



# Reference and preference: how does the brain scale subjective value?

John A. Clithero<sup>1,2\*</sup> and David V. Smith<sup>2,3</sup>

<sup>1</sup> Department of Economics, Duke University, Durham, NC, USA

<sup>2</sup> Center for Cognitive Neuroscience, Duke University, Durham, NC, USA

<sup>3</sup> Department of Psychology and Neuroscience, Duke University, Durham, NC, USA

\*Correspondence: john.clithero@duke.edu

## Commentary on

### The neurobiology of reference-dependent value computation.

by B. De Martino, D. Kumaran, B. Holt and R.J. Dolan (2009) *J. Neurosci.* 29, 3833–3842.

Would you be willing to pay more for a new television the same day you bought a new house? Would you be more likely to purchase that television if it was marked on sale? Most individuals respond affirmatively to questions of this form, likely because they use a “reference point” to help with evaluating their options. A common example is when ownership of an item affects its valuation. This is known as the “endowment effect”, demonstrated by individuals reporting a higher minimum willingness-to-accept (WTA) as a purchase price for an already-owned item than a maximum willingness-to-pay (WTP) to acquire the same item.

The importance of understanding reference-point phenomena like the endowment effect has been reflected in a sizeable number of related behavioral and theoretical papers. While some aspects of the endowment effect can be dissociated with behavioral data (Brenner et al., 2007), neuroscience can aid in delineating others. To this end, there have already been functional magnetic resonance imaging (fMRI) explorations of the endowment effect and its neurobiological underpinnings (Knutson et al., 2008; Weber et al., 2007). These studies, though, do not obtain fMRI data for within-subject comparisons of buying and selling the same items; a true measure of the endowment effect would demonstrate how a change in reference affected valuation (here, WTA and WTP). Given the large number of value comparisons that go into decisions, it might be the case that the brain determines

“relative” preferences using relative value scales instead of computing “absolute” preferences (Seymour and McClure, 2008). The presence of one or more reference points during construction of subjective values could lead to the WTA–WTP disparity.

A recent study by De Martino et al. (2009) seeks to fill this important gap, using fMRI to identify the neural valuation processes underlying the endowment effect. This goal, though seemingly addressing a specific economic phenomenon, reaches a broad set of neuroscience questions. Although the neural mechanisms of subjective value computation – particularly those involving specific subregions of the striatum and the prefrontal cortex – have been extensively studied, there remain questions about how these regions represent and integrate individual preferences. Studies also implicate additional regions (e.g., insular cortex and amygdala) in the modulation and computation of subjective value, making it plausible that this set of brain regions may both work in concert and in competition to create value signals. If the brain does employ relative value scaling through reference points, then, it could be implemented via any subset of these regions.

In order to isolate this part of value construction, De Martino et al. (2009) used a task that matches a proxy for subjective value, WTA and WTP, to a range of items: lottery tickets for money. Subjects were first given £36 and asked to choose one of two decks of lottery tickets that were identical save for their color (red or green). Subjects completed a task with five different trial types: (1) either naming selling prices for tickets from their endowed deck, (2) buying tickets from the non-endowed deck, (3) being forced to sell to the computer from their endowed deck at a pseudo-randomly picked price, (4) being forced to buy from the computer

at a pseudo-randomly picked price from the non-endowed deck, or (5) reporting the amount they felt the tickets from a third deck (yellow) were worth.

This is an elegant task for eliciting valuation differences. First, by presenting subjects with two sets of tickets that differ only in color, they made it possible for individuals to both buy and sell the same item while in the scanner. Second, unlike common binary choice tasks (e.g., accept/reject), the task obtained a reported value on each trial. Third, the computer trials served as their control condition: those trials should employ the same cognitive processes save for evaluating the gamble. However, it should be noted that subjects might still construct their value for a lottery on control trials (e.g., react to an unfair price on computer trials). Further research will be necessary to determine if the presence of mandatory choice leads to the absence of the relevant valuation processes.

De Martino et al. (2009) constructed two models to uncover indications of dissociable value computations. The first model identified which brain regions had blood-oxygenation-level dependent (BOLD) signal that correlated with WTA and WTP. There were several positive correlations, which dissociate orbitofrontal cortex (OFC) into medial (WTP) and lateral (WTA) components, and point to caudate for encoding both WTA and WTP. The second model used measures of  $\Delta$ -WTP ( $\Delta$ -WTA), the percentage change in the purchase (selling) price relative to the values reported for the corresponding yellow tickets. This model revealed a positive (negative) correlation between BOLD signal in ventral striatum (VSTR) and the posterior part of the putamen for  $\Delta$ -WTA ( $\Delta$ -WTP). De Martino et al. (2009) also reported an additional region of activation in insular cortex that demonstrated a negative correlation with  $\Delta$ -WTP.

Because brain regions were sensitive to both components of the endowment effect in the second model, De Martino et al. (2009) conducted a between-subjects regression to track neural correlates of the endowment effect across trials. The analysis uncovered another region of VSTR, indicating that individuals with a greater endowment effect ( $\Delta$ -WTA greater than  $\Delta$ -WTP) had increased VSTR signal. Knutson et al. (2008) also quantified susceptibility to the endowment effect, measured as differences in *post hoc* ratings of bought and sold products. This measure correlated significantly with signal from right anterior insula, a result that was attributed to the salience of potential loss of ownership. Given the significant differences in anterior insula for  $\Delta$ -WTP and  $\Delta$ -WTA, it is noteworthy that an analogous effect for individual differences was not found for insular cortex. The endowed goods were in the same reward modality as the exchange currency (money), so the loss of ownership may not have been as salient (Brenner et al., 2007).

Some other recent findings can aid in interpreting these results. One fMRI study of probabilistic choice (simple gambles) argued that the evaluation of risk is nonlinear in probabilities, and that this encoding takes place in VSTR in isolation of choice (Hsu et al., 2009). Additionally, VSTR has also recently been shown to encode a risk-independent measure of expected value, whereas lateral prefrontal cortex (LPFC) showed evidence for a subjective value signal that captured both expected value and risk (Tobler et al., 2009). Signal from part of the ventromedial prefrontal cortex, in a region more anterior to that found in De Martino et al. (2009), exhibits behavior that corresponds to money illusion, the common phenomenon in which individuals fail to evaluate money in terms of its actual purchasing power (Weber et al., 2009). And, as De Martino et al. (2009) noted, WTP for small food items correlates with BOLD signal from a more posterior region of medial OFC (Plassmann et al., 2007). These studies, as well as many others, point to choice parameters being encoded first in either subcortical structures or more posterior regions of the PFC, before a more integrative value signal is computed in anterior PFC. Thus, this gradient

of abstract representation may correspond to a gradient of different scales, leading to an appearance of reference-independence or reference-dependence. For example, an early parameter representation (e.g., probability) might have a specific reference point (e.g., 1/3) whereas a later value representation has a more complex reference point (e.g., available options).

The idea that different brain regions and different choice environments exhibit different scaling mechanisms is consistent with results from recent physiological recordings. Some work suggests that OFC neurons can maintain a single scale within a given behavioral context (Padoa-Schioppa and Assad, 2008). Simultaneous recordings of neurons in OFC, LPFC, and anterior cingulate cortex (ACC) revealed neurons in all three regions encoded cost, probability, and payoff (Kennerley et al., 2009). Single neurons that encoded two or more of these values were found in all three regions, and there were also neighboring neurons with opposite firing rate relationships (positive or negative) to one or more of the choice parameters. So, reference points might not only be region specific, but might also be specific to one or several choice parameters. Reference-dependence might therefore be a general concept that includes many factors, including choice history, and may thus be dynamic. This would likely not only implicate regions found by De Martino et al. (2009), but parts of the LPFC (Tobler et al., 2009) and ACC as well (Kennerley et al., 2009; Rushworth and Behrens, 2008).

Several paths for future research that investigate reference points and value scaling are apparent. De Martino et al. (2009) constrained the lottery outcomes to non-negative monetary amounts; a generalization of the current findings would include identifying the neural differences between evaluating a positive lottery and evaluating a negative lottery, as well as how buying and selling affect those differences (Brenner et al., 2007). The authors also claimed that the insula activity is in line with a risk prediction error (Preuschoff et al., 2008). This intuition would be strengthened by an additional manipulation of insula activation via risk changes within trial type, as all the lotteries here have a uniformly distributed payout of winnings. It would also serve

as a (further) test of which value signals are modulated by risk (Tobler et al., 2009).

By grounding an exploration of the endowment effect in terms of value computation, De Martino et al. (2009) show that identifying the neural components of subjective value construction can help in the understanding of a common behavioral phenomenon, the endowment effect. They break down the endowment effect for lottery tickets into two distinct value signals in dissociable brain regions and track individual differences in behavior using the fMRI data. Thus, while the literature establishes that many regions encode elements of subjective value, De Martino et al. (2009) demonstrate that these regions may have different reference points when constructing those components. Future studies can strive to identify both how the brain chooses reference points to implement value scales and how (and if) it subsequently abandons multiple scales to create a value signal that is common across a wide class of behavioral phenomena.

## ACKNOWLEDGEMENTS

We thank our mentor, Dr Scott Huettel, as well as Dr McKell Carter, Brandi Newell, and Vinod Venkatraman, for helpful comments and discussions regarding this paper.

## REFERENCES

- Brenner, L., Rottenstreich, Y., Sood, S., and Bilgin, B. (2007). On the psychology of loss aversion: possession, valence, and reversals of the endowment effect. *J. Consum. Res.* 34, 369–376.
- De Martino, B., Kumaran, D., Holt, B., and Dolan, R. J. (2009). The neurobiology of reference-dependent value computation. *J. Neurosci.* 29, 3833–3842.
- Hsu, M., Krajbich, I., Zhao, C., and Camerer, C. F. (2009). Neural response to reward anticipation under risk is nonlinear in probabilities. *J. Neurosci.* 29, 2231–2237.
- Kennerley, S. W., Dahmubed, A. F., Lara, A. H., and Wallis, J. D. (2009). Neurons in the frontal lobe encode the value of multiple decision variables. *J. Cogn. Neurosci.* 21, 1162–1178.
- Knutson, B., Wimmer, G. E., Rick, S., Hollon, N. G., Prelec, D., and Loewenstein, G. (2008). Neural antecedents of the endowment effect. *Neuron* 58, 814–822.
- Padoa-Schioppa, C., and Assad, J. A. (2008). The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. *Nat. Neurosci.* 11, 95–102.
- Plassmann, H., O'Doherty, J., and Rangel, A. (2007). Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *J. Neurosci.* 27, 9984–9988.
- Preuschoff, K., Quartz, S. R., and Bossaerts, P. (2008). Human insula activation reflects risk prediction errors as well as risk. *J. Neurosci.* 28, 2745–2752.

- Rushworth, M. F., and Behrens, T. E. (2008). Choice, uncertainty and value in prefrontal and cingulate cortex. *Nat. Neurosci.* 11, 389–397.
- Seymour, B., and McClure, S. M. (2008). Anchors, scales and the relative coding of value in the brain. *Curr. Opin. Neurobiol.* 18, 173–178.
- Tobler, P. N., Christopoulos, G. I., O'Doherty, J. P., Dolan, R. J., and Schultz, W. (2009). Risk-dependent reward value signal in human prefrontal cortex. *Proc. Natl. Acad. Sci. USA* 106, 7185–7190.
- Weber, B., Aholt, A., Neuhaus, C., Trautner, P., Elger, C. E., and Teichert, T. (2007). Neural evidence for reference-dependence in real-market-transactions. *Neuroimage* 35, 441–447.
- Weber, B., Rangel, A., Wibral, M., and Falk, A. (2009). The medial prefrontal cortex exhibits money illusion. *Proc. Natl. Acad. Sci. USA* 106, 5025–5028.
- Received: 11 June 2009; published online: 08 July 2009.
- Citation: *Front. Hum. Neurosci.* (2009) 3:11. doi: 10.3389/neuro.09.011.2009
- Copyright © 2009 Clithero and Smith. This is an open-access publication subject to an exclusive license agreement between the authors and the Frontiers Research Foundation, which permits unrestricted use, distribution, and reproduction in any medium, provided the original authors and source are credited.