

# Reward sharpens orientation coding independently on attention

Stefano Baldassi<sup>1</sup> and Claudio Simoncini<sup>2</sup>

<sup>1</sup> Department of Psychology, University of Florence, Via di San Salvi, 12, Complesso di San Salvi Padiglione 26, 50135 Firenze, Italy

<sup>2</sup> DyVA Group - Institut de Neurosciences Cognitives de la Méditerranée (CNRS), 31, Chemin Joseph Aiguier, 13402 Marseille cedex, France

Correspondence should be addressed to S.B. ([stefano.baldassi@unifi.it](mailto:stefano.baldassi@unifi.it))

## Abstract

Rewarding improves performance. Is it due to modulations of the output modules of the neural systems or are there mechanisms favoring more 'generous' inputs? Some recent study included V1 in the the circuitry of reward-based modulations, but the effects of reward can easily be confused with effects of attention. Here we address this issue with a psychophysical dual task to control attention while orientation sensitivity on targets associated to different levels of reward is measured. We found that different reward rates improve orientation discrimination and sharpen the internal response distributions. Data are unaffected by changing attentional load nor by dissociating the feature of the reward cue from the feature relevant for the task. This suggests that reward may act independently on attention by modulating the activity of early sensory stages, perhaps V1, through a SNR improvement of task-relevant channels. Reward acts like attention, but using separate channels.

*“Gold? Yellow, glittering, precious gold?  
No, Gods, I am no idle votarist! ...  
Thus much of this will make black white, foul fair,  
Wrong right, base noble, old young, coward valiant.  
... Why, this”* (Shakespeare in *Timon of Athens*, 1623)

The activity of the visual channels, both at the neuronal and at the overall behavioral level, can be modulated by virtue of several sources of influence. Many such modulatory activities depend on the global behavioral state of the organism, driven by cognitive, emotional or motivational factors. Since these states have a profound impact on the behavioral performance of the individuals, determining successes or failures of goal-directed behavior, their mechanisms of action have attracted the interests of psychologists, cognitive neuroscientists and neurophysiologists for long time.

Attention and reward are among the most studied modulating factors of behavior and have traditionally been investigated independently from one another. In general, as we refer to modulations of the sensory systems and of perceptual performance, the idea of attention is more often reflecting fast, short-term modulation based on exogenous or endogenous cues to bias the processing power toward specific spatial location or stimulus features. Reward of specific actions or classes of stimuli is instead investigated assuming that it exerts long-term effects on sensory channels; these effects would alter the learning process toward specific stimuli or classes of stimuli and/or responses. Moreover, selective attention has been often studied in its relations to changes of the early stages of the input-output flow of information processing, with an active effort of finding neural correlates of attention in the Visual Area V4<sup>1-3</sup>, V2<sup>3, 4</sup>, V1<sup>5, 6</sup> and as early as in the LGN<sup>7, 8</sup>. On the other hand, reward has been widely studied as a variable affecting the later stages, closer to mechanisms related to visual-motor transformations<sup>9</sup>, to the decision-making modules<sup>10, 11</sup> and to the overt behavior<sup>12</sup>. More recently a number of studies have shifted the focus backward in the attempt of seeking the effect of reward to purely sensory areas, opening new doors for re-framing the functional properties of the early visual modules<sup>13-15</sup>. However, since reward is a built-in feature of the neurophysiological paradigms of attention, focusing on the early effects of reward has the implicit risks of confusing the effects of reward with those of attention<sup>16</sup>. Moreover, recent proposals have raised the idea that perceptual performance can be modulated by reward through its action on the attentional system<sup>17</sup>, implying that attention has a monopoly over the modulation of perception.

Here we specifically address the functional liaison between attention and reward by investigating whether the probability of reward may become associated to a change of

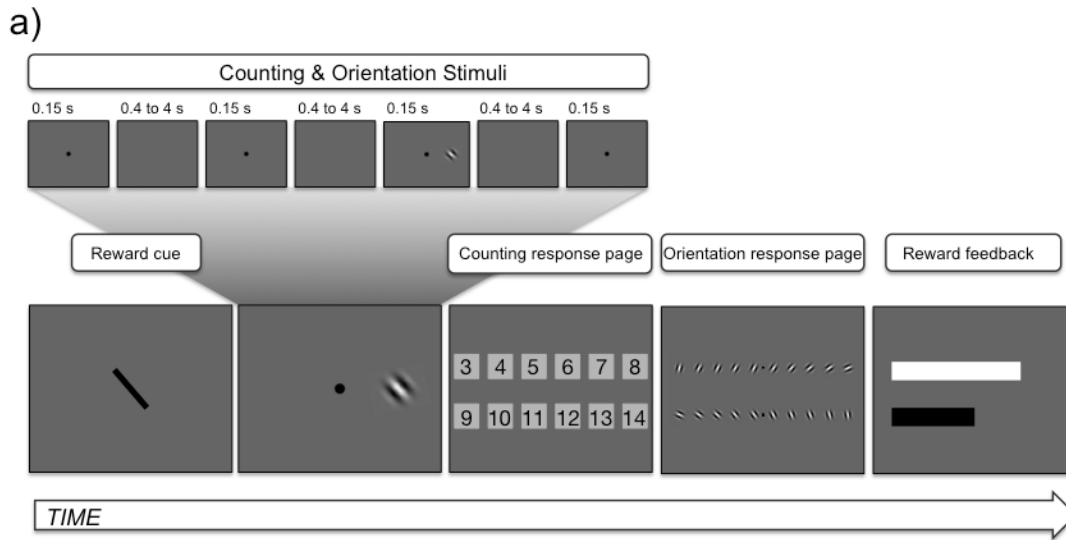
perceptual performance when attention is engaged in a concurrent task and learning is prevented by making the reward value associated to specific stimuli contingent on a trial-to-trial base. We have used a recently introduced psychophysical paradigm<sup>18</sup> to measure orientation discrimination acuity for a simple peripheral target (a task assumed to summon early mechanisms<sup>19, 20</sup>) and to obtain at once a quantitative estimate of the observer's noisy internal response distributions for any physical value of the target. Attention was controlled through the use of a concurrent task of varying difficulty, that has the key potential of showing independence of resources<sup>21</sup>, while learning could be excluded based on the fact that the same stimulus and the same response could be associated to one of two probabilities of obtaining reward (0.9 present/0.1 absent Vs 0.1 present/0.9 absent) unpredictably at each trial based on a precue (see Fig. 1). We found that a higher likelihood of earning credit to obtain a Scratch & Win ticket, a highly efficient and effective reward even in non-gamblers, improved performance. In particular, higher reward rates produced finer orientation acuity, as revealed by lower thresholds (about 50% decrease), and this was possibly due to a significant change of the channel's Signal-to-Noise Ratio (SNR), as revealed by sharper response distributions when the reward was more likely to be achieved. The reward-based modulation was unaffected by whether the orientation task was performed in isolation or it was competing with a central task of varying difficulty. Moreover, the effect was dissociated from the nature of the cue, as it remained stable when the cue was modulated in the colour domain and the task in the orientation domain.

Our results are coherent with the possibility that attention and reward may act independently to modulate perceptual performance and offer novel insights for studying reward and attention measuring their effect independently in the context of the same experimental paradigm.

## **RESULTS**

In the main task, observers were requested to perform a dual task based on the trial structure exemplified in Figure 1a that was leading to the reward pattern exemplified in Figure 1b. At the beginning of the trial a cue line, tilted 45° ClockWise (CW) or CounterClockWise (CCW) off-vertical was shown. It cued which axis of the peripheral stimulus yielded reward with a probability of 0.9 conditional to a correct identification of the tilt direction. Then the stimuli of the two tasks were shown. The attention loading task required observers to count a foveal disk (subtending 0.5° of visual angle) that was flashed for 100 ms a variable number of times (3 to 14 on a random base). In order to modulate the attentional load deputed to the central task, the contrast of the disk was adjusted to yield, in separate conditions, two levels of counting performance (~90% and

~60% of accuracy, respectively) and the interval between two consecutive disks was jittered within an interval varying between 0.4 and 4 seconds to avoid predictability about the timing of the upcoming disk. During the presence of one of the disks (except the first and the last) a peripheral target, synchronous with the disk, was displayed 7° of visual angle to the left or to the right of fixation. This was the reward-effective stimulus. The temporal onset and offset of the Gabor patch were aligned with onset and offset of the synchronous disk, to avoid attentional capture, and the duration of the display was short enough to prevent eye movements to the target. The oriented target was tilted randomly around one of two axes centered at -45° or +45° by a varying amount ( $\pm 2^\circ$  to  $\pm 32^\circ$  to yield a complete psychometric function), coinciding with the cue line shown at the beginning of the trial or orthogonal to it. If the axes of cue and stimulus coincided and a correct identification was made, i.e. a correct report of the direction of target tilt CW or CCW away from the reference, then there was a 90% chance of earning credit for obtaining a Scratch & Win ticket (High Reward Probability condition, HRP). If instead cue and stimulus were orthogonal (i.e., cue at -45° and stimulus at +45°, or viceversa), the same credit was obtained only in 10% of the correct identifications (Low Reward Probability, LRP condition). Wrong identifications voided the probability of reward. After the counting string disappeared, two response pages appeared, one containing the list of digits corresponding to the entire range of possible disk numbers, and observers had to click on the number they counted. The following response page contained Gabor probes representing the entire set of CW and CCW tilts for both the -45° and the +45° axis, (6 tilts x 2 directions x 2 axes), and observers were asked to click on the probe that matched more closely the perceived tilt. Responses were scored both binarily, as correct and wrong based on the direction of tilt of the clicked probe, to allow standard psychometric measures and independently for each probe to draw the histogram representing the distribution of reported tilts for each physical signal displayed<sup>18</sup>. We will call the two measures orientation *discrimination* and orientation *identification*, respectively. Trials in which counting was wrong were discarded from further analysis. At the end of each trial a feedback accumulation feedback page was displayed; it contained two bars, a white bar that was elongated if the outcome of the trial led to reward and a black bar that was elongated in the presence of a wrong identification. Unrewarded trials (in the presence of correct discrimination) were signaled by no change in either bars. The change to the bars was clearly visible to each subject. A Scratch & Win ticket was awarded at any 20 rewarded responses, then both bars were reset to 0.



b)

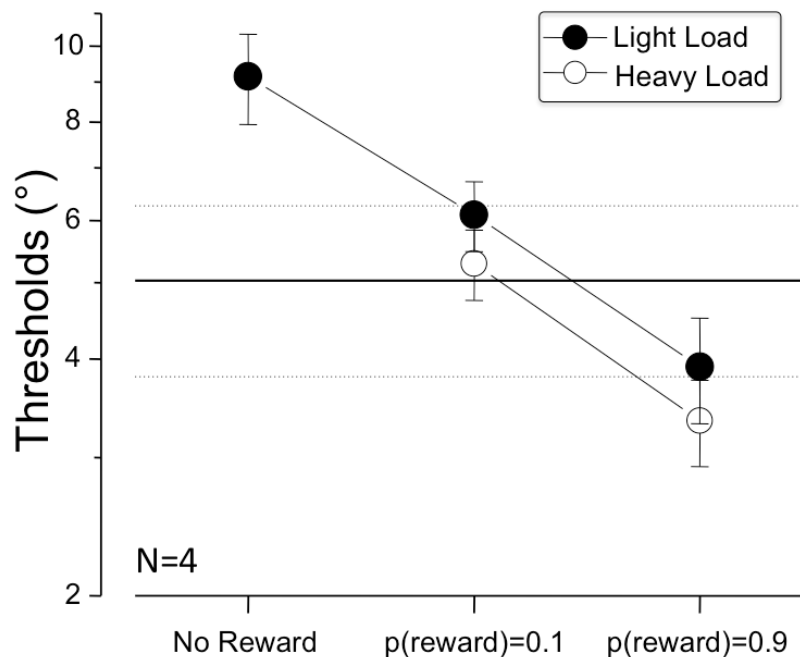
Rewardcue	rCCW  $p(\text{rCCW})=0.5$		rCW  $p(\text{CCW} \text{rCW})=0.5$	
Target	 $p(\text{CCW} \text{rCCW})=0.5$	 $p(\text{CW} \text{rCCW})=0.5$	 $p(\text{CCW} \text{rCW})=0.5$	 $p(\text{CW} \text{rCW})=0.5$
$p(\text{reward}   \text{correctid})$	0.9	0.1	0.1	0.9

**Figure 1.** Temporal structure of a trial (a) and reward patterns (b). **a)** A trial began with a foveal line ( $3^\circ$  long) displayed for 500 ms and tilted  $45^\circ$  ClockWise (CW) or CounterClockWise (CCW) from vertical. Immediately after the central attention loading task started. It consisted in the counting of a sequence of flashes (3 to 14 flashes on random base, 4 in the Figure) of foveal disks displayed for 100 ms with a random inter-disk interval (in the range of 0.4 to 4 s.). During the display of one of the disks, on a random base from the  $2^{\text{nd}}$  to the  $n^{\text{th}-1}$ , a peripheral oriented target was shown  $7^\circ$  to the left or to the right of fixation, with temporal onset and offset synchronous with the corresponding central disk. This was the reward-effective stimulus, a Gabor patch that was tilted CW or CCW relative to one of the two main oblique axes. Five hundred milliseconds after the disappearance of the last central disk the first response page was displayed; it contained all the digits of the range of possible disk numbers and subjects were asked to click with a mouse on the digit corresponding to the counted number of disks and to do within the limits of a brighter square containing the number. Subjects were clearly warned that wrong counting would have annulled the trial and that a new trial would have been added to the block. The next display contained the orientation identification and discrimination page. It contained 20 Gabor probes, one for each possible tilt around both the  $+45^\circ$  (upper line) and the  $-45^\circ$  axis (lower line). The five probes to the left, in each line, corresponded to CW tilts while the five to the right corresponded to CCW tilts. Observers had to click on the response probe that best matched the peripheral target, generating in

a single click, both an orientation discrimination response (based on the match of the CW vs CCW tilt with the stimulus) and an identification response (i.e the classification of the apparent tilt). Reward was granted based on the success in the discrimination task, independently on the precision of the identification. After the response, the last page of the trial sequence was shown. It contained a white and a black bar providing feedback about whether or not a trial led to reward, based on a visually salient size increase of the white or the black bar, respectively, and about the amount of rewarded identification needed to achieve another Scratch & Win ticket. The white bar was completed, and a ticket donated, after any 20 correct discriminations. **b)** Different lines of the table indicate, from top to bottom, the probability of each cue type, of each target type given the cue type, and the probability of earning reward given the combination of cues and targets. It has to be clear that: 1) there was an even probability (0.5/0.5) that any of the two cues were shown, 2) there was an even probability (0.5/0.5) that the target was tilted around the  $-45^\circ$  or the  $+45^\circ$  angle, but 3) the probability of earning a reward depended on whether the main axes of cue and target matched or not, according to a 0.9 vs 0.1 pattern, respectively.

**Main experiment.** We measured orientation discrimination thresholds and response distributions for three reward levels and two attentional load levels. A condition with no reward (Fig. 2 left points) was compared with two conditions in which correct orientation discrimination responses (i.e. selecting a CW angle when the target was tilted CW and CCW when the target was tilted CCW from the reference) yielded reward in 10% or in 90% of the trials (LRP and HRP, respectively), based on the match (or the mismatch) of the target's reference axis with the reward cue (Fig. 2, middle and left points). Attentional load to the central task was modulated by changing the contrast of the central dot from 80%, yielding to a counting accuracy of about 90%, to contrast threshold (based on preliminary measures), yielding to counting performances from 50% to 60% in different observers. Orientation discrimination thresholds for the dual task without reward (left point) were significantly higher than for any rewarded condition, fluctuating around  $9^\circ$ . In the presence of reward, average thresholds decreased substantially, spanning from about  $6^\circ$  to about  $3^\circ$  according to whether the credit was acknowledged 10% or 90% of the correct responses (Fig.2, middle and left points). The fact that even a small reward probability (Fig. 2, middle point) introduced a large change of performance may simply be interpreted as a general arousal effect.

The results obtained in the presence of reward are more specific to reveal the mechanisms of reward and its independence on attention, arousal and learning. Orientation discrimination thresholds were about 50% higher when the reward probability was equal to 0.1 (fig. 2, middle symbols) than when it was equal to 0.9 (fig. 2, right symbols). This difference was not affected by the attentional load to the central counting task, that remains stable across attentional conditions, suggesting that the difference obtained in the Light Load condition could not be attributed to residual attentional resources allocated peripherally in the HRP condition. It is noteworthy that this effect was obtained when the reference axis of the peripheral target was tilted in the same direction of the cue, as if the higher likelihood of achieving a reward improved the representation of the cued axes.



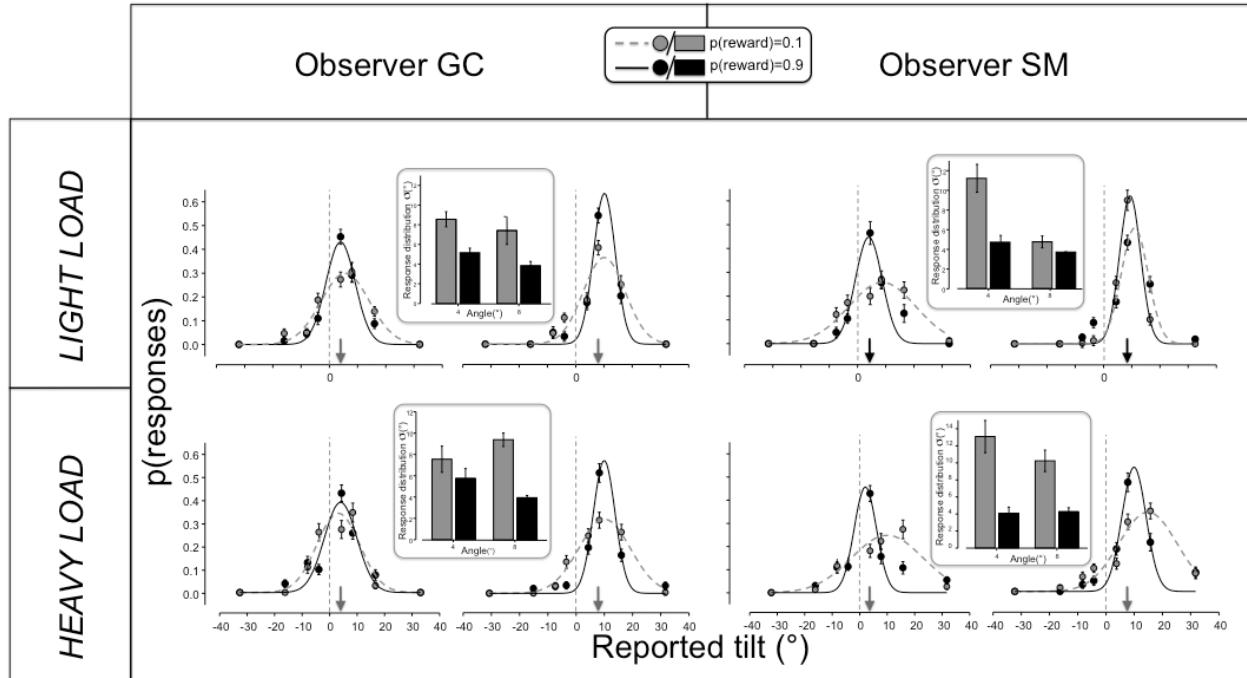
**Figure 2.** Average orientation discrimination thresholds (N=4), corresponding to the 75% correct point of the psychometric function, obtained scoring CW vs CCW responses as a function of the target tilt away from both the  $-45^\circ$  and the  $+45^\circ$  axis. The points represent the different reward conditions (from No Reward, left, to 0.9 probabilities of being rewarded, right) and different symbols represent different attentional conditions (Light Load, filled symbols, and Heavy Load, empty symbols). The straight horizontal line marks the average orientation discrimination threshold for the peripheral target alone, in the absence of attentional loading task and the two dotted lines above and below it represent the s.e.m.. Plotted data include only the analysis of trials in which the central task was successful (accurate counting). Moreover, the No Reward condition was measured in separated blocks, while the two rewarded patterns were interleaved within a block. The order of conditions (blocks) was shuffled throughout the experiment for all but the Heavy Load condition, executed later as a control experiment. Introducing a reward to correct orientation discrimination responses, though as unfrequent as to 1/10<sup>th</sup> of the correct responses (middle points), sets performance to a level comparable to when there was no central task, whereas highly frequent rewards show an additional advantage of the same magnitude (about a factor of 1.5). Importantly, the modulation of performance obtained by increasing the reward probability from 0.1 to 0.9 (middle and right points) is of the same amount across attentional condition, suggesting that the effect cannot be explained by the use of residual attentional resources allocated to the peripheral task.

We then inspected the response distributions to probe the nature of the mechanism triggered by the increasing probability of obtaining a reward. In particular, we compared the response distributions obtained by two observers (CG and SM) for the two target tilts closest to threshold, at  $4^\circ$  and  $8^\circ$ , as they are more informative for containing identification errors<sup>18</sup>. Each of the four panels of Fig. 3 reports two pairs of response distributions, for the LRP and the HRP condition, in gray and black respectively, and for the angle at  $4^\circ$  and  $8^\circ$  (pointed by the small gray arrows), to the left and to the right, respectively. The two observers are reported in the two columns, while the two

attentional loads, light and heavy, are reported in the two rows. The bar plots inside each panel plot the  $\sigma$  of the response distributions according to the same color code and spatial arrangement of the main graphs. The points in each graph show the proportion of responses to each response probe for any given physical tilt, with positive values reporting correct discrimination (i.e. CW for CW tilts and CCW for CCW tilts) and negative values indicating wrong discriminations (CW when CCW and viceversa). The smooth curves are Gaussian fits to the data-points, continuous black and dashed gray for the HRP and the LRP condition, respectively. The main result, clearly evident across observers and conditions, is that a higher likelihood of earning a bonus makes all the curves narrower and sharper, indicating a more reliable representation of the physical angle at the perceptual level. In the LRP condition the range of confusability over the orientation domain was substantially broader, as indicated by the differences in the  $\sigma$  of the Gaussian fits, that was significant for all but SM's Angle 8° condition under light attentional load to the central task. The suggestion that reward makes perception more veridical is confirmed by the position of the means (peaks,  $\mu$ ) of the response distributions. In the HRP condition, this parameter matches more closely the physical tilt of the stimulus in all cases, but more clearly in observer SM. The mispositioning of the distribution peaks to tilt values higher than the actual stimulus for the discrimination, is well known in literature as 'off-orientation looking'<sup>22, 23</sup>, i.e. the strategy of relying on orientation channel more tilted than the stimulus to optimize performance in orientation discrimination tasks.

Importantly, this effect takes place with a comparable strength in both the Light Load and the Heavy Load condition, as confirmed by the bar plots embedded in Fig.3, confirming that we can reduce drastically the possibility that the peripheral task can depend on 'residual attentional resources' saved from the central task demand and allocated to the peripheral task. Indeed, wrong counting made the  $p(\text{reward})=0$ , and the counting performance was around 60% in all observers in the Heavy Load task; therefore, as confirmed by personal reports, they always had to put a great attentional effort to keep their counting performance as high as they could.

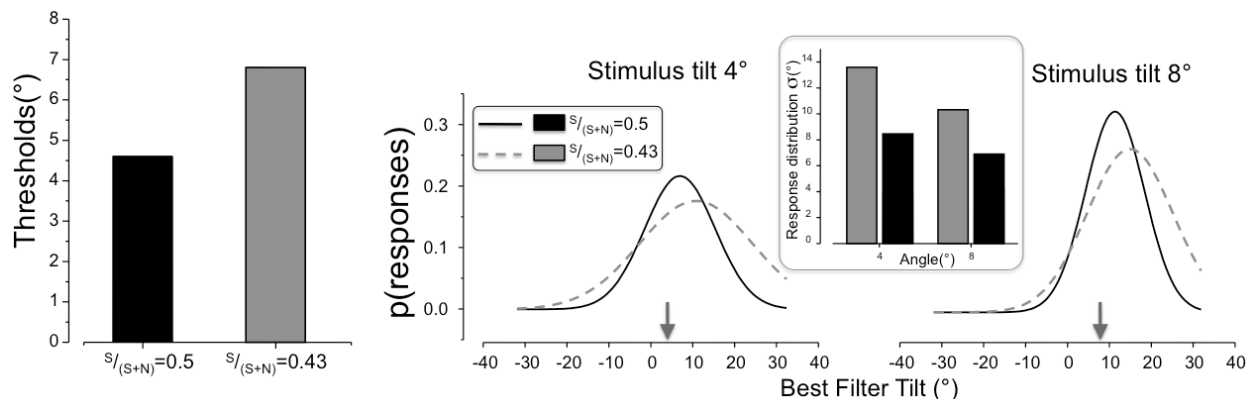




**Figure 3.** Response distributions obtained by two observers (CG, left, and SM, right) in the two attentional load conditions (Light Load, top, and Heavy Load, bottom). Each table cells reports two pairs of graphs, for the two physical angles around threshold ( $4^\circ$  and  $8^\circ$ , indicated by the gray arrows). Each pair is constituted by plots of the High Reward Probability condition (black symbols) and of the Low Reward Probability condition (gray symbols), fitted with gaussian pdfs (straight black and dashed gray, respectively). Specifically, the symbols represent the proportion of responses (clicks) to the different response probes for any given physical stimulus considered (collapsing CW and CCW tilts); the error bars of each symbol represent the s.e.m. of the estimate calculated by a bootstrap procedure<sup>24</sup>. Positive values report correct discriminations, negative values plot reported tilts for wrong discriminations. The framed bar plots show the  $\sigma$  of the gaussian fits with the confidence intervals of the estimates shown as error bars. The main effect, coherent with the threshold measurements shown in Fig. 2, is that the width of the distributions of the HRP condition is considerably narrower than the LRP condition in all conditions, as directly shown by the embedded bar plots. This implies a more precise representation of the target's orientation when the task was more likely rewarded. The second effect is that the peaks of the distributions are shifted toward tilt values larger than the physical angle, implying a general non-veridical representation of orientation (usually explained as off-orientation looking<sup>22,23</sup>); however, higher reward rates move the peaks toward a more veridical value close to the physical angle of the stimulus. The overall change in both the  $\mu$  and the  $\sigma$  of the response distributions indicates that reward improves greatly the aware representation of the orientation of a stimulus.

**Model.** In order to verify the possibility that the mechanism supporting the reward-based modulation of orientation discrimination were a reduction of SNR at an early level, we ran a MonteCarlo simulation using the same stimuli of our experiment that, at each trial, were convolved with a bank of noisy filters of optimal spatial frequency and phase. The filters' set was formed by selecting all the orientations that were used as stimuli and that could be selected in the response page. Each filter was perturbed by an independent source of noise that was recalculated at each iteration (trial) and whose amount was modulated in different runs. The sum of the squares of each pixel of the convolution

matrix was taken as a measure of response of each filter. The filter yielding maximum output simulated the magnitude matching probe selected at each iteration and was used to determine correct and wrong responses in the -simulated- orientation discrimination task. If, for example, in a given iteration a stimulus of  $4^\circ$  produced the maximum output in the  $-8^\circ$  filter, the latter angle was counted for the response distribution and the discrimination response was wrong. We ran 2000 trial for each of the 10 angles and used 4 Signal-to-Noise Ratios (SNR), from 0.5 to 0.35 (where lower numbers imply stronger noise). We reasoned that if our simple SNR hypothesis was correct, then we should be able to reproduce the results of our experiment, i.e the difference between the LRC and HRC condition could be reproduced by finding two appropriately different SNRs. Surprisingly, Figure 4 shows that this simple simulation reproduced very closely the entire pattern of results, both qualitatively and quantitatively. Thresholds increased from  $4.5^\circ$  to  $6.8^\circ$  when the SNR moved from 0.5 to 0.43. More importantly, the two noise levels reproduced very well the response distributions found empirically: decreasing SNR not only increased the  $\sigma$  of the distribution, but it also moved its peak in both the  $4^\circ$  and the  $8^\circ$  angle to tilt values larger than the stimulus tilt. Thus, the entire pattern of results of our experiment are well explained by the behavior of a simple model of orientation discrimination/identification whose decision rule is based on the maximum output of a bank of linear, noisy filters tuned to the possible signals.

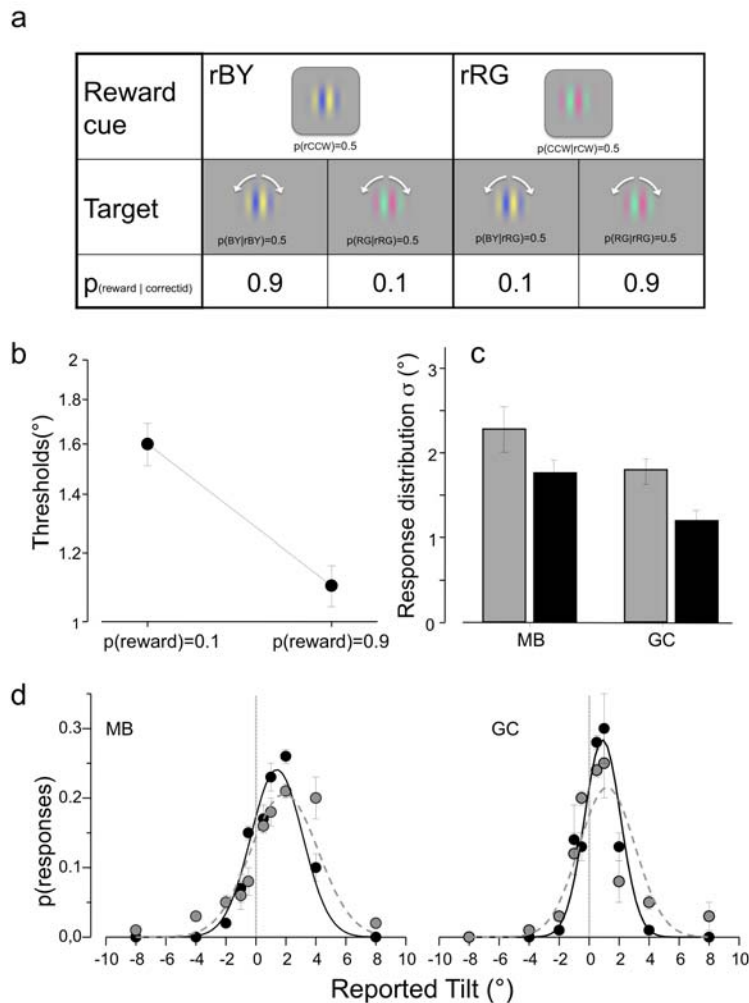


**Figure 4.** Simulated thresholds (left) and response distributions (right). The simulation compares at each trial the output of noisy filters having different tilts (in the range from  $-32^\circ$  to  $+32^\circ$  relative to a  $45^\circ$  axis) produced by convolving it with the stimuli used in the experiment and chooses the best filter, i.e. the one with the strongest response. Two SNRs are shown here (.5 and 0.43) whose values reproduce very well our data in the two reward probability conditions, The left panel shows the simulated thresholds, scaled by a factor of about 1.5 when the SNR was 0.43. This two SNR values generated the response distributions for the two stimulus angles  $4^\circ$  and  $8^\circ$  reported in the right panel. The simulation captures all the features of our data: increasing the SNR not only sharpens the distributions, decreasing their  $\sigma$  (shown by the embedded bar graph), but also it reduces the tilt-overestimation effect by moving the peaks ( $\mu$ ) toward the value of the physical angle.

**Feature-based attention.** In the main experiment we have modulated the attentional load by summoning the observers' attentional resources onto a central task at two levels of difficulty and have found that different attentional loads do not alter quantitatively nor qualitatively the results. This may imply that the modulatory channels of reward and attention act independently, even at early processing stages. However, there are evidence in literature that the mechanisms of spatial- and of feature-based attention rely on different mechanisms<sup>25</sup> and there is a possibility that they do not share the same resources. If this was the case, since the reward cue and the task rely on a unique feature (orientation), then our data could be alternatively explained as an effect of some sort of feature cueing depending on a priming effect of the reward cue to the subsequent orientation task, independently on the central counting task. In other word, the presence of a cue line tilted at  $+45^\circ$  or  $-45^\circ$  might have enhanced the representation of angles around that value at the expenses of the orthogonal tilts. We controlled this effect in two ways. In a first experiment we have cued both axis with a sort of X-like cue in which one oblique bar was black while the other was white and instructed the observer that an HRP trial was signaled by a match of the stimulus axis with the white line of the X cue, whereas an LRP trial was signaled by a match of the black line of the cue with the stimulus. The polarity of the two axis was randomly established and the results (not reported here) confirmed completely the trend obtained by visually showing only one of the cues.

However, to rule out more directly the possibility that feature-based attention played a role, we decided to carry a different experiment in which the feature of the cue and that of the task were independent. To this aim we signaled the reward probability through an association of *color* of cue and target by using oriented Gabor patches modulated around two independent color axes, Blue-Yellow (BY) or Red-Green (RG), while the task required an *orientation* judgment. The structure of the trial matched that of the main experiment and is summarized in Fig. 5a. The cue was either a BY or a RG patch, always vertical, and the peripheral target was tilted around vertical and could either match the color of the cue, yielding to a  $p(\text{reward})=0.9$  conditional to a correct identification, or be of the opposite color pattern, yielding to a  $p(\text{reward})=0.1$ . The response page contained 10 response probes, five for each side CW and CCW off-vertical, visualized in grayscale. The counting task did not vary from the first experiment. Figure 5b reports average thresholds of four observers (two of which new to the experiment) and shows clearly that even if the cue did not contain any information to bias the processing of orientation signals, a color coincidence between cue and target improved threshold by a factor of almost 1.5, consistently with the results of the main experiment. As expected, orientation sensitivity measured around the main axis was much finer than around the oblique axis and the response distributions were sharper.

Fig. 5c reports the  $\sigma$  of the response distributions, plotted in Fig. 5d, for two of the four observers (one new to the experiment). Again, distributions were sharper and the mean was more veridical when the chance of obtaining reward increased. We think this rules out convincingly the idea that the effects measured in this study were due to some sort of implicit priming provided by the tilted cue.



**Figure 5.** Feature based procedure (a) and results (b-d). (a) The reward cue was now a Gabor patch modulated either along the blue-yellow (rBY) or along the red-green (rRG) color axis. The task remained an orientation identification/discrimination task of either a color matching (HRP condition) or of a differently colored patch (LRP condition). For both colors, we asked to judge tilt offset from a unique, vertical reference axis. Orientation discrimination thresholds (b) of four observers follow closely the pattern shown by the main experiment and the model, decreasing by a factor of about 1.5 when the color of the cue and the target gratings matched. The response distributions (c and d) obtained in two

observers (MB and GC) by collapsing the two near-threshold tilts ( $1^\circ$  and  $2^\circ$ ) confirm the usual trend showing narrower  $\sigma$  and more veridical  $\mu$  in the HRP condition.

## DISCUSSION

In this study we provided converging measures to show that the precision of orientation judgment is modulated by the probability that a positive answer leads to a reward consisting in the offbeat and cost-efficient form of Scratch-and-Win lottery tickets. This occurs independently on whether or not attention is engaged elsewhere in space and takes place also when the reward cue provides no information for the response whatsoever. Perceptual learning or associations that extend over the span of a single trial cannot explain our results as the same stimuli and the same responses could be associated unpredictably to high or low reward probability. Rather, we have modulated the 'motivational' state<sup>26</sup> on a trial-to-trial base and found quick modulations of the perceptual representation of features encoded at an early stage, such as orientation. A reduced arousal could explain the difference between the unrewarded and the two reward conditions, but not the key difference between Low and High Reward Probability. In fact, the two reward schemes were interleaved within a trial and observers needed to keep their alertness high at least until the peripheral target appeared, as it implicitly signaled the level of reward probability of each given trial, and the data of the LRP condition confirm this by showing performances comparable to those obtained in the absence of central attentional task (horizontal line of Fig. 2). Moreover, the counting performance was unaffected by the reward probability (i.e. by the match between cue and target) and depended only on the central disk contrast, suggesting that the alertness was constant across conditions throughout the extent of a trial. Importantly, the data fit with a model that bases its behavior on the modulation of SNRs of early linear, noisy filters whose individual output is compared with a max rule to make a decision at each trial.

In summary, these data suggest that the probability of reward affects the SNR of individual orientation units at the early stages of the visual system independently on attention to the rewarded task and stimuli. The primary visual cortex (V1) is a good candidate for such an effect, as most of its cells have orientation-tuned receptive fields<sup>27, 28</sup> it has been evoked to account for psychophysical orientation discrimination<sup>19, 29</sup>, and it has been recently found to be modulated by the reward rate in animals<sup>13</sup> as well as by the reward history in human observers<sup>14, 15</sup>. This is coherent with recent accounts of perceptual learning in psychophysical hyperacuity tasks, which is explained by the action of feedback mechanisms acting onto the receptive fields properties of V1 neurons<sup>30</sup>. We have not studied the interocular transfer (as the study by Fahle<sup>30</sup> did),

but we have dealt with orientation channels consistently with the properties of orientation tuning of the primary visual cortex.

Platt and Glimcher<sup>31</sup> have observed LIP neurons, which projects direct feedback to V1<sup>32</sup>, with levels of activity positively correlated with the reward value of different stimuli independently on the motor factors. The reward value biases also caudate neurons speeding up saccadic latencies<sup>33</sup>. Similar structures may well be involved in our results. However, while these experiments set a constant (or encompassing many trials before switching) association between each stimulus and the amount of reward associated, our experiment overtook this by showing reward effects based on a trial-to-trial, unpredictable coincidence between a cue and the target stimulus. The present findings are novel and we think that they open many questions on the physiological mechanisms and anatomical circuitries of reward, that since very recently were assumed not to involve primary sensory areas at all<sup>34</sup>. The peculiar feature of our task of relying on trial-wide effects makes it different from recent studies showing reward-based modulation in V1<sup>13-15</sup>, as in those cases the modulation depends on the reward history associated to each stimulus while in our case integrating past trials does not provide any additional cue to succeed in the task and earn reward.

The direct involvement of early sensory stages within the network of reward-related neuromodulatory activities, and in particular the involvement of dopaminergic activity in our results may fit with the presence of D1 receptors in the striate cortex<sup>35</sup>. Fast, phasic response of dopamine neurons has been found with reward probabilities lower than one but not when the reward was always acknowledged<sup>34,36</sup>. Further research using similar behavioral paradigms in animals may shed light on this question.

These results provide insights on the basic computations performed by the elementary visual channels involved in the tasks. A point to solve would be that of discerning whether the SNR modulation is due to some form of gain control<sup>37</sup> or to a mechanism of noise reduction<sup>38</sup>. The simulation performed cannot operate such a distinction as what we actually change is the balance of the signal vs that of the noise. On the other hand, even though the net behavioral effect of reward consists in making the representation of the target tilt sharper (i.e. less noisy), our experiment was not aimed to dissociate the separate contribution of reward to signal and noise. We are currently running new experiments in order to observe how reward affects the perceived contrast<sup>37</sup> and how external noise impacts performance across reward and attentional conditions, in order to address directly this issue.

Another interesting result of our study lies in the reduction of the 'off-orientation looking' effect of orientation discrimination<sup>22, 23</sup> with high reward rates. It seems that the reward based modulation makes orientation discrimination more efficient by allowing the use of matched filters (i.e., better tuned to the physical signal) that in neutral conditions would

be performing less efficiently because of a negative trade off between signal and noise associated to this specific task. In other words, off-orientation looking would occur when lower SNRs would cause the internal response distributions to ‘invade’ the negative side -corresponding to wrong discrimination- too often, then the system mediates by using a channel that is less optimal but more certain about the tilt side. When similar top-down modulations intervene by reducing the spread of the responses to the given signal (that is increasing the SNR), then the system recognizes the improvement and selects the best matching filter for the orientation discrimination task.

It has been argued that most of the findings on the perceptual and decisional modulations by reward are contaminated by some form of visual attention, and that reward and attention cannot be easily disentangled empirically<sup>16</sup>. We think that as long as attention is operationally defined as the limited amount of resources available to process task-relevant information, being thus withdrawn by more primary tasks (such as our counting task), our study can provide a useful methodological template for both single neuron and brain imaging studies aimed to disentangling the two behavioral factors.

What may be the general implications of our findings? At a more general level we found that when one’s performance is rewarded, this not only will affect the output of goal-directed behavior, as it is intuitive to happen, but it will also improves the quality of the input on which motor responses are based. In other words, to make an example, the archer will succeed not only for a cleaner adjustment of his arch, but also because the target is better seen. This in turn has several implications in several professional fields in which sensory-based performance is fundamental but motivation may be variable, in training and education, and in more recreational fields such as sport.

## **ACKNOWLEDGMENTS**

We thank Nicoletta Berardi, Preeti Verghese and David Burr for helpful comments on the manuscript. This research has been supported by the European Research Council Advanced Grant 229445 “STANIB” and by the Italian MIUR “PRIN 2007”.

## **AUTHOR CONTRIBUTIONS**

S.B. conceived the study, designed the experiment and the model, analyzed the data and wrote the paper. C.S. designed the experiment, collected and analyzed the data.

## METHODS

**Observers.** A total of six observers participated in this research. Two of them to the main experiment alone, two to the feature-based attention experiment and two to both experiments. They were undergraduates of the Faculty of Psychology of the University of Florence, all naïve to the purpose of the study. They were also non-gamblers based on the criterion according to which subjects involved in gambling activities (including purchase of lottery tickets) more than once a month were excluded. They all had normal or correct-to-normal vision. Three of the subjects completed 600 trials for condition to measure thresholds, three completed 2000 to 2400 trials per condition to achieve a reliable sample size to measure both thresholds and response distributions.

**Apparatus and stimuli.** Stimuli were created on a G4 Power Macintosh using the Psychophysics Toolbox v. 2.55<sup>39, 40</sup> and displayed on a 17" gamma-corrected CRT monitor (Mitsubishi Diamond Pro) with average luminance equal to 29 cd/m<sup>2</sup>. Stimuli of the central task were disks with a diameter of 0.5° of visual angle flashed foveally for 100 ms. In the two attentional conditions the contrast was varied from a level of 80% (Light Load condition), at which the stimulus was well visible, to a level ranging from 4% to 8% (Heavy Load condition) in different subjects based on preliminary contrast detection thresholds measures. The peripheral task's stimuli were Gabor patches (2 cpd sinusoidal gratings vignettted by a 2d gaussian modulation of contrast with a space constant of 0.5°) displayed at a contrast of 80% at an eccentricity of 7° to the left or to the right of fixation. The peripheral stimulus' onset and offset was synchronous with the corresponding disk.

**Design and procedure.** Each trial began with a display showing a foveal line subtending 3° of visual angle, visualized for 500 ms and tilted 45° ClockWise (CW) or CounterClockWise (CCW) from vertical. In the Feature-based attention experiment the line was replaced by a Gabor patch equal to the target but modulated along the Red-Green or the Blue-Yellow axes, randomly. Immediately after the central attention loading task started. It consisted in the counting of a sequence (3 to 14 randomly, 4 in Fig. 1) of foveal disks displayed for 100 ms with a random inter-disk interval (in the range of 0.4 to 4 s.) to avoid predictability of upcoming counting stimuli. During the display of one of the disks, on a random base from the 2<sup>nd</sup> to the  $n^{th-1}$ , the peripheral oriented target was shown. This was the reward-effective stimulus, a Gabor patch tilted CW or CCW relative to one of the two oblique axis (e.g., in a trial with a 4° tilt, CCW was either a -49° or a +41° angle, while CW was either a -41° or a +49° angle) of varying angle ( $\pm 2^\circ$  to  $\pm 32^\circ$  in the main experiment and  $\pm 0.5^\circ$  to  $\pm 16^\circ$  in the Feature Based control experiment). 500 milliseconds after the disappearance of the last central disk the first response page was displayed; it contained all the digits of the range of possible disk numbers and subjects were asked to click with a mouse on the digit corresponding to the counted number of disks and to do within the limits of a brighter square containing the number. Wrong counting voided the trial; for any voided trial a new trial was inserted at the end of the block. The second response page allowed the orientation identification and discrimination response. It contained 20 Gabor probes in two lines, one for each possible tilt around both the +45° (upper line) and the -45° axis (lower line). The five probes to the left, in each line, corresponded to CW tilts while the five to the right corresponded to CCW tilts. Observers were asked to click on the response probe that best matched the peripheral target and primed about the fact that their response was scored as correct or wrong based on the match of the CW vs CCW tilt between stimulus and response and that clicking on the wrong probes (wrong identification) but in the right side (correct discrimination) yielded reward anyway. The final page of the trial contained a white and a black bar providing feedback about whether or not that given trial led to reward, based on a visually salient lengthening of the white or the black bar, respectively, and about the amount of rewarded identification needed to achieve another Scratch & Win ticket. The white bar was completed, and a ticket donated, after any 20 correct identifications.

**Reward pattern.** In the Rewarded conditions the probability of achieving credit for the Scratch-and-Win ticket was equal to 0.9 (HRP condition) if the main axes of cue and stimulus (in the main experiment) or if their colour (in the feature-based control) coincided. In the opposite case reward was granted with a probability of 0.1 (LRP condition). These reward probabilities were conditional to correct orientation



discriminations; wrong discriminations gave no reward. HRP and LRP trials were fully randomized, thus observers could not predict the reward probability until the peripheral target was shown.

**Data analysis.** Data were analyzed separately for orientation discrimination and identification. *Orientation discrimination* data formed psychometric functions fitted by cumulative normal *cdf*. Each function was bootstrapped<sup>24</sup> and refitted 100 times and the threshold was calculated (75% accuracy of the fitted function) for each bootstrap sample in order to have a reliable estimate of the threshold and its standard error. Orientation identification data fed response distributions (histograms representing the proportion of reported, or perceived tilt in the presence of a given physical tilt). We generated one distribution for each physical angle used in the experiment and bootstrapped it 100 times to estimate the reliability of individual points. Each bootstrap sample was fitted with a normal *pdf* in order to estimate the stability of the parameters measured ( $\mu$  and  $\sigma$ ).

## REFERENCES

1. Ghose, G. & Maunsell, J. Spatial summation can explain the attentional modulation of neuronal responses to multiple stimuli in area V4. *J Neurosci* **28**, 5115-5126 (2008).
2. McAdams, C. & Maunsell, J. Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *J Neurosci* **19**, 431-441 (1999).
3. Reynolds, J., Chelazzi, L. & Desimone, R. Competitive mechanisms subserve attention in macaque areas V2 and V4. *J Neurosci* **19**, 1736-1753 (1999).
4. Fang, F., Boyaci, H. & Kersten, D. Border ownership selectivity in human early visual cortex and its modulation by attention. *J Neurosci* **29**, 460-465 (2009).
5. Kamitani, Y. & Tong, F. Decoding seen and attended motion directions from activity in the human visual cortex. *Curr Biol* **16**, 1096-1102 (2006).
6. Watanabe, T., *et al.* Attention-regulated activity in human primary visual cortex. *J Neurophysiol* **79**, 2218-2221 (1998).
7. McAlonan, K., Cavanaugh, J. & Wurtz, R.H. Guarding the gateway to cortex with attention in visual thalamus. *Nature* **456**, 391-394 (2008).
8. McAlonan, K., Cavanaugh, J. & Wurtz, R.H. Attentional modulation of thalamic reticular neurons. *J Neurosci* **26**, 4444-4450 (2006).
9. Schultz, W., Tremblay, L. & Hollerman, J. Reward processing in primate orbitofrontal cortex and basal ganglia. *Cereb Cortex* **10**, 272-284 (2000).
10. Hampton, A. & O'doherty, J. Decoding the neural substrates of reward-related decision making with functional MRI. *Proc Natl Acad Sci U S A* **104**, 1377-1382 (2007).
11. Glimcher, P. & Rustichini, A. Neuroeconomics: the consilience of brain and decision. *Science* **306**, 447-452 (2004).
12. Behrens, T., Woolrich, M., Walton, M. & Rushworth, M. Learning the value of information in an uncertain world. *Nat Neurosci* **10**, 1214-1221 (2007).
13. Shuler, M. & Bear, M. Reward timing in the primary visual cortex. *Science* **311**, 1606-1609 (2006).
14. Serences, J.T. Value-based modulations in human visual cortex. *Neuron* **60**, 1169-1181 (2008).
15. Seitz, A., Kim, D. & Watanabe, T. Rewards evoke learning of unconsciously processed visual stimuli in adult humans. *Neuron* **61**, 700-707 (2009).
16. Maunsell, J. Neuronal representations of cognitive state: reward or attention? *Trends Cogn Sci* **8**, 261-265 (2004).
17. Della Libera, C. & Chelazzi, L. Visual selective attention and the effects of monetary rewards. *Psychol Sci* **17**, 222-227 (2006).
18. Baldassi, S., Megna, N. & Burr, D. Visual clutter causes high-magnitude errors. *PLoS Biol* **4**, e56 (2006).
19. Regan, D. & Beverley, K. Postadaptation orientation discrimination. *J Opt Soc Am A* **2**, 147-155 (1985).

20. Bradley, A., Skottun, B., Ohzawa, I., Sclar, G. & Freeman, R. Visual orientation and spatial frequency discrimination: a comparison of single neurons and behavior. *J Neurophysiol* **57**, 755-772 (1987).
21. Alais, D., Morrone, C. & Burr, D. Separate attentional resources for vision and audition. *Proc Biol Sci* **273**, 1339-1345 (2006).
22. Mareschal, I., Dakin, S. & Bex, P. Dynamic properties of orientation discrimination assessed by using classification images. *Proc Natl Acad Sci U S A* **103**, 5131-5136 (2006).
23. Solomon, J.A. Noise reveals visual mechanisms of detection and discrimination. *J Vis* **2**, 105-120 (2002).
24. Efron, B. & Tibshirani, R.J. *An Introduction to the Bootstrap (Monographs on Statistics and Applied Probability)* (Chapman & Hall/CRC, 1994).
25. Baldassi, S. & Verghese, P. Attention to locations and features: different top-down modulation of detector weights. *J Vis* **5**, 556-570 (2005).
26. Kawagoe, R., Takikawa, Y. & Hikosaka, O. Expectation of reward modulates cognitive signals in the basal ganglia. *Nat Neurosci* **1**, 411-416 (1998).
27. Hubel, D.H. & Wiesel, T.N. Receptive fields and functional architecture of monkey striate cortex. *J Physiol* **195**, 215-243 (1968).
28. Hubel, D.H., Wiesel, T.N. & Stryker, M.P. Anatomical demonstration of orientation columns in macaque monkey. *J Comp Neurol* **177**, 361-380 (1978).
29. Ringach, D.L. Tuning of orientation detectors in human vision. *Vision Res* **38**, 963-972 (1998).
30. Fahle, M. Perceptual learning: A case for early selection. *Journal of Vision* **4**, 879-890 (2004).
31. Platt, M.L. & Glimcher, P.W. Neural correlates of decision variables in parietal cortex. *Nature* **400**, 233-238 (1999).
32. Barone, P., Batardiere, A., Knoblauch, K. & Kennedy, H. Laminar distribution of neurons in extrastriate areas projecting to visual areas V1 and V4 correlates with the hierarchical rank and indicates the operation of a distance rule. *J Neurosci* **20**, 3263-3281 (2000).
33. Lauwereyns, J., Watanabe, K., Coe, B. & Hikosaka, O. A neural correlate of response bias in monkey caudate nucleus. *Nature* **418**, 413-417 (2002).
34. Schultz, W. Multiple reward signals in the brain. *Nature reviews Neuroscience* **1**, 199-207 (2000).
35. Eickhoff SB, Rottschy C & K., Z. Laminar distribution and co-distribution of neurotransmitter receptors in early human visual cortex. *Brain Struct Funct.* (2007).
36. Mirenowicz, J. & Schultz, W. Importance of unpredictability for reward responses in primate dopamine neurons. *J Neurophysiol* **72**, 1024-1027 (1994).
37. Carrasco, M., Ling, S. & Read, S. Attention alters appearance. *Nat Neurosci* **7**, 308-313 (2004).
38. Lu, Z., Lesmes, L. & Doshier, B. Spatial attention excludes external noise at the target location. *Journal of vision* **2**, 312-323 (2002).
39. Brainard, D. The Psychophysics Toolbox. *Spat Vis* **10**, 433-436 (1997).
40. Pelli, D. The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis* **10**, 437-442 (1997).

