Layer-specific ultra-high field 7T fMRI showing that threat anticipation is mediated by the pulvinar input to the superficial layers of primary visual cortex (V1)

Ai Koizumi[1,2]*, Minye Zhan[3], Hiroshi Ban[1,2], Ikuhiro Kida[1,2], Federico de Martino[3], Beatrice de Gelder[3]*, Kaoru Amano[1,2]

1. Center for Information and Neural Networks (CiNet), National Institute of Information and Communications Technology (NICT), Address: 1-4 Yamadaoka, Suita City, Osaka, 565-0871, Japan
2. Graduate School of Frontier Biosciences, Osaka University, Address: 1-3 Yamadaoka, Suita City, Osaka, 565-0871, Japan
3. Department of Cognitive Neuroscience, Maastricht University, Address: Oxfordlaan 55, 6229 EV Maastricht, The Netherlands
*Corresponding to: Ai Koizumi (bellkoizumi@gmail.com) & Beatrice de Gelder (b.degelder@maastrichtuniversity.nl)

Supporting Information

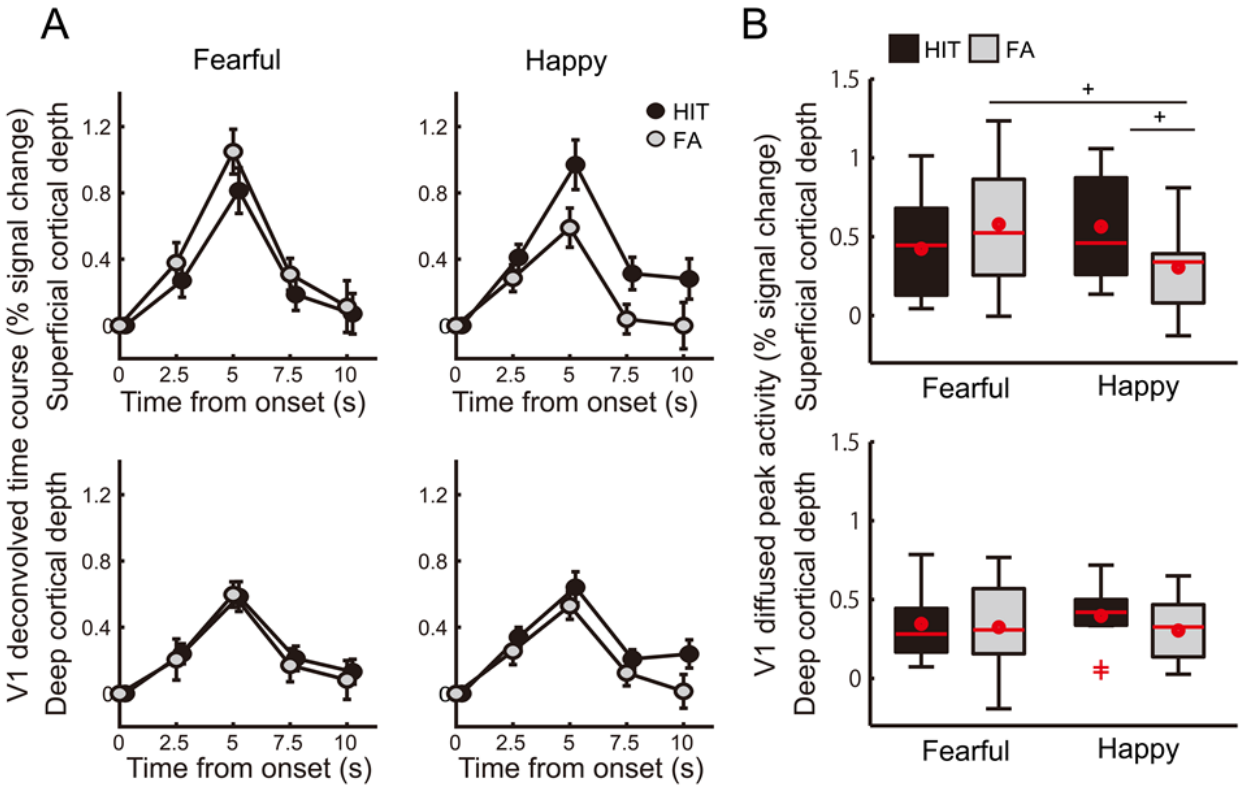

Fig S1. V1 activity at superficial (top row) and deep cortical depths (bottom row). **A**. Deconvolved time course of V1 for the HIT and false alarm (FA) trials during the fearful face detection task and happy face detection task (left and right panels, respectively). Error bar indicates standard error of mean. **B**. V1 temporal peak activity, which was estimated by averaging the time points across a more diffuse time window (2.5 to 7.5 s from the target face onsets) than the estimates described in the Main text (single time point at 5 s). Importantly, the results with this temporally diffused peak activity were qualitatively similar to those described in the Main text, and there was a significant second-order interaction between percept type, emotion, and cortical depth ($F(1,10) = 6.896$, $p = 0.025$). Box plot shows upper (75%) and lower (25%) quartiles with median (red line) and mean (red dot), with whisker showing maximum and minimum value. Outliers (outside of ± 2.7 standard deviations within a distribution for a given condition) are shown with red crosses. +$p < .10$. Related to **Fig 2**.

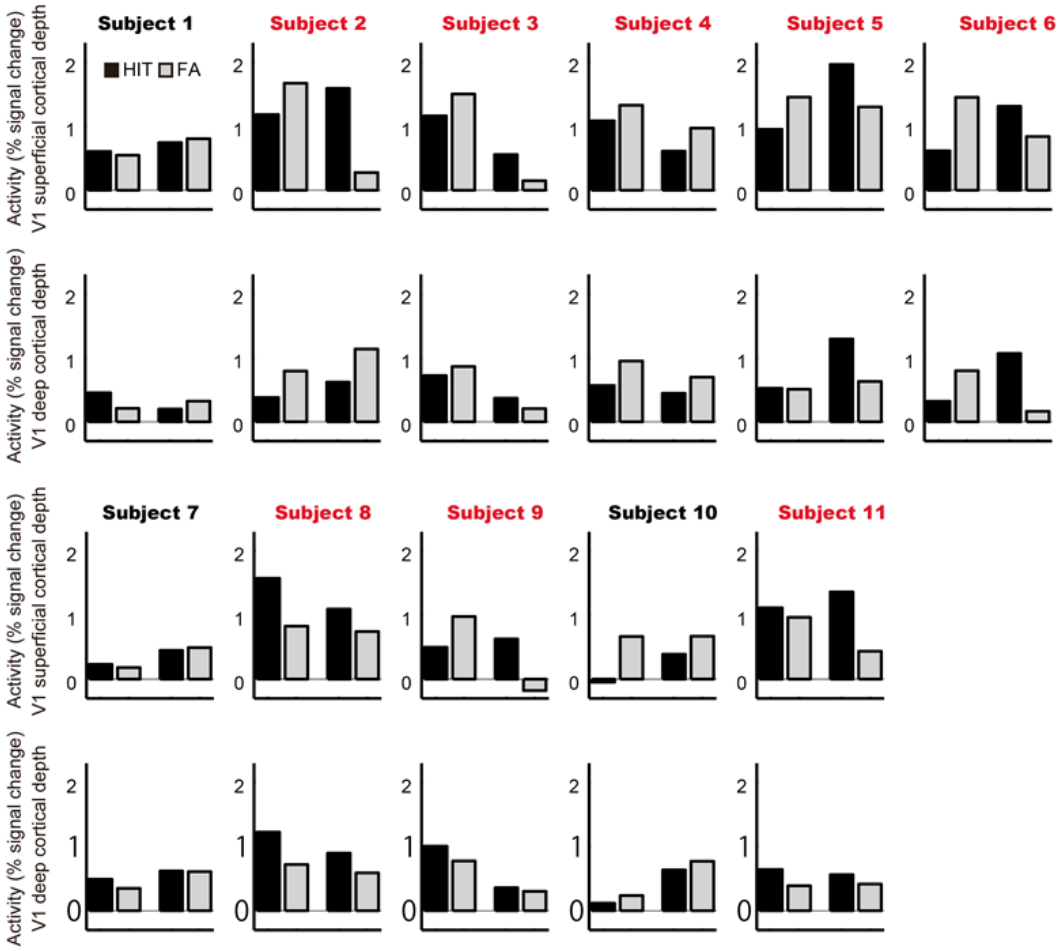

Fig S2. V1 activity in each participant at superficial (top row) and deep cortical depths (bottom row). For each subject, deconvolved time course of V1 during the HIT and false alarm (FA) trials during the fearful face detection task and happy face detection task (left and right panels, respectively) are shown. Subjects who showed numerically larger activity on FA trials of a fearful face relative to FA trials of a happy face, which is in line with the pattern of group level result shown in **Fig 2C**, are highlighted with red fonts (8 out of 11 participants). Related to **Fig 2**.

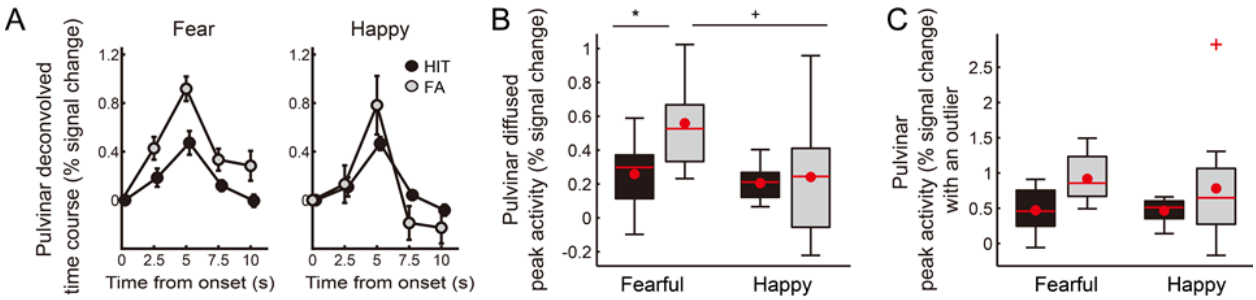

Fig S3. Pulvinar activity. **A**. Deconvolved time course for the pulvinar for HIT and false alarm (FA) trials during the detection tasks with fearful and happy face targets (left and right panels, respectively). Error bar indicates standard error of mean. **B**. Similar to the analysis of V1 activity (**Fig S1B**), mean temporal peak activity of the pulvinar was estimated with a more diffuse time window (2.5 to 7.5 s from the target face onsets) than the estimates described in the Main text (5 s). With this more diffuse time window, a repeated-measures ANOVA revealed a significant main effect of percept type (HIT/FA) ($F(1, 10) = 8.089$, $p = 0.017$) and of emotion (fearful/happy) ($F(1,10) = 7.041$, $p = 0.024$), which did not significantly interact with each other ($F(1, 10) = 1.835$, $p = 0.205$). Post-hoc t-tests were conducted for demonstrative purposes and revealed that the diffuse peak activity of the pulvinar was significantly larger on FA trials than on HIT trials of fearful faces ($t(10) = - 2.503$, $p = 0.031$, $d = 0.5$), while such activity did not significantly differ between HIT and FA trials of happy faces ($t(10) = - 0.339$, $p = 0.741$, n.s.). However, this difference between the tasks should be treated as indicative, because the interaction in the omnibus ANOVA was non-significant. Here, data from all participants were included as there was no longer an extreme outlier when peak activity was estimated with the diffused time window. **C**. Instead of the diffused time window shown in B, mean temporal peak activity of the pulvinar here was estimated with the non-diffused time window (5 s from the target face onsets) as in the Main text (**Fig 3**). However, unlike **Fig 3** in the Main text, the results shown here include a participant with an extreme outlier (shown in red cross) whose activity on FA of a happy face was more than 5 standard deviations above the group mean across conditions. The results of ANOVA with this outlier participant remained qualitatively similar to that reported in the Main text. That is, the pulvinar showed enhanced activity in FA trials relative to HIT trials, especially during the fearful face detection task (fearful face, $t(10) = -2.944$, $p = 0.015$, $d = 0.8$; happy face, $t(10) = -1.230$, $p = 0.247$, n.s.), although there was no significant interaction between percept type and emotion ($F(1,10) = 0.155$, $p = 0.702$). There was no significant difference between FA trials of the fearful face detection task and that of the happy face detection tasks. In B and C, box plot shows upper (75%) and lower (25%) quartiles with median (red line) and mean (red dot), with whisker showing maximum and minimum value. *$p < 0.05$, +$p < 0.10$ in planned post-hoc tests. Related to **Fig 3**.

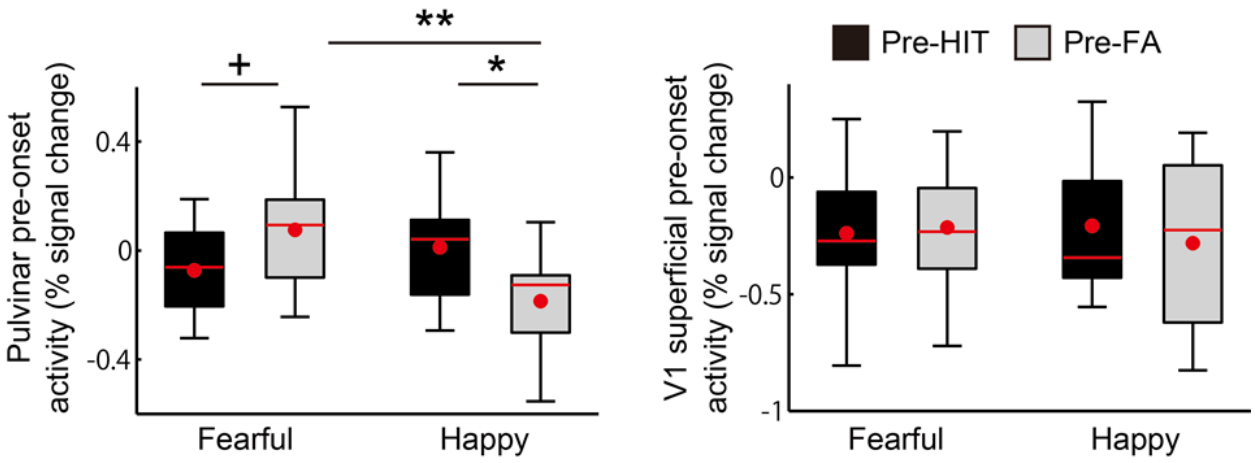

Fig S4. Prior to the onsets of target faces, there was already enhanced activity in the pulvinar preceding the false alarm (FA) trial of a fearful face compared with a happy face (left panel). A repeated measures ANOVA with two factors of percept type (pre-HIT/pre-FA) and emotion (fearful/happy) revealed a significant interaction ($F(1,10) = 10.094$, $p = 0.010$). Post-hoc analyses revealed that pulvinar activity was significantly greater on FA trials of fearful faces compared with happy faces in the pre-onset period ($t(10) = 3.392$, $p = 0.007$, $d = 1.1$). During the same pre-onset period, while pulvinar activity was significantly larger on HIT than on FA trials of happy faces ($t(10) = 2.386$, $p = 0.038$, $d = 0.7$), there was a non-significant opposite trend such that activity was relatively greater on FA than on HIT trials of fearful faces ($t(10) = -2.147$, $p = 0.057$, n.s.). Meanwhile, in the same pre-onset period, there was not yet any differential activity in the V1 superficial cortical depth (right panel), with no significant main effects or interaction between percept type and emotion ($ps > 0.50$, n.s.). These results suggest that anticipatory activity in the pulvinar modulated the response of the V1 superficial cortical depth triggered by the subsequent onset of a face target shown in **Fig 2D**. Box plot shows upper (75%) and lower (25%) quartiles with median (red line) and mean (red dot), with whisker showing maximum and minimum value. $^{**}p < 0.01$, $^{*}p < 0.05$, $^{+}p < 0.10$. Related to **Figs 2D and 4**.

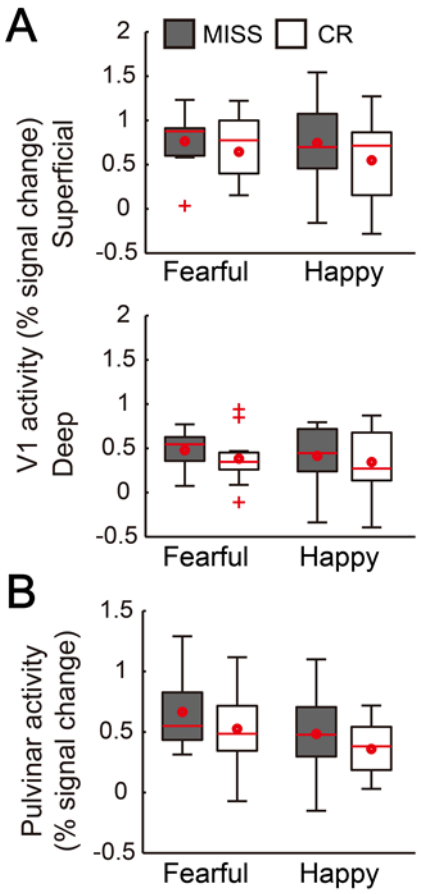

Fig S5. The results of a control analysis showing no differential activity in the pulvinar and V1 on MISS versus Correct rejection (CR) trials. We examined whether similar results on FA trials of fearful faces (**Figs 2D and 3B**) may be present on MISS trials, where a fearful face was presented but was not detected. While the results for FA trials of fearful faces may reflect enhanced top-down processing, other non-mutually exclusive possibilities are worth considering. Specifically, the results for FA trials may reflect a mere mismatch between sensory input and reported percept. Although this possibility is unlikely considering the fact that such results did not generalise to FA trials of happy faces, there remains a possibility that such input-to-percept mismatch may evoke certain neural activity only in anticipation of threat cues. We therefore examined whether the patterns of results we obtained from the contrast between FA and HIT trials may be also obtained from the contrast between MISS and correct rejection (CR) trials, where neutral faces were perceived with or without mismatched sensory input, respectively. **A.** Unlike the contrast of V1 activity between the HIT and false alarm (FA) trials shown in **Fig 2**, there was no differential activity in V1 between the MISS and CR trials, regardless of emotion and cortical depth (a non-significant second-order interaction: $F(1,10) = 1.251$, $p = 0.290$). While there was generally greater activity on MISS trials relative to CR trials that was not specific to any facial emotion and potentially reflected sensory inputs of salient emotional faces albeit undetected, there was no significant main effect of percept type ($F(1,10) = 2.573$, $p = 0.140$). **B.** Likewise, there was no difference in pulvinar activity

between the MISS and CR trials regardless of the facial emotion (a non-significant interaction between percept type (MISS/CR) and emotion: $F(1,10) = 0.013$, $p = 0.912$). Although we saw a trend for pulvinar activity to be larger for the fearful face detection task than for the happy face detection task, there was no significant main effect of emotion ($F(1,10) = 2.365$, $p = 0.155$). Box plot shows upper (75%) and lower (25%) quartiles with median (red line) and mean (red dot), with whisker showing maximum and minimum value. An outlier (outside of ± 2.7 standard deviations within a distribution for a given condition) is shown with a red cross. Related to **Figs 2D** and **3B**.

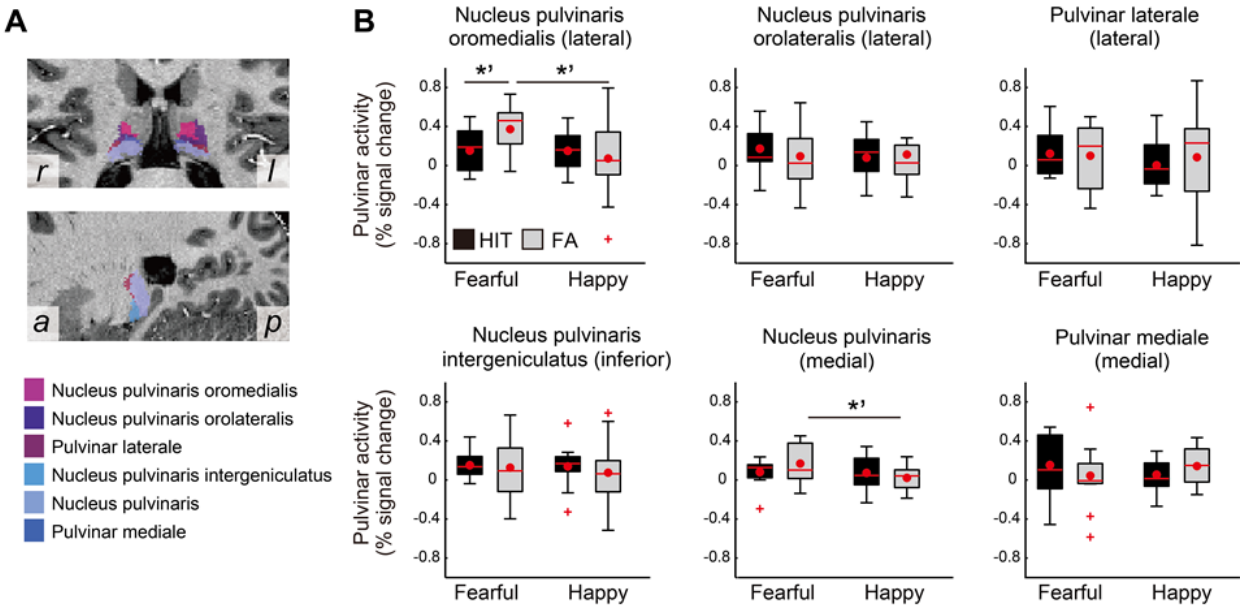

Fig S6. Activity of the pulvinar on FA and HIT trials in each of its subregions. **A.** The subregions composing the lateral, inferior, and medial portions of the pulvinar were defined based on a histological atlas (Chakravarty et al., 2006). **B**. Similarly to the results in the Main text (**Fig 3**) with the functionally defined pulvinar ROIs, there was significantly larger activity on the FA trials of a fearful face relative to the FA trials of a happy face in two subregions, namely the nucleus pulvinaris oromedialis (lateral) and nucleus pulvinaris (medial). However, there was no significant interaction between emotion (fearful/happy) and condition (HIT/FA) in these two subregions (nucleus pulvinaris oromedialis: $F(1,10) = 4.496$, $p = 0.060$; nucleus pulvinaris: $F(1,10) = 2.327$, $p = 0.158$). Box plot shows upper (75%) and lower (25%) quartiles with median (red line) and mean (red dot), with whisker showing maximum and minimum value. An outlier (outside of ± 2.7 standard deviations within a distribution for a given condition) is shown with a red cross. *$p < 0.05$ in planned post-hoc tests. r: right, l: left, a: anterior, p: posterior. Related to **Fig 3**.

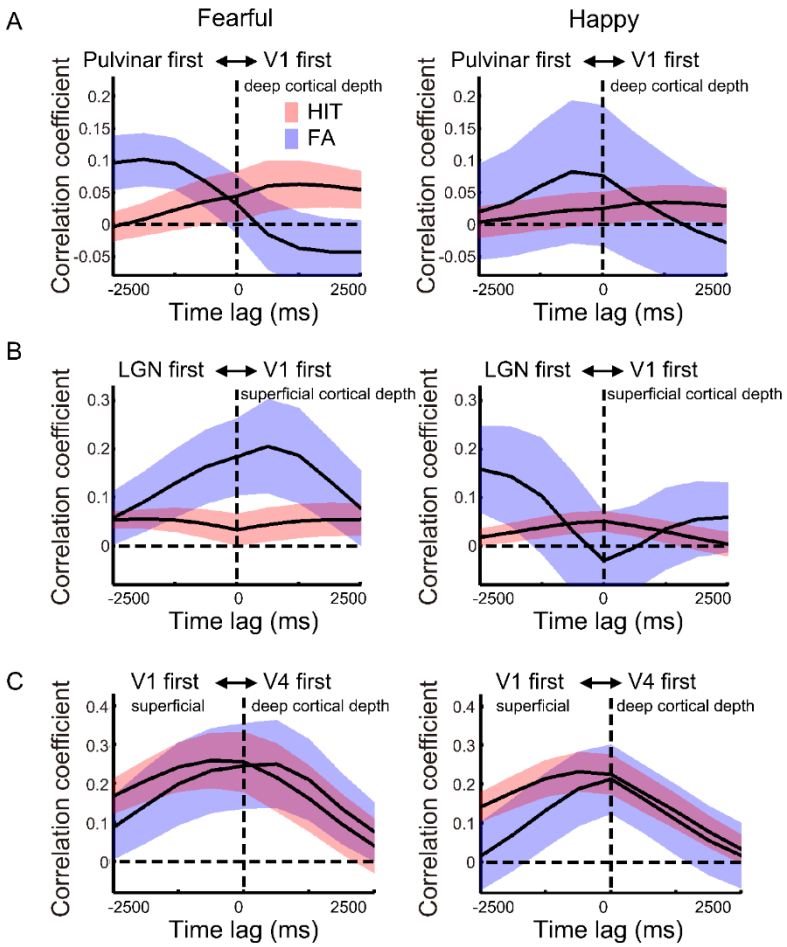

Fig S7. The results of control cross correlation analyses on the HIT and false alarm (FA) trials of the fearful face detection task (left panels) and the happy face detection task (right panels). **A.** The time-shifted cross correlation between the time courses for the pulvinar and V1 deep cortical depth. **B.** The correlation between the lateral geniculate nucleus (LGN) and V1 superficial cortical depth. **C.** The correlation between the V4 deep cortical depth and V1 superficial cortical depth. Although the V1 superficial layer (L1) is known to receive feedback from the deep layer (L6) of higher visual areas (Shipp, 2003, Shipp 2007), there was no increase in correlation between the V4 deep cortical depth and V1 superficial cortical depth. There was no significant difference between the HIT and FA trials in any of these paired brain areas. Colored area indicates standard error of mean. Related to **Fig 4**.

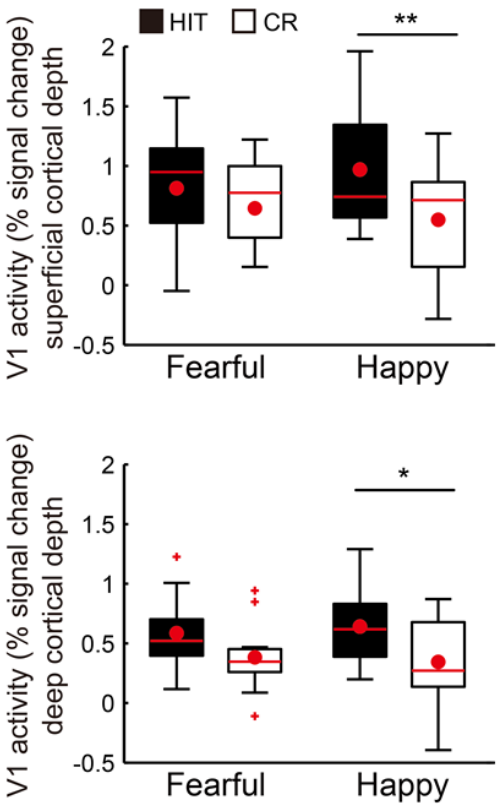

Fig S8. V1 activity at superficial (top) and deep cortical depths (bottom) on HIT and correct rejection (CR) trials. There was a significant main effect of condition (HIT/CR) ($F(1,10) = 10.527$, $p = .009$), which did not significantly interact with neither emotion (fearful/happy) nor cortical depth (superficial/deep). Planned t-tests revealed significantly larger activity on HIT than CR trials of a happy face at both the superficial cortical depth ($t(10) = 4.603$, $p = .001$, $d = 3.3$) and the deep cortical depth ($t(10) = 2.839$, $p = .018$, $d = 1.3$), which demonstrates that signals at the V1 deep cortical depth has sufficient power to detect activity difference across conditions. *$p < 0.05$, **$p < 0.01$. Box plot shows upper (75%) and lower (25%) quartiles with median (red line) and mean (red dot), with whisker showing maximum and minimum value. An outlier (outside of ± 2.7 standard deviations within a distribution for a given condition) is shown with a red cross. Related to **Fig 2**.