more minor size information displayed below. This would be analogous to what we see in many consumer product scenarios: prices often appear underneath products, and product features (e.g., a car or a jewelry piece) are often listed in order of importance. The violation of this expectation by displaying sizes on top in C2 may have led to idiosyncratic behavioral reactions by different subjects: for instance, the wide dispersion in search patterns that we observed.

Our results have important implications for understanding and influencing choice. Our findings provide new insight for economic theory by demonstrating that exogenous factors that are uncorrelated with value, such as location, can affect the information integration process and, through it, choice. As a result, systematic biases in attentional fixations could lead to deficits in decision-making. Modeling the visual fixation process could be used to better understand and correct for such deficits. Moreover, such biases have a number of applications in business and marketing, including store display arrangement, product attribute emphasis or the organization of features on a website. Higher-placed attributes for a given item will receive more attention and thus greater weighting in an overall value for the item. This explains why prices are typically much less salient in location and visual aspect compared to other, more attractive attributes.

Further research extending our study to multi-item choice with a larger number of attributes would yield further insights and confirm whether the biases present in our data generalize to more complex settings. In addition, such an extension would be a more realistic approximation to the types of complex decisions that consumers actually face. In addition, our results do not rule out alternative models of how attentional fixations might interact with choice. A systematic investigation and comparison of other models is another important topic for future research.

REFERENCES

Armel, K.C., Beaumel, A. and Rangel, A. (2008) Biasing simple choices by manipulating relative visual attention. *Journal of Judgment and Decision Making* 3, 396-403.

Busemeyer, J. R. and Diederich, A. (2002). Survey of decision field theory. *Mathematical Social Sciences, 43*, 345-370.

Busemeyer, J. R. and Townsend, J. T. (1993). Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review, 100*(3), 432-459.

Camerer, C.F., Johnson, E.J., Rymon, T. and Sen, S. (1993) "Cognition and framing in sequential bargaining for gains and losses," in *Frontiers of Game Theory*, ed. Ken Binmore, Alan Kirman, and Piero Tani, 27-47. Cambridge, MA: MIT Press.

Caplin, A., Dean, M. and Martin, M. (2011) Search and satisficing. *American Economic Review* 101 (7): 2899-2922.

Chandon, P., Hutchinson, W., Bradlow, E. and Young, S. (2008) Measuring the value of point-of-purchase marketing with commercial eye-tracking data. In *Visual Marketing: From Attention to Action*, ed. Michel Wedel and Rik Pieters, 225-258. New York: Lawrence Erlbaum Associates.

Gabaix, X., Laibson, D., Moloche, G. and Weinberg, S. (2006) Costly information acquisition: experimental analysis of a boundedly rational model. *American Economic Review* 96 (4): 1043-1068.

Gold, J.I. and Shadlen, M.N. (2001) Neural computations that underlie decisions about sensory stimuli. *Trends Cogn. Sci.* 5, 10-16.

Gold, J.I., and Shadlen, M.N. (2002) Banburisms and the brain: decoding the relationship between sensory stimuli, decisions and reward. *Neuron* 36, 299-308.

Johnson, E., Camerer, C., Sen, S. and Rymon, T. (2002) Detecting failures of backward induction: monitoring information search in sequential bargaining. *Journal of Economic Theory* 104 (1): 16-47.

Knoepfle, D.T., Wang, J.T. and Camerer, C. (2009) Studying learning in games using eye-tracking. *Journal of the European Economic Association* 7 (2-3): 388-98.

Krajbich, I., Armel, C. and Rangel, A. (2010) Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience* 13 (10): 1292-1298.

Krajbich, I. and Rangel, A. (2011) Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proceedings of the National Academy of Science* 108 (33): 13852-13857.

Lohse, G. (1997) Consumer eye movement patterns on yellow pages in advertising. *Journal of Advertising* 26 (1): 61-73.

Mazurek, M.E., Roitman, J.D., Ditterich, J. and Shadlen, M.N. (2003) A role for neural integrators in perceptual decision making. *Cerebral Cortex* 13, 1257-1269.

Payne, J.W., Bettman, J.R. and Johnson, E.J. (1993) *The Adaptive decision maker*. New York: Cambridge University Press.

Ratcliff, R. (1978) A theory of memory retrieval. *Psychol. Rev.* 85, 59-108.

Reutskaja, E., Nagel, R., Camerer, C.F. and Rangel, A. (2011) Search dynamics in consumer choice under time pressure: an eye-tracking study. *American Economic Review* 101 (2): 900-926.

Roe, R.M., Busemeyer, J.R. and Townsend, J.T. (2001) Multialternative decision field theory: a dynamic connectionist model of decision making. *Psychological Review* 108 (2): 370-392.

Russo, E. and Rosen, L. (1975) An eye-fixation analysis of multi-alternative choice. *Memory and Cognition* 3 (3): 267-276.

Shimojo, S., Simion, C., Shimojo, E. and Sheier, C. (2003) Gaze bias both reflects and influences preference. *Nature Neuroscience* 6, 1317-1322.

Stone, M. (1960) Models for choice-reaction time. *Psychometrika* 25, 251-260.

Van der Lans, R., Pieters, R. and Wedel, M. (2008) Competitive brand salience. *Marketing Science* 7 (5): 922-931.

Wang, J.T., Spezio, M. and Camerer, C. (2010) Pinocchio's pupil: using eye-tracking and pupil dilation to understand truth telling and deception in sender-receiver games. *American Economic Review* 100 (3): 984-1007.

APPENDIX

SUPPLEMENTARY FIGURES

Correlation between poster design and size ratings. Due to the construction of the choice set, the design ratings were highly negatively correlated with the size ratings (correlation = -0.4957, SE = 0.0073). Fig. S1 below shows a histogram of the correlation between poster and size ratings by subject.

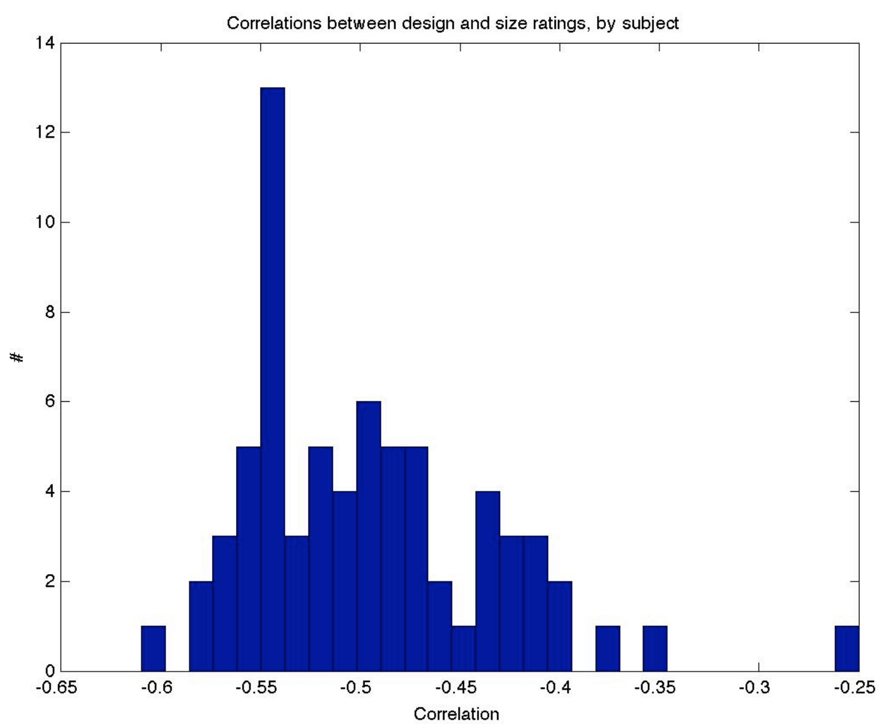
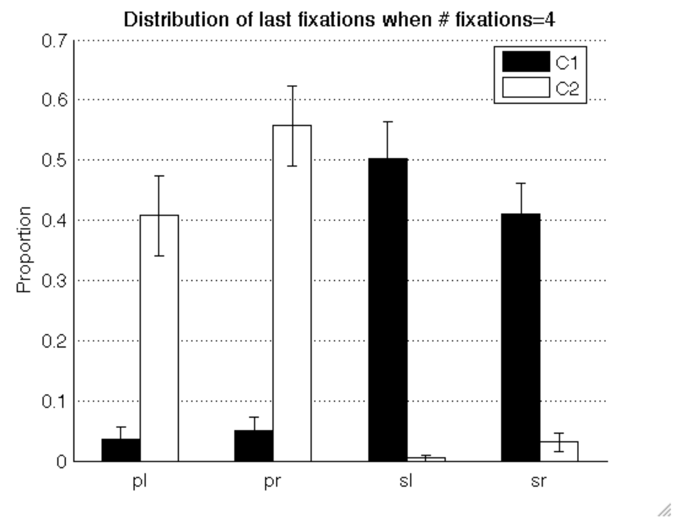


Fig. S2 (a) Distribution of the last fixation by attribute and condition when the total number of fixations in the trial is 4. **(b)** Distribution of the last fixation by attribute and condition when the total number of fixations in the trial is 5.



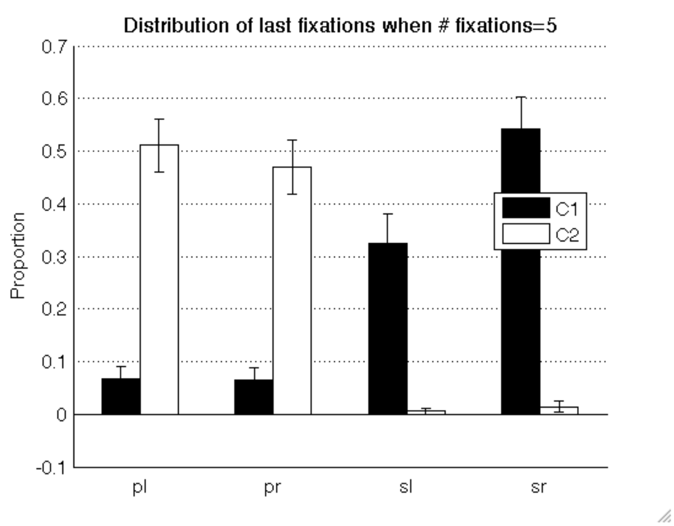
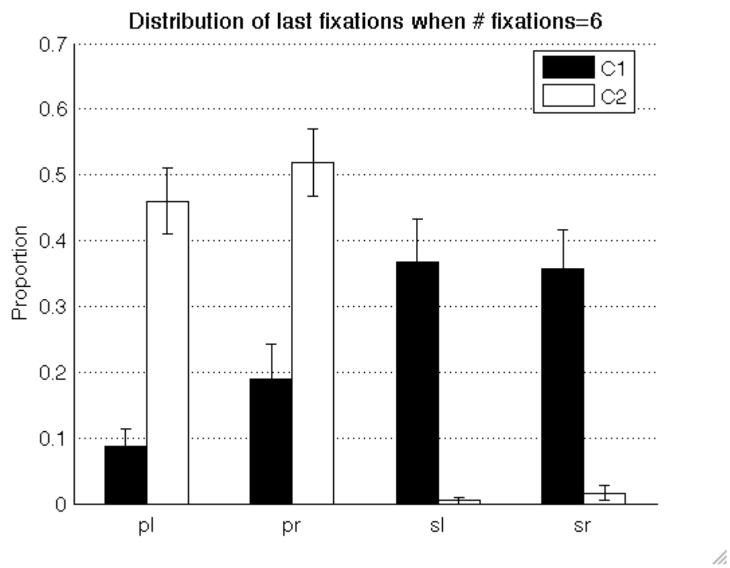
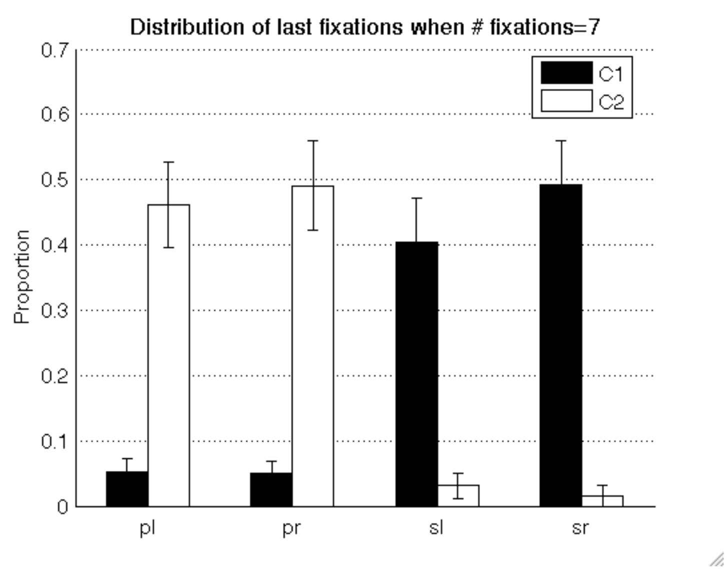


Fig. S2 (c) Distribution of the last fixation by attribute and condition when the total number of fixations in the trial is 6. **(d)** Distribution of the last fixation by attribute and condition when the total number of fixations in the trial is 7.





EXPERIMENTAL INSTRUCTIONS

Poster task

In this study, we are interested in your choices between posters. There are no right or wrong answers; we are really interested in what you think. Art posters come in different sizes and can be of different kinds. In this task you will see pictures of 7 different art posters by well known artists. We will present you with pairs of different kinds of posters that are of different sizes.

In a few minutes we will ask you to pick which of two poster/size combinations you would rather own. You will actually be mailed one of the posters based upon your choice. We will pick one pair of choices that you make randomly, and send you the poster you prefer, so your choices will matter. First we will ask you to rate several different kinds of poster and sizes, then you move on to the main choices.

For this study you need a screen size of at least 1000x700 pixels.

When you click the button, a new window will be opened with the task. Please do not close the current window. After the task you will return to this window again.

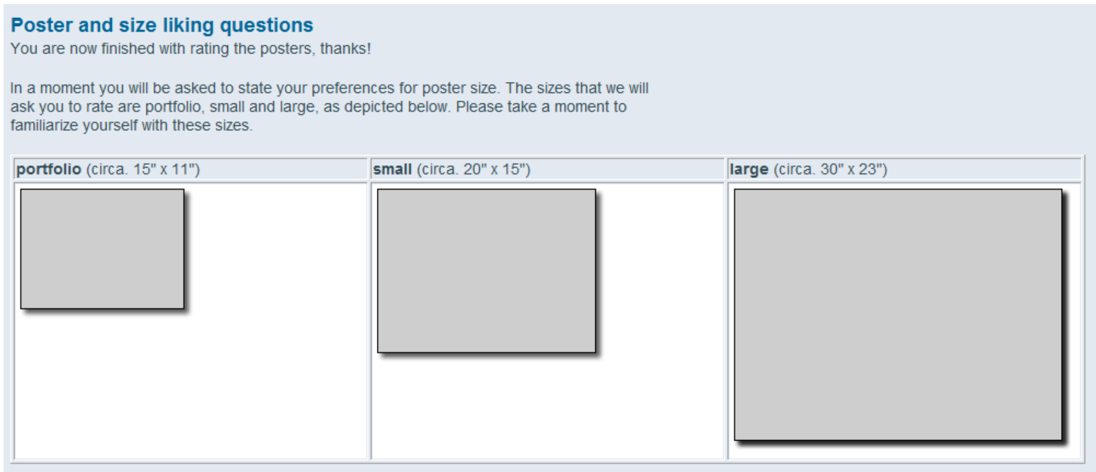
Design rating

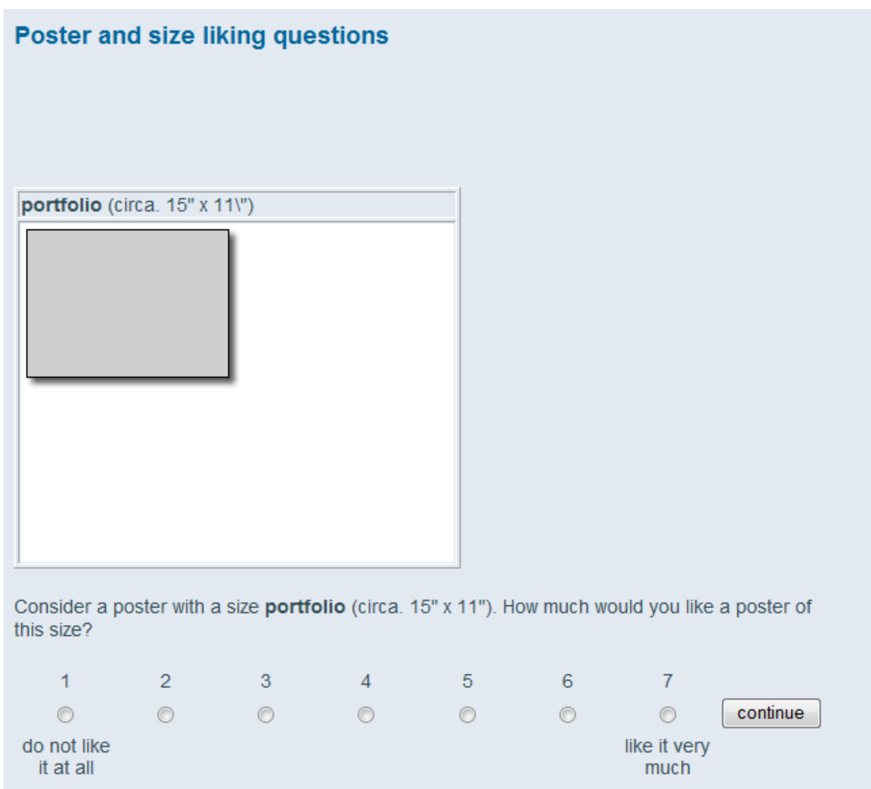
Poster and size liking questions

How much do you like this poster?

1	2	3	4	5	6	7	
○	○	○	○	○	○	○	continue
do not like it at all						like it very much	

Size rating


Poster and size liking questions
You are now finished with rating the posters, thanks!
In a moment you will be asked to state your preferences for poster size. The sizes that we will ask you to rate are portfolio, small and large, as depicted below. Please take a moment to familiarize yourself with these sizes.
portfolio (circa. 15" x 11")
small (circa. 20" x 15")
large (circa. 30" x 23")


Poster and size liking questions
portfolio (circa. 15" x 11\")
Consider a poster with a size portfolio (circa. 15" x 11"). How much would you like a poster of this size?
1
2
3
4
5
6
7
continue
do not like it at all
like it very much

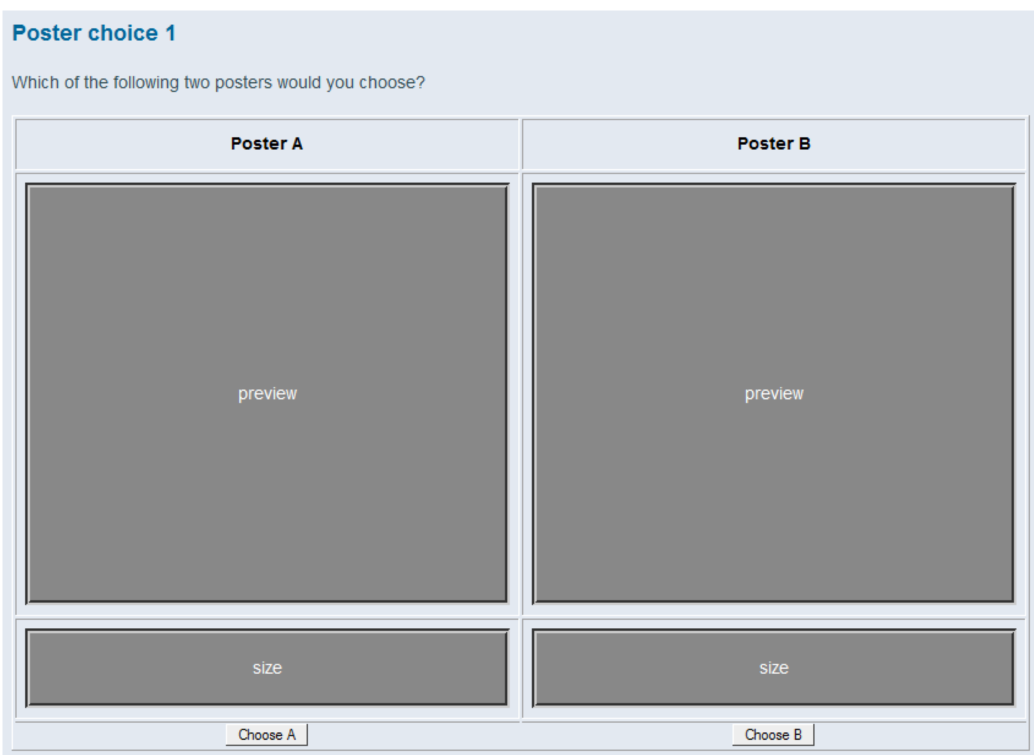
Choice introduction

Thanks for rating the sizes.

On the following pages, we will present you with 30 choices between two posters of various sizes. You can view the poster and their sizes by moving the mouse over the boxes on the screen.

You will actually be mailed one of the posters based upon your choice. We will pick one pair of choices that you make randomly, and send you the poster you prefer, so your choices will matter.

Choice trial



Feedback screen for randomly selected choice

Thanks!

We randomly selected **choice 5** from the series of choices you just made. In this trial, you chose between:

Poster A	Poster B
	
large (circa. 30 inch x 23 inch)	**small** (circa. 20 inch x 15 inch)

You chose poster B. We will mail this poster to you in a few weeks.

Enter a valid US mail address below to which we can send the poster: