Pairing stimulation of the VTA with an unattended visual category:

- Improves perception in a category-specific way
- Changes representations in early and late visual cortex, and higher-order cortical and subcortical regions

Electrical stimulation of the macaque ventral tegmental area drives category-selective learning without attention.

Sjoerd R Murris^{1,2,5}, John T Arsenault^{1,2,5}, Rajani Raman^{1,2}, Rufin Vogels^{1,2}, Wim Vanduffel^{1,2,3,4,*}

¹Laboratory for Neuro- and Psychophysiology, Department of Neurosciences, KU Leuven Medical School, Leuven 3000, Belgium

²Leuven Brain Institute, KU Leuven, Leuven 3000, Belgium

³Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA 02129, USA

⁴Department of Radiology, Harvard Medical School, Boston, MA 02144, USA

⁵These authors contributed equally

* Address correspondence to: Wim Vanduffel, KU Leuven, Laboratory for Neuro- and Psychophysiology, Campus Gasthuisberg, O&N 2, Herestraat 49-bus 1021, 3000, Leuven, Belgium. Telephone: +32 16 33 06 66.

Lead contact: Wim Vanduffel. Email: wvanduffel@mgh.harvard.edu.

SUMMARY

Perception improves by repeated practice with visual stimuli, a phenomenon known as visual perceptual learning (VPL). The interplay of attentional and neuromodulatory reward signals is hypothesized to cause these behavioral and associated neuronal changes, although VPL can occur without attention (i.e. task-irrelevant VPL). In addition, task-relevant VPL can be category-selective for

simple attended oriented stimuli. Yet, it is unclear whether category-selective task-irrelevant VPL occurs, nor which brain centers mediate underlying forms of adult cortical plasticity. Here, we show that pairing subliminal complex visual stimuli (faces and bodies) with electrical microstimulation of the ventral tegmental area (VTA-EM) causes category-selective task-irrelevant VPL. These perceptual improvements are accompanied by fMRI signal changes in early and late visual and frontal areas, as well as the cerebellum, hippocampus, claustrum and putamen. In conclusion, Pavlovian pairing of unattended complex stimuli with VTA-EM causes category-selective learning accompanied by changes of cortical and subcortical neural representations in macaques.

INTRODUCTION

To survive in a challenging and dynamic environment, we benefit from our ability to uncover visual patterns that signal prospective rewards. It has been shown that repeated practice with stimuli can improve their detection and discrimination, a phenomenon referred to as visual perceptual learning (VPL) (Doshier and Lu, 2017; Maniglia and Seitz, 2018; Sagi, 2011; Seitz, 2017; Vogels, 2010). By improving an individual's ability to perceive reward-predicting cues, VPL offers a process through which visual information can be better utilized to increase future rewards. Theories postulate that VPL is driven by the interaction of selective attention and neuromodulatory reward signals (Roelfsema et al., 2010; Watanabe and Sasaki, 2015). Top-down attention is thought to regulate which features are learned, by enhancing or suppressing activity in visual cortex for relevant and irrelevant features, respectively (Ahissar and Hochstein, 1993; Gutnisky et al., 2009). In contrast, rewards are assumed to trigger neuromodulatory signals capable of regulating neuroplasticity (Bromberg-Martin et al., 2010; Kim et al., 2015). Despite the apparent synergistic roles of reward and attention signals during VPL, task-irrelevant VPL (TI-VPL) studies indicate that neuromodulatory signals driven by reward may be sufficient to drive learning and associated plasticity (Seitz et al., 2009; Shibata et al., 2011a; Shibata et al., 2009).

49 TI-VPL studies showed that subjects can improve their perception of features, without conscious
50 knowledge of these features (Kim et al., 2015; Seitz and Watanabe, 2009). In addition to providing a
51 mechanism for the implicit learning of rewarded cues, TI-VPL offers a powerful approach to study the
52 influence of neuromodulatory signals in VPL. This is because TI-VPL is most evident when, during cue-
53 reward pairing, attention is not oriented toward learned stimuli (Seitz and Watanabe, 2009; Shibata et
54 al., 2011b). TI-VPL experiments examining the role of neuromodulators should explicitly test the
55 potentially correlated effect of attention.

56

57 Rodent studies have shown that pairing stimulation of dopaminergic and cholinergic centers with
58 auditory stimuli can both enhance their neuronal representation (Bao et al., 2001; Kilgard and
59 Merzenich, 1998; Martins and Froemke, 2015) and improve perception (Carcea and Froemke, 2013;
60 Froemke et al., 2013). Nonetheless, due to large differences between rodent and primate visual
61 (Huberman and Niell, 2011) and reward systems (Berger et al., 1991) it is difficult to predict how these
62 changes are related to VPL of rewarded cues in primates. To bridge the gap between **(1)** human studies,
63 using complex psychophysical paradigms to dissect the contribution of reward from attention in VPL,
64 and **(2)** rodent studies elucidating the causal mechanisms behind sensory plasticity, we recently
65 examined the role of the primate ventral tegmental area (VTA), a dopaminergic center, in
66 neuroplasticity and TI-VPL (Arsenault and Vanduffel, 2019). We demonstrated that electrical
67 microstimulation of the VTA (VTA-EM) could drive neuroplasticity in extrastriate visual cortex, and
68 learning of task-irrelevant stimuli. Importantly, while this work demonstrated that the association of
69 an irrelevant stimulus and VTA-EM was sufficient to drive TI-VPL, these experiments did not explicitly
70 test whether attention was oriented toward the irrelevant stimuli. Therefore, it remains to be shown
71 whether activation of a primate neuromodulatory center can drive VPL in the absence of stimulus-
72 directed attention.

73

74 In addition, our earlier work examined TI-VPL using simple grating and motion stimuli (Arsenault and
75 Vanduffel, 2019), which resonates with most VPL studies using simple features. Nonetheless, learning
76 about more complex visual stimuli, such as faces or bodies, is crucial to identify potential rewards in
77 natural environments, because such specific categories often hold higher ecological value. A recent
78 study found evidence for category-selective VPL by demonstrating that VPL can transfer across simple
79 features within the same but not across categories (Tan et al., 2019). Despite this, it remains to be
80 shown whether the association of a cue and neuromodulatory activity is also sufficient to drive
81 category-selective learning. Moreover, complex visual categories, like faces and bodies, are
82 preferentially represented in more anterior patches along the temporal lobe than simpler stimuli
83 (Fisher and Freiwald, 2015; Popivanov et al., 2012; Tsao et al., 2008). Since anterior IT also receives
84 stronger dopaminergic innervation than more posterior visual regions (Berger et al., 1991; Berger et
85 al., 1988), it is conceivable that face and body representations may be more susceptible to dopamine-
86 driven VPL.

87

88 Based on these considerations, we designed a paradigm to test whether TI-VPL and associated
89 neuroplasticity can be driven by VTA-EM in the absence of stimulus-directed attention. The
90 experiments were designed to tightly control attention onto a relevant task for the monkeys (i.e. an
91 orthogonal color discrimination task) and thereby avoid attention onto the task-irrelevant cues during
92 stimulus-VTA-EM pairing. We used both a stimulus category that was paired with VTA-EM (faces or
93 bodies) and a control category, not paired with VTA-EM (again faces or bodies). Furthermore, the
94 capacity of these stimuli to drive temporal lobe activity allowed us to examine how category selective
95 temporal lobe representations were affected by pairing with VTA-EM.

96

97

RESULTS

We hypothesized that pairing images of faces and bodies (henceforth referred to as face and body cues) with VTA-EM, in the absence of directed attention, would generate category selective learning and neuroplasticity. To test this, we designed a three-part experiment consisting of: **(1) a pre-association phase:** in which behavior and/or physiology was initially measured, **(2) a cue-VTA-EM association phase:** in which face or body cues were paired with VTA-EM while attention was oriented onto an orthogonal color task at a different visual field location, and **(3) a post-association phase:** in which behavior and/or physiology was re-measured to identify VTA-EM induced changes.

Establishing parathreshold cue parameters

In the main experiment, we tested whether the monkeys' capacity to discriminate the profile orientation of either face or body cues in interleaved blocks of +/- 400 trials (**profile task, Figure 1B, Video S1**) improved after VTA-EM pairing. We used an orthogonal color discrimination task (**color task, Figure 1C, Video S2**) to orient attention away from the cues associated with VTA-EM. Prior to the main experiment, we determined signal-to-noise ratios (SNRs) for face and body cues to avoid attentional capture (Anderson et al., 2011) and to render the cues suitable for TI-VPL. The ideal stimuli for TI-VPL are strong enough to drive activity but sufficiently weak to avoid attentional suppression (Roelfsema et al., 2010), at so-called parathreshold intensities (Tsushima et al., 2008).

Consequently, we generated cues with SNRs that were **(1)** near discrimination threshold on a profile task using an **SNR test**, and **(2)** inconspicuous enough that their presence did not divert attention from an orthogonal color task using an **interference test (Figure 1A)**. Importantly, if attentional shifts were observed during the interference test, we repeated this test with lower SNRs to ensure that face and body cues did not draw attention away from the color task.

To determine the discrimination thresholds, we first ran the SNR test (**Figure 1A**), a variant of the profile task (**Figure 1B**). Monkeys were trained on the profile task to report the profile orientation (left- or rightward) of face and body cues (**Figures S1**, STAR Methods, **Video S1**). To reduce the salience of both cues, a dynamically refreshed, white noise background (60 Hz, luminance equated to face/body cues) was presented continuously throughout each trial. Therefore, the presentation of face and body profile cues can be thought of as a brief insertion of shape and texture information, the amount dependent on SNR, into a constant stream of dynamic visual noise.

We fit a cumulative Gauss sigmoid model to the animals' performance (**Figure 1D**) (Wichmann and Hill, 2001) to determine the SNR associated with threshold profile discrimination (70% correct). For M1, subsequent interference tests indicated that when face and body cues were displayed at this threshold SNR, color task performance was impaired (ANOVA, main effect of cue, $F(2, 246) = 19.351$, $p = 1.57 \times 10^{-8}$). Therefore, we changed our criterion for M1 to a lower, but still parathreshold, performance level (60%). Utilizing these performance levels (M1 - 60%, M2 - 70%), the sigmoids were used to determine the parathreshold SNR levels for each monkey (**Figure 1D**, M1: faces - 9.6%, bodies - 17.8%; M2: faces - 10.3%, bodies - 20.1%).

Attention was not oriented toward irrelevant cues

Next, we examined whether the cues diverted attention from the color task (**Figure 1C**), using a variant of the color task, deemed the interference test (**Figure 1A**, **Video S2**). The color target was displayed in the upper visual field (VF) while the task-irrelevant face or body cue was shown in the lower quadrant of the LVF or RVF at the parathreshold SNR as determined in the SNR test. Trials without task-irrelevant cues were included as a baseline measure of color discrimination performance. In addition, the same dynamic noise as in the profile task was presented throughout the color task to reduce the saliency of the cues.

Importantly, to exclude attentional confounds (Maunsell, 2004), we needed to confirm that the interference tests had the capacity to reveal possible attentional shifts from the color targets. We took the sensitivity to cue duration as a measure for attentional shifts from the color target to the irrelevant cue, at high and parathreshold SNR (**Video S2**).

When high SNR cues were presented, color discrimination performance was significantly reduced (**Table S1**, ANOVA, main effect of cue, $F(2, 359) = 1.37$, $p = 0.01$). Thus, the interference test has the capacity to detect attentional shifts from the color task. Using parathreshold SNR levels, we observed that cue duration significantly affected color discrimination performance (**Figure 1E**, **Table S1**, ANOVA, $F(4, 374) = 67.424$, $p = 6.36 \times 10^{-43}$). For durations surrounding the parathreshold cue duration, higher durations improved performance (Tukey's HSD; level 2 vs. level 0, $p < 10^{-7}$; level 2 vs. level 1, $p = 8.10 \times 10^{-7}$) while shorter durations reduced it (Tukey's HSD; level 2 vs. level 4, $p = 3.55 \times 10^{-3}$). This indicates that color task performance at parathreshold cue durations (level 2) also contained the dynamic range needed to observe potential attentional interference. Yet, the presence of parathreshold cues did not affect color discrimination performance (**Table S1**, ANOVA, main effect of cue, $F(2, 374) = 0.26$, $p = 0.77$). Thus, parathreshold face and body cues do not divert attention from the color task.

Pairing task-irrelevant face and body cues with VTA-EM

After the parathreshold face and body cues were established, we began five rounds of the main experiment. During each round, we measured behavior (profile task) and/or physiology (through fMRI) before and after pairing face or body cues with VTA-EM, to assess how cue-VTA-EM associations affect performance in the profile task and neural responses. Round one consisted of five phases that occurred in chronological order: **(1)** pre-association behavior, **(2)** pre-association fMRI, **(3)** cue-VTA-EM association, **(4)** post-association fMRI, and **(5)** post-association behavior (**Figure S2**, STAR Methods). The goal of this first round was to examine how both physiology (fMRI activity) and behavior (profile task) changed after cue-VTA-EM association. Rounds two through five of the main experiment focused

on the behavioral effects of cue-VTA-EM association (**Figure S3**, STAR Methods). These rounds therefore consisted of three phases: **(1)** pre-association behavior, **(2)** cue-VTA-EM association and **(3)** post-association behavior.

The cue-VTA-EM association paradigm was identical to the color task (**Figure 1C**) employed during interference tests, with the exception that only one cue duration was used (M1: 100ms, M2: 150 ms) and one of the categories (faces or bodies) in one of the VFs was paired with VTA-EM (**Figure 2A**). During pairing trials, we employed forward conditioning to simulate reward signals (Arsenault et al., 2014) and to generate strong associations between the paired category and VTA-EM (Bermudez and Schultz, 2014; Gallistel and Gibbon, 2000). Accordingly, VTA-EM (100 Hz, 200 ms duration) began immediately after the presentation of the paired cue (**Figure 2A**). During both fMRI and behavioral post-association phases, we intermixed cue-VTA-EM association and the post-association fMRI/behavior days to maximize the effects of pairing (**Figure 2B** and **S2-3**). Therefore, each post-association fMRI/behavioral day was preceded by a day of cue-VTA-EM association. Finally, an interference test was performed each Friday to confirm that color task performance remained sensitive to attentional shifts from the color target to the task-irrelevant cues.

Attention is not oriented toward cues before, during and after VTA-EM association sessions

We first used the interference test days to confirm that color task performance remained sensitive to cue duration throughout the main experiment (**Figure 2C**). We found a significant effect of cue duration (**Table S2**, ANOVA, $F(4, 1106) = 191.83$, $p = 6.29 \times 10^{-125}$) with monkeys performing worse at shorter durations (Tukey's HSD; level 2 vs. level 0, $p < 10^{-7}$; level 2 vs. level 1, $p = 10^{-7}$) and better with longer durations (Tukey's HSD; level 2 vs. level 3, $p < 10^{-7}$; level 2 vs. level 4, $p < 10^{-7}$). Yet, we did not find a general effect of task irrelevant cues (**Table S2**, ANOVA, $F(2, 1106) = 0.3176$, $p = 0.73$), nor a specific effect at parathreshold cue durations (**Table S2**, level 2, Friedman's test, $\chi^2 = 1.18$, $p = 0.55$), nor an effect of the VF in which the cues were presented relative to the placement of the electrodes (ANOVA,

F(2, 1106) = 0.094, $p = 0.76$). Importantly, we also examined color discrimination performance during the cue-VTA-EM association days (**Figure 2D**). Again, no significant difference in color performance was found (Friedman's test, $\chi^2 = 1.64$, $p = 0.80$), indicating that attentional allocation in the color task during VTA-EM trials did not differ from trials without VTA-EM. In addition, color discrimination performance did not change as a function of hemifield (RVF or LVF) in which the cues were presented (ANOVA: F(2,1187) = 0.78, $p = 0.51$), arguing against an attentional bias. Thus, the color task remained sensitive to possible shifts in attention and attention did not shift from the color target to task-irrelevant parathreshold cues.

Category selective effects of cue-VTA-EM association: behavior

Next, we tested whether VTA-EM-cue pairing generated category-selective changes in profile discrimination behavior. We first examined how profile discrimination sensitivity (d' , STAR Methods) changed (post vs. pre cue-VTA-EM association) as a function of SNR level (**Figure 3A**). At the highest SNR, we found an interaction of pairing and category (linear mixed effects (LME) model parameter estimates: SNR 5 = 0.41, bootstrapped 95% CI_{Bonf}: (0.25, 0.56), see **Table S3** and STAR methods), with the paired category exhibiting stronger learning after VTA-EM association than the control category. This effect was consistent, with both monkeys exhibiting stronger learning for the paired category in 4 out of 5 rounds (**Figure 3A**, right inset) and the individual monkeys both showing a significant interaction of pairing and category (LME: M1 = 0.52 (0.29, 0.74), M2 = 0.28 (0.07, 0.50), **Table S3**). We also found a general improvement in performance at SNR levels 2 - 4 (LME: SNR 2 = 0.21 (0.06, 0.35), SNR 3 = 0.33 (0.17, 0.47), SNR 4 = 0.14 (0.00, 0.28)) although these category-*nonspecific* improvements may result from mere task exposure. Because task exposure was equal for the control and the paired category, it cannot explain a category selective improvement at SNR 5 and suggests an additional, more interesting effect of VTA-EM-association. A detailed rendering of the d' data per round and per monkey further substantiates these findings (**Figure S4**).

Because both monkeys had VTA electrodes implanted in the right hemisphere (RH), we hypothesized that VTA-EM induced learning may be larger for stimuli presented in the LVF (**Figure 3B**). We found a trend for the interaction of pairing and VF at SNR level 3 (LME: SNR 3 = 0.13 (-0.04, 0.30)) suggesting that sensitivity improvements depended on the VF in which the cues were presented. Next, we ran separate LME models for each VF. At group level, sensitivity improvements for the paired category were observed at the highest SNR level for both VFs (LME: LVF – SNR 5 = 0.56 (0.33, 0.80), RVF – SNR 5 = 0.25 (0.04, 0.47)). Interestingly, this effect was only significant for individual animals in the LVF (M1 = 0.74 (0.40, 1.06), M2 = 0.36 (0.04, 0.68)). An additional isolated category-specific improvement was found at a lower SNR level, restricted to the RVF (LME: RVF – SNR 2 = 0.25 (0.01, 0.50), but this effect was driven by a single animal (M1 = -0.07 (-0.37, 0.24), M2 = 0.62 (0.19, 1.06); **Figure S4**). We also found a robust increase in sensitivity in the LVF over rounds, independent of the location of the cue-VTA-EM association (**Table S3**). Overall, robust sensitivity improvements were found for the paired category at the highest SNR level, indicating that the effect of pairing at parathreshold SNR levels generated learning at higher SNR levels.

Behavioral responses to stimuli associated with higher incentive salience have been shown to be faster (Watanabe et al., 2001a). Therefore, we examined how changes in reaction time (RT; post vs. pre) varied across categories and SNR levels (**Figure 3C**). At group level, a significant interaction of pairing and category was found at SNR levels 1 and 2 (LME: SNR 1 = -5.06 (-8.77, -1.33), SNR 2 = -6.92 (-10.5, -3.38)) with VTA-EM association reducing RT more strongly for the paired category. At SNR level 2, each animal displayed a significant interaction (LME: M1 = -5.43 (-7.54, -3.30), M2 = -8.62 (-16.1, -1.47)). Therefore, while only a small trend for category-selective improvements in sensitivity was observed at SNR levels 1 and 2 (**Figure 3A**), when combined with consistent reductions in RT observed at the same SNR levels, these behavioral changes result in more efficient task performance at low SNR levels.

Finally, we examined whether pairing of VTA-EM with a spatial location (LVF or RVF) biased the animals' responses during the profile task. Although the saccadic responses were never directly paired

with VTA-EM, changes in incentive salience of the paired spatial location may be observed during subsequent responses during the profile task. To test this, we compared saccade bias during the profile task before and after VTA-EM association (c-criterion, **Figure 3D**, STAR Methods). At group level, we found a significant effect of pairing on the c-criterion at SNR levels 0-2 (LME parameter estimates: SNR 0 = -0.20 (-0.30, -0.10), SNR 1 = -0.16 (-0.25, -0.06), SNR 2 = -0.09 (-0.17, 0.00), see **Table S3**) with an increased saccade bias toward the VF associated with VTA-EM. This effect was consistent for each monkey, with SNR level 0 showing the strongest effect (M1 = -0.14 (-0.29, 0.02), M2 = -0.27 (-0.14, -0.39), see **Table S3**), indicating that when information was absent, the bias to make a saccade toward the paired VF was strongest.

Category selective effects of cue-VTA-EM association: physiology

To examine how neural representations of face and body cues changed after pairing, we acquired fMRI responses to these cues before and after the cue-VTA-EM association phase of round 1 (**Figure S2**). To avoid the influence of attention on face and body cues, monkeys performed the orthogonal color discrimination task while these fMRI responses were measured. Therefore, the design of fMRI days was identical to the cue-VTA-EM association days with the exception that VTA-EM was not delivered, and fMRI activity was acquired (STAR Methods).

We fit a 2nd-level linear model (LM) to the per-run beta weights (STAR Methods) of the categorical fMRI response (paired vs. control category) and tested the effect of pairing (**Figure 4A**). Significantly enhanced responses to the paired category (LM, $p < 0.005$, cluster size = 10 voxels) were found, after pairing, in visual regions (early visual cortex and upper bank of the superior temporal sulcus (STS)) and subcortical regions involved in associative learning (putamen and hippocampus) within the hemisphere ipsilateral to the electrodes. Contralaterally, increased responses to the paired category were also found in the cerebellum and prefrontal cortex while decreased responses were found in the orbito-frontal cortex, anterior cingulate and claustrum. To examine the consistency of these effects

across animals, we fitted separate LMs for each animal and performed a conjunction analysis across animals. To avoid false positives, we only presented activation clusters in the group analysis that were also significantly modulated in the conjunction analysis (**Figure S6A,C-D**, LM conjunction, $p < 0.05$, cluster size 10, STAR Methods).

Effects of VTA-EM pairing on fMRI response within the temporal lobe

Based on the capacity of face and body cues to drive activity along the ventral stream and the higher dopaminergic innervation of anterior temporal regions (Berger et al., 1991; Berger et al., 1988), we hypothesized that cue-VTA-EM association may generate plasticity in more anterior IT regions than we previously observed with simpler stimuli (Arsenault and Vanduffel, 2019). Within the RH (ipsilateral to the electrodes), we found an increased response to the paired category (LM, effect of pairing, Bonferroni corr., $p < 0.005$) within posterior (mm = -17, -16, -13, -12, -7) and anterior (mm = +26, +27) visual cortical positions while the categorical response at intermediate positions (mm = -3, -2) was suppressed (**Figure 4B**).

In the left hemisphere (LH), we found a similar pattern with increased responses to the paired category in anterior (+23, +24) positions while the categorical response was suppressed at intermediate (-4, +1, +6) positions. While posterior enhancements of the categorical response were restricted to the RH, anterior modulations were bilateral. The hemisphere-specific enhancement in early visual cortex, along with the stronger behavioral improvements for cues presented in the LVF (projecting to the RH) indicates that for some features the effect of cue-VTA-EM pairing was more pronounced in brain regions ipsilateral to the stimulated VTA.

We hypothesized that visual regions that are preferentially responsive to the paired category will exhibit increased responses to the paired category (e.g. face stimuli in face-selective patches). To test this, we identified patches that preferentially responded to faces, bodies and non-face/non-body objects using an independent localizer experiment (**Figure S6B**, STAR methods, (Popivanov et al.,

2012)). However, a ROI analysis did not reveal significant modulations of the categorical response in any of the face, body and object patches of either animal (LME: **Table S4-5**).

To directly test the hypothesis that regions representing the paired category experienced stronger enhancements in the paired category response, we combined data across the network of patches (referred to as functional networks) that responded to the same category (face, body and object selective cortical regions; **Figure 5A**). In addition, we combined data from specific functional networks across animals based on whether patches were responsive to the paired (yellow lines, M1 = bodies, M2 = faces), control (M1 = faces, M2 = bodies) or object categories. We then fit a LM and examined whether changes to the categorical response after pairing with VTA-EM differed across functional networks. No significant interaction of pairing and functional networks was found (LME: **Table S5**, STAR Methods).

Next, we tested whether the change in general visual responses (cues vs. fixation) differed among the face, body, and object selective networks (**Figure 5A**). Surprisingly, we found that visual responses were enhanced in both the paired and the object category patch networks (**Figure 5B**) with a larger increase in visual responses after pairing within the paired category (LME: 0.43 (0.06, 0.80), see **Table S5**) and the object category (LME: 0.52, (0.14, 0.90)) networks relative to the control category network. This indicates that there was a general enhancement in visual activity after pairing that was selective for the paired and object selective IT patches (**Table S4**: visual non-specific model). RH patches tended to be more affected than LH patches. Based on these fMRI findings in conjunction with the behavioral results indicating that VTA-EM induced changes were more pronounced ipsilateral to the stimulating electrode, we examined whether changes in visual responses differed between hemispheres (**Figure 5C**). We found a significant interaction of pairing and hemisphere (LME: 0.65 (0.36, 0.95), **Table S5**) indicating that fMRI responses in category selective patches of the hemisphere with the VTA electrodes increased more after pairing. A further dissection of the fMRI results shows that this hemispheric bias was present also in each animal, yet the individual contribution of each of the patch networks was

more variable (**Table S5**). Thus, VTA-EM-induced plasticity is most pronounced ipsilateral to the site of VTA-EM.

DISCUSSION

We demonstrated that pairing images of faces and bodies with VTA-EM drives category-selective learning and neuroplasticity in primates. Moreover, during cue-VTA-EM association, attention was tightly controlled and monitored to ensure it was oriented onto the orthogonal color task instead of the cues. Therefore, these results indicate that pairing the activation of a dopaminergic neuromodulatory center with visual stimuli is sufficient for visual perceptual learning and associated plasticity in the absence of stimulus directed attention.

During the cue-VTA-EM association phase, the face and body cues were presented at parathreshold intensities to reduce their salience and to avoid drawing attention away from the orthogonal color task. When examining how cue-VTA-EM association affected the profile task, however, we tested performance over a range of SNR levels ranging from pure noise (0% SNR) to high SNR levels (faces 50%, bodies 60%). In contrast to earlier work demonstrating TI-VPL at parathreshold SNR levels (Seitz and Watanabe, 2003; Watanabe et al., 2001b), category-selective improvements in sensitivity were observed at the highest SNR levels (**Figure 3A**). Nonetheless, these sensitivity improvements were consistent across rounds and animals.

As monkeys are performing the discrimination task over extended periods of time, one needs to control for mere task and/or stimulus exposure. The difference in sensitivity (d') between the paired and control condition (**Figure 3**) actually reflects the relative change in perceptual improvement between stimulated and control category, which is most informative. One would have expected that

if the perceptual changes are invariant to category (e.g., reflecting some general increase in the tolerance to noise), then VTA-EM induced changes in d' should be equal for paired and control categories. This was not the case for the highest SNR (**Figure 3A**), indicative for a category-selective effect of VTA-EM.

While sensitivity is an important component of performance, overall performance can be better viewed as a trade-off between speed and sensitivity. The reduction in RT, while sensitivity remained stable (at low SNR levels), indicates that overall performance improved after VTA-EM association at parathreshold SNRs. These performance improvements at parathreshold SNR levels are in agreement with increases in sensitivity for corresponding stimulus intensities in human studies (Seitz and Watanabe, 2003; Watanabe et al., 2001b). It is unclear why increases in sensitivity are expressed at higher SNR levels, while changes in RT are more evident at lower SNR. Possibly, parallel mechanisms are at work advancing sensitivity or speed which are competing at different SNR levels to optimize task efficiency. In addition to RT changes, we also observed a significant response bias for cues presented at parathreshold SNR levels and at noise level (0% SNR). This response bias was typified by an increased preference to make saccades toward the visual field that had been associated with VTA-EM. While stronger biases are often observed when sensory evidence is weak (Fechner et al., 1966; Sridharan et al., 2014), note that we observed a change in bias resulting from either unilateral stimulation, or the association of VTA-EM with a cue at a particular location. This finding suggests that preference for reward associated locations can be acquired even while attention is oriented toward other positions and onto other tasks.

We also measured how fMRI responses to parathreshold SNR cues changed after VTA-EM association. Interestingly, most of the areas, showing either VTA-EM induced increased (putamen, dIPFC) or

decreased (OFC, ACC) fMRI responses to the paired category, are regions typically implicated in processing value-related signals (Bartra et al., 2013). In addition, increased fMRI responses to the paired category were also found in the hippocampus. Intriguingly, both human and non-human primate studies of perceptual learning have shown that hippocampal training effects are correlated with behavioral improvements suggesting the hippocampus may play a direct role in VPL (Baraduc et al., 2019; Guggenmos et al., 2015). We also measured an increased response to the paired category in the cerebellum contralateral to the stimulating electrode, approximately in the left crus. While the cerebellum is traditionally associated with motor control and motor learning, it also plays a role in perception (Baumann et al., 2015; Nguyen et al., 2017). For example, activation of the crus has been associated with motion perception and VPL in humans (Baumann and Mattingley, 2010; Sokolov et al., 2010). In addition, diffusion tensor imaging suggests that the crus is anatomically connected to the right STS which would allow access to the ventral visual stream (Sokolov et al., 2014), which is consistent with the predominantly ipsilateral VTA-EM effects we observed.

Evidence for the location of cortical plasticity in VPL along areas in the visual ventral processing stream is varied, and may depend on specific task requirements (Watanabe and Sasaki, 2015). VPL studies have found evidence supporting both early visual regions (Schoups et al., 2001; Watanabe et al., 2002), mid-stage visual regions (Adab et al., 2014; Adab and Vogels, 2011; Arsenault and Vanduffel, 2019; Yang and Maunsell, 2004) and even high-order decision making regions (Diaz et al., 2017). In our study, cue-VTA-EM enhanced responses to the paired category in both early visual cortex and anterior temporal regions. Increased responses in early visual cortex were restricted to the hemisphere ipsilateral to the stimulating electrodes while more anterior plasticity was found bilaterally. We did not find enhanced category-selective responses in posterior IT cortex as we previously did with both simple gratings and moving random dot stimuli (Arsenault and Vanduffel, 2019). Conversely, increases in functional activity in early visual and anterior temporal cortex established for complex face and body

cues here, were absent in our previous study (see **Figure 6** for a comparison across studies). Three main factors that differed between the experiments were **(1)** more complex stimuli were used (faces and bodies vs. gratings and moving random dots), **(2)** the stimulus duration was reduced (100-150 ms vs. 500 ms) to allow tighter control of attention, and **(3)** in the present experiment, fMRI acquisition was alternated with days of cue-VTA-EM pairing sessions. It is conceivable that the complexity of the subliminal visual stimuli determines where plasticity occurs.

Interestingly, meta-analyses of subliminal face presentations in humans showed that subliminal face presentations consistently activate primary visual cortex (Brooks et al., 2012). In addition, the human fusiform face area, a region that shows similar face selectivity to anterior temporal regions in monkeys (Yovel and Freiwald, 2013), is also activated by subliminal faces (Jiang and He, 2006). Thus, it is reasonable to assume that weak, unattended face and body stimuli preferentially generated weak visual activity within these regions. In turn, this could provide the activity needed to support plasticity within these early and later nodes in the visual system since an interaction is required between retinally-driven activity and a neuromodulatory signal to drive TI-VPL (Roelfsema et al., 2010). Finally, since the anterior temporal regions receive stronger dopaminergic innervations (Berger et al., 1988; Brown et al., 1979; Jacob and Nienborg, 2018), these regions are likely more susceptible to dopamine driven plasticity. Taken together these differences may explain why relatively stronger plasticity was observed in these anterior regions.

The widescale pattern of VTA efferent connections in primates is not well known. One study examining the laterality of mesostriatal connections found that 13% of connections were contralateral while the remaining connections were ipsilateral (Francois et al., 1984). Together with our recent anterograde tracer study targeting the VTA (Zubair et al., 2021), this suggests that the primate dopaminergic midbrain regions are more strongly connected with ipsilateral structures. Recent evidence, however, has shown that contralateral projections of the rodent VTA do have functional significance, albeit to a

smaller extent than ipsilateral projections (Fox et al., 2016). In fact, previously we have confirmed the impact of contralateral projections when directly combining VTA-EM with fMRI (Murriss et al, 2020). Note, however, that these findings revealed reduced ipsilateral fMRI sensitivity, which can most probably be attributed to an asymmetry in the temporal SNR when stimulating inside the scanner (**Table S6**). Without such an asymmetry, and thus in agreement with the preferential contribution of ipsilateral connections from the VTA, strongest behavioral effects of VTA-EM association in our present and previous VPL studies (Arsenault & Vanduffel 2019) were observed in the visual field projecting to the hemisphere with VTA electrodes. Moreover, category-selective and general enhancements of the visual responses were observed bilaterally (**Figure 4**), yet more pronounced in functional patch networks ipsilateral to the stimulating electrodes (**Figure 5C**). Although behavioral and physiological data were inevitably collected at different sessions and under different conditions within this TI-VPL study, it is interesting to note that the combined results support the hypothesis that VTA activation affects ipsilateral areas more than contralateral ones.

Recently, it has been shown that VPL of a trained orientation transferred to untrained orientations when both belong to the same predefined category, but not to different categories (Tan et al., 2019). This category-induced transfer of VPL is position dependent. This led to a hypothetical model which accounts for both location-specific VPL for features outside the context of visual categories and location-independent category learning (Tan et al., 2019). In line with these findings, we found category-selective learning effects induced by VTA-EM at high SNR levels (**Figure 3A**). At this high level, more category-specific information is available to the animals, which we suspect makes them more receptive for a transfer of learning from low to high SNR levels induced by VTA-EM for low SNR cues. Moreover, a comparison of behavioral data with categorization performance of machine learning–based classifiers using low-level features revealed that the animals require substantially more information than the classifiers, particularly at low SNR levels (**Fig. S5**). It remains speculative, however, whether animals use a combination of high versus low level cues. The aforementioned hypothetical

model further suggests that specific category units are trained at higher stages of the visual system, which subsequently activate lower location-specific input stages via top-down signals, which in turn is consonant with the observed modulations of activity in both early and higher order areas (**Figure 4A**). Corroborating the transfer of VPL within specific categories (Tan et al., 2019), we found increased stimulus induced activity after cue-VTA-EM (compared to the activity measured in the pre-cue-VTA-EM phase) in cortical patches selective for the paired but not the unpaired control category (**Figure 5B**). These effects were stronger in the category-selective patches in the hemisphere ipsilateral to the stimulated VTA (**Figure 5C**), in line with the position dependence of category-induced transfer of VPL. Future research is required to understand why object-selective parts of extrastriate cortex also show increased stimulus-driven activity after cue-VTA-EM (**Figure 5B**).

Taken together our findings demonstrate the role of reward-related neuromodulatory structures in driving both perceptual learning and motivated behavior toward temporally correlated sensory features. These findings support a compelling mechanism that may underpin advantageous behavioral changes toward important but unattended visual elements in the environment.

ACKNOWLEDGMENTS

The authors thank C. Fransen, A. Coeman, P. Kayenbergh, I. Puttemans, C. Ulens, A. Hermans, G. Meulemans, M. Depaep, W. Depuydt and S. Verstraeten for technical and administrative support. This work received funding from KU Leuven C14/17/109; Fonds Wetenschappelijk Onderzoek-Vlaanderen (FWO-Flanders) G0D5817N, G0B8617N, G090714N, G0E0520N, and Odysseus G0007.12; the European Union's Horizon 2020 Framework Programme for Research and Innovation under Grant Agreement No 945539 (Human Brain Project SGA3) and from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 856495).

470

471 **AUTHOR CONTRIBUTIONS**

472 Conceptualization, J.T.A., S.R.M., R. V., and W.V., Methodology, S.R.M., J.T.A., and W.V., Investigation,
473 S.R.M., J.T.A. and R.R., Writing – original draft, J.T.A., S.R.M., W.V., Writing - review & editing S.R.M.,
474 J.T.A., R.V., W.V., Funding acquisition, J.T.A., R.V. and W.V., Supervision, W.V. and R.V.

475

476 **DECLARATION OF INTERESTS**

477 The authors declare no competing interests.

478 John T. Arsenault is currently employed by Beacon Health Options, Boston, MA 02109, USA.

479

480 **MAIN FIGURE TITLES AND LEGENDS**

481 **Figure 1. Establishing parathreshold cue parameters.**

482 **(A)** Parathreshold cue SNR was determined by a two-step process. Parathreshold SNR was established
483 with the SNR test, followed by the interference test. **(B)** Overview profile discrimination task: faces or
484 bodies of 5 different SNR levels were presented at the lower diagonal in the right or left visual field at
485 3° eccentricity (duration = M1: 100 ms, M2: 150 ms). Profile orientation was reported by a saccade. **(C)**
486 Orthogonal color task with simultaneously presented task-irrelevant face/body cue (example: RVF_{face}).
487 A fixation period was followed by a color target (33-500 ms) with or without a task-irrelevant cue.
488 Target color was reported with a hand response. **(D)** Performance on the SNR test as a function of SNR
489 levels for faces (left) and bodies (right). Grey dots in the graph represent the mean profile
490 discrimination performance per run (M1 faces: $n=17$, bodies: $n=18$; M2 faces: $n=12$, bodies: $n=11$). A
491 cumulative Gauss sigmoid (in black) was fit to the performance data and the coordinate associated
492 with parathreshold performance (dotted blue line; M1 - 60%, M2 - 70%, see Methods). **(E)** Mean color

discrimination with (body, face) and without (baseline) task-irrelevant cues. Task-irrelevant cues were shown at parathreshold (left; M1: n=16, M2: n=9) and high (right; M1: n=15; M2: n=11) SNR levels. Shaded error = SEM across runs.

Figure 2. Pairing task-irrelevant face and body cues with VTA-EM.

(A) Chronic electrodes were placed in the right VTA, visualized here by T1-weighted MR-images. Cue-VTA-EM association paradigm: one category at one visual field position (example: LVF_{body}) was paired with VTA-EM. The remaining cues (here: LVF_{face}, RVF_{body}, RVF_{face}) were not paired with VTA-EM. All task-irrelevant cue types were equiprobable (parathreshold SNR; M1: 100 ms, M2: 150 ms; 9 conditions: 2 categories, 2 visual fields, 2 orientations and 1 no cue condition). During paired trials, VTA-EM (200 ms, 100 Hz, STAR Methods) began at cue offset. (B) A single round of the main experiment alternated between a cue-VTA-EM association and a behavioral testing day. Each week ended with an interference test day. Each monkey performed five VTA-EM association rounds with the paired category and visual field changing each round. (C) Mean color discrimination performance per run as a function of cue duration measured during interference tests (M1, n = 47; M2, n = 28). (D) Mean color discrimination performance per run during cue-VTA-association sessions (M1, n = 144; M2, n = 94). Shaded error bars = SEM across runs.

Figure 3. Category-selective effects of cue-VTA-EM association: Behavior.

(A) Mean change (post – pre, Δ) in face/body profile discrimination sensitivity (d-prime). Right inset depicts the mean change (post - pre) in sensitivity of each round performed by each monkey (n = 5 rounds/monkey) at the highest SNR level. (B) Mean change in sensitivity for paired and control categories presented in the LVF or RVF. (C) Mean difference in RT for paired and control categories. (D) Mean change in response bias (c-criterion). Data points are the mean behavioral parameter of each

condition calculated using 400 trial bins (M1, pre: n=151, post: n=152; M2: pre: n=128, post: n=129).
Shaded error bars = SEM across bins of each condition.

Figure 4. Category-selective effects of cue-VTA-EM association: Physiology.

(A) Second-level linear model fit to the per-run β -weights (M1 pre: n=49, post: n=63; M2 pre: n=51, M2 post: n=40) of the categorical response (paired vs. control category, STAR Methods). The effect of pairing (post vs. pre) on the categorical response displayed on coronal slices in D99 space ($p < 0.005$, cluster size 10 voxels). Abbreviations in bold indicate anatomical regions of the activation clusters. CB = Cerebellum; V1 = primary visual cortex; Hipp = hippocampus; Pu = putamen (striatum); CM = caudal medial, belt region of the auditory cortex; ACC = anterior cingulate cortex; TAa = STS dorsal bank area; Cl = claustrum; OFC = orbitofrontal cortex; dlPFC = dorsolateral prefrontal cortex. **(B)** Cortical patches responsive to faces and bodies (STAR Methods) presented on a flattened representation of LH and RH. Black overlaid lines indicate a priori defined equally spaced divisions in interaural distance along the anterior-posterior (AP) axis. The effect of pairing on categorical responses, within face and body responsive regions, as a function of position along the AP axis divisions for 2 hemispheres. Thick line denotes group level parameter estimates. Shaded error bars denote 95% confidence interval of parameter estimates. Thin lines with circles (M1) and squares (M2) denote individual animals' parameter estimates. Stars denote a significant, group level effect of pairing (Bonferroni corrected across positions, $p < 0.005$) that was consistent across animals (red and blue stars indicate positive and negative beta value, respectively). Anatomical regions: TG = temporal pole area; TE = anterior inferior temporal area TE; TAa = dorsal bank STS area; IPa = area in the fundus STS; PGa = dorsal bank STS area ; TPO = temporal-parietal-occipital area; TEO = posterior inferior temporal cortical area TEO; MT = medial temporal area (V5); FST = fundus of the STS; V4 = visual area 4; V3v = visual area 3, ventral part; V3d = visual area 3, dorsal part; V2 = visual area 2; V1 = primary visual cortex.

Figure 5. Enhancement of visual activity selective for category representations.

(A) Face (blue), body (green) and object (brown) responsive patches defined in individual animals on the flattened cortical surface. Regions responsive to the paired category (M1 = Body, M2 = Face) are shown with yellow lines while those responsive to the control category (M1 = Face, M2 = Body) are not. **(B)** Mean change (post – pre) in the stimulus response (cues vs. fixation) after cue-VTA-EM association within patches responsive to paired, control and object categories. **(C)** Mean change in the stimulus response within all patches in LH and RH. The lightning bolt indicates the electrode positioning (RH). Mean beta weights in **(B)** and **(C)** were computed from per run beta weights (M1 pre: n=49, post: n=63; M2 pre: n=51, M2 post: n=40). Error bars in **(B)** and **(C)** = SEM.

Figure 6. VTA-EM induced cortical plasticity for face and body cues compared with gratings.

(A) Flatmap (RH) with fMRI activity from the present study (Post vs Pre (paired vs control category), conjunction analysis: $p < 0.0025$). Retinotopic coordinates (unpublished probabilistic maps derived from 13 animals participating in Zhu and Vanduffel, 2019) are displayed (+: upper VF; -: lower VF; white dashed line = horizontal meridian), together with eccentricity (colored lines) and polar angle (black dashed lines). The bottom right image provides a schematic of the stimuli. **(B)** Flatmap (RH) with fMRI activity from our previous study (Arsenault & Vanduffel., 2019). Conventions as in (A), but gratings as stimuli. **(C)** Functional ROI-based comparison between TI-VPL studies. Functionally defined ROIs based on the categorical response (Post vs Pre (paired vs control category)); green). ROIs based on the gratings (Post vs Pre (paired vs control orientation); blue). Both sets of ROIs are derived from a conjunction analysis using the same threshold criteria ($p < 0.0025$; cluster size > 10) and are overlaid on a flatmap of D99 (RH). Graphs show mean t-score for both the categorical (green) and orientation responses (blue). Note that plasticity occurred at different cortical locations when different cues were paired with VTA-EM.

Table 1
(A) Group data

Cluster (#)	Hemisphere	Region	Coordinates (mm)	K _E (#)	p _{uc} (mean)	p _{uc} (peak)
1	Left	Brainstem	-7.25, -19.75, -11.5	27	0.0015	2.4*10 ⁻⁴
2	Right	TAa	25.75, -0.75, -10.5	23	0.0010	1.0*10 ⁻⁴
3	Right	Cerebellum	10.75, -28.75, -10.5	13	0.0016	3.4*10 ⁻⁴
4	Left	Cerebellum	-17.25, -35.75, -9.5	14	0.0017	4.0*10 ⁻⁴
5	Right	Clastrum	14.75, 1.25, -5.5	25	0.0012	1.3*10 ⁻⁴
6	Right	DG	12.75, -13.75, -5.5	26	0.0008	1.9*10 ⁻⁵
7	Left	Cerebellum	-8.25, -31.75, -5.5	22	0.0019	8.0*10 ⁻⁴
8	Left	RPB	-26.25, -3.75, -2.5	23	0.0018	7.6*10 ⁻⁴
9	Right	12o	17.75, 12.25, -2.5	20	0.0016	4.1*10 ⁻⁴
10	Right	Striatum	13.75, -7.75, 2.5	38	0.0014	1.5*10 ⁻⁴
11	Left	46v	-12.25, 14.25, 6.5	28	0.0010	4.1*10 ⁻⁵
12	Right	CM	17.75, -13.75, 5.5	14	0.0017	3.0*10 ⁻⁴
13	Right	24c	5.75, 14.25, 6.5	18	0.0009	7.0*10 ⁻⁵
14	Right	Striatum	7.75, 0.25, 12.5	34	0.0016	3.0*10 ⁻⁴
15	Right	36p	-0.25, -1.75, 12.5	21	0.0016	5.2*10 ⁻⁴
16	Right	V1	18.75, -36.75, 15.5	11	0.0020	6.2*10 ⁻⁴
17	Left	F2	-12.25, 3.25, 19.5	12	0.0018	1.8*10 ⁻⁴

(B) Conjunction data

Cluster (#)	Hemisphere	Region	Coordinates (mm)	K _E (#)	p _{uc} (mean)	p _{uc} (peak)
1	Right	PACs	5.75, -0.75, -13.5	10	0.0489	0.0340
2	Right	Cerebellum	5.75, -26.75, -11.5	13	0.0451	0.0241
3	Left	CA1	-17.25, -8.75, -12.5	11	0.0591	0.0450
4	Right	TAa	25.75, -1.75, -9.5	28	0.0412	0.0102
5	Right	Clastrum	14.75, 2.25, -7.5	24	0.0470	0.0182
6	Left	Cerebellum	-8.25, -31.75, -5.5	36	0.0461	0.0195
7	Right	DG	13.75, -12.75, -4.5	11	0.0527	0.0232
8	Left	14r	-1.25, 25.25, -4.5	10	0.0308	0.0014
9	Left	V3v	-15.25, -32.75, -1.5	15	0.0511	0.0237
10	Right	V3v	13.75, -29.75, -2.5	18	0.0582	0.0400
11	Right	12o	17.75, 12.25, -2.5	21	0.0449	0.0147
12	Right	Striatum	13.75, -7.75, 2.5	28	0.0347	0.0126
13	Left	10mr	-3.25, 26.25, 4.5	11	0.0489	0.0180
14	Left	46v	-12.25, 14.25, 6.5	31	0.0350	0.0036
15	Right	CM	17.75, -13.75, 5.5	14	0.0441	0.0123
16	Left	MT	-17.25, -24.75, 9.5	11	0.0468	0.0300
17	Right	36p	-0.25, -1.75, 12.5	27	0.0402	0.0120
18	Right	F2	10.75, 1.25, 13.5	17	0.0526	0.0118
19	Right	V1	18.75, -34.75, 15.5	17	0.0565	0.0310

(A-B) Overview of (de-)activation clusters derived from the **(A)** group (same analysis as in **Figure 4A**), and **(B)** conjunction analyses (same analysis as in **Figure S6A**) throughout the brain. Clusters are ranked

(#) from the most ventral to the most dorsal activations. Coordinates indicate the peak voxel of each cluster within D99 space. K_E = cluster size, p_{uc} (mean) = mean uncorrected p-value of an entire activation cluster, p_{uc} (peak) = p-value of the most significant voxel within a cluster.

STAR METHODS: TEXT

Resource Availability:

Lead contact

Further questions regarding the methodology and resource sharing should be directed and will be fulfilled by the Lead Contact, Wim Vanduffel (wvanduffel@mgh.harvard.edu)

Materials availability

The study did not produce new materials

Data and Code availability

The behavioral datasets are available at Mendeley Data, V1, doi: 10.17632/dkzkg79c77.1. Associated custom-made scripts to generate the main figures and statistical outcomes in this study are available in the same online repository. Raw data can be obtained upon request.

KEY RESOURCE TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, Peptides, and Recombinant Proteins		
Molday ION	BioPAL Inc.; Worcester, USA	http://www.biopal.com/molday-ion.htm
Deposited Data		
behavioral data	Sjoerd Murris, John Arsenault, Wim Vanduffel	Mendeley Data, V1, doi: 10.17632/dkzkg79c77.1.
Experimental Models: Organisms/Strains		

Rhesus macaque (<i>Macaca mulatta</i>)	Biomedical Primate Research Centre (BPRC), the Netherlands	https://www.bprc.nl/en
Software and Algorithms		
MATLAB 2017a	MathWorks	RRID: SCR_001622
Python Programming Language	Python	RRID: SCR_008394
R Project for Statistical Computing	R Project	RRID: SCR_001905
SPM 12	SPM, University College London, UK	RRID: SCR_007037
FSL	FMRIB, WIN, Oxford, UK	RRID: SCR_002823
ANTS – Advanced Normalization Tools	ANTS, University of Pennsylvania, USA	RRID: SCR_004757
AFNI – Analysis of Functional Images	AFNI, NIMH, USA	RRID: SCR_005927
GLMdenoise	NITRC, USA	RRID: SCR_014116
Psignifit toolbox	Eberhard Karls Universität Tübingen: Wichmann & Hill, 2001a	https://github.com/wichmann-lab/psignifit/wiki
Other		
DS 8000	World Precision Instruments	RRID: SCR_008593
DLS 100	World Precision Instruments	RRID: SCR_008593
Microwire Brush Array (MBA)	Microprobes, National Institutes of Health	https://www.microprobes.com/products/multichannel-arrays/mba
Eye Tracking Device	ISCAN Inc.	http://iscaninc.com/

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Two adult male rhesus macaque monkeys (*Macaca mulatta*) participated in the current study (Monkey 1: ~6.5 Kg, 7 years old; Monkey 2: ~8.5 Kg, 8 years old). The animals are housed in enclosures of the KU Leuven Medical School with a natural day-night cycle and at least one other cage companion. On weekdays, dry food is provided *ad libitum* and monkeys can obtain water during experiments until satiated. During weekends, the animals get water and a mixture of fruits and vegetables. The animals have continuous access to toys and other enrichment devices. Prior to experiments, both monkeys underwent surgery to place an MR-compatible headset attached to the skull by ceramic screws, covered by a layer of dental cement (Vanduffel et al., 2001). In a second subsequent surgery monkeys were implanted with micro-brush electrode arrays (Microprobes; Gaithersburg, USA), consisting of 34

micro-wires (25-50 μm in diameter) (Arsenault et al., 2014; McMahon et al., 2014). Both animals have participated in previous studies in which we tested the impact of electrical microstimulation of the VTA on behavior and during contrast-agent enhanced fMRI (Arsenault and Vanduffel, 2019; Murris et al., 2020). All surgical and experimental protocols are in agreement with institutional (KU Leuven Medical School: Ethische Commissie Dierproeven), national and European guidelines (Directive 2010/63/EU).

METHOD DETAILS

Behavioral paradigm: Profile task

Both animals were trained on a profile discrimination task in which they had to indicate the profile orientation (left- or rightward) of fifty human faces and fifty headless human bodies, generated using FaceGen (faces) and MakeHuman (bodies) software (for an overview of these cues see **Figure S1**). At the onset of the trial the monkey had to fixate at a white square (0.14×0.14 visual degrees) in the center of the screen with a background of dynamic white noise. The noise background was refreshed at a frequency of 60 Hz and the background luminance was equated to the luminance of the face/body stimuli ($\sim 66.7 \text{ cd/m}^2$). After a variable amount of time (1500 - 2500 ms), monkeys were presented with a face or body cue in the lower quadrant of either the LVF (225 degrees) or the RVF (315 degrees) at a distance of three visual degrees. These images were presented for a short duration (M1: 100 ms, M2: 150 ms), during which the monkeys had to maintain fixation. Following cue offset, the monkeys reported the profile orientation of the cue with a saccade to targets on either the left or right of the fixation point. Saccades to the left denoted leftward orientation, while saccades to the right denoted a rightward orientation. The monkeys were subsequently rewarded with juice for correct responses. After initial training at high SNRs, all profile discrimination days tested profile discrimination at a range of SNR levels with face and body cues being presented during alternating runs (**Figure 1B, Video S1**). The SNR levels were calculated as the percentage of randomly selected pixels from the face or body cue that were presented. All other pixels on the screen consisted of dynamic background noise refreshed at 60 Hz. Six levels of SNR were used for faces (0, 5, 10, 15, 25 and 50) and six alternative SNR

levels for bodies (0, 7, 15, 25, 45, 65). We used higher SNR levels for body images, since pilot data indicated that it was more difficult for the two animals to discriminate the orientation of these images.

Behavioral paradigm: Color task

Monkeys were trained to perform a color discrimination task (**Figure 1C, Video S2**). Throughout the trials of the color task, monkeys were required to keep their gaze on a white fixation square (0.14 x 0.14 visual degrees). After a variable time interval (1500 - 2500 ms), two stimuli appeared simultaneously: **(1)** a color target consisting of small colored box (0.22 x 0.22 visual degrees) located 7.5 visual degrees above the central fixation square and **(2)** a task-irrelevant face or body cue located 3 visual degrees in the lower diagonal of the LVF or RVF (same positions as in the profile task). Both the color target and the irrelevant cue were presented for the same duration. After these stimuli were presented and while maintaining fixation, the monkeys had to discriminate whether the color target was more blue or more red. They responded by interrupting an infrared signal with either their left (blue) or right hand (red). The monkeys received juice for correct answers. The color values of the boxes were adjusted on a run-by-run basis to keep the monkeys' performance in a range between 70-80%. In addition, trials without irrelevant cues were intermixed (1 in 9 trials) so that color task performance could be compared.

Behavioral paradigm: Cue durations

To ensure that the interference test was sensitive, a cue duration associated with performance on the color task above chance but below ceiling level was needed. Accordingly, we performed pilot experiments in which we tested color task performance over a range of color target durations for each animal. Color targets durations of 100 ms and 150 ms were associated with an adequately sensitive

performance level (below ceiling but above chance) for M1 and M2, respectively. These cue durations were used in subsequent experiments.

Behavioral paradigm: Interference test

The interference test was a variant of the color task used to confirm that the color task was sensitive enough to detect whether attention was drawn away from the color target (**Video S2**). To achieve this, the influence of cue duration on color task performance was assessed. Accordingly, during the interference tests, we tested color task performance at a range of cue durations. These cue durations were tuned to the performance of each monkey (M1: 33, 66, 100, 200 and 400 ms; M2: 50, 100, 150, 300 and 500 ms). This allowed performance at the parathreshold duration (M1: 100ms, M2: 150 ms) used in the main experiment to be compared to performance at shorter and longer durations. In addition, trials with task-irrelevant cues (presented at parathreshold SNR levels) and trials without the presence of a task-irrelevant cue (1 in 9 trials) were randomly intermixed to assess whether the task-irrelevant cues also affected color task performance.

Behavioral paradigm: Establishing parathreshold cue parameters

We determined parathreshold cue parameters based on two criteria: **(1)** performance on the profile task was in a parathreshold range (near threshold level discrimination) and **(2)** when task-irrelevant cues were shown simultaneously with the color targets, color task performance was not affected. These criteria were tested using a two-step process (**Figure 1A**). In the SNR test, we measured the monkeys' performance on the profile task as a function of SNR. This was then used to determine the SNR levels associated with parathreshold performance. In the interference test, we tested whether irrelevant cues affected color task performance when they were presented at the parathreshold SNR level. If color task performance was affected then the SNR level was lowered and the interference test was repeated until

color task performance was no longer affected. These parathreshold cues found with this process were used during all phases of all the rounds of the main experiment (rounds 1-5, M1: faces – 9.6%, bodies – 17.8%, M2: faces – 10.3%, bodies - 20.1%).

VTA electrical microstimulation

A micro-brush electrode array (Bondar et al., 2009) was chronically implanted into the VTA of both animals (Arsenault et al., 2014). The microwires of each electrode array consisted of a mix of 25 & 50 μ m diameter wires. The microwires were attached to a 36-pin omnetics connector and were uniformly cut 5 mm past the microfil tube at the recording end. For M2, we performed unipolar electrical microstimulation using the stimulating electrodes in the VTA and a low impedance ground wire implanted below the skull. For M1 we performed bipolar stimulation across two electrodes in the VTA. For both animals, we attached an eight-channel digital stimulator (DS8000, World Precision Instruments; Sarasota, USA) connected to stimulus isolators (DLS100, World Precision Instruments; in current mode) to the chronically implanted electrodes in the VTA of the monkeys. Electrical stimulation events were triggered by custom scripts in Python (Python; Beaverton, USA), that were simultaneously responsible for behavioral and visual paradigms. Stimulation epochs lasted for 200 ms and consisted of 20 equally spaced biphasic pulses of 0.4 ms (100 Hz). Stimulation parameters were based on our previous study in which we successfully demonstrated the induction of VPL through VTA-EM (Arsenault & Vanduffel, 2019).

Behavioral paradigm: Cue-VTA-EM association

During cue-VTA-EM association days monkeys performed the color task with parathreshold task-irrelevant face or body cues (cue duration; M1: 100 ms, M2: 150 ms) being shown concurrently with the color target. One of two categories (face or body) displayed in one of two VFs (RVF or LVF) was paired with VTA-EM (rendering four possible paired stimulus configurations: RVF_{Face}, LVF_{Face}, RVF_{Body},

LVF_{Body}). For the paired condition, VTA-EM immediately followed the task-irrelevant cues presentation with stimulation parameters as described above (VTA electrical microstimulation).

Design: Main experiment (round 1)

Round 1 of the main experiment aimed to examine both the behavioral (profile discrimination) and physiological (fMRI) changes resulting from pairing task-irrelevant face and body images with VTA-EM ($n = 2$, 2 monkeys x 1 round/monkey). This round was comprised of five phases (**Figure S2**). First, the monkeys performed the profile discrimination task for a series of days during the pre-association behavioral phase. Next, a series of pre-association fMRI days were acquired during the pre-association fMRI phase. Following this, the cue-VTA-EM association phase began in which the monkeys performed a series of successive cue-VTA-EM association days after which cue-VTA-EM association days were intermixed with post-association fMRI and behavior days. Therefore, all post-association fMRI and profile discrimination days performed during the post-association phases were preceded by a day of cue-VTA-EM association (see **Figure S2** for further details).

Design: Main experiment (rounds 2 – 5)

Rounds 2-5 of the main experiment aimed to further examine the behavioral effects of cue-VTA-EM association on profile discrimination ($n = 8$, 2 monkeys X 4 rounds/monkey). This portion of the main experiment consisted of three distinct phases (pre-association behavior, cue-VTA-EM association, post-association behavior, see **Figure S3**). In the first phase (pre-association behavior), five days of the profile discrimination task were used to test the animal's ability to discriminate the profile orientation of face and body cues across a range of SNR levels. For rounds 2 through 5, the previous rounds post-association behavioral phase was used as the current round's pre-association behavioral phase. This was possible because the category and VF paired with VTA-EM were alternated between rounds. Following the pre-association behavioral phase, the cue-VTA-EM association phase began. To maximize potential behavioral improvements from cue-VTA-EM association, we intermixed the post-association

behavioral phase with the cue-VTA-EM association phase (**Figure 2B** and **S3**). Therefore, each week began with a cue-VTA-EM association day (Day 1) followed by a post-association behavior day (Day 2). This was repeated on Days 3 and 4. On Fridays (Day 5), the interference test was performed by the animals to confirm that the color task remained sensitive to potential shifts in attention. For each of these rounds (rounds 2-5) five cue-VTA-EM association days and five profile discrimination days (post-association behavioral phase) were acquired. Each monkey (M1 and M2) performed four rounds using this design.

fMRI acquisition

All functional MRI experiments used the contrast agent Molday (BioPAL Inc.; Worcester, USA) that was injected before each session at a dose proportional to the weight of the monkey (9-11 mg/Kg). Functional images were acquired in a 3 Tesla horizontal bore scanner (PrismaFit, Siemens; Erlangen, Germany) with an echo-planar imaging (EPI) sequence and interleaved sampling acquisition (50 horizontal slices, TE = 13 ms, TR = 2 s, flip angle: 84°, voxel size: 1.25 x 1.25 x 1.2 mm). An external transmitter and 8-channel receiver-coils were fitted around the monkey's head before positioning him in the isocentre of the magnet. To confirm whether the external coils were functional and to aid co-registration with anatomical images, we took a separate T2* weighted Gradient Echo (GRE) scan at the onset of each day (TR = 5 ms, TE = 2 ms, voxel size = 1.25 x 1.25 x 1.2 mm).

fMRI experimental design: pre- and post-association fMRI (round 1)

fMRI responses to face and body images were measured before (pre-association fMRI phase; M1: 5 days, 49 runs; M2: 5 days, 63 runs) and after (post-association fMRI phase; M1: 5 days, 51 runs; M2: 6 days, 40 runs) the cue-VTA-EM association phase during round 1. During the cue-VTA-EM association phase (situated in between the pre- and post-association fMRI phases) the monkey performed the color discrimination task while one category in one VF was paired with VTA stimulation (M1: RVF_{Body}, M2: LVF_{Face}). The experimental design during individual fMRI runs was identical to the cue-VTA-EM

association days with the exception that no VTA-EM was delivered. In addition, a fixation condition was included during which monkeys performed the color task without additional cues. For each of the runs (500 TRs, 1000 s), nine conditions (2 VFs x 2 categories x 2 orientations + 1 fixation/no cue) were randomly presented (duration M1: 100 ms, M2: 150 ms) every 1500 - 2500 ms. In the scanner the color task was presented on a translucent screen placed 57 cm from the monkey's eyes, using a Barco LCD Projector (Barco; Kortrijk, Belgium). We ensured that the visual stimuli presented in the scanner were equated to the behavioral set-ups, where behavior and cue-VTA-EM association days were performed. To achieve this, the RGB-to-luminance relationships were measured in the scanner and set-up with a CS-100 Chroma Meter (Konica Minolta; Ramsey, USA). This information was then used to adapt the visual stimuli (gamma correction) such that mean background and mean face/body image luminance values were equated (~ 66.7 cd/m²).

fMRI experimental design: Face/body localizers

Prior to the VTA-EM association blocks and associated behavioral paradigms, we localized representations of face, body, and object categories using an established fMRI block design (Popivanov et al., 2012). fMRI blocks (20 s duration) for the face, body or object conditions consisted of grayscale images (500 ms per image, monkey and human faces, monkey and human bodies) presented at different orientations. Blocks of inanimate objects and fruits (500 ms per image, 20 s block duration) of comparable aspect ratios and luminance values were shown within the same runs. During a run, each block of sequential images was presented twice in a pseudo-randomized order without inter stimulus intervals.

Classifier Implementation: Pre-processing

A stimulus set from the profile task was used as inputs to the classifiers, with the notable exception that the images were static, single frames of the actual dynamic noise stimuli. For each stimulus category (faces and bodies), 18,000 images (164 x 164 pixels) were generated from a combination of

300 base images (Overview in **Figure S1**: 50 individuals x 2 views x 3 sizes) at 6 SNR levels (0, 5, 10, 15, 25 and 50), for which we generated 10 random instances of signal to noise. For each of these images, a noise background was randomly added from a set of 50 different backgrounds. The pixels (values and locations) of the background image were replaced with pixels randomly drawn from a base image according to the SNR value. Finally, pixel values (0-255) in the generated images were linearly rescaled to values between 0 and 1.

QUANTIFICATION AND STATISTICAL ANALYSIS

Behavioral metrics

Behavior of the monkeys in the profile discrimination task was quantified in three different ways: discrimination performance was assessed by **(1)** the sensitivity index d-prime (d'), **(2)** the bias index c-criterion and **(3)** by RTs. For the color discrimination task, analysis was focused on percent of correct trials (% correct). To compute % correct, only trials in which the monkey responded with a saccade were considered. From these trials, % correct was calculated as the number of correct trials divided by the sum of correct and incorrect trials:

$$(1) \% \text{ correct} = N_{\text{correct}} / (N_{\text{correct}} + N_{\text{incorrect}}) * 100$$

D-prime (d') and c-criterion were used to measure profile discrimination sensitivity and bias, respectively. These summary statistics were computed separately for each combination of VF and category (LVF_{face} , LVF_{body} , RVF_{face} , RVF_{body}) at each SNR level. Importantly, although VTA-EM was never directly associated with a saccade direction, VTA-EM was associated with a specific spatial location (LVF or RVF). Consequently, to determine whether the VF associated with VTA-EM biased the profile orientation animals reported, we defined hit rate and false alarm rate relative to the spatial location associated with VTA-EM. Therefore, when the LVF was associated with VTA-EM, hit rate (H) was defined as the proportion of correct leftward saccades out of all left oriented cue presentations [$H = P(\text{left}_{\text{saccade}} | \text{left}_{\text{orientation}})$] while the (F) false response rate was the proportion of incorrect leftward

saccades from all right oriented cue presentations $[(F) = P(\text{left}_{\text{saccade}} | \text{right}_{\text{orientation}})]$. When the RVF was associated with VTA-EM the formula for hit rate $[(H) = P(\text{right}_{\text{saccade}} | \text{right}_{\text{orientation}})]$ and false alarm $[(H) = P(\text{right}_{\text{saccade}} | \text{left}_{\text{orientation}})]$ were inverted. Using these definitions of hit and false-alarm rate, D-prime was calculated as the difference between the z-scores of the hit rate (H) and the false alarm rate (F).

$$(2) d' = z(H) - z(F)$$

The c-criterion, which indicated the bias to report the saccade toward the spatial location associated with VTA-EM, was calculated as the average of the z-scored hit rate and the false alarm rate multiplied by negative 1 and can be formulated as follows:

$$(3) C = -1 * (z(H) + z(F)) / 2$$

Finally, the RT was measured as the time interval (in milliseconds) between the onset of the visual stimulus and the monkey indicating his choice by making a saccade to one of the two equally sized reporting windows ('left' or 'right').

Behavioral analysis: Interference test

To determine whether the presence of the task-irrelevant cues affected color discrimination performance, we fit a LM to the performance (% correct) of the animals per run (+/- 400 trials in which the monkey answered). The factors used were cue duration (levels 0, 1, 2, 3 or 4, from shortest to longest), category (face, body, or baseline), the interaction of cue duration and category, and monkey (M1 or M2) to control for differences in performance across animals. Separate models were constructed for parathreshold and high SNR tests to determine cue parameters (**Figure 1E**) and for the interference tests performed during the cue-VTA-EM phase (**Figure 2D**). Before fitting the LMs, the dependent variable (% correct) was transformed by squaring to better meet the assumptions of normal residuals and homoscedasticity. After the models were fit, a Shapiro-Wilk test was performed on the residuals to test for normality. No significant difference from normality was found for any models suggesting that the assumption of normally distributed residuals was met (see **Tables S1** and **S2**).

Levene's test was also performed to test whether the variances differed across categories. For the interference test performed before the main experiment (**Figure 1E**), no significant difference from equal variance was found indicating that the assumption of homoscedasticity was met (see **Table S1**). In contrast a significant inequality of variances was found for the cue-VTA-EM association phase (see **Table S2**). Accordingly, a series of non-parametric Friedman's tests were run testing the effect of category for each cue duration. This supported the results of the ANOVA, with no significant effect of category found at any cue duration ($p > 0.09$, uncorrected for multiple comparisons, see **Table S2**).

Behavioral analysis: Main experiment (data partitioning)

Data from the pre- and post-association behavioral phases of each round and for each monkey were grouped into bins of 400 trials, which corresponds approximately to the number of trials in a block in which the monkey performed the profile tasks (before starting with a new block). From these bins d-prime and c-criterion were calculated for each combination of VF and category (LVF_{paired} , LVF_{control} , RVF_{paired} , RVF_{control}) while mean RT was calculated for each combination of VF, category and orientation ($LVF_{\text{paired_leftward}}$, $LVF_{\text{control_leftward}}$, $RVF_{\text{paired_leftward}}$, $RVF_{\text{control_leftward}}$, $LVF_{\text{paired_rightward}}$, $LVF_{\text{control_rightward}}$, $RVF_{\text{paired_rightward}}$, $RVF_{\text{control_rightward}}$).

Behavioral analysis: Main experiment (LME models)

To assess how behavior was affected by cue-VTA-EM association we fit random-intercept LME models (LME4 package: (Bates et al., 2014)) at each SNR level and for each behavioral metric collected (d-prime, c-criterion, and RT) across the 5 rounds of the main experiment (see **Table S3** for exact model specifications). For group models, the monkeys (M1 and M2) and the rounds (1-5) each monkey performed were used as a random effects factor. In the LME models for the separate animals, only the rounds were used as a random effects factor. Visual inspection of the residuals from several of these models indicated that assumptions of normality and homoscedasticity were not met. Accordingly, we treated the skewed residuals and heteroscedasticity as a nuisance and used bootstrapped confidence

intervals to determine significance under the assumption that the data provided a representative sample of the two monkeys change in behavior after cue-VTA-EM association (Pek et al., 2018). Parametric bootstrapping was performed using 5.000 simulations (LME4) and percentiles for the 95% confidence intervals (alpha = 0.05) were Bonferroni corrected for multiple comparisons across SNR levels (n=6):

$$\text{Bonferronicorrected} \left(\frac{\alpha}{2}, 1 - \frac{\alpha}{2} \right) = \left(\frac{\frac{\alpha}{2} * 1}{n}, 1 - \frac{\frac{\alpha}{2} * 1}{n} \right)$$

The parameter estimates and Bonferroni corrected 95% confidence intervals for the group analyses are shown in **Table S3**. In addition, to assess whether significant effects found in the group analyses were consistent across animals, separate models were made for each monkey. The parameter estimates of these models and their bootstrapped 95% confidence intervals are also reported in **Table S3**. Although we provide a full overview of all the models (both on a group and individual level) in the supplementary material, we restrict our discussion in the results to effects that were found to be significant at the group level and those that exhibited parameter estimates for the individual monkeys that showed the same effect direction.

fMRI analysis: Pre-processing

Images were first re-orientated and then slice timing correction was performed (AFNI; NIMH, USA: (Cox, 1996)). A non-rigid, slice-by-slice realignment within runs and an affine realignment between within day runs was performed to correct for motion (Kolster et al., 2009). A representative mean functional image was then non-rigidly co-registered to the D99 atlas space (Reveley et al., 2017) using ANTs (Advanced Normalization Tools; University of Pennsylvania, USA: (Avants et al., 2011)) following a three-step procedure: **(1)** The mean functional image was co-registered to a GRE image acquired in the same session. **(2)** That GRE image was co-registered to a T1 anatomical from the same monkey. **(3)** This anatomical was registered to the D99 atlas space. These transformations were then applied to each 3D fMRI time-series acquired within the same session. In a final pre-processing step, we spatially

878 smoothed the images in FSL (FMRIB Software Library; University of Oxford, UK: (Smith et al., 2004))
879 employing an isotropic Gaussian kernel (FWHM: 4 mm).

880

881

882 **fMRI analysis: Pre- and post-association fMRI phases**

883 Subsequently, regressors of interest were made using the onsets and durations of each condition to
884 generate a boxcar model that was convolved with the iron oxide response function (Leite et al., 2002).
885 A fixed-effect analysis was then performed on each run by fitting a general linear model (GLM) using
886 high-pass filtering (128 s), an AR(1) model and restricted maximum likelihood estimation (SPM12;
887 London, UK (Worsley and Friston, 1995)). In addition to the regressors of interest, a single motion
888 regressor (change in correlation with first image in time-series) and multiple noise regressors
889 (generated by GLMdenoise (Kay et al., 2013)) were added as covariates of no interest. After this first-
890 level analysis, we ran a 2nd-level LM analysis on the per run beta images of the categorical response
891 (paired category vs. control category). The LM was fit with factors VF (LVF or RVF), orientation (leftward
892 or rightward), pairing (pre-association or post-association) and monkey (M1 or M2) with pairing being
893 the factor of interest. In addition, separate LMs were fit to the data from each monkey that was
894 identical to the group analysis except the monkey factor was omitted. A conjunction analysis was then
895 performed on the results from the individual animals. Based on pre-determined statistical criteria,
896 activation clusters from the group ($p < 0.002$, cluster size > 10 voxels) and conjunction ($p < 0.005$,
897 cluster size > 10) analysis were isolated and considered significant when present in both analyses.

898

899 **fMRI analysis: face/body localizers**

900 After collecting and pre-processing (see previous section) localizer data from both subjects (M1: 54
901 runs, 3 days; M2: 38 runs, 2 days) per run beta maps were generated for contrasts that identifying
902 regions responsive to faces (faces vs bodies/inanimate objects), bodies (bodies vs faces/inanimate
903 objects) and objects (inanimate objects vs faces/bodies). A 2nd level analysis t-test was then performed

across runs to define category responsive patches for the group and individuals (FWE correction, $p < 0.05$). The results of this analysis were then projected onto flatmaps using FreeSurfer (FreeSurfer; Charlestown, USA) and temporal lobe patches were selected (**Figure S4B and S5A**).

fMRI analysis: Categorical response along temporal lobe

Analysis of changes in the categorical response after pairing was restricted to a large ROI of cortical regions responsive to faces and bodies. In early visual cortex, this ROI was defined using a 2nd Level LM ($\beta\text{-value}_{VIS} \sim \text{pairing} + \text{monkey} + \text{category} + \text{orientation} + \text{VF}$) of the visual response (cue vs. fixation, conveyed by $\beta\text{-values}$). We then selected all voxels in visual cortical regions that displayed a significant effect of VF (LM, main effect of VF, $p < 0.05$). While the underlying data used to generate this part of the ROI was the same as the subsequent analysis of the categorical response, this analysis avoided circularity because **(1)** we looked at the stimulus response not the categorical response and **(2)** we used the main effect of VF to define voxels that showed a differential response when all task-irrelevant cues were shown in the LVF compared to the RVF. Therefore, this part of the ROI was not biased toward finding changes in the categorical response after pairing. The temporal lobe portion of the ROI was generated using the group analysis of the independent localizer experiment. All significant temporal lobe voxels that responded preferentially to faces, bodies or objects were selected (2nd level t-test, FWE, $p < 0.05$).

After the ROI was generated, we collected the per run beta weights of the categorical response at each voxel within the ROI for all combinations of VF and orientation from each monkey. For each position along the anterior to posterior axis, with greater than five voxels present for a given ROI and hemisphere, a LM of the categorical response ($\beta\text{-value}_{CAT} \sim \text{VF} + \text{orientation} + \text{pairing} + \text{monkey}$) was fit. Separate LMs were also fit for the data from each animal and the monkey factor was omitted. For each hemisphere, the group level p values were Bonferroni corrected for the number of positions

tested (LH = 46 positions, RH = 46 positions) and positions that were significant on the group level ($p < 0.005$) and that displayed the same effect direction in both animals were reported.

fMRI analysis: Face, body and object patch ROI analysis

Face, body and object responsive patches were defined in each monkey from the independent localizer experiment (FWE, $p < 0.05$, **Figure S4B**). For the analysis of the categorical response, the mean per run beta estimate of the categorical response (paired category vs. control category) for each ROI was calculated for all combinations of VF and orientation. An LM was then fit to the categorical response ($\beta\text{-value}_{CAT} \sim VF + \text{orientation} + \text{pairing}$). For the analysis of the stimulus response, the mean per run beta estimate of the visual response (cue vs. fixation) for each ROI was calculated for all combinations of VF and orientation and category. An LM was then fit to the data (visual response $\sim VF + \text{orientation} + \text{category} + \text{pairing}$). For both analyses, visual inspection indicated that several of the LMs did not meet the assumptions of homoscedasticity and normally distributed residuals. Therefore we, used bootstrapping (5.000 simulations) to estimate the 95% confidence intervals of the parameters.

fMRI analysis: Face, body and object patch network analysis

Face, body and object responsive ROIs were generated for each monkey using an independent localizer (FWE, $p < 0.05$). We then combined data within each patch network (group of patches that responded to the same category), by computing the mean over the per run beta weights across all patches. In addition, we classified patch networks based upon whether the category they respond to was the paired category (M1: body, M2: face) or the control category (M1: face, M2: body).

An LM was then fit to the categorical response (categorical response $\sim VF + \text{orientation} + \text{monkey} + \text{patch_hemisphere} + \text{pairing} \times \text{patch_category} + \text{pairing} \times \text{patch_hemisphere}$) and the visual response (visual response $\sim VF + \text{orientation} + \text{monkey} + \text{category} + \text{patch_hemisphere} + \text{pairing} \times \text{patch_category} + \text{pairing} \times \text{patch_hemisphere}$). For both analyses visual inspection indicated that the residuals were

not normally distributed. Consequently, we computed 95% confidence intervals using bootstrapping and determined significance using these estimates ($p < 0.05$). For significant effects of interest (pairing x patch_category (object), pairing x patch_category (paired), pairing x patch_hemisphere) found in the group analyses, the results of LM fit to each monkey's data are also provided in **Table S5**.

Classifier Implementation: Statistics

To investigate at which SNR levels successful profile discrimination could be achieved, classifiers were trained and tested on the pixels of the stimulus set. First, we tested the classification accuracy of a linear support vector machine (SVM) on pixel values of the images, in which the regularization parameter C was set to 1. Two alternative training schemes were implemented, which we refer to as Model I and Model II. In Model I, SVMs were trained and tested separately on images belonging to one of the six SNR values, using five-fold cross-validation. Profile discrimination accuracy (in %) at each of these SNR levels was computed for the test data. In Model II, SVMs were trained on images of all SNRs, except those that did not contain any information (i.e. 0 % SNR), using five-fold cross-validation. Subsequent testing of the images was performed at each SNR-level separately.

In addition to classification based on pixel-level SVMs, we also added a classification based on the first convolutional layer of the DCNN AlexNet (Krizhevsky et al., 2012) to our analysis, which represents low-level features akin to those of V1. We employed an implementation of AlexNet that was pre-trained on ImageNet data, imported in MATLAB from the Caffe Model Zoo (<https://github.com/BVLC/caffe/wiki/Model-Zoo>). Pre-processing of the input images consisted of rescaling the images (to 227 x 227 pixels) and normalization to the ImageNet requirements (Deng et al., 2009). We subsequently registered the response of layer-1 units to the images. These responses were standardized (z-score) and analyzed in the same way as pixel gray levels using SVMs. All

classification analyses were performed in MATLAB. Separate toolboxes were used for implementation of the SVMs (Statistics and Machine Learning Toolbox) and DCNNs (The Deep Learning Toolbox).

SUPPLEMENTAL VIDEO AND EXCEL TABLE TITLES AND LEGENDS

Video S1 (Related to Figure 1)

Examples of profile task trials with bodies and faces.

Video S2 (Related to Figure 1)

Examples of different color task trials with high and parathreshold face and body stimuli and distractors.

REFERENCES

Adab, H.Z., Popivanov, I.D., Vanduffel, W., and Vogels, R. (2014). Perceptual learning of simple stimuli modifies stimulus representations in posterior inferior temporal cortex. *Journal of Cognitive Neuroscience* 26, 2187-2200.

Adab, H.Z., and Vogels, R. (2011). Practicing coarse orientation discrimination improves orientation signals in macaque cortical area v4. *Current Biology* 21, 1661-1666.

Ahissar, M., and Hochstein, S. (1993). Attentional control of early perceptual learning. *Proc Natl Acad Sci USA* 90, 5718-5722.

Anderson, B.A., Laurent, P.A., and Yantis, S. (2011). Value-driven attentional capture. *Proc Natl Acad Sci USA* 108, 10367-10371.

Arsenault, J.T., Rima, S., Stemmann, H., and Vanduffel, W. (2014). Role of the primate ventral tegmental area in reinforcement and motivation. *Current Biology* 24, 1347-1353.

1001 Arsenault, J.T., and Vanduffel, W. (2019). Ventral midbrain stimulation induces perceptual learning and
 1002 cortical plasticity in primates. *Nat Comm* 10, 3591.
 1003 Athalye, V.R., Santos, F.J., Carmena, J.M., and Costa, R.M. (2018). Evidence for a neural law of effect.
 1004 *Science* 359, 1024-1029.
 1005 Avants, B.B., Tustison, N.J., Song, G., Cook, P.A., Klein, A., and Gee, J.C. (2011). A reproducible
 1006 evaluation of ANTs similarity metric performance in brain image registration. *NeuroImage* 54, 2033-
 1007 2044.
 1008 Bao, S., Chan, V.T., and Merzenich, M.M. (2001). Cortical remodelling induced by activity of ventral
 1009 tegmental dopamine neurons. *Nature* 412, 79-83.
 1010 Baraduc, P., Duhamel, J.R., and Wirth, S. (2019). Schema cells in the macaque hippocampus. *Science*
 1011 363, 635-639.
 1012 Bartra, O., McGuire, J.T., and Kable, J.W. (2013). The valuation system: a coordinate-based meta-
 1013 analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage* 76,
 1014 412-427.
 1015 Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R.H.B., Singmann, H., Dai, B., and Eigen, C.
 1016 (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67, 1-48.
 1017 Baumann, O., Borra, R.J., Bower, J.M., Cullen, K.E., Habas, C., Ivry, R.B., Leggio, M., Mattingley, J.B.,
 1018 Molinari, M., Moulton, E.A., *et al.* (2015). Consensus paper: the role of the cerebellum in perceptual
 1019 processes. *Cerebellum* 14, 197-220.
 1020 Baumann, O., and Mattingley, J.B. (2010). Scaling of neural responses to visual and auditory motion in
 1021 the human cerebellum. *J Neurosci* 30, 4489-4495.
 1022 Berger, B., Gaspar, P., and Verney, C. (1991). Dopaminergic innervation of the cerebral cortex:
 1023 unexpected differences between rodents and primates. *Trends Neurosci* 14, 21-27.
 1024 Berger, B., Trottier, S., Verney, C., Gaspar, P., and Alvarez, C. (1988). Regional and laminar distribution
 1025 of the dopamine and serotonin innervation in the macaque cerebral cortex: a radioautographic study.
 1026 *J Comp Neurol* 273, 99-119.

1027 Bermudez, M.A., and Schultz, W. (2014). Timing in reward and decision processes. *Philos Trans R Soc*
1028 *Lond B Biol Sci* 369, 20120468.

1029 Bondar, I.V., Leopold, D.A., Richmond, B.J., Victor, J.D., and Logothetis, N.K. (2009). Long-term stability
1030 of visual pattern selective responses of monkey temporal lobe neurons. *PLoS One* 4, e8222.

1031 Bromberg-Martin, E.S., Matsumoto, M., and Hikosaka, O. (2010). Dopamine in motivational control:
1032 rewarding, aversive, and alerting. *Neuron* 68, 815-834.

1033 Brooks, S.J., Savov, V., Allzen, E., Benedict, C., Fredriksson, R., and Schioth, H.B. (2012). Exposure to
1034 subliminal arousing stimuli induces robust activation in the amygdala, hippocampus, anterior
1035 cingulate, insular cortex and primary visual cortex: a systematic meta-analysis of fMRI studies.
1036 *Neuroimage* 59, 2962-2973.

1037 Brown, R.M., Crane, A.M., and Goldman, P.S. (1979). Regional distribution of monoamines in the
1038 cerebral cortex and subcortical structures of the rhesus monkey: concentrations and in vivo synthesis
1039 rates. *Brain Res* 168, 133-150.

1040 Carcea, I., and Froemke, R.C. (2013). Cortical plasticity, excitatory-inhibitory balance, and sensory
1041 perception. *Prog Brain Res* 207, 65-90.

1042 Censor, N., Sagi, D., and Cohen, L.G. (2012). Common mechanisms of human perceptual and motor
1043 learning. *Nat Rev Neurosci* 13, 658-664.

1044 Chubykin, A.A., Roach, E.B., Bear, M.F., and Shuler, M.G. (2013). A cholinergic mechanism for reward
1045 timing within primary visual cortex. *Neuron* 77, 723-735.

1046 Cox, R.W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance
1047 neuroimages. *Computers and Biomedical Research* 29, 162-173.

1048 Deluca, C., Golzar, A., Santandrea, E., Lo Gerfo, E., Estocinova, J., Moretto, G., Fiaschi, A., Panzeri, M.,
1049 Mariotti, C., Tinazzi, M., and Chelazzi, L. (2014). The cerebellum and visual perceptual learning:
1050 evidence from a motion extrapolation task. *Cortex* 58, 52-71.

1051 Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L. (2009) ImageNet: A Large-Scale Hierarchical
1052 Image Database. *IEEE conference on computer vision and pattern recognition*, 248-255.

1053 Diaz, J.A., Queirazza, F., and Philiastrides, M.G. (2017). Perceptual learning alters post-sensory
 1054 processing in human decision-making. *Nat Hum Behav* 1, 1-9.
 1055 Doshier, B., and Lu, Z.L. (2017). Visual Perceptual Learning and Models. *Annual Review of Vision Science*
 1056 3, 343-363.
 1057 Fechner, G.T., Adler, H.E., Howes, D.H., and Boring, E.G. (1966). *Elements of psychophysics* (New York:
 1058 Holt, Rinehart & Winston).
 1059 Fisher, C., and Freiwald, W.A. (2015). Whole-agent selectivity within the macaque face-processing
 1060 system. *PNAS* 112, 14717-14722.
 1061 Fox, M.E., Mikhailova, M.A., Bass, C.E., Takmakov, P., Gainetdinov, R.R., Budygin, E.A., and Wightman,
 1062 R.M. (2016). Cross-hemispheric dopamine projections have functional significance. *PNAS* 113, 6985-
 1063 6990.
 1064 Francois, C., Percheron, G., and Yelnik, J. (1984). Localization of Nigrostriatal, Nigrothalamic and
 1065 Nigrotectal Neurons in Ventricular Coordinates in Macaques. *Neuroscience* 13, 61-76.
 1066 Froemke, R.C., Carcea, I., Barker, A.J., Yuan, K., Seybold, B.A., Martins, A.R., Zaika, N., Bernstein, H.,
 1067 Wachs, M., Levis, P.A., *et al.* (2013). Long-term modification of cortical synapses improves sensory
 1068 perception. *Nat Neurosci* 16, 79-88.
 1069 Gallistel, C.R., and Gibbon, J. (2000). Time, rate, and conditioning. *Psychol Rev* 107, 289-344.
 1070 Guggenmos, M., Rothkirch, M., Obermayer, K., Haynes, J.D., and Sterzer, P. (2015). A hippocampal
 1071 signature of perceptual learning in object recognition. *J Cogn Neurosci* 27, 787-797.
 1072 Gutnisky, D.A., Hansen, B.J., Iliescu, B.F., and Dragoi, V. (2009). Attention alters visual plasticity during
 1073 exposure-based learning. *Curr Biol* 19, 555-560.
 1074 Huberman, A.D., and Niell, C.M. (2011). What can mice tell us about how vision works? *Trends Neurosci*
 1075 34, 464-473.
 1076 Irvine, D.R.F. (2018). Auditory perceptual learning and changes in the conceptualization of auditory
 1077 cortex. *Hear Res* 366, 3-16.

1078 Jacob, S.N., and Nienborg, H. (2018). Monoaminergic Neuromodulation of Sensory Processing. *Front*
1079 *Neural Circuits* 12: 51.

1080 Jiang, Y., and He, S. (2006). Cortical responses to invisible faces: dissociating subsystems for facial-
1081 information processing. *Curr Biol* 16, 2023-2029.

1082 Kay, K.N., Rokem, A., Winawer, J., Dougherty, R.F., and Wandell, B.A. (2013). GLMdenoise: A fast,
1083 automated technique for denoising task-based fMRI data. *Frontiers in Neuroscience* 7, 247.

1084 Kilgard, M.P., and Merzenich, M.M. (1998). Cortical map reorganization enabled by nucleus basalis
1085 activity. *Science* 279, 1714-1718.

1086 Kim, D., Ling, S., and Watanabe, T. (2015). Dual mechanisms governing reward-driven perceptual
1087 learning. *F1000Res* 4, 764.

1088 Kolster, H., Mandeville, J.B., Arsenault, J.T., Ekstrom, L.B., Wald, L.L., and Vanduffel, W. (2009). Visual
1089 Field Map Clusters in Macaque Extrastriate Visual Cortex. *J Neurosci* 29, 7031-7039.

1090 Krizhevsky, A., Sutskever, I., and Hinton, G.E. (2012) ImageNet Classification with Deep Convolutional
1091 Neural Networks. *Advances in Neural Information Processing Systems* 25, 2.

1092 Kumar, S., and Vogels, R. (2019). Body Patches in Inferior Temporal Cortex Encode Categories with
1093 Different Temporal Dynamics. *J Cogn Neurosci* 31, 1699-1709.

1094 Leite, F.P., Tsao, D., Vanduffel, W., Fize, D., Sasaki, Y., Wald, L.L., Dale, A.M., Kwong, K.K., Orban, G.A.,
1095 Rosen, B.R., *et al.* (2002). Repeated fMRI using iron oxide contrast agent in awake, behaving macaques
1096 at 3 Tesla. *NeuroImage* 16, 283-294.

1097 Liu, C.H., Coleman, J.E., Davoudi, H., Zhang, K., and Hussain Shuler, M.G. (2015). Selective activation of
1098 a putative reinforcement signal conditions cued interval timing in primary visual cortex. *Curr Biol* 25,
1099 1551-1561.

1100 Maniglia, M., and Seitz, A.R. (2018). Towards a whole brain model of Perceptual Learning. *Curr Opin*
1101 *Behav Sci* 20, 47-55.

1102 Martins, A.R., and Froemke, R.C. (2015). Coordinated forms of noradrenergic plasticity in the locus
 1103 coeruleus and primary auditory cortex. *Nat Neurosci* 18, 1483-1492.
 1104 Maunsell, J.H. (2004). Neuronal representations of cognitive state: reward or attention? *Trends Cogn*
 1105 *Sci* 8, 261-265.
 1106 McMahon, D.B.T., Jones, A.P., Bondar, I.V., and Leopold, D.A. (2014). Face-selective neurons maintain
 1107 consistent visual responses across months. *PNAS* 111, 8251-8256.
 1108 Murris, S.R., Arsenault, J.T., and Vanduffel, W. (2020). Frequency- and State-Dependent Network
 1109 Effects of Electrical Stimulation Targeting the Ventral Tegmental Area in Macaques. *Cereb Cortex* 30,
 1110 4281-4296.
 1111 Nguyen, V.T., Sonkusare, S., Stadler, J., Hu, X., Breakspear, M., and Guo, C.C. (2017). Distinct Cerebellar
 1112 Contributions to Cognitive-Perceptual Dynamics During Natural Viewing. *Cereb Cortex* 27, 5652-5662.
 1113 Pek, J., Wong, O., and Wong, A.C.M. (2018). How to address non-normality: A taxonomy of approaches,
 1114 reviewed, and illustrated. *Frontiers in Psychology* 9, 2104.
 1115 Pinsk, M.A., Doniger, G.M., and Kastner, S. (2004). Push-pull mechanism of selective attention in
 1116 human extrastriate cortex. *Journal of Neurophysiology* 92, 622-629.
 1117 Popivanov, I.D., Jastorff, J., Vanduffel, W., and Vogels, R. (2012). Stimulus representations in body-
 1118 selective regions of the macaque cortex assessed with event-related fMRI. *NeuroImage* 63, 723-741.
 1119 Reed, A., Riley, J., Carraway, R., Carrasco, A., Perez, C., Jakkamsetti, V., and Kilgard, M.P. (2011). Cortical
 1120 map plasticity improves learning but is not necessary for improved performance. *Neuron* 70, 121-131.
 1121 Reveley, C., Gruslys, A., Ye, F.Q., Glen, D., Samaha, J., Russ, B.E., Saad, Z., Seth, A.K., Leopold, D.A., and
 1122 Saleem, K.S. (2017). Three-dimensional digital template atlas of the macaque brain. *Cerebral Cortex*
 1123 27, 4463-4477.
 1124 Roelfsema, P.R., van Ooyen, A., and Watanabe, T. (2010). Perceptual learning rules based on
 1125 reinforcers and attention. *Trends Cogn Sci* 14, 64-71.
 1126 Sagi, D. (2011). Perceptual learning in Vision Research. *Vision Res* 51, 1552-1566.

1127 Schiltz, C., Bodart, J.M., Dubois, S., DeJardin, S., Michel, C., Roucoux, A., Crommelinck, M., and Orban,
 1128 G.A. (1999). Neuronal mechanisms of perceptual learning: changes in human brain activity with
 1129 training in orientation discrimination. *Neuroimage* 9, 46-62.
 1130 Schoups, A., Vogels, R., Qian, N., and Orban, G. (2001). Practising orientation identification improves
 1131 orientation coding in V1 neurons. *Nature* 412, 549-553.
 1132 Seitz, A.R. (2017). Perceptual learning. *Curr Biol* 27, 631-636.
 1133 Seitz, A.R., Kim, D., and Watanabe, T. (2009). Rewards Evoke Learning of Unconsciously Processed
 1134 Visual Stimuli in Adult Humans. *Neuron* 61, 700-707.
 1135 Seitz, A.R., and Watanabe, T. (2003). Psychophysics: Is subliminal learning really passive? *Nature* 422,
 1136 36.
 1137 Seitz, A.R., and Watanabe, T. (2009). The phenomenon of task-irrelevant perceptual learning. *Vision*
 1138 *Res* 49, 2604-2610.
 1139 Serences, J.T. (2008). Value-based modulations in human visual cortex. *Neuron* 60, 1169-1181.
 1140 Shibata, K., Watanabe, T., Sasaki, Y., and Kawato, M. (2011). Perceptual learning incepted by decoded
 1141 fMRI neurofeedback without stimulus presentation. *Science* 334, 1413-1415.
 1142 Shibata, K., Yamagishi, N., Ishii, S., and Kawato, M. (2009). Boosting perceptual learning by fake
 1143 feedback. *Vision Res* 49, 2574-2585.
 1144 Shuler, M.G., and Bear, M.F. (2006). Reward timing in the primary visual cortex. *Science* 311, 1606-
 1145 1609.
 1146 Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E.J., Johansen-Berg, H.,
 1147 Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., *et al.* (2004). Advances in functional and
 1148 structural MR image analysis and implementation as FSL. *NeuroImage* 23, 208-219.
 1149 Sokolov, A.A., Erb, M., Grodd, W., and Pavlova, M.A. (2014). Structural loop between the cerebellum
 1150 and the superior temporal sulcus: evidence from diffusion tensor imaging. *Cereb Cortex* 24, 626-632.
 1151 Sokolov, A.A., Gharabaghi, A., Tatagiba, M.S., and Pavlova, M. (2010). Cerebellar engagement in an
 1152 action observation network. *Cereb Cortex* 20, 486-491.

1153 Sridharan, D., Steinmetz, N.A., Moore, T., and Knudsen, E.I. (2014). Distinguishing bias from sensitivity
 1154 effects in multialternative detection tasks. *J Vis* 14, 16.
 1155 Tan, Q., Wang, Z., Sasaki, Y., and Watanabe, T. (2019). Category-Induced Transfer of Visual Perceptual
 1156 Learning. *Curr Biol* 8, 1374-1378.
 1157 Tsao, D.Y., Moeller, S., and Freiwald, W.A. (2008). Comparing face patch systems in macaques and
 1158 humans. *PNAS* 105, 19514-19519.
 1159 Tsushima, Y., Sasaki, Y., and Watanabe, T. (2006). Greater disruption due to failure of inhibitory control
 1160 on an ambiguous distractor. *Science* 314, 1786-1788.
 1161 Tsushima, Y., Seitz, A.R., and Watanabe, T. (2008). Task-irrelevant learning occurs only when the
 1162 irrelevant feature is weak. *Curr Biol* 18, R516-517.
 1163 Vaina, L.M., Belliveau, J.W., des Roziers, E.B., and Zeffiro, T.A. (1998). Neural systems underlying
 1164 learning and representation of global motion. *PNAS* 95, 12657-12662.
 1165 Vanduffel, W., Fize, D., Mandeville, J.B., Nelissen, K., Van Hecke, P., Rosen, B.R., Tootell, R.B.H., and
 1166 Orban, G.A. (2001). Visual motion processing investigated using contrast agent-enhanced fMRI in
 1167 awake behaving monkeys. *Neuron* 32, 565-577.
 1168 Vogels, R. (2010). Mechanisms of visual perceptual learning in macaque visual cortex. *Top Cogn Sci* 2,
 1169 239-250.
 1170 Watanabe, M., Cromwell, H.C., Tremblay, L., Hollerman, J.R., Hikosaka, K., and Schultz, W. (2001a).
 1171 Behavioral reactions reflecting differential reward expectations in monkeys. *Exp Brain Res* 140, 511-
 1172 518.
 1173 Watanabe, T., Nanez, J.E., and Sasaki, Y. (2001b). Perceptual learning without perception. *Nature* 413,
 1174 844-848.
 1175 Watanabe, T., Nanez, J.E., Sr., Koyama, S., Mukai, I., Liederman, J., and Sasaki, Y. (2002). Greater
 1176 plasticity in lower-level than higher-level visual motion processing in a passive perceptual learning task.
 1177 *Nat Neurosci* 5, 1003-1009.

1178 Watanabe, T., and Sasaki, Y. (2015). Perceptual Learning: Toward a Comprehensive Theory. *Annu Rev*
 1179 *Psychol* 66, 197-221.
 1180 Wichmann, F.A., and Hill, N.J. (2001). The psychometric function: I. Fitting, sampling, and goodness of
 1181 fit. *Percept Psychophys* 63, 1293-1313.
 1182 Worsley, K.J., and Friston, K.J. (1995). Analysis of fMRI time-series revisited — Again. *NeuroImage* 2,
 1183 173-181.
 1184 Yang, T., and Maunsell, J.H. (2004). The effect of perceptual learning on neuronal responses in monkey
 1185 visual area V4. *J Neurosci* 24, 1617-1626.
 1186 Yovel, G., and Freiwald, W.A. (2013). Face recognition systems in monkey and human: are they the
 1187 same thing? *F1000Prime Rep* 5, 10.
 1188 Adab, H.Z., Popivanov, I.D., Vanduffel, W., and Vogels, R. (2014). Perceptual learning of simple stimuli
 1189 modifies stimulus representations in posterior inferior temporal cortex. *J Cogn Neurosci* 26, 2187-
 1190 2200.
 1191
 1192 Adab, H.Z., and Vogels, R. (2011). Practicing coarse orientation discrimination improves orientation
 1193 signals in macaque cortical area v4. *Curr Biol* 21, 1661-1666.
 1194
 1195 Ahissar, M., and Hochstein, S. (1993). Attentional control of early perceptual learning. *Proc Natl Acad*
 1196 *Sci U S A* 90, 5718-5722.
 1197
 1198 Anderson, B.A., Laurent, P.A., and Yantis, S. (2011). Value-driven attentional capture. *Proc Natl Acad*
 1199 *Sci U S A* 108, 10367-10371.
 1200
 1201 Arsenault, J.T., Rima, S., Stemmann, H., and Vanduffel, W. (2014). Role of the primate ventral
 1202 tegmental area in reinforcement and motivation. *Current Biology* 24, 1347-1353.
 1203
 1204 Arsenault, J.T., and Vanduffel, W. (2019). Ventral midbrain stimulation induces perceptual learning and
 1205 cortical plasticity in primates. *Nature Communications* 10, 1-13.
 1206
 1207 Avants, B.B., Tustison, N.J., Song, G., Cook, P.A., Klein, A., and Gee, J.C. (2011). A reproducible
 1208 evaluation of ANTs similarity metric performance in brain image registration. *NeuroImage*.
 1209
 1210 Bao, S., Chan, V.T., and Merzenich, M.M. (2001). Cortical remodelling induced by activity of ventral
 1211 tegmental dopamine neurons. *Nature* 412, 79-83.
 1212

1213 Baraduc, P., Duhamel, J.R., and Wirth, S. (2019). Schema cells in the macaque hippocampus. *Science*
1214 363, 635-639.

1215
1216 Bartra, O., McGuire, J.T., and Kable, J.W. (2013). The valuation system: a coordinate-based meta-
1217 analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage* 76,
1218 412-427.

1219
1220 Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R.H.B., Singmann, H., Dai, B., and Eigen, C.
1221 (2014). Package "lme4".

1222
1223 Baumann, O., Borra, R.J., Bower, J.M., Cullen, K.E., Habas, C., Ivry, R.B., Leggio, M., Mattingley, J.B.,
1224 Molinari, M., Moulton, E.A., *et al.* (2015). Consensus paper: the role of the cerebellum in perceptual
1225 processes. *Cerebellum* 14, 197-220.

1226
1227 Baumann, O., and Mattingley, J.B. (2010). Scaling of neural responses to visual and auditory motion in
1228 the human cerebellum. *J Neurosci* 30, 4489-4495.

1229
1230 Berger, B., Gaspar, P., and Verney, C. (1991). Dopaminergic innervation of the cerebral cortex:
1231 unexpected differences between rodents and primates. *Trends Neurosci* 14, 21-27.

1232
1233 Berger, B., Trottier, S., Verney, C., Gaspar, P., and Alvarez, C. (1988). Regional and laminar distribution
1234 of the dopamine and serotonin innervation in the macaque cerebral cortex: a radioautographic study.
1235 *J Comp Neurol* 273, 99-119.

1236
1237 Bermudez, M.A., and Schultz, W. (2014). Timing in reward and decision processes. *Philos Trans R Soc*
1238 *Lond B Biol Sci* 369, 20120468.

1239
1240 Bondar, I.V., Leopold, D.A., Richmond, B.J., Victor, J.D., and Logothetis, N.K. (2009). Long-term stability
1241 of visual pattern selective responses of monkey temporal lobe neurons. *PLoS One* 4, e8222.

1242
1243 Bromberg-Martin, E.S., Matsumoto, M., and Hikosaka, O. (2010). Dopamine in motivational control:
1244 rewarding, aversive, and alerting. *Neuron* 68, 815-834.

1245
1246 Brooks, S.J., Savov, V., Allzen, E., Benedict, C., Fredriksson, R., and Schioth, H.B. (2012). Exposure to
1247 subliminal arousing stimuli induces robust activation in the amygdala, hippocampus, anterior
1248 cingulate, insular cortex and primary visual cortex: a systematic meta-analysis of fMRI studies.
1249 *Neuroimage* 59, 2962-2973.

1250
1251 Brown, R.M., Crane, A.M., and Goldman, P.S. (1979). Regional distribution of monoamines in the
1252 cerebral cortex and subcortical structures of the rhesus monkey: concentrations and in vivo synthesis
1253 rates. *Brain Res* 168, 133-150.

1254
1255 Carcea, I., and Froemke, R.C. (2013). Cortical plasticity, excitatory-inhibitory balance, and sensory
1256 perception. *Prog Brain Res* 207, 65-90.

1257
1258 Cox, R.W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance
1259 neuroimages. Computers and Biomedical Research.

1260
1261 Diaz, J.A., Queirazza, F., and Philastides, M.G. (2017). Perceptual learning alters post-sensory
1262 processing in human decision-making. Nat Hum Behav 1.

1263
1264 Doshier, B., and Lu, Z.L. (2017). Visual Perceptual Learning and Models. Annu Rev Vis Sci 3, 343-363.

1265
1266 Fechner, G.T., Adler, H.E., Howes, D.H., and Boring, E.G. (1966). Elements of psychophysics (New York,:
1267 Holt).

1268
1269 Fisher, C., and Freiwald, W.A. (2015). Whole-agent selectivity within the macaque face-processing
1270 system. Proc Natl Acad Sci U S A 112, 14717-14722.

1271
1272 Francois, C., Percheron, G., and Yelnik, J. (1984). Localization of Nigrostriatal, Nigrothalamic and
1273 Nigrotectal Neurons in Ventricular Coordinates in Macaques. Neuroscience 13, 61-76.

1274
1275 Froemke, R.C., Carcea, I., Barker, A.J., Yuan, K., Seybold, B.A., Martins, A.R., Zaika, N., Bernstein, H.,
1276 Wachs, M., Levis, P.A., *et al.* (2013). Long-term modification of cortical synapses improves sensory
1277 perception. Nat Neurosci 16, 79-88.

1278
1279 Gallistel, C.R., and Gibbon, J. (2000). Time, rate, and conditioning. Psychol Rev 107, 289-344.

1280
1281 Guggenmos, M., Rothkirch, M., Obermayer, K., Haynes, J.D., and Sterzer, P. (2015). A hippocampal
1282 signature of perceptual learning in object recognition. J Cogn Neurosci 27, 787-797.

1283
1284 Gutnisky, D.A., Hansen, B.J., Iliescu, B.F., and Dragoi, V. (2009). Attention alters visual plasticity during
1285 exposure-based learning. Curr Biol 19, 555-560.

1286
1287 Huberman, A.D., and Niell, C.M. (2011). What can mice tell us about how vision works? Trends Neurosci
1288 34, 464-473.

1289
1290 Jacob, S.N., and Nienborg, H. (2018). Monoaminergic Neuromodulation of Sensory Processing. Front
1291 Neural Circuits 12, 51.

1292
1293 Jiang, Y., and He, S. (2006). Cortical responses to invisible faces: dissociating subsystems for facial-
1294 information processing. Curr Biol 16, 2023-2029.

1295
1296 Kay, K.N., Rokem, A., Winawer, J., Dougherty, R.F., and Wandell, B.A. (2013). GLMdenoise: A fast,
1297 automated technique for denoising task-based fMRI data. Frontiers in Neuroscience.

1298

1299 Kilgard, M.P., and Merzenich, M.M. (1998). Cortical map reorganization enabled by nucleus basalis
1300 activity. *Science* 279, 1714-1718.

1301
1302 Kim, D., Ling, S., and Watanabe, T. (2015). Dual mechanisms governing reward-driven perceptual
1303 learning. *F1000Res* 4, 764.

1304
1305 Kolster, H., Mandeville, J.B., Arsenault, J.T., Ekstrom, L.B., Wald, L.L., and Vanduffel, W. (2009). Visual
1306 Field Map Clusters in Macaque Extrastriate Visual Cortex. *Journal of Neuroscience* 29, 7031-7039.

1307
1308 Leite, F.P., Tsao, D., Vanduffel, W., Fize, D., Sasaki, Y., Wald, L.L., Dale, A.M., Kwong, K.K., Orban, G.A.,
1309 Rosen, B.R., *et al.* (2002). Repeated fMRI using iron oxide contrast agent in awake, behaving macaques
1310 at 3 Tesla. *NeuroImage* 16, 283-294.

1311
1312 Maniglia, M., and Seitz, A.R. (2018). Towards a whole brain model of Perceptual Learning. *Curr Opin*
1313 *Behav Sci* 20, 47-55.

1314
1315 Martins, A.R., and Froemke, R.C. (2015). Coordinated forms of noradrenergic plasticity in the locus
1316 coeruleus and primary auditory cortex. *Nat Neurosci* 18, 1483-1492.

1317
1318 Maunsell, J.H. (2004). Neuronal representations of cognitive state: reward or attention? *Trends Cogn*
1319 *Sci* 8, 261-265.

1320
1321 McMahon, D.B.T., Jones, A.P., Bondar, I.V., and Leopold, D.A. (2014). Face-selective neurons maintain
1322 consistent visual responses across months. *Proceedings of the National Academy of Sciences* 111,
1323 8251-8256.

1324
1325 Murris, S.R., Arsenault, J.T., and Vanduffel, W. (2020). Frequency- and State-Dependent Network
1326 Effects of Electrical Stimulation Targeting the Ventral Tegmental Area in Macaques. *Cereb Cortex*.

1327
1328 Nguyen, V.T., Sonkusare, S., Stadler, J., Hu, X., Breakspear, M., and Guo, C.C. (2017). Distinct Cerebellar
1329 Contributions to Cognitive-Perceptual Dynamics During Natural Viewing. *Cereb Cortex* 27, 5652-5662.

1330
1331 Pek, J., Wong, O., and Wong, A.C.M. (2018). How to address non-normality: A taxonomy of approaches,
1332 reviewed, and illustrated. In *Frontiers in Psychology*.

1333
1334 Popivanov, I.D., Jastorff, J., Vanduffel, W., and Vogels, R. (2012). Stimulus representations in body-
1335 selective regions of the macaque cortex assessed with event-related fMRI. *NeuroImage*.

1336
1337 Reveley, C., Gruslys, A., Ye, F.Q., Glen, D., Samaha, J., Russ, B.E., Saad, Z., Seth, A.K., Leopold, D.A., and
1338 Saleem, K.S. (2017). Three-dimensional digital template atlas of the macaque brain. *Cerebral Cortex*
1339 27, 4463-4477.

1340
1341 Roelfsema, P.R., van Ooyen, A., and Watanabe, T. (2010). Perceptual learning rules based on
1342 reinforcers and attention. In *Trends in Cognitive Sciences*.

1343
1344 Sagi, D. (2011). Perceptual learning in Vision Research. In Vision Research.

1345
1346 Schoups, A., Vogels, R., Qian, N., and Orban, G. (2001). Practising orientation identification improves
1347 orientation coding in V1 neurons. *Nature* 412, 549-553.

1348
1349 Seitz, A.R. (2017). Perceptual learning. In *Current Biology*.

1350
1351 Seitz, A.R., Kim, D., and Watanabe, T. (2009). Rewards Evoke Learning of Unconsciously Processed
1352 Visual Stimuli in Adult Humans. *Neuron*.

1353
1354 Seitz, A.R., and Watanabe, T. (2003). Psychophysics: Is subliminal learning really passive? *Nature* 422,
1355 36.

1356
1357 Seitz, A.R., and Watanabe, T. (2009). The phenomenon of task-irrelevant perceptual learning. *Vision*
1358 *Res* 49, 2604-2610.

1359
1360 Shibata, K., Watanabe, T., Sasaki, Y., and Kawato, M. (2011a). Perceptual learning incepted by decoded
1361 fMRI neurofeedback without stimulus presentation. *Science*.

1362
1363 Shibata, K., Watanabe, T., Sasaki, Y., and Kawato, M. (2011b). Perceptual Learning Incepted by
1364 Decoded fMRI Neurofeedback Without Stimulus Presentation. *Science* 334, 1413-1415.

1365
1366 Shibata, K., Yamagishi, N., Ishii, S., and Kawato, M. (2009). Boosting perceptual learning by fake
1367 feedback. *Vision Res* 49, 2574-2585.

1368
1369 Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E.J., Johansen-Berg, H.,
1370 Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., *et al.* (2004). Advances in functional and
1371 structural MR image analysis and implementation as FSL. In *NeuroImage*.

1372
1373 Sokolov, A.A., Erb, M., Grodd, W., and Pavlova, M.A. (2014). Structural loop between the cerebellum
1374 and the superior temporal sulcus: evidence from diffusion tensor imaging. *Cereb Cortex* 24, 626-632.

1375
1376 Sokolov, A.A., Gharabaghi, A., Tatagiba, M.S., and Pavlova, M. (2010). Cerebellar engagement in an
1377 action observation network. *Cereb Cortex* 20, 486-491.

1378
1379 Sridharan, D., Steinmetz, N.A., Moore, T., and Knudsen, E.I. (2014). Distinguishing bias from sensitivity
1380 effects in multialternative detection tasks. *J Vis* 14.

1381
1382 Tan, Q., Wang, Z., Sasaki, Y., and Watanabe, T. (2019). Category-Induced Transfer of Visual Perceptual
1383 Learning. *Current Biology*.

1384

1385 Tsao, D.Y., Moeller, S., and Freiwald, W.A. (2008). Comparing face patch systems in macaques and
1386 humans. *Proc Natl Acad Sci U S A* 105, 19514-19519.

1387
1388 Tsushima, Y., Seitz, A.R., and Watanabe, T. (2008). Task-irrelevant learning occurs only when the
1389 irrelevant feature is weak. *Curr Biol* 18, R516-517.

1390
1391 Vanduffel, W., Fize, D., Mandeville, J.B., Nelissen, K., Van Hecke, P., Rosen, B.R., Tootell, R.B.H., and
1392 Orban, G.A. (2001). Visual motion processing investigated using contrast agent-enhanced fMRI in
1393 awake behaving monkeys. *Neuron* 32, 565-577.

1394
1395 Vogels, R. (2010). Mechanisms of visual perceptual learning in macaque visual cortex. *Top Cogn Sci* 2,
1396 239-250.

1397
1398 Watanabe, M., Cromwell, H.C., Tremblay, L., Hollerman, J.R., Hikosaka, K., and Schultz, W. (2001a).
1399 Behavioral reactions reflecting differential reward expectations in monkeys. *Exp Brain Res* 140, 511-
1400 518.

1401
1402 Watanabe, T., Nanez, J.E., and Sasaki, Y. (2001b). Perceptual learning without perception. *Nature* 413,
1403 844-848.

1404
1405 Watanabe, T., Nanez, J.E., Sr., Koyama, S., Mukai, I., Liederman, J., and Sasaki, Y. (2002). Greater
1406 plasticity in lower-level than higher-level visual motion processing in a passive perceptual learning task.
1407 *Nat Neurosci* 5, 1003-1009.

1408
1409 Watanabe, T., and Sasaki, Y. (2015). Perceptual Learning: Toward a Comprehensive Theory. In SSRN.

1410
1411 Wichmann, F.A., and Hill, N.J. (2001). The psychometric function: I. Fitting, sampling, and goodness of
1412 fit. *Percept Psychophys* 63, 1293-1313.

1413
1414 Worsley, K.J., and Friston, K.J. (1995). Analysis of fMRI time-series revisited — Again. *NeuroImage* 2,
1415 173-181.

1416
1417 Yang, T., and Maunsell, J.H. (2004). The effect of perceptual learning on neuronal responses in monkey
1418 visual area V4. *J Neurosci* 24, 1617-1626.

1419
1420 Yovel, G., and Freiwald, W.A. (2013). Face recognition systems in monkey and human: are they the
1421 same thing? *F1000Prime Rep* 5, 10.

1422
1423 Zubair, M. Murris, S.M, Isa, K., Onoe, H., Koshimizu, Y., Kobayashi, K., Vanduffel, W., and Isa, T. (2021)
1424 Divergent whole brain projections from the ventral midbrain in macaques. *Cerebral Cortex* (In press).

1425

(C)

(A) Figure 2 **(C)** **(B)** **(D)**

Figure 3

Figure 5

Figure 6

Supplementary Material

Electrical stimulation of the macaque ventral tegmental area drives category-selective learning without attention. *Sjoerd R Murris, John T Arsenault, Rajani Raman, Rufin Vogels, Wim Vanduffel*

Figure S1. Overview of the original visual stimuli, before being imbedded in dynamic noise at different SNR levels. Related to STAR methods.

(A) The 50 body identities were first created through MakeHuman software (The MakeHuman team; USA) using the randomization option to alter the morphology of each body. In addition, we manually selected different postures for the bodies. We subsequently cropped the heads of the images and converted them to grayscale. Each of these body identities has six variants: three sizes (88x127 pixels, 101x145 pixels, 114x164 pixels) at two profile orientations (rightward orientation: mirror image of the default leftward image). **(B)** In a similar fashion, we generated 50 different face identities. These faces were first generated through FaceGen software (Singular Inversions Inc; Toronto, Canada), which has an inbuilt option to randomize face characteristics. As with the bodies we converted the images to grayscale and created six individual variants for each identity; three different sizes (88x127pixels, 101x145 pixels, 114x164 pixels) and two profile orientations. Altogether the stimuli set consisted of 50 (identities) x 3 (sizes) x 2 (orientations) = 300 images of bodies and 50 (identities) x 3 (sizes) x 2 (orientations) = 300 images of faces.

Figure S2. Detailed design of main experiment (Round 1: fMRI+behavior). Related to STAR methods.

The design of the main experiment for round 1 was comprised of 5 phases (listed chronologically; pre-association behavior, pre-association fMRI, cue-VTA-EM association, post-association fMRI, post-association behavior). The pre-association behavior consisted of a series of profile discrimination days. The pre-association fMRI phase consisted of a series of pre-association fMRI days. The cue-VTA-EM association phase began with a series of successive cue-VTA-EM association days after which cue-VTA-EM association sessions were intermixed with post-association fMRI and profile discrimination days. The post-association fMRI phases consisted of alternating days of cue-VTA-EM association and post-association fMRI. The post-association behavioral phases consisted of alternating days of cue-VTA-EM association and profile discrimination. The boxes to the right of the design indicate the number of days (n) acquired during each phase.

Figure S3. Detailed design of main experiment (Rounds: 2-5, behavior only). Related to STAR methods.

The design of the main experiment followed a hierarchical structure (from highest to lowest, right to left in figure; round -> phase -> day). Each round (top level of design) contained a pre-association behavioral phase, a cue-VTA-EM association phase and a post-association behavioral phase. The pre-association phase consisted of 5 days of profile discrimination. During rounds 2 through 5, the post-association behavioral phase of the previous round (light blue, dotted lines) was used as the current rounds of the pre-association behavioral phase. This was possible because the paired category and visual field were alternated between rounds. The cue-VTA-EM association phase consisted of 5 days of cue-VTA-EM association (red, lines) and the post-association behavioral phase consisted of 5 days of profile discrimination (purple, lines). These phases were intermixed such that cue-VTA-EM association and profile discrimination days were alternated between. In addition, every profile discrimination day was preceded by a cue-VTA-EM day. The figure depicts all of the data collection that contributes to round 3 (shown in its entirety) and how it relates to round 2.

Figure S4. Behavioral data (d' and reaction time) of individual monkeys and rounds. Related to Figure 3. (A) Mean face/body profile discrimination sensitivity (d') of all experimental rounds (1-5) and

animals (M1, M2). Changes in mean sensitivity before (pre) and after (post) are connected with lines of which the grey levels illustrate the SNR level. **(B)** Mean changes (post-pre, Δ) after cue-VTA-EM association for paired and control categories of each round performed by each monkey at all six tested SNR levels (0-5). Superimposed bold lines in green and red, respectively indicate an increase or decrease in sensitivity in the paired versus the control category over rounds and monkeys. **(C)** Mean reaction times for face/body profile discrimination of all experimental rounds (1-5) and animals (M1, M2). Changes in reaction time before (pre) and after (post) are connected with lines, that also illustrate the SNR levels. Note that the range and variability in reaction times differed substantially between the two animals. **(D)** Mean changes (post-pre, Δ) after cue-VTA-EM association per round separated by SNR level (0-5) and subject (M1, M2). Considering the profound difference in absolute reaction times between the two animals, changes over rounds here are displayed per subject. Superimposed bold lines in green and red respectively indicate increased and decreases in the mean reaction time for the paired versus the control category over rounds. C = control category, P = paired category.

Figure S5. Comparative profile orientation discrimination performance by an SVM classifiers using pixel gray levels or convolutional layer 1 (Alexnet) activations as input. Related to Figure 3. (A) Cross-validated mean accuracy of SVM classification performance for pixel gray levels at each of the tested SNRs for face and body images. In model I the classifier was trained for each SNR level separately. The inset shows the weights of the SVM hyperplane in pixel space at an example SNR level (SNR = 5%), clearly revealing low level, luminance information. (B) An SVM classifier applied to the same dataset but the images from all SNR levels (excluding 0% SNR) were used to train the classifier, which was subsequently tested on the separate SNR levels (Model II). (C) An SVM trained on layer 1 activations of the DCNN Alexnet, rendering similar results as for the pixel gray levels. As in (A) training separately at each SNR level. (D) SVM using layer 1 activations as input using the same approach as in (B) in which the classifier was trained on all SNR levels combined. Shaded areas correspond to 95% confidence intervals (bootstrapping: $n = 1000$ iterations).

Additional discussion: Although of less importance for our specific questions, the observed increase in sensitivity *within* a category also merits attention and calls into question what information the animals rely on when their perceptual abilities are selectively enhanced. One possibility is that low-level features are used, that differ consistently between right- and left oriented profiles within a specific category (e.g., contrast gradients). Secondly, the animals may rely on higher level information captured in the profiles. In an attempt to distinguish between these possibilities, we compared the performance to that of a machine learning algorithm applied on the same stimulus set presented to the monkeys (**Figure S5**, STAR Methods): a linear support vector machine (SVM) classifier based on (1) simple pixel-luminance values of the different stimuli, and (2) the output of the first layer of a deep convolutional neural network (DCNN, Alexnet). Contrary to the gradually rising psychometric curves of the monkeys at increasing SNR levels (**Figure 1D**), both classifiers show already close to perfect accuracy at an SNR level of 5%. The ability of a linear classifier to discriminate between right- and left profiles suggests that a simple linear weighting of low-level features is sufficient for high accuracy performance. Performance levels in the monkeys, in particular at low SNR levels, did not remotely approach that of the classifiers, which suggests that they may use also higher order cues for their *within*-category profile discrimination.

(A) Effect of VTA-EM-cue pairing (conjunction analysis)

(B) Face, body and object patches

Figure S6. Consistent category-selective fMRI signal changes after cue-VTA-EM association. Body, Face and object patches. Related to Figure 4.

(A) A second-level linear model was fit to the per run beta weights (M1 pre: n=49, post: n=63; M2 pre: n=51, M2 post: n=40) of the categorical response (paired vs. control category, see Methods) for each animal separately. A conjunction analysis was then performed to determine consistent modulations across animals. The effect of pairing (post-association vs. pre-association) on the categorical response is displayed on representative coronal slices in D99 space (conjunction, $p < 0.05$, cluster size 10 voxels).

(B) Face, body and object responsive temporal lobe patches were defined using an independent localizer experiment (Popivanov et al., 2012). Per run beta maps (M1: 54 runs, 3 days; M2: 38 runs, 2 days) were generated for contrasts to identify face (faces vs bodies/inanimate objects), body (bodies vs faces/inanimate objects) and object (inanimate objects vs faces/bodies) responsive regions. A 2nd level t-test was then performed across runs to define category responsive temporal lobe patches (FWE correction, $p < 0.05$). Here the category responsive patches are projected onto flattened cortical representations of the left and right hemisphere.

Related to Figure 1E: % correct							
Model: % correct ~cue_duration*distractor							
Effect	df	Sum Sq.	Mean Sq.	F-value	p	η_p^2	η_p^2 90% CI
monkey	1	0.05	0.05	2.75	0.10	0.01	(0.00, 0.03)
cue_duration	4	5.29	1.32	67.42	6.36x10⁻⁴³*	0.42	(0.35, 0.47)
distractor	2	0.01	0.01	0.26	0.77	0.00	(0.00, 0.01)
cue_duration*distractor	8	0.22	0.03	1.43	0.18	0.03	(0.00, 0.04)
residuals	374	7.33	0.02				
monkey	1	0.03	0.03	1.66	0.19	0.00	(0.00, 0.02)
cue_duration	4	6.11	1.53	77.82	1.79x10⁻⁴⁷*	0.46	(0.40, 0.51)
distractor	2	0.18	0.09	4.63	0.01*	0.03	(0.00, 0.05)
cue_duration*distractor	8	0.22	0.03	1.37	0.21	0.03	(0.00, 0.04)
residuals	374	7.05	0.02				

	Test	df	F-value	W	p
Parathresh. SNR	Shapiro-Wilk			1.00	0.80
	Levene's Test	2	1.77		0.17
High SNR	Shapiro-Wilk			0.99	0.25
	Levene's Test	2	2.08		0.13

Table S1. Overview of the ANOVAs applied to color discrimination performance during interference tests (cue parameter definition). Related to Figure 1.

ANOVA results from interference tests at parathreshold (top) and high SNR (bottom) when determining initial parathreshold cue parameters. η_p^2 refers to partial eta squared measure of effect size. η_p^2 90% CI refers to the 90% confidence intervals of the partial eta squared. Note the strong

significant effect of cue duration in both paradigms. The effect of category (i.e. faces/bodies against no cue) is only apparent for the high SNR, but not for parathreshold SNR cues. In addition, there is no significant effect when taking the individual monkeys as a factor. The lower panel contains normality tests for both paradigms. A Shapiro-Wilk test was used to test the normality of the residuals for each ANOVA. Levene's test was used to test for the homoscedasticity of each ANOVA.

Related to Figure 2C: % correct

Model: % correct ~cue_duration*distractor

Effect	Group		η_p^2	η_p^2 90% CI	M1	M2
	F	p			p	p
monkey	10.59	1*10^{-3*}	0.01	(0.00, 0.02)		
cue_duration	191.83	6x10^{-125*}	0.40	(0.37, 0.44)	2*10^{-81*}	1*10^{-41*}
distractor	0.32	0.73	0.00	(0.00, 0.00)	0.80	0.78
cue_duration*distractor	1.35	0.21	0.01	(0.00, 0.01)	0.75	0.91

Friedmans' Test	df	χ^2	p
cue_duration 0	2	4.78	0.09
cue_duration 1	2	0.27	0.87
cue_duration 2	2	1.18	0.55
cue_duration 3	2	4.20	0.12
cue_duration 4	2	3.91	0.14

Normality Tests	df	F-value	W	p
Shapiro-Wilk			1.00	0.13
Levene's Test	2	9.52		7.95 x 10^{-5*}

Table S2. ANOVA of color discrimination performance during interference tests (cue-VTA-EM association phase). Related to Figure 2.

ANOVA results from interference tests performed during the cue-VTA-EM association phase (at parathreshold SNR levels for face and body cues). η_p^2 refers to partial eta squared measure of effect size. η_p^2 90% CI refers to the 90% confidence intervals of the partial eta squared. Separate ANOVA outcomes for data from the two individual monkeys are provided in the right side of the table. In addition, a non-parametric Friedman's test was performed for each cue duration on the group data, resulting in no significant effects of category. χ^2 refers to the chi squared value from the Friedman's test. A Shapiro-Wilk test was used to test the normality of the residuals for each ANOVA. Levene's test was used to test for the homoscedasticity of each ANOVA.

Related to Figure 3A: d-prime

Model: $d\text{-prime} \sim \text{visual_field} + \text{pairing} \times \text{category} + \text{pairing} \times \text{visual_field} + (1|\text{round})$

SNR	Effect	Group		M1		M2	
		Param.	95% CI _{Bonf}	Param.	95% CI _{Bonf}	Param.	95% CI _{Bonf}
0	pairing	-0.02	(-0.13, 0.09)	0.08	(-0.08, 0.24)	-0.13	(-0.29, 0.03)
1		0.01	(-0.11, 0.12)	0.02	(-0.14, 0.17)	-0.00	(-0.18, 0.17)
2		0.21*	(0.06, 0.35)	0.34*	(0.15, 0.52)	0.06	(-0.17, 0.29)
3		0.33*	(0.17, 0.47)	0.32*	(0.13, 0.51)	0.34*	(0.11, 0.57)
4		0.14*	(0.00, 0.28)	0.11	(-0.09, 0.32)	0.16**	(-0.03, 0.34)
5		-0.11	(-0.25, 0.03)	-0.25*	(-0.42, -0.01)	0.02	(-0.16, 0.21)
0	pairing x category	0.03	(-0.10, 0.16)	-0.14	(-0.33, 0.04)	0.23*	(0.04, 0.41)
1		0.04	(-0.09, 0.19)	0.03	(-0.15, 0.21)	0.06	(-0.13, 0.26)
2		0.09	(-0.07, 0.27)	-0.24*	(-0.44, -0.03)	0.47*	(0.21, 0.74)
3		-0.04	(-0.20, 0.13)	-0.16	(-0.37, 0.06)	0.10	(-0.16, 0.37)
4		0.09	(-0.08, 0.25)	-0.02	(-0.25, 0.22)	0.20*	(0.00, 0.41)
5		0.41*	(0.25, 0.56)	0.52*	(0.29, 0.74)	0.28*	(0.07, 0.50)
0	pairing x visual_field	-0.01	(-0.14, 0.12)	-0.02	(-0.20, 0.16)	-0.01	(-0.20, 0.19)
1		-0.02	(-0.16, 0.11)	0.03	(-0.15, 0.20)	-0.08	(-0.27, 0.12)
2		0.08	(-0.09, 0.23)	0.07	(-0.14, 0.27)	0.08	(-0.18, 0.33)
3		0.13**	(-0.04, 0.30)	0.10	(-0.12, 0.31)	0.16	(-0.11, 0.43)
4		0.03	(-0.12, 0.19)	-0.02	(-0.25, 0.22)	0.07	(-0.13, 0.27)
5		0.01	(-0.14, 0.16)	0.04	(-0.27, 0.21)	0.08	(-0.14, 0.30)

Related to Figure 3B, left panel: d-prime, LVF

Model: $d\text{-prime} \sim \text{pairing} \times \text{category} + (1|\text{round})$

SNR	Effect	Group		M1		M2	
		Param.	95% CI _{Bonf}	Param.	95% CI _{Bonf}	Param.	95% CI _{Bonf}
0	pairing	0.01	(-0.13, 0.14)	0.13	(-0.04, 0.30)	-0.13	(-0.31, -0.05)
1		0.02	(-0.12, 0.15)	-0.01	(-0.17, 0.17)	0.04	(-0.15, 0.24)
2		0.30*	(0.15, 0.45)	0.43*	(0.22, 0.64)	0.14	(-0.10, 0.39)
3		0.36*	(0.18, 0.53)	0.36*	(0.12, 0.58)	0.36*	(0.10, 0.64)
4		0.16*	(0.01, 0.32)	0.17**	(-0.05, 0.38)	0.16**	(-0.05, 0.36)
5		-0.18*	(-0.35, -0.01)	-0.33*	(-0.57, -0.09)	-0.01	(-0.24, 0.23)
0	pairing x category	-0.03	(-0.20, 0.16)	-0.24*	(-0.50, -0.00)	0.22	(-0.04, 0.48)
1		0.02	(-0.16, 0.20)	0.09	(-0.33, 0.14)	-0.04	(-0.32, 0.23)
2		-0.07	(-0.29, 0.15)	-0.41*	(-0.70, -0.13)	0.32	(0.02, 0.66)
3		-0.05	(-0.28, 0.19)	-0.19	(-0.51, 0.14)	0.12	(-0.23, 0.50)
4		0.06	(-0.16, 0.26)	-0.09	(-0.39, 0.21)	0.23	(-0.06, 0.51)
5		0.56*	(0.33, 0.80)	0.74*	(0.40, 1.06)	0.36*	(0.04, 0.68)

Related to Figure 3B, right panel: d-prime, RVF

Model: $d\text{-prime} \sim \text{pairing} \times \text{category} + (1|\text{round})$

SNR	Effect	Group		M1		M2	
		Param.	95% CI _{Bonf}	Param.	95% CI _{Bonf}	Param.	95% CI _{Bonf}
0	pairing	-0.03	(-0.17, 0.10)	0.05	(-0.13, 0.22)	-0.13	(-0.33, 0.07)
1		0.02	(-0.12, 0.17)	0.02	(-0.17, 0.20)	0.02	(-0.18, 0.23)
2		0.04	(-0.13, 0.22)	0.17	(-0.05, 0.39)	-0.12	(-0.41, 0.19)
3		0.18*	(0.01, 0.33)	0.18	(-0.03, 0.41)	0.15	(-0.10, 0.40)
4		0.08	(-0.10, 0.25)	0.07	(-0.17, 0.32)	0.08	(-0.15, 0.29)
5		-0.05	(-0.21, 0.11)	-0.05	(-0.29, 0.16)	-0.04	(-0.28, 0.17)
0	pairing x category	0.09	(-0.10, 0.28)	-0.04	(-0.29, 0.20)	0.24	(-0.03, 0.53)
1		0.07	(-0.13, 0.27)	-0.03	(-0.30, 0.24)	0.17	(-0.10, 0.45)
2		0.25*	(0.01, 0.50)	-0.07	(-0.37, 0.24)	0.62*	(0.19, 1.01)
3		-0.03	(-0.25, 0.20)	-0.13	(-0.44, -0.17)	0.09	(-0.27, 0.44)
4		0.12	(-0.11, 0.35)	0.07	(-0.28, 0.41)	0.18	(-0.11, 0.48)
5		0.25*	(0.04, 0.47)	0.29**	(-0.02, 0.62)	0.20	(-0.10, 0.52)

Related to Figure 3C: reaction time

Model: $\text{reaction time} \sim \text{visual_field} + \text{pairing} \times \text{category} + (1|\text{round})$

SNR	Effect	Group		M1		M2	
		Param.	95% CI _{Bonf}	Param.	95% CI _{Bonf}	Param.	95% CI _{Bonf}
0	pairing	5.18*	(2.60, 7.73)	0.29	(-1.24, 1.74)	10.8*	(5.58, 16.4)
1		5.64*	(2.95, 8.31)	2.08*	(0.57, 3.64)	9.71*	(4.71, 15.4)
2		4.06*	(1.50, 6.65)	2.16*	(0.55, 3.71)	6.24*	(0.95, 11.2)
3		1.36	(-0.62, 3.34)	1.65	(-0.13, 3.42)	1.05	(-2.68, 4.82)
4		1.61*	(0.07, 3.14)	1.72*	(0.07, 3.34)	1.47	(-1.40, 4.31)
5		2.25*	(0.72, 3.60)	1.81*	(0.14, 3.36)	2.73*	(0.36, 5.15)
0	pairing x category	-1.28	(-4.54, 2.43)	-3.57*	(-5.68, -1.52)	1.78	(-5.51, 9.03)
1		-5.06*	(-8.77, -1.35)	-5.69*	(-7.74, -3.55)	-4.30	(-12.0, 3.25)
2		-6.92*	(-10.5, -3.38)	-5.43*	(-7.54, -3.30)	-8.62*	(-16.1, -1.47)
3		-1.68	(-4.43, 1.07)	-4.68*	(-7.01, -1.97)	1.57	(-3.60, 6.70)
4		-0.31	(-2.42, 1.82)	2.20	(-4.43, 0.07)	1.90	(-2.10, 5.93)
5		-0.20	(-2.00, 1.74)	-0.41	(-2.54, 1.73)	0.05	(-3.17, 3.24)

Related to Figure 3D: c-criterion

Model: $c\text{-criterion} \sim \text{visual_field} + \text{pairing} \times \text{category} + (1|\text{round})$

SNR	Effect	Group		M1		M2	
		Param.	95% CI _{Bonf}	Param.	95% CI _{Bonf}	Param.	95% CI _{Bonf}
0	pairing	-0.20*	(-0.30, -0.10)	-0.14**	(-0.29, 0.02)	-0.27*	(-0.39, -0.14)
1		-0.16*	(-0.25, -0.06)	-0.09**	(-0.23, 0.05)	-0.23*	(-0.11, -0.35)
2		-0.09*	(-0.17, -0.00)	-0.10**	(-0.22, 0.02)	-0.08**	(-0.20, 0.05)
3		-0.06**	(-0.14, 0.02)	-0.11**	(-0.22, 0.01)	-0.00	(-0.11, 0.10)
4		-0.03	(-0.10, 0.04)	-0.06	(-0.16, 0.05)	-0.01	(-0.08, 0.08)
5		-0.04	(-0.12, 0.04)	-0.03	(-0.16, 0.11)	-0.05	(-0.13, 0.03)

0	pairing	0.02	(-0.12, 0.17)	-0.01	(-0.22, 0.21)	0.06	(-0.12, 0.23)
1	x	0.02	(-0.10, 0.16)	-0.01	(-0.21, 0.18)	0.07	(-0.09, 0.23)
2	category	0.01	(-0.11, 0.12)	0.03	(-0.14, 0.20)	-0.01	(-0.17, 0.16)
3		-0.02	(-0.13, 0.09)	0.03	(-0.14, 0.19)	-0.08	(-0.23, 0.07)
4		-0.02	(-0.11, 0.08)	-0.03	(-0.18, 0.11)	-0.00	(-0.12, 0.11)
5		-0.02	(-0.13, 0.09)	-0.05	(-0.23, 0.14)	0.01	(-0.10, 0.12)

Related to Figure 3A-C: d-prime

Model: $d\text{-prime}_{\text{CONTROL}} \sim v f_{\text{abs}} + v f_{\text{stim}} + \text{pairing} + (1|\text{round})$

		Group		M1		M2	
SNR	Effect	Param.	95% CI _{Bonf}	Param.	95% CI _{Bonf}	Param.	95% CI _{Bonf}
0	pairing	0.01	(-0.14, 0.12)	0.13	(-0.06, 0.30)	-0.12	(-0.31, 0.08)
1		0.02	(-0.07, 0.15)	-0.04	(-0.23, 0.14)	0.08	(-0.12, 0.28)
2		0.31*	(0.15, 0.46)	0.35*	(0.14, 0.55)	0.15	(-0.10, 0.41)
3		0.37*	(0.21, 0.55)	0.35*	(0.11, 0.56)	0.36*	(0.10, 0.61)
4		0.19*	(0.02, 0.36)	0.18	(-0.07, 0.41)	0.23*	(0.02, 0.44)
5		-0.09	(-0.25, 0.07)	-0.18	(-0.29, 0.04)	0.06	(-0.15, 0.28)
0	visual_field (LVF)	-0.05	(-0.17, 0.08)	-0.02	(-0.20, 0.16)	-0.09	(-0.29, 0.10)
1		0.09	(-0.04, 0.22)	0.12	(-0.06, 0.32)	0.00	(-0.20, 0.20)
2		0.52*	(0.37, 0.67)	0.69*	(0.49, 0.89)	0.14	(-0.11, 0.39)
3		0.76*	(0.59, 0.93)	1.11*	(0.87, 1.33)	0.27*	(0.01, 0.52)
4		0.61*	(0.46, 0.78)	0.92*	(0.67, 1.17)	0.31*	(0.09, 0.53)
5		0.31*	(0.14, 0.47)	0.48*	(0.23, 0.71)	0.19	(-0.04, 0.42)
0	visual_field (stim)	-0.02	(-0.11, 0.08)	-0.01	(-0.14, 0.12)	-0.00	(-0.14, 0.13)
1		0.03	(-0.07, 0.12)	-0.02	(-0.15, 0.11)	0.10	(-0.04, 0.24)
2		0.15*	(0.04, 0.25)	-0.03	(-0.18, 0.12)	0.44*	(0.26, 0.61)
3		0.11	(-0.00, 0.24)	0.07	(-0.10, 0.24)	0.31*	(0.12, 0.48)
4		-0.03	(-0.14, 0.09)	0.07	(-0.10, 0.24)	-0.03	(-0.18, 0.13)
5		0.02	(-0.09, 0.14)	0.12	(-0.05, 0.29)	-0.05	(-0.21, 0.11)

Related to Figure 3A-C: d-prime

Model: $d\text{-prime}_{\text{PAIRED}} \sim v f_{\text{abs}} + v f_{\text{stim}} + \text{pairing} + (1|\text{round})$

		Group		M1		M2	
SNR	Effect	Param.	95% CI _{Bonf}	Param.	95% CI _{Bonf}	Param.	95% CI _{Bonf}
0	pairing	-0.01	(-0.15, 0.12)	-0.10	(-0.29, 0.09)	0.10	(-0.10, 0.30)
1		0.04	(-0.09, 0.19)	0.11	(-0.07, 0.30)	-0.01	(-0.21, 0.19)
2		0.20*	(0.04, 0.36)	0.08	(-0.11, 0.29)	0.44*	(0.20, 0.71)
3		0.25*	(0.08, 0.43)	0.15	(-0.07, 0.37)	0.44*	(0.17, 0.70)
4		0.17*	(0.01, 0.34)	0.06	(-0.18, 0.29)	0.30*	(0.09, 0.51)
5		0.28*	(0.12, 0.44)	0.27*	(0.05, 0.49)	0.27*	(0.04, 0.49)
0	visual_field (LVF)	-0.10	(-0.23, 0.03)	-0.09*	(-0.29, 0.10)	-0.08	(-0.28, 0.11)
1		0.08	(-0.04, 0.20)	0.24*	(0.05, 0.43)	-0.06	(-0.27, 0.14)
2		0.30*	(0.16, 0.46)	0.53*	(0.33, 0.73)	0.23	(-0.49, 0.02)
3		0.69*	(0.52, 0.85)	1.04*	(0.83, 1.27)	0.42*	(0.15, 0.67)
4		0.57*	(0.42, 0.73)	0.80*	(0.57, 1.04)	0.30*	(0.09, 0.52)
5		0.48*	(0.33, 0.63)	0.64*	(0.42, 0.86)	0.24*	(0.01, 0.46)
0	visual_field (stim)	0.02	(-0.07, 0.11)	0.05	(-0.09, 0.19)	-0.01	(-0.15, 0.13)
1		-0.06	(-0.15, 0.03)	0.03	(-0.10, 0.16)	-0.10	(-0.25, 0.04)
2		-0.24*	(-0.13, -0.35)	0.01	(-0.13, 0.16)	-0.47*	(-0.65, -0.30)

3	-0.19*	(-0.07, -0.31)	0.04	(-0.12, 0.19)	-0.35*	(-0.54, -0.16)
4	-0.03	(-0.13, 0.08)	0.02	(-0.15, 0.19)	0.00	(-0.14, 0.16)
5	-0.10*	(-0.21, 0.00)	-0.14*	(-0.31, -0.03)	0.02	(-0.13, 0.17)

Table S3. LME model results for profile task behavioral metrics. Related to Figure 3.

LME models were fit to d-prime data from both visual fields (Figure 3A), the LVF (Figure 3B, left) and the RVF (Figure 3B, right) separately at each SNR level. LME models were also fit for reaction times (Figure 3C) and the c-criterion (Figure 3D) at each SNR level. ‘Param.’ refers to the parameter estimate of the LME model. The complete factor labels are visual field (LVF), pairing (post), and category (paired). ‘95% CI_{Bonf}’ refers to the 95% confidence intervals of the parameter, Bonferroni corrected for multiple comparisons across tested SNR levels. LME models of each dataset are applied to joint data from both monkeys in a group analysis and in the two animals separately (M1 and M2). Values in bold with an added asterisk denote significant effects of interests, in which the conservative Bonferroni corrected bootstrapped confidence interval does not include zero. The double asterisk denotes interesting additional trends addressed in the results section.

Reversed LME models partitioned by category, exploring VF effects. Two LME models were fit to d-prime data from the control and paired category separately at each SNR level over rounds. ‘Param.’ refers to the parameter estimate of the LME model. The complete factor labels are pairing (post), visual_field_{abs} (LVF), and visual_field_{stim} (stim). Conventions as above.

Related to Figure S4B: β-values (fMRI)							
Category-selective model: β -value _{CAT} ~ pairing							
		M1			M2		
Hemi.	Patch Category	Patch Name	Param.	95% CI	Patch Name	Param.	95% CI
left	paired	ASB	-0.20	(-1.20, 0.72)	AM	0.41	(-0.61, 1.49)
		MSB	0.60	(-0.11, 1.39)	MF	-0.17	(-1.71, 1.39)
					ML	0.25	(-1.28, 1.71)
					PL	-0.63	(-1.47, 0.17)
	control	AF	0.39	(-0.64, 1.42)	ASB	1.40	(-0.40, 3.30)
		AM	0.51	(-0.58, 1.60)	MSB	-0.02	(-1.00, 1.09)
		MF	0.11	(-0.70, 0.99)			
		ML	0.32	(-0.79, 1.40)			
		PL	-0.00	(-0.78, 0.73)			
	object	Anterior	-0.41	(-1.47, 0.51)	Anterior	-0.18	(-1.43, 1.05)
		Middle fundus	-0.57	(-1.82, 0.76)	Middle fundus	-0.51	(-1.99, 0.98)
		Middle lateral	-0.37	(-1.30, 0.67)	Posterior	-0.21	(-1.42, 0.95)
		posterior	-0.23	(-1.10, 0.67)	Posterior 2	0.26	(-0.93, 1.33)
right	paired	ASB	0.84	(-0.50, 2.19)	AM	-0.05	(-0.99, 0.88)
		MSB	0.17	(-0.68, 1.06)	MF	-0.45	(-1.79, 0.71)
					ML	-0.15	(-1.22, 0.81)
					PL	-0.21	(-0.93, 0.44)

control	AF	0.33	(-0.54, 1.20)	ASB	0.20	(-1.51, 1.86)	
	AM	0.22	(-0.72, 1.19)	MSB	-0.16	(-1.26, 0.90)	
	MF	0.86	(-0.02, 1.75)				
	ML	0.10	(-0.63, 0.81)				
	PL	-0.22	(-1.02, 0.55)				
object	Anterior	-0.24	(-1.51, 0.90)	Anterior	-0.10	(-1.35, 1.18)	
	Middle fundus	0.39	(-0.88, 1.61)	Middle fundus	0.45	(-0.95, 1.63)	
	Middle lateral	0.39	(-0.79, 1.63)	Posterior	0.40	(-0.60, 1.37)	
	Posterior	0.01	(-0.75, 0.79)	Posterior 2	-0.20	(-1.30, 0.98)	
	Posterior 2	0.06	(-0.91, 1.07)				
Visual non-specific model: $\beta\text{-value}_{vis} \sim \text{pairing}$							
	M1			M2			
Hemi.	Patch Category	Patch Name	Param.	95% CI	Patch Name	Param.	95% CI
left	paired	ASB	0.03	(-0.71, 0.80)	AM	0.53	(-0.26, 1.34)
		MSB	0.21	(-0.30, 0.74)	MF	-0.18	(-1.30, 0.92)
					ML	0.38	(-0.64, 1.38)
					PL	0.56*	(0.01, 1.14)
control	AF	-1.23*	(-1.91, -0.47)	ASB	-2.33*	(-3.58, -1.06)	
	AM	0.54	(-0.13, 1.26)	MSB	-1.15*	(-1.83, -0.44)	
	MF	1.09*	(0.52, 1.67)				
	ML	-0.20	(-0.92, 0.59)				
	PL	0.42	(-0.17, 1.01)				
object	Anterior	-2.07*	(-2.74, -1.40)	Anterior	1.49*	(0.61, 2.41)	
	Middle fundus	0.58	(-0.33, 1.45)	Middle fundus	-0.20	(-1.20, 0.81)	
	Middle lateral	1.57*	(0.84, 2.31)	Posterior	0.66	(-0.07, 1.37)	
	Posterior	-0.17	(-0.78, 0.40)	Posterior 2	1.18*	(0.31, 2.04)	
right	paired	ASB	1.73*	(0.79, 2.60)	AM	-1.19*	(-1.79, -0.60)
		MSB	0.80*	(0.26, 1.35)	MF	1.55*	(0.52, 2.55)
					ML	1.36*	(0.64, 2.07)
					PL	0.01	(-0.46, 0.50)
control	AF	0.05	(-0.53, 0.65)	ASB	1.95*	(1.02, 3.02)	
	AL	-0.07	(-0.71, 0.59)	MSB	0.22	(-0.66, 1.07)	
	AM	-0.11	(-0.77, 0.49)				
	MF	0.95*	(0.42, 1.50)				
	ML	0.36	(-0.16, 0.86)				
PL	0.32	(-0.28, 0.88)					
object	Anterior	0.12	(-0.75, 0.96)	Anterior	-1.18*	(-2.28, -0.29)	
	Middle fundus	1.45*	(0.55, 2.33)	Middle fundus	1.42*	(0.51, 2.23)	
	Middle lateral	1.45*	(0.50, 2.30)	Posterior	-0.86*	(-1.49, -0.20)	
	Posterior	0.69*	(0.14, 1.19)	Posterior 2	0.48	(-0.38, 1.35)	
	Posterior 2	1.06*	(0.33, 1.74)				

Table S4. LM results of the effect of pairing on fMRI responses in category responsive patches. Related to Figure S4.

(top) An LM was fit to the mean per run beta weights of the categorical response ($\beta\text{-value}_{CAT}$: paired category vs control category) for each category responsive ROI. The effect of pairing on the categorical response is reported. **(bottom)** An LM was fit to the mean per run beta weights of the stimulus response ($\beta\text{-value}_{VIS}$: visual stimulus vs fixation) for each category responsive ROI. The effect of pairing on the stimulus response is reported. ROIs were defined using an independent localizer (Figure S4B, STAR Methods). For all tables: ‘Hemi.’ In the first column indicates the hemisphere in which the ROI is located. ‘patch category’ in the second column denotes the category (bodies, faces or objects) the patch responds to and whether it was paired with VTA-EM (paired or control). Patch name denotes the specific patch label in terms of anatomical positioning (Figure S4B). ‘Param.’ denotes the

parameter estimate for pairing (post). '95% CI' denotes the 95% confidence interval of the parameter as determined by bootstrapping. Values in bold with an added asterisk indicate either a significant increase (in green) or decrease (in red) after VTA-EM pairing ($p < 0.05$). The face and body patches are according to the widely accepted nomenclature as also indicated in Fig 4SB. The Object patches are given tentative (irrelevant) names, according to their position along IT cortex (anterior, posterior, fundus or lateral convexity of STS).

Related to Figure 5: β -values (fMRI)

(A) Category-selective model (**Group_{CAT}**): $\beta\text{-value}_{CAT} \sim \text{visual_field} + \text{orientation} + \text{monkey} + \text{hemisphere} + \text{pairing} \times \text{patch_category} + \text{pairing} \times \text{hemisphere}$

(B) Visual non-specific model (**Group_{VIS}**): $\beta\text{-value}_{VIS} \sim \text{visual_field} + \text{orientation} + \text{category} + \text{monkey} + \text{hemisphere} + \text{pairing} \times \text{patch_category} + \text{pairing} \times \text{hemisphere}$

Effect	Group _{CAT}		Group _{VIS}	
	Param.	95% CI	Param.	95% CI
visual_field	0.10	(-0.12, 0.31)	0.06	(-0.09, 0.21)
Orientation	0.14	(-0.07, 0.35)	0.04	(-0.10, 0.19)
Monkey	-0.12	(-0.33, 0.09)	0.00	(-0.15, 0.15)
category (paired)			0.12	(-0.02, 0.27)
Hemisphere	0.08	(-0.24, 0.39)	0.28	(0.06, 0.50)
Pairing	0.28	(-0.18, 0.74)	-0.33	(-0.66, -0.02)
patch_category (type)	-0.10	(-0.48, 0.29)	0.15	(-0.12, 0.42)
patch_category (paired)	0.04	(-0.38, 0.44)	0.25	(-0.02, 0.53)
pairing x patch_category (type)	-0.38	(-0.90, 0.15)	0.43*	(0.06, 0.80)
pairing x patch_category (paired)	-0.15	(-0.70, 0.41)	0.52*	(0.14, 0.90)
pairing x hemisphere	0.01	(-0.42, 0.42)	0.65*	(0.36, 0.95)

(C) Visual non-specific model: significant group effects in individual animals

Effect	M1		M2	
	Param.	95% CI	Param.	95% CI
pairing*patch_category (type)	0.20	(-0.20, 0.60)	0.70*	(0.08, 1.31)
pairing*patch_category (paired)	0.42	(-0.01, 0.88)	0.70*	(0.11, 1.28)
pairing*hemisphere	0.69*	(0.34, 1.06)	0.71*	(0.24, 1.16)

(D) Underlying statistics (beta-values) for the individual animals in Figure 5B and 5C

		M1			M2		
Patch	Hemisphere	$\Delta\beta$	SE	p	$\Delta\beta$	SE	P
Paired	Both	0.69	0.18	1*10⁻⁴	0.38	0.15	0.01
	Left	0.12	0.14	0.62	0.32	0.23	0.15

	Right	1.26	0.12	3*10⁻⁶	0.43	0.19	0.02
Control	Both	0.27	0.09	0.004	-0.33	0.25	0.21
	Left	0.28	0.14	0.04	-1.74	0.37	3*10⁻⁸
	Right	0.25	0.12	0.04	1.09	0.35	0.002
Objects	Both	0.52	0.13	7*10⁻⁵	0.37	0.16	0.02
	Left	-0.02	0.19	0.91	0.78	0.23	7*10⁻⁴
	Right	0.95	0.18	1*10⁻⁷	-0.04	0.22	0.87

Table S5. LM results of the effect of pairing within category responsive patch networks. Related to Figure 5.

(A) For each network of patches (face, body and object) the mean per run beta weight of the categorical response was calculated for all combinations of visual field and profile orientation. For the group level analysis, data from patches responding to the paired category (M1: Body, M2: Face) were combined and those representing the control category (M1: Face, M2: Body) were also combined. An LM was then fit to per run beta weight of the categorical response (paired vs. control category). **(B)** For each network of patches the mean per run beta weight of the stimulus response was calculated for all combinations of visual field, orientation and category. For the group level analysis, data from patches responding to the paired category (M1: Body, M2: Face) were combined and those representing the control category (M1: Face, M2: Body) were also combined. An LM was then fit to per run beta weight of the categorical response (paired vs. control category). For both models, significance was determined using the bootstrapped 95% confidence intervals of the parameters ($p < 0.05$). For both models, the effects of interests were the interactions of pairing and patch category and the interaction pairing and patch hemisphere; as indicated by values in bold with an asterisk. **(C)** For all significant effects found in the group analysis, the results of the individual animals are shown. 'Param.' refers to the parameter estimate from the LME. '95% CI' refers to the bootstrapped 95% confidence intervals. **(D)** Mean differences in beta weights for each of the network patches (paired, control or object) in the two individual animals, underlying the data points in **Figure 5C**. The standard error (SE) was calculated based on a summation of the standard error before (pre) and after (post) VTA-EM-pairing: $\Delta SE = \sqrt{SE_{PRE}^2/n_{PRE} + SE_{POST}^2/n_{POST}}$. Two-sided student's t-tests reveal to what extent the differences in beta weights are significantly different from zero. Significant p-values that survive a Bonferonni correction for multiple comparisons are highlighted in bold.

(A)

tSNR levels between hemispheres (<i>present study</i>)						
<i>Pre</i>	M _{left}	SD _{left}	M _{right}	SD _{right}	p-value	Δ M _{left- M} _{right}
Monkey D (49 runs)	26.52	1.61	27.08	2.01	0.34	-0.56
Monkey T (63 runs)	22.05	1.40	21.96	1.53	0.73	0.09
<i>Post</i>	M _{left}	SD _{left}	M _{right}	SD _{right}	p-value	Δ M _{left- M} _{right}
Monkey D (51 runs)	21.95	1.40	21.56	1.07	0.14	0.39
Monkey T (40 runs)	21.00	2.42	20.72	1.56	0.22	0.29

(B)

tSNR levels between hemispheres (<i>Murris et al, 2020: Supplement 3</i>)						
<i>Awake</i>	M _{left}	SD _{left}	M _{right}	SD _{right}	p-value*	Δ M _{left- M} _{right}
Monkey D (60 runs)	32.71	3.27	25.36	1.51	$1.63 * 10^{-11}$	7.35
Monkey T (60 runs)	25.54	2.32	21.99	1.94	$1.63 * 10^{-11}$	3.54
<i>Anesthetized</i>	M _{left}	SD _{left}	M _{right}	SD _{right}	p-value*	Δ M _{left- M} _{right}
Monkey D (30 runs)	26.91	1.77	26.18	2.30	0.0082	0.73
Monkey T (30 runs)	27.12	2.39	25.82	3.04	$1.73 * 10^{-6}$	1.30

Table S6 (related to Figure 4-5): A quantification of inter-hemispheric differences of the temporal signal-to-noise ratio (tSNR) in our present study as compared to a previous VTA-EM-fMRI study (Murris et al, 2020). **(A)** Means (M) and standard deviations (SD) of tSNR levels across runs are given for both monkeys in the fMRI sessions prior to (Pre) and after (Post) VTA-EM association. Subscript notations of ‘left’ or ‘right’ respectively indicate tSNR for the contra- and ipsilateral hemisphere to the stimulation electrode. P-values in the sixth column are obtained from a Wilcoxon Signed Rank Test, indicating that none of the interhemispheric differences in the current study are significant ($p>0.05$). In addition, the right most column shows how small the differences in tSNR between hemispheres are. **(B)** Conventions as in (A), except that sessions from VTA-EM-fMRI in awake and anesthetized animals are being compared here. The inter-hemispheric differences are considerably stronger compared to those in the present study, in particular those of the awake fMRI data, as evidenced by significant p-values from the Wilcoxon Signed Rank Test ($p<<0.05$).

[Click here to access/download](#)

Supplemental Videos and Spreadsheets
FaceBody_profile_1280x720.mp4

[Click here to access/download](#)

Supplemental Videos and Spreadsheets
FaceBody_color_1280x720.mp4