

# Value Is in the Eye of the Beholder: Early Visual Cortex Codes Monetary Value of Objects during a Diverted Attention Task

Andrew S. Persichetti, Geoffrey K. Aguirre, and Sharon L. Thompson-Schill

## Abstract

■ A central concern in the study of learning and decision-making is the identification of neural signals associated with the values of choice alternatives. An important factor in understanding the neural correlates of value is the representation of the object itself, separate from the act of choosing. Is it the case that the representation of an object within visual areas will change if it is associated with a particular value? We used fMRI adaptation to measure the neural similarity of a set of novel objects before and after participants learned to associate monetary values with the objects. We used a range of both positive and negative values to allow us to distinguish effects of behavioral

salience (i.e., large vs. small values) from effects of valence (i.e., positive vs. negative values). During the scanning session, participants made a perceptual judgment unrelated to value. Crucially, the similarity of the visual features of any pair of objects did not predict the similarity of their value, so we could distinguish adaptation effects due to each dimension of similarity. Within early visual areas, we found that value similarity modulated the neural response to the objects after training. These results show that an abstract dimension, in this case, monetary value, modulates neural response to an object in visual areas of the brain even when attention is diverted. ■

## INTRODUCTION

The neural representation of a visual stimulus must code many dimensions, and so, the similarity space of a set of objects is multidimensional, even in a single brain region. For example, the similarity of the neural responses in V1 reflects both stimulus orientation and spatial frequency (Mazer, Vinje, McDermott, Schiller, & Gallant, 2002), and the similarity of neural codes in V4 reflects both stimulus color and shape (Roe et al., 2012). The voxel-level BOLD response can reflect multiple stimulus dimensions that are coded, at the neural level, either independently or conjointly (Drucker, Kerr, & Aguirre, 2009).

For these basic visual dimensions and beyond, it is clear that neural responses early in the visual pathway are shaped by learning, both of categorical boundaries along visual stimulus dimensions (Folstein, Palmeri, & Gauthier, 2012) and of abstract information about objects (e.g., biological class structure of living things; Connolly et al., 2012). The goal of the current study is to explore whether (and where) neural responses to novel visual stimuli reflect the abstract but behaviorally relevant variable of value. We show that early in the visual processing pathway nonvisual information is coded in the neural response, even when (a) the abstract dimension is orthogonal to all visual dimensions, (b) the abstract dimension is

newly learned, and (c) responses are measured during a task that makes no reference to the abstract (value) dimension.

Many recent experiments have sought to identify neural signals associated with the values of choice alternatives (see Bartra, McGuire, & Kable, 2013, for a review). It has been suggested that the process of choosing between items is a two-staged process, in which values are first assigned to each option and then compared to yield a choice (Levy, Lazzaro, Rutledge, & Glimcher, 2011; Kable & Glimcher, 2009). This two-staged process of choice suggests that the process of tracking values of items is independent of choosing between them (Lebreton, Jorge, Michel, Thirion, & Pessiglione, 2009). Several studies have shown brain responses that engage automatically to different kinds of valuations, including monetary value (Tallon-Baudry, Meyniel, & Bourgeois-Gironde, 2011), facial attractiveness (e.g., Chatterjee, Thomas, Smith, & Aguirre, 2009), houses and paintings (e.g., Lebreton et al., 2009), consumer goods (e.g., Levy et al., 2011), and faces that have learned associations to monetary values (Rothkirch, Schmack, Schlagenhaut, & Sterzer, 2012). These results address if values are stored separately from a choice task, but their use of familiar objects makes it difficult to disentangle the value of a stimulus from its cultural significance and familiarity (Rangel, Camerer, & Montague, 2008; Erk, Spitzer, Wunderlich, Galley, & Walter, 2002). In Rothkirch et al. (2012), a baseline measure of brain response to the face stimuli before value learning is





each participant. The stimuli were presented using a Sanyo (Moriguchi, Japan) SXGA 4200 lumens projector with a Buhl long-throw lens for rear projection onto Mylar screens, which participants viewed through a mirror mounted on the head coil.

## Data Analysis

Image preprocessing and analyses were conducted using FMRIB Software Library (Smith et al., 2004). The first eight TRs from each scan were removed before analysis. The data were smoothed with a FWHM Gaussian kernel of 5 mm. The functional images were aligned to the middle image of the time series with MCFLIRT (Jenkinson, Bannister, Brady, & Smith, 2002) and then transformed to standard Montreal Neurological Institute space. Within-subject statistical models were created using a general linear model (Figure 1D). Experimental conditions were convolved with a canonical hemodynamic response function, and spikes caused by head motion were included as covariates in the model. Beta estimates from the model were then averaged across each ROI.

When using the continuous carry-over approach to measure neural adaptation, the relationship of each stimulus to the prior stimulus forms the basis of the covariates (Aguirre, 2007). In our design, we used both positive and negative values to distinguish changes in responses due to the actual value of the objects from those due to the behavioral salience of each object (i.e., large vs. small values, as reflected by the absolute value, or what we will call the unsigned value). With this in mind, we modeled the distance between the value of each stimulus and the prior stimulus in two ways: One covariate models the “signed effect,” with nine different transition sizes, such that the stimulus associated with wins of \$10 was maximally different from the one associated with losses of \$10; the second covariate models the “unsigned effect,” with five transition sizes such that those two stimuli (positive and negative \$10) would be maximally similar. The inclusion of the unsigned covariate was added to the analysis to demonstrate that our effect of interest, namely, the signed adaptation effect in EVC, was about value per se and not simply behavioral salience.

Our model included covariates for signed adaptation, unsigned adaptation, shape, color, blank trials, and trials after a blank trial. When using the continuous carry-over approach to measure adaptation, a trial in the adaptation covariates represents the difference along a given dimension between the present stimulus and the one that precedes it. Therefore, we included trials that followed blank screens as a covariate of no interest in the model because a blank screen trial does not have a value to serve as a comparison for the stimulus it precedes. For each participant, the five main task runs were first modeled individually and then combined using a higher-level fixed effects model. Data were then combined across participants using a random effects model.

We report analyses in five ROIs. We included two functionally defined ROIs in visual cortex, EVC and LOC, as described above. Although we were primarily interested in the effects of value on object representations in EVC (and potentially LOC), we included value-related ROIs to provide a point of reference for our findings in EVC. For example, if we find that the signed value covariate does explain variance in EVC after training, how does that compare with more traditionally value-related ROIs? Is there dissociation between EVC and these other regions, or are the effects similar across these different cortical regions? Because there is currently not an agreed-upon way to functionally localize value-related brain areas, we defined value-related ROIs based on the anatomical coordinates reported in a meta-analysis of fMRI experiments that examined subjective value (Bartra et al., 2013); we defined ROIs in ventromedial pFC (VMPFC), dorsomedial pFC (DMPFC), and striatum.

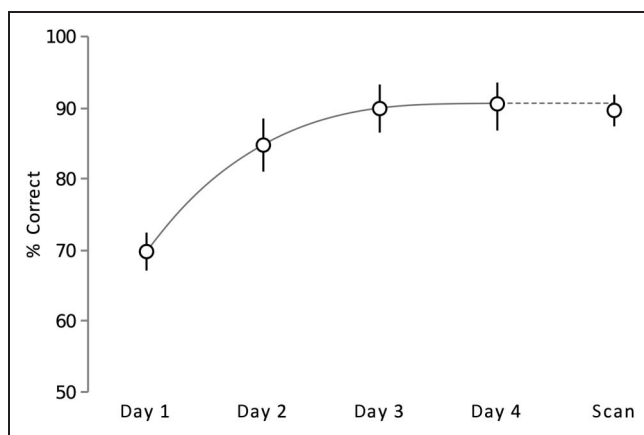
## RESULTS

### Behavioral Results

During each day of the training and at the end of the second scan, participants performed a choice task that required them to use their knowledge of the values of the objects to optimize their winnings. Participants chose the object with the greater associated reward more frequently from the first training day (mean  $\pm$  SEM correct = 69.7  $\pm$  2.7%) to the final training day (mean  $\pm$  SEM correct = 90.1  $\pm$  3.4%). The scores at the end of training were above chance (50%) for all participants (Figure 3) and significantly different than the scores on Day 1 ( $t(12) = 9.7, p < .001$ ). Participants were also able to identify the values associated with the stimuli during the alternate choice task (i.e., object vs. \$1) at the end of the second scan (mean  $\pm$  SEM correct = 89.7  $\pm$  2.3%).

### fMRI Results

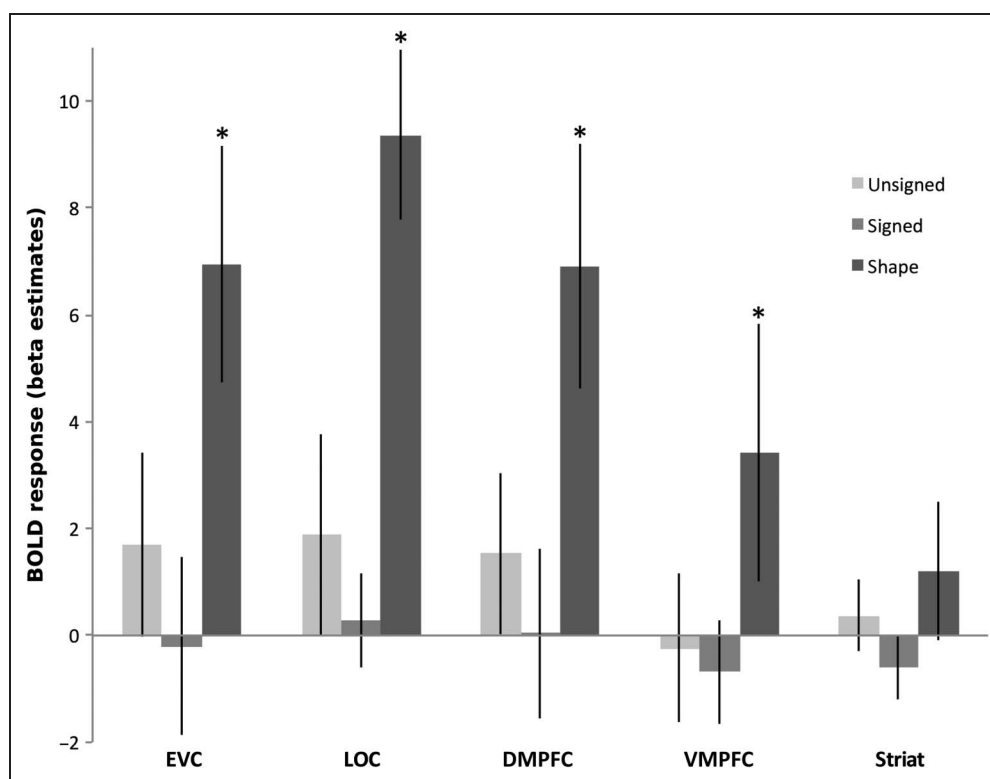
We measured the degree to which the learned value of an object modulates the neural response to the presentation of subsequent objects of greater or lesser value. This neural adaptation effect could be manifested as proportional to the signed value of an object, meaning that an object associated with a loss of \$10 would be treated as maximally different from an object associated with a gain of \$10. Alternatively, neural populations could encode the absolute value of an object and thus reflect behavioral salience. In this case, an object associated with either a gain or a loss of \$10 would be treated as maximally different from an object with a small relative value (e.g., \$2.50). We tested for both of these possible effects in each of five ROIs. In addition, we examined these effects before and after learning of object value.



**Figure 3.** Value learning across participants. Accuracy during each training session and the in-scanner choice task. Each session is the mean of all participants' accuracy during each session. Error bars represent  $\pm 1 SE$ . Fit line is a third-order polynomial.

The primary question of this experiment is whether nonvisual, value information is coded in the neural response early in the visual processing pathway even when it is orthogonal to all visual dimensions and the responses are measured during a task that makes no reference to the abstract (value) dimension. As expected, before training, no significant adaptation effect related to either signed or unsigned value was found in any ROI (Figure 4). In contrast, significant neural adaptation related to shape similarity was found in multiple regions.

**Figure 4.** Brain regions showing adaptation effects during the prescan. The adaptation effect for shape was significant in EVC, LOC, DMPFC, and VMPFC. Error bars represent  $\pm 1 SE$ .

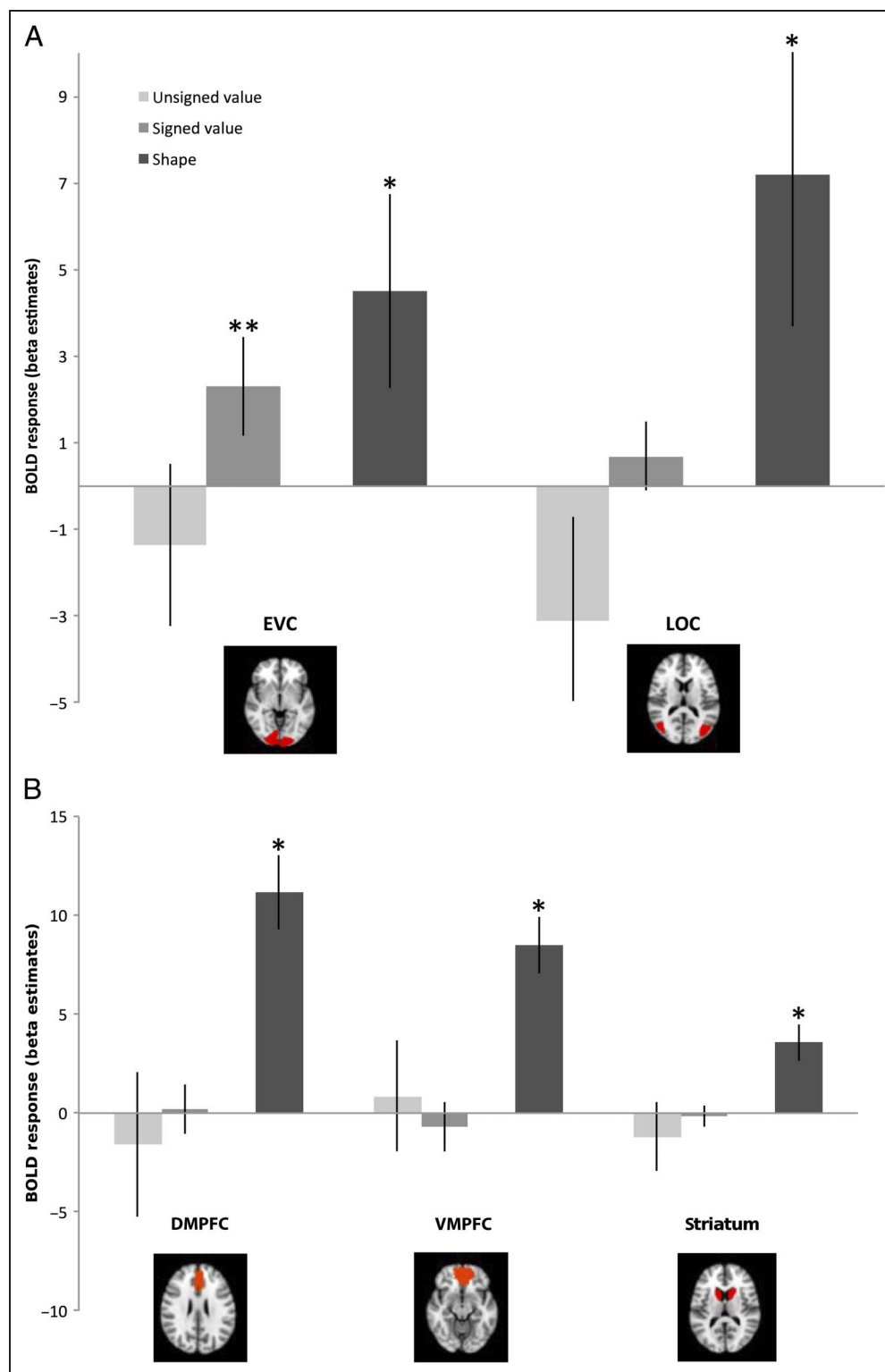


This indicates that object value and shape were unrelated before training.

After training, we observed an adaption effect for the signed model in EVC ( $t(12) = 2.87, p < .02$ ; Figure 5A). This effect was greater than that observed during the pre-training scan (marginally significant interaction,  $t(12) = 2.11, p = .057$ ). Signed value was not associated with adaptation in any of the other ROIs, nor did we find significant effects with the unsigned model in any of our ROIs.

Prior studies of value representation have examined the overall magnitude of neural response associated with item value (Bartra et al., 2013; Levy et al., 2011; Tusche, Bode, & Haynes, 2010). We tested for similar item effects in our data by determining if the magnitude of the BOLD fMRI signal was proportional to object value, independent of the similarity of value to the preceding or following stimuli. We did not observe a significant item effect for signed value in any of our ROIs during either scan session. Given that prior studies have shown monotonic increase of activity in pFC and striatum as a function of stimulus value during value-related decision tasks, we also measured these effects during the value task at the end of the second scan session. We analyzed the data from these scans in the same way as the data from the main task. Whereas we did not find adaptation effects of value during this task, we did find direct effects of relative value in DMPFC ( $t = 2.77, p < .02$ ) but not in striatum or VMPFC (both  $ps > .1$ ). Although the inability to measure a significant response to value in striatum and VMPFC may seem odd, it could be because participants did not expect

**Figure 5.** (A) Visual brain regions showing adaptation effects during the postscan. The adaptation effect for shape was significant in LOC ( $t(12) = 2.96, p = .01$ ) and EVC ( $t(12) = 2.31, p < .05$ ) during prescan and postscan. The adaptation effect for the signed model was significant in EVC ( $t(12) = 2.87, p = .01$ ) during the postscan. Error bars represent  $\pm 1 SE$ . (B) Frontal brain regions showing adaptation effects during the postscan. The adaptation effect for shape was significant in DMPFC ( $t(12) = 3.05, p = .01$ ) and VMPFC ( $t(12) = 3.02, p = .01$ ) and marginally significant in the striatum ( $t(12) = 2.06, p = .06$ ) during prescan and postscan. Error bars represent  $\pm 1 SE$ .



a reward based on their performance in this task (i.e., participants were aware that we were simply testing their knowledge of the values).

Finally, we examined the adaptation effect of shape similarity. Although this effect has no direct bearing on the goals of this study, we report it here as it establishes the sensitivity of our ROIs to object similarity, beyond

value. We observed an adaptation effect of shape similarity in LOC ( $t(12) = 2.96, p < .02$ ), EVC ( $t(12) = 2.31, p < .05$ ), DMPFC ( $t(12) = 3.05, p = .01$ ), and VMPFC ( $t(12) = 3.02, p = .01$ ) and a marginal effect of shape similarity in striatum ( $t(12) = 2.06, p = .06$ ; Figure 5B). There were no reliable differences in the magnitude of the shape adaptation effect pretraining versus posttraining in

any of these ROIs. We also tested for neural adaptation related to the sequential effect of color change. This effect was not significant during either scan in the EVC (both  $p$ s > .3) or the LOC (both  $p$ s > .2).

In summary, EVC showed adaptation to shape before and after training and adaptation to value after training. In contrast, LOC, DMPFC, and VMPFC showed adaptation to shape (at both time points) but no adaptation to value.

## DISCUSSION

We scanned participants while they viewed a set of nine novel objects that varied across two visual dimensions (color and shape) and one orthogonal, nonvisual dimension (monetary value). We measured fMRI adaptation to characterize the neural similarity of these novel objects before and after a training period during which participants associated monetary values with the objects. We found a recovery from adaptation along the monetary value dimension in EVC only after training. We interpret this as evidence that object representations in the visual cortex are affected by value learning even when value is orthogonal to visual features and even when participants are engaged in a perceptual task unrelated to object value.

This value effect may be related to other effects of learning that have been reported in the visual system. Folstein et al. (2012) mention that in previous studies in which shape spaces were created by morphing complex objects, the dimensions that define those spaces are unclear and may not have existed before category learning (Gureckis & Goldstone, 2008; Goldstone & Steyvers, 2001). The authors further posit that because the objects that differ along dimensions relevant to the learned categories are more perceptually discriminable after learning, category learning may create representations of those relevant object dimensions. In this study, we believe that a similar process is at work. Specifically, learning the abstract property of value has modified the representations within certain assemblies of color- and shape-responsive neurons in EVC to better discriminate along the value dimension.

Our results go beyond the evidence for attention effects of value in the visual system (Serences, 2008), or what others have called visual perceptual learning (Sasaki, Nanez, & Watanabe, 2009). These studies have shown that reward history associated with a specific feature of a set of stimuli will modulate responses to that feature in EVC and out into higher-level visual areas. Our results add to this literature by showing that even when reward history is orthogonal to the low-level features of objects, neural activity in EVC tracks the value of each stimulus. Previous studies have suggested that coupling reward with visual stimuli may improve performance during perceptual tasks (Nomoto et al., 2010; Serences, 2008; Engelmann & Pessoa, 2007; Pessiglione et al., 2006).

Whereas prior studies of the representation of value have examined the bulk neural response evoked by the stimulus, we measured here neural adaptation induced by the similarity of object values. We used neural adaptation to characterize the similarity space represented within separate neural populations. We chose this method because it allowed for us to manipulate and measure each stimulus dimension separately and to observe neural sensitivity to each stimulus dimension across our ROIs. We did not observe a modulation of overall level of neural activity driven by object value in any of our ROIs. Interestingly, we also found that brain regions commonly found to be associated with value (DMPFC and VMPFC) encoded the shapes of objects, but not their values, indicating that when value is irrelevant to the present task, these regions will instead track task-relevant dimensions of the presented stimuli. This finding is not surprising because it has been reported that regions of frontal cortex, including dorsal and medial PFC, play a crucial role in the active biasing of task-relevant processes against strong competing alternatives (Chadick, Zanto, & Gazzaley, 2014; Chadick & Gazzaley, 2011; Miller & Cohen, 2001). Furthermore, evidence from several species suggests that the striatum contributes directly to decision-making, action selection, and initiation (Green, Biele, & Heekeren, 2012; Balleine, Delgado, & Hikosaka, 2007). This finding suggests that the process of tracking values occurs early in perception and that more frontal brain regions, involved in executive functions, are not recruited in value representation unless a choice between items is necessary.

Taken together, our findings that value learning modified the representations within certain neurons in EVC to better discriminate along the value dimension and that neurons in brain regions commonly found to be associated with value are not tuned to the value dimension when it is irrelevant to the task suggest that the sensory system plays a large role in the valuation process in the brain. This top-down facilitation in visual areas serves to assist in learning and in disambiguating input data, which can lead to faster RTs and better accuracy in important situations (Hsieh, Vul, & Kanwisher, 2010; Nomoto et al., 2010; Bar, 2003). If neurons in executive brain regions are not tuned to dimensions irrelevant to task demands, then where does value information reside when not in use? Our data support the idea that neural tuning in EVC, as measured by fMRI adaptation, contains information that could facilitate faster processing during value-based decision tasks. However, the lack of a main effect of value in EVC suggests that value is represented in visual cortex differently from putative "value regions." In this study, we did find an item-level effect of value in DMPFC after training, indicating that a different representation of value is associated with this brain region.

An interesting topic for further examination is if altered neural representations in EVC give rise to altered performance on simple perceptual tasks. Speeded discrimination judgments might reflect the value of the objects,

suggesting that learning to associate a value with an object influences perceptual similarity, in addition to neural similarity.

In conclusion, we found that neural populations in EVC encode value information. The response in EVC was measured during a demanding task that diverted attention away from the object value. The stimuli used were novel objects with no prior reward history, cultural significance, or familiarity, and the value dimension was orthogonal to the low-level visual features of the stimuli. These findings suggest that the learned value of objects penetrates to the earliest level of cortical sensory representation.

## Acknowledgments

This work was funded by National Institutes of Health grant R01EY021717. We would like to thank the Kable laboratory at the University of Pennsylvania for their insightful comments.

Reprint requests should be sent to Andrew S. Persichetti, Department of Psychology, Emory University, 36 Eagle Row, Room 410, Atlanta, GA 30322, or via e-mail: [apersic@emory.edu](mailto:apersic@emory.edu).

## REFERENCES

- Aguirre, G. K. (2007). Continuous carry-over designs for fMRI. *Neuroimage*, *35*, 1480–1494.
- Aguirre, G. K., Mattar, M. G., & Magis-Weinberg, L. (2011). de Bruijn cycles for neural decoding. *Neuroimage*, *56*, 1293–1300.
- Arsenault, J. T., Nelissen, K., Jarraya, B., & Vanduffel, W. (2013). Dopaminergic reward signals selectively decrease fMRI activity in primate visual cortex. *Neuron*, *77*, 1174–1186.
- Balleine, B. W., Delgado, M. R., & Hikosaka, O. (2007). The role of the dorsal striatum in reward and decision-making. *Journal of Neuroscience*, *27*, 8161–8165.
- Bar, M. (2003). A cortical mechanism for triggering top-down facilitation in visual object recognition. *Journal of Cognitive Neuroscience*, *15*, 600–609.
- Bartra, O., McGuire, J. T., & Kable, J. W. (2013). The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage*, *76*, 412–427.
- Benson, N. C., Butt, O. H., Brainard, D. H., & Aguirre, G. K. (2014). Correction of distortion in flattened representations of the cortical surface allows prediction of V1–V3 functional organization from anatomy. *PLoS Computational Biology*, *10*, e1003538.
- Chadick, J. Z., & Gazzaley, A. (2011). Differential coupling of visual cortex with default or frontal-parietal network based on goals. *Nature Neuroscience*, *14*, 830–832.
- Chadick, J. Z., Zanto, T. P., & Gazzaley, A. (2014). Structural and functional differences in medial prefrontal cortex underlie distractibility and suppression deficits in ageing. *Nature Communications*, *5*, 4223.
- Chatterjee, A., Thomas, A., Smith, S. E., & Aguirre, G. K. (2009). The neural response to facial attractiveness. *Neuropsychology*, *23*, 135–143.
- Connolly, A. C., Guntupalli, J. S., Gors, J., Hanke, M., Halchenko, Y. O., Wu, Y.-C., et al. (2012). The representation of biological classes in the human brain. *The Journal of Neuroscience*, *32*, 2608–2618.
- Drucker, D. M., Kerr, W. T., & Aguirre, G. K. (2009). Distinguishing conjoint and independent neural tuning for stimulus features with fMRI adaptation. *Journal of Neurophysiology*, *101*, 3310–3324.
- Engelmann, J. B., & Pessoa, L. (2007). Motivation sharpens exogenous spatial attention. *Emotion*, *7*, 668–674.
- Erk, S., Spitzer, M., Wunderlich, A. P., Galley, L., & Walter, H. (2002). Cultural objects modulate reward circuitry. *NeuroReport*, *13*, 2499–2503.
- Folstein, J. R., Palmeri, T. J., & Gauthier, I. (2012). Category learning increases discriminability of relevant object dimensions in visual cortex. *Cerebral Cortex*, *23*, 814–823.
- Goldstone, R. L., & Steyvers, M. (2001). The sensitization and differentiation of dimensions during category learning. *Journal of Experimental Psychology: General*, *130*, 116–139.
- Green, N., Biele, G. P., & Heekeren, H. R. (2012). Changes in neural connectivity underlie decision threshold modulation for reward maximization. *Journal of Neuroscience*, *32*, 14942–14950.
- Grill-Spector, K., & Malach, R. (2001). fMR-adaptation: A tool for studying the functional properties of human cortical neurons. *Acta Psychologica*, *107*, 293–321.
- Gureckis, T. M., & Goldstone, R. L. (2008). The effect of the internal structure of categories on perception. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (p. 843). Austin, TX: Cognitive Science Society.
- Hsieh, P.-J., Vul, E., & Kanwisher, N. (2010). Recognition alters the spatial pattern of fMRI activation in early retinotopic cortex. *Journal of Neurophysiology*, *103*, 1501–1507.
- Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage*, *17*, 825–841.
- Kable, J. W., & Glimcher, P. W. (2009). The neurobiology of decision: Consensus and controversy. *Neuron*, *63*, 733–745.
- Lebreton, M., Jorge, S., Michel, V., Thirion, B., & Pessiglione, M. (2009). An automatic valuation system in the human brain: Evidence from functional neuroimaging. *Neuron*, *64*, 431–439.
- Levy, I., Lazzaro, S. C., Rutledge, R. B., & Glimcher, P. W. (2011). Choice from non-choice: Predicting consumer preferences from blood oxygenation level-dependent signals obtained during passive viewing. *The Journal of Neuroscience*, *31*, 118–125.
- Mazer, J. A., Vinje, W. E., McDermott, J., Schiller, P. H., & Gallant, J. L. (2002). Spatial frequency and orientation tuning dynamics in area V1. *Proceedings of the National Academy of Sciences, U.S.A.*, *99*, 1645–1650.
- Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167–202.
- Nomoto, K., Schultz, W., Watanabe, T., & Sakagami, M. (2010). Temporally extended dopamine responses to perceptually demanding reward-predictive stimuli. *Journal of Neuroscience*, *30*, 10692–10702.
- Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*, *442*, 1042–1045.
- Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience*, *9*, 545–556.
- Roe, A. W., Chelazzi, L., Connor, C. E., Conway, B. R., Fujita, I., Gallant, J. L., et al. (2012). Toward a unified theory of visual area V4. *Neuron*, *74*, 12–29.



- Rothkirch, M., Schmack, K., Schlagenhaut, F., & Sterzer, P. (2012). Implicit motivational value and salience are processed in distinct areas of orbitofrontal cortex. *Neuroimage*, *62*, 1717–1725.
- Sasaki, Y., Nanez, J. E., & Watanabe, T. (2009). Advances in visual perceptual learning and plasticity. *Nature Reviews Neuroscience*, *11*, 53–60.
- Seitz, A. R., Kim, D., & Watanabe, T. (2009). Rewards evoke learning of unconsciously processed visual stimuli in adult humans. *Neuron*, *61*, 700–707.
- Serences, J. T. (2008). Value-based modulations in human visual cortex. *Neuron*, *60*, 1169–1181.
- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*, *23*(Suppl. 1), S208–S219.
- Stanisor, L., van der Togt, C., Pennartz, C. M. A., & Roelfsema, P. R. (2013). A unified selection signal for attention and reward in primary visual cortex. *Proceedings of the National Academy of Sciences, U.S.A.*, *110*, 9136–9141.
- Tallon-Baudry, C., Meyniel, F., & Bourgeois-Gironde, S. (2011). Fast and automatic activation of an abstract representation of money in the human ventral visual pathway. *PLoS One*, *6*, e28229.
- Tusche, A., Bode, S., & Haynes, J.-D. (2010). Neural responses to unattended products predict later consumer choices. *The Journal of Neuroscience*, *30*, 8024–8031.