Cross-modal effects of value on perceptual acuity and stimulus encoding

Arezoo Pooresmaeili[a,1,2], Thomas H. B. FitzGerald[b,1], Dominik R. Bach[a,b,c], Ulf Toelch[a], Florian Ostendorf[a], and Raymond J. Dolan[a,b]

[a]Berlin School of Mind and Brain, Humboldt University of Berlin, 10099 Berlin, Germany; [b]Wellcome Trust Centre for Neuroimaging, Institute of Neurology, University College London, London WC1N 3BG, United Kingdom; and [c]Hospital of Psychiatry, University of Zurich, 8032 Zurich, Switzerland

Cross-modal interactions are very common in perception. An important feature of many perceptual stimuli is their reward-predicting properties, the utilization of which is essential for adaptive behavior. What is unknown is whether reward associations in one sensory modality influence perception of stimuli in another modality. Here we show that auditory stimuli with high-reward associations increase the sensitivity of visual perception, even when sounds and reward associations are both irrelevant for the visual task. This increased sensitivity correlates with a change in stimulus representation in the visual cortex, indexed by increased multivariate decoding accuracy in simultaneously acquired functional MRI data. Univariate analysis showed that reward associations modulated responses in regions associated with multisensory processing in which the strength of modulation was a better predictor of the magnitude of the behavioral effect than the modulation in classical reward regions. Our findings demonstrate a value-driven cross-modal interaction that affects perception and stimulus encoding, with a resemblance to well-described modulatory effects of attention. We suggest that multisensory processing areas may mediate the transfer of value signals across senses.

reward value | sensory discrimination | audiovisual

The world is structured such that objects or events that cause sensations in one sensory modality influence those in another modality, a mechanism that underlies the near-ubiquitous phenomenon of multisensory interaction (1). This phenomenon has been the focus of a large and burgeoning theoretical and experimental literature (1–4). One feature of environmental stimuli important for adaptive behavior is their rewarding or reward-predicting properties. Surprisingly, we know little about how reward associations in one sensory modality influence processing in other modalities. In particular, whether an association with increased reward, known to increase the accuracy of perceptual processing within that sensory modality (5–10), can transfer between modalities for concurrently presented stimuli is unclear. This is important both because of what it reveals about the nature of cross-modal associations and because it might constitute an important perceptual mechanism in its own right. For example, one can easily imagine the importance of increased sensitivity to change in the visual scene to the concurrent sound of a tiger roaring or the mating call of a conspecific.

To study cross-modal transfer of value, we designed a visual orientation discrimination task in which visual stimuli were presented concurrently with one of two arbitrary pure tones previously associated with different levels of monetary reward. Critically, these tones were task-irrelevant and bore no relationship either to the orientation of the visual stimulus or to the outcome of the trial. (No feedback was presented about the accuracy of perceptual judgments, and performance on the orientation discrimination task was not related to the payment subjects received at the end of the trial.) We hypothesized that psychophysical measures of visual orientation judgment (d' and performance accuracy) would show improvement in the presence of a sound with high (compared with low) reward association. Furthermore, simultaneous and spatially overlapping stimuli are more likely to have a common source and

thus a shared reward association. Previous studies have shown that interactions between sensory modalities are strongest when stimuli are presented simultaneously and at contiguous spatial locations (1, 2, 11–13), especially at the earliest stages of processing (2, 11). On this basis, we hypothesized that the effect of rewarded sounds on vision should be strongest when visual and auditory stimuli overlap in time and in space. Owing to the low temporal resolution of functional MRI (fMRI) signals, here we manipulated only the spatial overlap of stimuli while keeping their temporal alignment constant (with simultaneous presentation).

We had two specific questions regarding the neuronal underpinnings of any observed behavioral effects, which we addressed in simultaneously acquired fMRI data. First, we were interested in whether better behavioral performance owing to the presence of high-reward sounds was accompanied by a more differentiated stimulus representation in early visual areas, as assessed by classification accuracy in a multivariate pattern analysis (MVPA). Second, we were interested in whether this effect involved multimodal cortical areas known to play a role in the integration of audiovisual information, such as the superior temporal sulcus (STS) (14–16), or instead solely involved processing in classical reward-related areas, such as the ventromedial prefrontal cortex (vmPFC) (17, 18) or ventral striatum (19, 20).

Our first question draws on evidence demonstrating a nonuniform spatial distribution of reward effects across the visual cortex (21) and a sharpening of sensory representations through suppression of redundant signals elicited by reward/information predictive signals (10, 22). This evidence suggests that reward

Significance

Reward-predicting signals could be acquired through any of our different sensory modalities, but should be used by other senses to achieve fast and accurate behavior. How reward information is communicated across different sensory modalities is unknown. We demonstrate that sounds associated with high rewards increase the sensitivity of vision, even when sounds and their reward associations are task-irrelevant. Multivariate analysis of the simultaneously acquired functional MRI data revealed that high-reward sounds increased the accuracy of stimulus representation in the visual cortex. Multisensory regions of the temporal cortex were modulated by sound values, and the strength of this modulation was correlated with the change in visual acuity. Our results demonstrate a value-driven cross-modal interaction that affects early stages of sensory processing and involves multisensory areas.

information might be reflected in the spatial pattern, rather than the absolute magnitude, of blood-oxygen-level–dependent (BOLD) responses. The second question is based on two possible mechanistic schemes of how reward impacts on perception, where reward could exert its effect either directly on perceptual (here visual) processing or via intermediate and presumably task-specific stages (here via regions involved in integrating auditory and visual information). Although both schemes predict a change in response magnitude or spatial pattern of BOLD responses in the visual cortex, only the latter predicts involvement of specific cross-modal areas. We provide evidence that supports task-specific influences of reward on perception.

Results

Behavioral Results. After subjects were familiarized with sounds and their respective rewards (Fig. 1*A*), they performed a visual orientation discrimination task in the presence of previously rewarded sounds (Fig. 1*B* and *Materials and Methods*). To test the effect of a sound's reward association on visual orientation discrimination, we first examined the time course of the performance difference between Gabors + high-reward sounds and Gabors + low-reward sounds (Fig. 1*C*). High-rewarded sounds led to a better performance in orientation discrimination task, with the maximum effect seen at the beginning of the experiment and lower but steady levels in later trials. Average discrimination sensitivity, measured by d' and accuracy, measured by percent correct rate, were higher for the Gabors accompanied by high-reward sounds (Fig. 1*D*). Repeated-measures ANOVA, with d' as the dependent factor and reward and spatial congruency as the independent factor, revealed a significant main effect of reward ($F_{1,23} = 5.64$, $P = 0.02$). The interaction between reward and spatial congruence was not significant ($F_{1,23} = 1.15$, $P = 0.295$). Planned pairwise comparisons showed that both the high-reward–congruent (HC) and high-reward–incongruent (HIC) conditions had a significantly higher d' value than the low-reward–congruent (LC) condition ($P = 0.01$ for comparison of HC vs. LC, $P = 0.044$ for comparison of HIC vs. LC; none of the other pairwise comparisons was significant, $P > 0.05$). We obtained similar results when percent correct rates were compared ($P = 0.03$ for comparison of HC vs. LC, paired t test; all other pairwise comparisons were nonsignificant).

Fig. 1*E* shows the time course of the behavioral effect of sound and rewards inside the scanner. As time progressed, there was a marked decrease in the effect of reward, culminating in a reversal of the effect (i.e., lower performance for high-reward sounds) during the last few trials. This effect, in which the "extinction" of responses to a conditioned stimulus ultimately leads to a behavioral reversal after repeated exposure to nonreinforced conditioned stimulus, is well described in the conditioning literature (23). Extinction occurred only inside the scanner, most likely owing to either the differences between the scanning and behavioral testing environments or, more likely, the longer sessions for scanning.

Because we were interested in the initial (nonextinguished) effect of rewarded sounds on visual discrimination, we discarded these last trials (~3 miniblocks of data, a total of 48 trials out of 288 trials, corresponding to 24 trials of each reward level; Fig. 1*E*). Fig. 1*F* shows d' values and percentage correct of the remaining data. Repeated-measures ANOVA showed a significant main effect of reward on d' ($F_{1,19} = 4.89$, $P = 0.03$), but the interaction between reward and spatial congruency was not significant ($F_{1,19} = 0.71$, $P = 0.41$). In a pairwise comparison of high rewards and low rewards, a significant effect of reward was present only when Gabor and sound were spatially congruent, for both d' ($P = 0.03$ for HC vs. LC, paired t test) and percent correct ($P = 0.004$ for HC vs. LC, paired t test). All other pairwise comparisons of d' and percent correct were nonsignificant. Herein we use d' as a measure of performance. Before reward training, behavioral performance for the two sounds did not differ (Fig. S1 *A* and *B*). Cross-modal value had no effect on reactions times (Fig. S1 *C* and *D*).

The largest and most parsimonious effect of cross-modal reward was observed when stimuli in two modalities were spatially
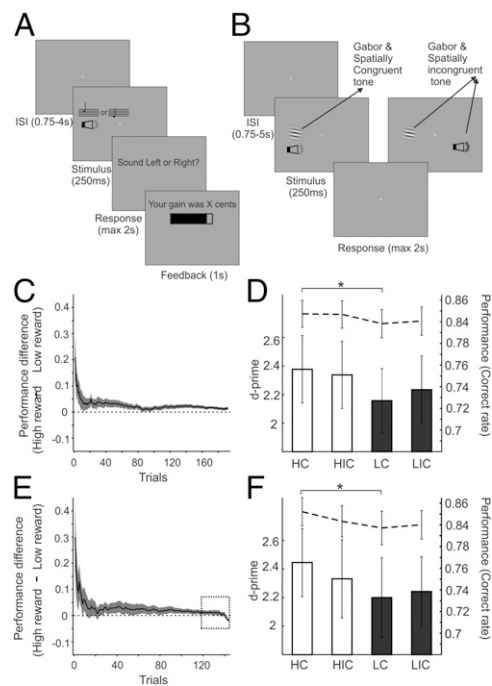

Fig. 1. Behavioral paradigm and results. (*A*) Auditory reward conditioning trials. Subjects determined whether a pure tone was played from the left side or the right side. Two tone frequencies (160 and 380 Hz, counterbalanced across subjects) were consistently paired with high or low monetary reward. Subjects learned reward pairings while performing the localization task. (*B*) Visual orientation discrimination task in the presence of sounds. Simultaneous with the Gabor presentation, a sound was played either from the same side as the Gabor or from the other side (congruent and incongruent, respectively). Subjects indicated the tilt orientation of a Gabor stimulus (clockwise or counterclockwise relative to the horizontal meridian). A "pretest" block with this task was recorded before subjects learned the reward associations. Thereafter, they were trained with the sounds (as shown in *A*) for ~50 trials. Our main experimental blocks were then recorded with interleaved short blocks of the reward training and orientation judgment tasks. (*C*) Cumulative performance difference between high-reward and low-reward stimuli, outside the scanner. Shaded area represents ± SEM. (*D*) Average psychophysical performance. Bars depict d', and the curve corresponds to the correct rate, outside the scanner. (*E*) Same as C, inside the scanner. The rectangle shows that in last trials, the effect was diminished and even reversed. These trials ($n = 24$ for each reward level) were not included in our analysis. (*F*) Same as D, inside the scanner. Error bars represent SEM.

congruent (i.e., in HC vs. LC). However, even when the two stimuli were at different locations, performance was marginally better for high-reward stimuli, which explains why we did not find a significant interaction between reward and congruence. This finding indicates that our observed effects are not strictly spatially specific, and we believe it is likely that spatially nonspecific mechanisms, such as general arousal/alertness, play a role as well. Nevertheless, we note that spatially unspecific mechanisms (e.g., alertness/arousal) are not specific to cross-modal value transfer and could occur for any rewarding stimulus. These considerations aside, our main interest in the present study was related to cross-modal effects at early stages of cortical processing. For this reason, we use the contrast between HC and LC as a proxy for the cross-modal effect of value on behavioral orientation discrimination, and in our analyses this contrast will be the main contrast of interest.

Modulation of Visual Cortex by Cross-Modal Reward Value. Fig. 2 shows the results of pattern classification analysis in individually defined visual areas (defined based on an orthogonal contrast; i.e., responses to the left visual hemifield > responses to the right hemifield and vice versa, for defining visual areas in the right and left hemispheres, respectively; *Materials and Methods*), as well as

the average response magnitude in the same regions of interest (ROIs) (Fig. 2 *A* and *B*). Classification accuracies were highest for the HC condition [average accuracies: 58.6% for HC, 53.4% for HIC, 51% for LC, and 52.6% for the low-reward–incongruent (LIC) condition; $P = 0.002$, 0.01, 0.38, and 0.14, respectively, for comparison with chance, i.e., 50%, paired t test]. Repeated-measures ANOVA with accuracy as the dependent factor and reward and spatial congruence as independent factors revealed a significant main effect of reward ($F_{1,16} = 4.77$, $P = 0.044$) and a significant interaction between reward and congruence ($F_{1,16} = 4.55$, $P = 0.048$). Pairwise comparisons showed that the effect of reward was significant only when the sound and Gabor were spatially congruent ($P = 0.007$, HC vs. LC, paired t test). This effect, greater accuracy in HC compared with LC, was significantly correlated with the difference in behavioral d' between these two conditions ($r = 0.61$, $P = 0.009$). Correlation between classification accuracy of the visual cortex and behavioral performance in all other pairwise conditions was nonsignificant (HC-LIC: $r = -0.02$, $P = 0.91$; HIC-LC: $r = 0.11$, $P = 0.66$; HIC-LIC: $r = -0.07$, $P = 0.76$). The average response magnitude of the visual cortex was not affected by cross-modal value ($P > 0.05$ for all, for main effect or interaction with reward and pairwise comparisons).

We replicated these results when eye position offsets were included in our generalized linear models (GLMs) for 11 subjects for whom eye-tracking data were available (*SI Materials and Methods* and Fig. S2). These results show that the value associated with the sounds affects the accuracy of orientation coding in the visual cortex.
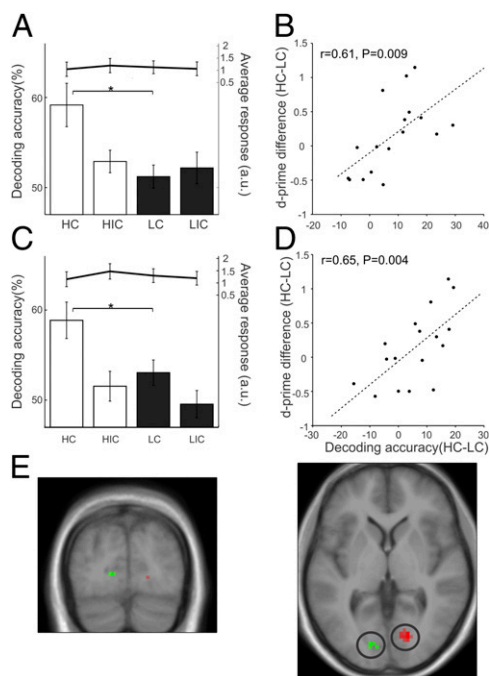
We conducted a number of additional tests to verify these results. First, we ensured that the differential effect of the two sounds on visual orientation coding was related to a difference in reward value as opposed to any difference in their physical attributes (frequency or perceived amplitude). To this end, we repeated our classification analysis for the data of the pretest block, in which subjects were not yet familiarized with the sound values. As shown in Fig. S2, in this pretest block, classification accuracies did not differ between the two sounds.

Second, we replicated our results using the same ROIs for all subjects (Fig. 2 *C* and *D*). These ROIs were defined based on the orthogonal contrast of responses to the left > responses to the right for the right visual cortex and vice versa for the left visual cortex at the group level, and were masked by anatomical ROIs of V1 and V2 (*Materials and Methods*). Classification accuracies were highest for the HC condition (average 59% for HC, 51.5% for HIC, 53% for LC, and 49% for LIC). Repeated-measures ANOVA revealed a significant main effect of reward ($F_{1,16} = 7.05$, $P = 0.01$), but a nonsignificant effect of congruence and the interaction term. In pairwise comparisons, accuracies were significantly different between the HC and LC conditions ($P = 0.03$ for comparison of HC and LC, paired t test). The greater accuracy in HC compared with LC was significantly correlated with the behavioral effect ($r = 0.65$, $P = 0.004$).

Third, our results were further corroborated by a searchlight analysis in which accuracies were independently computed for each voxel (Fig. 2*E* and *Materials and Methods*). This analysis revealed two clusters, one in each hemisphere (demarcated by dashed circles in Fig. 3*C*), within the same mask used for the second analysis in which classification accuracies were significantly higher for contralateral HC compared with LC stimuli. The activation of the right cluster (peak x, y, z: 18, −73, −2) was significant when corrected for multiple comparisons ($P < 0.05$, small-volume familywise error-corrected). Thus, the same stimulus-specific regions of cortex demonstrated a reward-related change in stimulus representation when classification was independently assessed for each voxel. In addition to these clusters in early visual areas, congruent rewards increased the classification accuracies in a number of other regions, encompassing visual- as well as memory-, attention-, and reward-related areas (Table S2).

Representation of Value Within Reward-Sensitive Regions and Cross-Modal Areas. Based on our specific hypotheses as outlined earlier, we tested the effect of cross-modal value on activations within three anatomically defined ROIs corresponding to the ventral striatum, vmPFC, and cross-modal areas (Fig. 3). Note that our aim here was to compare two specific hypotheses, rather than to exhaustively test a large number of different brain regions, and as such, we present the present results as exploratory rather than definitive. All of these areas exhibited significantly greater activation for the Gabor + high-reward stimuli compared with the Gabor + low-reward stimuli ($P < 0.05$ for all, Wilcoxon signed-rank test), but only in cross-modal areas [i.e., STS/superior temporal gyrus (STG)] was the effect correlated with a behavioral change in d' ($r = 0.51$, $P = 0.03$).

To rule out the possibility that this correlation is driven by potential outliers, we performed additional robust regression analyses (using the robustfit algorithm in MATLAB with the Huber weighting function). This analysis showed a significant correlation between the effect size in STS/STG and change in d' ($t = 2.14$, df = 15, $P < 0.05$). Robust regression analysis also confirmed that effect size in vmPFC and striatum was not significantly correlated with behavioral change in d' ($P > 0.05$ for all). Formal model comparison using the Bayes information criterion (BIC) (*SI Materials and Methods*) strongly favored the model containing data from the STS (BIC: STS, −26.49; striatum, −31.58; vmPFC, −31.49). The posterior probability of the STS model was 0.99. Within the context of our limited model space, this finding provides strong evidence in favor of the hypothesis that the behavioral effects that we observed are crucially mediated by

Fig. 2. Effect of cross-modal value on visual areas. (*A*) Classification accuracies (bars) and average responses of individually defined visual ROIs (line). (*B*) Correlation between differences in classification accuracy and differences in d' values for HC vs. LC. Error bars represent SEM. (*C*) Same as *A*, computed in group-level ROIs. Identical ROIs were used for all subjects. (*D*) Correlation between difference in classification accuracy and difference in d' values for HC vs. LC in group-level, visual ROIs. (*E*) Searchlight results masked with group-level visual ROIs (as used for *C* and *D*), showing two clusters in right (red) and left (green) visual cortex. For demonstration purposes, data are thresholded at $P = 0.05$, uncorrected. The right cluster had significant activation when corrected for multiple comparison ($P < 0.05$, small-volume FWE corrected within ipsilateral visual ROI).
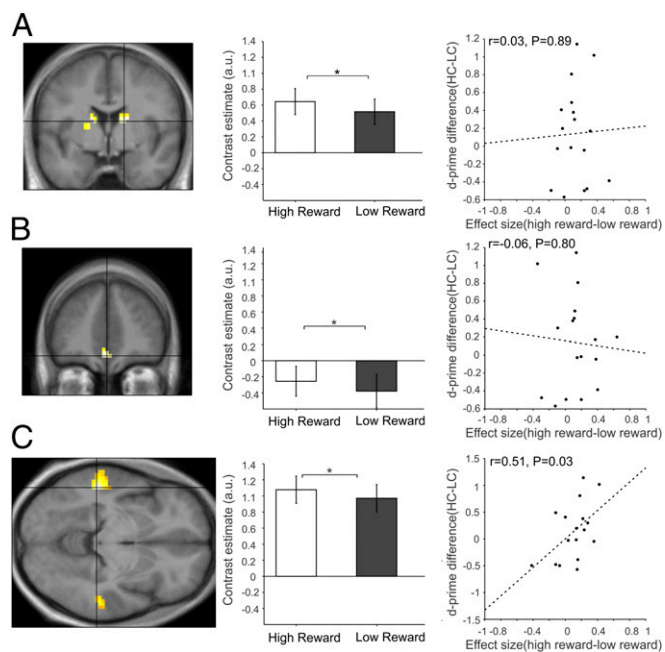
Fig. 3. Effect of cross-modal value on reward network and cross-modal areas. (*A*) Contrast between high-reward and low-reward conditions, masked by anatomically defined ROI of ventral striatum. *P* < 0.005, uncorrected for display purposes (cursor at peak voxel: x, y, z = 15, −1, 13). (*Center*) Bar graphs showing the contrast estimates for the high-reward vs. low-reward conditions. (*Right*) Scatterplot showing the correlation of striatal effect size with the behavioral effect (HC-LC). (*B*) Same as in *A*, for the vmPFC ROI (cursor at peak voxel: x, y, z = 0, 44, −14). (*C*) Same as in *A* for the STS/STG ROI (cursor at peak voxel: x, y, z = −51, −34, −8). The correlation between STS/STG and *d′* difference was robust to potential outliers and remained significant (*P* < 0.05) when we used a least squares fitting procedure that minimizes the effects of outliers (Robustfit in MATLAB); see *Results* for details. Error bars represent SEM. *P < 0.05, Wilcoxon signed-rank test.

multisensory processing areas. Importantly, before subjects learned the reward associations, STS/STG showed no difference in responses to high-reward stimuli and low-reward stimuli (Fig. S3; *P* > 0.05, Wilcoxon signed-rank test), indicating that a modulation by sound could not be related to any difference in the sounds' inherent physical attributes (frequency or perceived amplitude).

We next tested whether the modulation of STS/STG after reward training was correlated with classification accuracies obtained from ipsilateral visual areas, an indication that feedback from cross-modal areas affects primary sensory areas (13). We found a significant correlation between STS/STG response modulation (i.e., high reward vs. low reward) and classification accuracies in visual areas (*r* = 0.50, *P* = 0.037). This correlation was still significant when a robust regression method was used (*t* = 2.13, df = 15, *P* < 0.05).

The results of our whole-brain analysis of the main contrast of high reward vs. low reward and high reward congruent vs. low reward congruent (HC-LC) indicate that in addition to reward and cross-modal regions, a number of areas classically involved in motor planning and executive functions (e.g., supplementary motor areas) were modulated by sound values (Table S1). Thus, our univariate analysis results show that during the orientation discrimination task, sound value is represented in reward processing and executive control networks, as well as in cross-modal areas. Activation within the STS/STG was correlated with the magnitude of the behavioral effect (*SI Results*), suggesting that this area may play a role in mediating the effect of cross-modal reward associations.

Discussion

Reward associations influence perception (5–8, 10). Here we show that these influences extend across different sensory modalities. Sounds previously paired with monetary reward modulated visual perception, leading to better orientation discrimination when a task-irrelevant sound with high-reward associations was presented concurrently with the visual stimulus. There was also a suggestion that the effect of reward was strongest when the auditory and visual stimuli were spatially congruent, although the differences between congruent and incongruent conditions were not statistically significant. Stimulus-specific regions of early visual areas contained a more accurate representation of stimulus orientation when high-reward conditions (signaled by a task-irrelevant auditory stimuli), as opposed to low-reward or incongruent conditions, were presented. This effect was correlated with an increase in orientation sensitivity for high-reward stimuli. In the same regions of the visual cortex, the magnitude of responses did not differ between the two reward conditions, suggesting that reward modulates perception, at least in part, through an increase in the distinctness of neuronal representations rather than through a homogenous response enhancement. Responses within cross-modal areas (i.e., STS/STG) were significantly affected by the reward value of sounds, and this effect was highly correlated with both behavioral and BOLD correlates of visual orientation sensitivity, suggesting that these multimodal areas may underlie the effect of rewarded sound on vision.

The beneficial effect of reward information on sensory processing is in line with the results of a host of previous studies (5–8, 10, 21, 24, 25). These studies have demonstrated that in early sensory areas, the same neurons that process sensory information are modulated by reward (5, 10, 26) and thereby influence perception from the earliest stages of cortical processing. Our findings extend this previous work by showing that task-irrelevant cues presented in another sensory modality modulate perceptual processing according to their reward-predicting properties.

Simultaneously presented sounds can influence visual perception (1, 2) with visual detection or discrimination typically enhanced in the presence of sounds (27–29). Maximal enhancement occurs when visual stimuli are near the detection/discrimination threshold (28), the auditory and visual stimuli overlap in time and in space (13, 29, 30) and sounds carry biologically important information (27, 29, 31). We show that value can modulate audiovisual interactions (and, we hypothesize, multisensory interactions more generally), an effect congruent with ideas on the Bayesian brain (32–34), in which a core idea is that the brain embodies previous beliefs about the structure (i.e., statistics) of its environment. Specifically, here brief temporally (and perhaps spatially) contiguous stimuli are likely to be associated, and thus to share motivational properties. Therefore, visual stimuli have increased importance or salience (and hence presumably attention; ref. 35) conferred on them by concurrently presented high-reward sounds (34, 36). In this broader context, our results have implications for approaches in which basic perceptual and cognitive processes are reappraised in the light of the expectations of the brain about the causal structure of its environment, which here we extend to the domain of motivation.

We suggest that the cross-modal effects that we describe herein are triggered by the reward associations of a task-irrelevant stimulus in another sensory modality, and provide preliminary evidence that a sound's reward information is transmitted to the visual cortex in the absence of visual stimuli, when attentional effects are minimal (Fig. S4). However, we note that it has proven difficult, if not impossible, to distinguish between reward and attention effects (37) in many studies, and in the present study as well. Attentional effects across sensory modalities are well described for both exogenous spatial attention (38) and object-based attention (39). The perceptual benefits of attention observed in these previous studies are similar to our effects but with several key differences. Exogenous attention [as in, e.g., Störmer et al. (38)] is mediated by the sudden appearance of a stimulus in a certain spatial location. This type of attention by itself cannot explain our effects, because exogenous

cues will improve the processing of any stimulus that appears at the cued location. Nonetheless, we find that whereas discrimination sensitivity was increased for HC, it was strongly decreased for LC, which could not be predicted by exogenous attentional effects except when modulated by reward associations. In this case, reward could modulate the "saliency" of exogenous cues, and thus a high-reward sound could be more effective at summoning attention to a specific spatial location. Endogenous attention, that is, the voluntary shift of attention to a certain location (40), might be triggered by rewarded sounds and could be then transferred to vision (39). This possibility is unlikely, however, given that classical endogenous attention effects need time to develop [a minimum of 300 ms (40)], whereas in the present study, sounds and Gabor stimuli were presented briefly and simultaneously. Future studies are needed to test whether the cross-modal value effect can occur when the influence of attention is minimized, as well as to map the temporal profile of these effects, which could point to the possible underlying mechanisms.

Although mean changes in activity levels are often considered the signature of perceptual benefits of attention and reward, some studies have shown that reward- and information-predictive cues might suppress redundant visual responses and thus sharpen the representation of relevant visual information (10, 22). We did not find that incidental reward signals result in enhancement of visual response amplitudes; instead, our data support an emerging account in which reward/information cues increase the signal-to-noise ratio so as to improve perception. However, it is also possible that although responses to visual stimuli in the reward condition were larger in magnitude, they were also shorter in duration, and thus produced no clear differences in the BOLD contrast. Another possibility is that cross-modal value modulates subthreshold neuronal responses and thus does not produce a net response enhancement. This is an interesting topic for future study, and would be well suited to neuroimaging techniques with a high temporal resolution, such as EEG and magnetoencephalography.

We found reward associations in a number of regions classically associated with value. Among these, only activity in the STS/STG, an area of known importance in cross-modal interactions (14–16), was correlated with the behavioral effect, and was a better predictor of behavior when a formal model comparison was performed. This suggests a model in which reward associations, either generated in regions typically associated with value or reflected in local processing in the sensory cortex (8, 26), influence sensory processing in other modalities via cross-modal association areas, such as the STS/STG. This suggestion is in line with a task-specific role of reward, in which the specific intervening processing stages involved in a task, not the final effectors alone, are informed and influenced by value-related information. Whether activity in cross-modal areas plays a causal role in increased perceptual sensitivity remains to be established.

The present study has addressed the question of whether task-irrelevant reward associations in one sensory modality improve the processing of temporally congruent stimuli presented in another modality. We have shown that they do, and have presented evidence indicating that increases in perceptual accuracy are reflected in the discriminability of neuronal representations, very likely mediated by activity in cross-modal areas. This finding provides insight into the expectations that the brain has about the causal structure of its environment, and sheds light on a mechanism that in and of itself is likely important for adaptive behavior in ecological contexts.

Materials and Methods

Participants. Forty-four subjects participated in either a behavioral study outside the scanner (n = 24; 14 females; mean age, 25 y; range, 22–33 y) or a combined behavioral and fMRI experiment inside the scanner (n = 20; 11 females; mean age, 24 y; range, 23–30 y). The data for three subjects from the fMRI experiments were analyzed only in terms of behavioral effects, owing to technical problems with image acquisition. Each subject provided oral and written consent for his or her participation. The study was approved by the local Ethics Committee of Berlin Charité University Hospital.

Stimuli and Tasks. We used a Pavlovian conditioning paradigm to familiarize the subjects with sounds and their associated rewards (Fig. 1A). The pairing of reward (high or low) with the sound frequency (160 Hz or 380 Hz) was counterbalanced across subjects. The visual orientation discrimination task (Fig. 1B) was adopted from one described in a previous study (29). In this task, after an initial fixation of variable duration, a Gabor stimulus was briefly (250 ms) presented at a parafoveal location, and subjects reported its tilt orientation (i.e., two-alternative forced choice, clockwise or counterclockwise relative to the horizontal meridian). The tilt of the Gabor from horizontal was set to each subject's discrimination threshold (75% performance). Subjects were instructed to maintain fixation throughout the trials (fixation circle diameter 1°, eye tracking described in *SI Materials and Methods*). At the same time as presentation of the visual stimulus, a tone was played either from the same side (spatially congruent) or opposite side (spatially incongruent) as the Gabor. The tone had either a low frequency (160 Hz) or high frequency (380 Hz) and was uninformative regarding the visual stimulus orientation. This provided for a 2 × 2 factorial design with sound identity (high reward and low reward) and its spatial congruence (congruent and incongruent) as the two independent factors. No feedback regarding correct performance was provided to the subjects. We collected a total of 199–288 trials of the audiovisual task, corresponding to 72 trials for each of the four experimental conditions. To avoid extinction, short miniblocks of sound localization and orientation judgment tasks were interleaved (*SI Materials and Methods*).

Univariate Analysis of the fMRI Data. Details of data acquisition and preprocessing are provided in *SI Materials and Methods*. To identify cortical regions modulated by sound rewards, we created a GLM with eight main regressors. These regressors included responses to HC, HIC, LC, and LIC conditions, modeled separately for left and right visual hemifields. Because reward conditioning and orientation discrimination miniblocks occurred alternately and in close temporal proximity to each other, four additional regressors were included to account for reward conditioning, so as to model two types of sound–reward pairs (high and low reward) on each side (left and right). Finally, two additional regressors were included to model the instruction display presented at the beginning of each miniblock (one for sound miniblocks and one for audiovisual miniblocks). Thus, the full GLM model included 12 regressors, each of which was modeled as a stick function at the onset of the corresponding stimulus on each trial, convolved with a hemodynamic response function (41).

Our behavioral analysis revealed that the effect of sounds on orientation discrimination performance decayed toward the end of the experimental session, with a reversal of the effect in the last three miniblocks (Fig. 1E). Thus, we included an additional regressor in our GLM to model the effect of the last reversal trials modeled with a boxcar function that covered all of the events after the third audiovisual miniblock of the last run (and thus undivided to different conditions). Six session-specific motion parameters were modeled as covariates of no interest. To test for regional-specific condition effects, we used linear contrasts for each subject and each condition (first-level analysis). These contrasts included the main effect of reward [(HC + HIC) − (LC + LIC)] pooled across congruency conditions and the contrast between HC and LC conditions. The resulting contrast images were entered into a second-level analysis, and significance was assessed by one-sample t tests. Whole-brain results were thresholded at $P < 0.001$ (uncorrected, $k = 10$; Table S1).

To define the visually responsive ROIs, we identified areas activated by contralateral visual stimuli, showing higher activation in the contrast of left Gabor vs. right Gabor and vice versa as in our univariate, first-level GLMs. This contrast thus highlights right (higher activation for left Gabors vs. right Gabors) and left (higher activation for left Gabors vs. right Gabors) visual cortices for each individual subject. Note that this contrast (left vs. right Gabors) is orthogonal to both reward and spatial congruence (fully randomized across the two visual fields), and thus the activity of selected visual ROIs is not directionally biased to show a difference between any of the factors of interest.

We also reproduced our results with a second set of visual ROIs that were identical for all of the subjects. Here we used the second-level contrast of left Gabor vs. right Gabor to identify the right and left visual areas at the group level. We then masked these activations with anatomically defined ROIs of areas V1 and V2 constructed based on the anatomy toolbox of SPM (42). This enabled us to identify visually responsive regions that matched the tissue probability maps of early visual areas known to be involved in low-level processing of stimulus orientation (43). Ideally, we also could have used detailed retinotopic mapping techniques to precisely map the topographic location of cortical areas and to identify our visual ROIs in relation to these topographic maps, but in the interest of time, we decided not to do so. For regions outside the visual cortex, ROIs were selected anatomically. The ventral striatum ROI consisted of the bilateral caudate, putamen, and globus pallidus, the vmPFC ROI comprised the bilateral gyrus rectus and medial

orbitofrontal gyrus, and the STS/STG ROI composed bilateral superior and middle temporal gyri as defined in the Anatomical Automatic Labeling atlas (44). The effect sizes of these ROIs (Fig. 3) were computed by averaging the contrast estimates across all voxels of an anatomical ROI.

MVPA of fMRI Data. We performed a MVPA to assess the effect of rewarded sounds on a trial-by-trial pattern classification accuracy of early visual areas (details provided in *SI Materials and Methods*). We used several approaches for the MPVA. In the first approach (Fig. 2 *A* and *B*), we performed a pattern classification analysis in individually defined visually responsive ROIs. The input of the support vector machine classifiers consisted of *t*-values of every trial for all voxels of a certain ROI. Eight classifiers were constructed, four for each hemisphere, to decode the tilt orientation (clockwise or counterclockwise) in the HC, HIC, LC, and LIC conditions of the contralateral visual field

(for a total of eight classifiers). The accuracies were then averaged across the two hemispheres (Fig. 2 *A* and *B*).

We also performed two supplementary analyses (Fig. 2 *C* and *D*). First, we used the same visual ROIs for all of the subjects, minimizing the possibility of subject-specific selection bias. Then we used a whole-brain searchlight method (*SI Materials and Methods*), which avoids ad hoc voxel selection (45). Note that here a whole-brain analysis was performed, and visual ROIs were subsequently used for small-volume correction.

1. Stein BE, Meredith MA (1993) *The Merging of the Senses* (MIT Press, Cambridge, MA).
2. Meredith MA, Stein BE (1983) Interactions among converging sensory inputs in the superior colliculus. *Science* 221(4608):389–391.
3. Driver J, Noesselt T (2008) Multisensory interplay reveals crossmodal influences on "sensory-specific" brain regions, neural responses, and judgments. *Neuron* 57(1):11–23.
4. Noppeney U (2012) Characterization of multisensory integration with fMRI: Experimental design, statistical analysis, and interpretation. *The Neural Bases of Multisensory Processes*, eds Murray MM, Wallace MT (CRC Press, Boca Raton, FL).
5. Polley DB, Heiser MA, Blake DT, Schreiner CE, Merzenich MM (2004) Associative learning shapes the neural code for stimulus magnitude in primary auditory cortex. *Proc Natl Acad Sci USA* 101(46):16351–16356.
6. Serences JT (2008) Value-based modulations in human visual cortex. *Neuron* 60(6):1169–1181.
7. Pleger B, Blankenburg F, Ruff CC, Driver J, Dolan RJ (2008) Reward facilitates tactile judgments and modulates hemodynamic responses in human primary somatosensory cortex. *J Neurosci* 28(33):8161–8168.
8. Weil RS, et al. (2010) Rewarding feedback after correct visual discriminations has both general and specific influences on visual cortex. *J Neurophysiol* 104(3):1746–1757.
9. Brosch M, Selezneva E, Scheich H (2011) Representation of reward feedback in primate auditory cortex. *Front Syst Neurosci* 5:5.
10. Arsenault JT, Nelissen K, Jarraya B, Vanduffel W (2013) Dopaminergic reward signals selectively decrease fMRI activity in primate visual cortex. *Neuron* 77(6):1174–1186.
11. Meredith MA, Stein BE (1986) Spatial factors determine the activity of multisensory neurons in cat superior colliculus. *Brain Res* 365(2):350–354.
12. Spence C, Pavani F, Driver J (2004) Spatial constraints on visual-tactile cross-modal distractor congruency effects. *Cogn Affect Behav Neurosci* 4(2):148–169.
13. Noesselt T, et al. (2007) Audiovisual temporal correspondence modulates human multisensory superior temporal sulcus plus primary sensory cortices. *J Neurosci* 27(42):11431–11441.
14. Calvert GA, Campbell R, Brammer MJ (2000) Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Curr Biol* 10(11):649–657.
15. Beauchamp MS, Argall BD, Bodurka J, Duyn JH, Martin A (2004) Unraveling multisensory integration: Patchy organization within human STS multisensory cortex. *Nat Neurosci* 7(11):1190–1192.
16. Werner S, Noppeney U (2010) Superadditive responses in superior temporal sulcus predict audiovisual benefits in object categorization. *Cereb Cortex* 20(8):1829–1842.
17. Hollerman JR, Tremblay L, Schultz W (2000) Involvement of basal ganglia and orbitofrontal cortex in goal-directed behavior. *Prog Brain Res* 126:193–215.
18. Gottfried JA, O'Doherty J, Dolan RJ (2003) Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science* 301(5636):1104–1107.
19. Schultz W, Apicella P, Ljungberg T, Romo R, Scarnati E (1993) Reward-related activity in the monkey striatum and substantia nigra. *Prog Brain Res* 99:227–235.
20. Elliott R, Dolan RJ, Frith CD (2000) Dissociable functions in the medial and lateral orbitofrontal cortex: Evidence from human neuroimaging studies. *Cereb Cortex* 10(3):308–317.
21. Goltstein PM, Coffey EB, Roelfsema PR, Pennartz CM (2013) In vivo two-photon Ca2$^+$ imaging reveals selective reward effects on stimulus-specific assemblies in mouse visual cortex. *J Neurosci* 33(28):11540–11555.
22. Kok P, Jehee JF, de Lange FP (2012) Less is more: Expectation sharpens representations in the primary visual cortex. *Neuron* 75(2):265–270.
23. Holland PC (1979) Differential effects of omission contingencies on various components of Pavlovian appetitive conditioned responding in rats. *J Exp Psychol Anim Behav Process* 5(2):178–193.
24. David SV, Fritz JB, Shamma SA (2012) Task reward structure shapes rapid receptive field plasticity in auditory cortex. *Proc Natl Acad Sci USA* 109(6):2144–2149.
25. Stănişor L, van der Togt C, Pennartz CM, Roelfsema PR (2013) A unified selection signal for attention and reward in primary visual cortex. *Proc Natl Acad Sci USA* 110(22):9136–9141.
26. Fitzgerald TH, Friston KJ, Dolan RJ (2013) Characterising reward outcome signals in sensory cortex. *Neuroimage* 83:329–334.
27. Romei V, Murray MM, Cappe C, Thut G (2009) Preperceptual and stimulus-selective enhancement of low-level human visual cortex excitability by sounds. *Curr Biol* 19(21):1799–1805.
28. Noesselt T, et al. (2010) Sound-induced enhancement of low-intensity vision: Multisensory influences on human sensory-specific cortices and thalamic bodies relate to perceptual enhancement of visual detection sensitivity. *J Neurosci* 30(41):13609–13623.
29. Leo F, Romei V, Freeman E, Ladavas E, Driver J (2011) Looming sounds enhance orientation sensitivity for visual stimuli on the same side as such sounds. *Exp Brain Res* 213(2-3):193–201.
30. Marchant JL, Ruff CC, Driver J (2012) Audiovisual synchrony enhances BOLD responses in a brain network including multisensory STS while also enhancing target-detection performance for both modalities. *Hum Brain Mapp* 33(5):1212–1224.
31. Arrighi R, Marini F, Burr D (2009) Meaningful auditory information enhances perception of visual biological motion. *J Vis* 9(4):1–7.
32. Friston K (2005) A theory of cortical responses. *Philos Trans R Soc Lond B Biol Sci* 360(1456):815–836.
33. Knill DC, Pouget A (2004) The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends Neurosci* 27(12):712–719.
34. Fiser J, Berkes P, Orbán G, Lengyel M (2010) Statistically optimal perception and learning: From behavior to neural representations. *Trends Cogn Sci* 14(3):119–130.
35. Dayan P, Kakade S, Montague PR (2000) Learning and selective attention. *Nat Neurosci* 3 Suppl:1218–1223.
36. Simoncelli EP, Olshausen BA (2001) Natural image statistics and neural representation. *Annu Rev Neurosci* 24:1193–1216.
37. Maunsell JH (2004) Neuronal representations of cognitive state: Reward or attention? *Trends Cogn Sci* 8(6):261–265.
38. Störmer VS, McDonald JJ, Hillyard SA (2009) Cross-modal cueing of attention alters appearance and early cortical processing of visual stimuli. *Proc Natl Acad Sci USA* 106(52):22456–22461.
39. Busse L, Roberts KC, Crist RE, Weissman DH, Woldorff MG (2005) The spread of attention across modalities and space in a multisensory object. *Proc Natl Acad Sci USA* 102(51):18751–18756.
40. Carrasco M (2011) Visual attention: The past 25 years. *Vision Res* 51(13):1484–1525.
41. Friston KJ, Jezzard PJ, Turner R (1994) Analysis of functional MRI time-series. *Hum Brain Mapp* 1:153–171.
42. Eickhoff SB, et al. (2005) A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage* 25(4):1325–1335.
43. Hubel DH, Wiesel TN (1974) Sequence regularity and geometry of orientation columns in the monkey striate cortex. *J Comp Neurol* 158(3):267–293.
44. Tzourio-Mazoyer N, et al. (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15(1):273–289.
45. Kriegeskorte N, Goebel R, Bandettini P (2006) Information-based functional brain mapping. *Proc Natl Acad Sci USA* 103(10):3863–3868.