nature neuroscience

Responses of pulvinar neurons reflect a subject's confidence in visual categorization

Yutaka Komura[1,2], Akihiko Nikkuni[1], Noriko Hirashima[1], Teppei Uetake[1] & Aki Miyamoto[1]

When we recognize a sensory event, we experience a confident feeling that we certainly know the perceived world 'here and now'. However, it is unknown how and where the brain generates such 'perceptual confidence'. Here we found neural correlates of confidence in the primate pulvinar, a visual thalamic nucleus that has been expanding markedly through evolution. During a categorization task, the majority of pulvinar responses did not correlate with any 'perceptual content'. During an opt-out task, pulvinar responses decreased when monkeys chose 'escape' options, suggesting less confidence in their perceptual categorization. Functional silencing of the pulvinar increased monkeys' escape choices in the opt-out task without affecting categorization performance; this effect was specific to the contralateral visual target. These data were supported by a theoretical model of confidence, indicating that pulvinar activities encode a subject's certainty of visual categorization and contribute to perceptual confidence.

When we are aware of a sensory event, our subjective experience has two aspects[1–6]: one involves perceptual content (for example, an upward-moving red ball) and the other refers to perceptual confidence (for example, a confident feeling that we have actually perceived it). Studies over last few decades have suggested that sensory awareness emerges from the thalamocortical complex[7–10], and the study we describe here is focused on the visual thalamus. The primate visual thalamus has two regions[11–14]: the lateral geniculate nucleus (LGN; the relay from the retina to the primary visual cortex) and the pulvinar, the largest thalamic area, which has expanded during primate evolution and connects with multiple visual cortices. Based on anatomical connectivity, we hypothesized that these two regions of the visual thalamus have different roles in the two aspects of a subject's visual experience. To examine their functional roles, we first recorded the single-unit activities while monkeys performed the two tasks: a perceptual categorization task, which we used to evaluate perceptually experienced content, and an opt-out task[15–17], which we used to explore the subjects' confidence levels. To test whether the LGN and pulvinar are functionally linked to perceptual content and confidence, we pharmacologically inactivated these regions.

In a recent study, a theoretical model of confidence was described, in which a confidence value is calculated by measuring the distance between the perceived stimulus value and the category boundary[18]. That model produces a close fit to neural responses of the orbitofrontal cortex during a categorization task[18,19]. We found that this model also accounted for the pulvinar responses observed in our experiments and the monkeys' psychometric function during an opt-out task. Furthermore, the trial-by-trial firing rates of pulvinar neurons predicted the monkeys' upcoming behaviors. These neural and behavioral data together with the computational modeling indicated that the pulvinar responses likely reflected confidence in perceptual categorization. In addition, inactivation of the pulvinar confirmed that the representation of confidence in the pulvinar was behaviorally used while the monkeys performed the opt-out task.

RESULTS

Behavioral task and performance

In the perceptual categorization task (**Fig. 1a**), the color of the first fixation point (either red or green) indicated which colored cloud of dots would be the target (target color). After a delay, the random-dots stimulus (RDS) appeared for 500 ms. Each dot had a color (red or green) and moving direction (upward or downward). After another delay, the monkey was required to report the direction of the moving cloud of dots of the 'target color', either downward or upward by touching either the left or right bar, respectively. Difficulty of the task was manipulated by a stimulus matrix in which the ratio of color-motion pairings changed while the ratio of each individual color or motion remained constant (**Fig. 1b**). When the proportion of red dots moving downward ('red-down pairing ratio') approached 50%, the stimulus ambiguity increased. In **Figure 1c** we show the behavioral result from the initial assessment of two monkeys (for each monkey's performance, see **Supplementary Fig. 1**). The monkey selected a bar to touch according to its decision about the perceived content from the four categories (that is, red-down, red-up, green-down and green-up). For example, in the leftmost stimulus shown in **Figure 1b**, the monkey made a decision about the motion direction (up or down) in the red or green task–relevant trials, respectively (**Fig. 1c**). The proportion of the monkey's bar choice varied as a function of the stimulus matrix, indicating that the monkey's visual categorization fluctuated in relation to the stimulus ambiguity.

[1]Systems Neuroscience, Human Technology Research Institute, National Institute of Advanced Industrial Science and Technology, Tsukuba, Japan. [2]Precursory Research for Embryonic Science and Technology, Japan Science and Technology Agency, Saitama, Japan. Correspondence should be addressed to Y.K. (komura-y@aist.go.jp).

Received 26 December 2012; accepted 5 April 2013; published online 12 May 2013; doi:10.1038/nn.3393

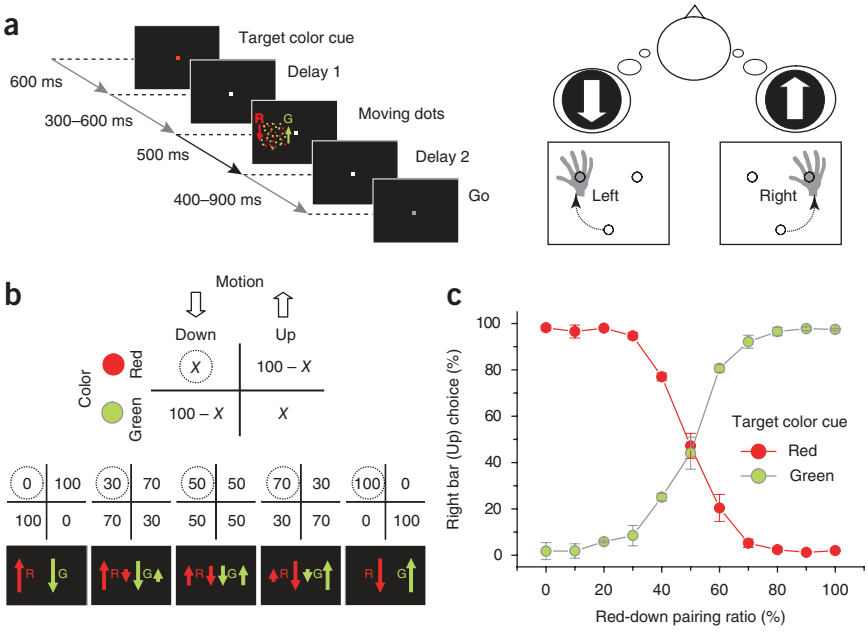


Figure 1 Perceptual categorization task. (**a**) Sequence of visual events in a trial. At the beginning of each trial, the color of the fixation point indicated which colored dots the monkey was to judge: red (R) or green (G). After the monkeys fixated, the RDS was presented on the screen. After another delay, the fixation point dimmed, signaling the monkey to report the perceived category by touching the left or right bar when the dots of task-relevant color moved downward or upward, respectively. Hues have been modified for presentation purposes. (**b**) A stimulus matrix determined the RDS characteristics, by changing only color-motion pairing ratios, holding the sum ratios (100%) of individual visual features constant. Note that the number of the red (or green) dots is always the same, and the number of the upward (or downward) moving dots is the same as well. The values surrounded by the dotted circles in the stimulus matrices are the percentage of red dots moving downward ('red-down pairing ratio'). (**c**) Proportion of right bar choice (mean ± s.d.) by two monkeys as a function of the red-down pairing ratio (*X* value surrounded by the dotted circle in **b**) for red and green (target color) task–relevant trials.

Pulvinar responses during a categorization task

The response examples of one pulvinar neuron are shown in **Figure 2a**. The neuron exhibited early transient and late sustained responses during the perceptual categorization task. The late responses decreased in a graded manner as stimulus ambiguity increased. To quantify the relationship between the neural responses and stimulus parameters, we calculated two linear regressions for firing rates versus the lower and higher ratios (0%, 30%, 45% and 55%, 70% and 100%) of red-down pairing, separately (**Fig. 2b**; equation (1) in Online Methods). The early responses (average firing rates 50–150 ms after onset of the stimulus) were not affected by changes in the red-down pairing ratios (permutation test, $P > 0.51$), whereas the regression slopes of the late responses (average firing rates of 400–500 ms after the onset of stimulus) on the lower and higher pairing ratios were negative and positive, respectively ($P < 0.001$). We recorded 618 single neurons with visual responses (83 LGN neurons and 535 pulvinar neurons) and found that the response magnitude of 163 neurons exhibited V-shaped modulations like those shown in **Figure 2b** (for characterization of the responses of LGN and pulvinar neurons, see Online Methods). Notably, all of the neurons with V-shaped response modulations were located in the pulvinar. We averaged the instantaneous regression slopes across the 163 neurons (**Fig. 2c**) to visualize the time course of the development of the regression slopes at the population level. We then estimated the onset of the response modulation by finding the earliest time of ten consecutive 1-ms bins for which the regression slopes were significantly different from the baseline (Wilcoxon signed rank test, $P < 0.05$). In the red task–relevant trials, onsets of response modulation were 191 ms and 178 ms for the lower and higher ratios of red-down pairing, respectively. In the green task–relevant trials, onsets of response modulation were 207 ms and 164 ms

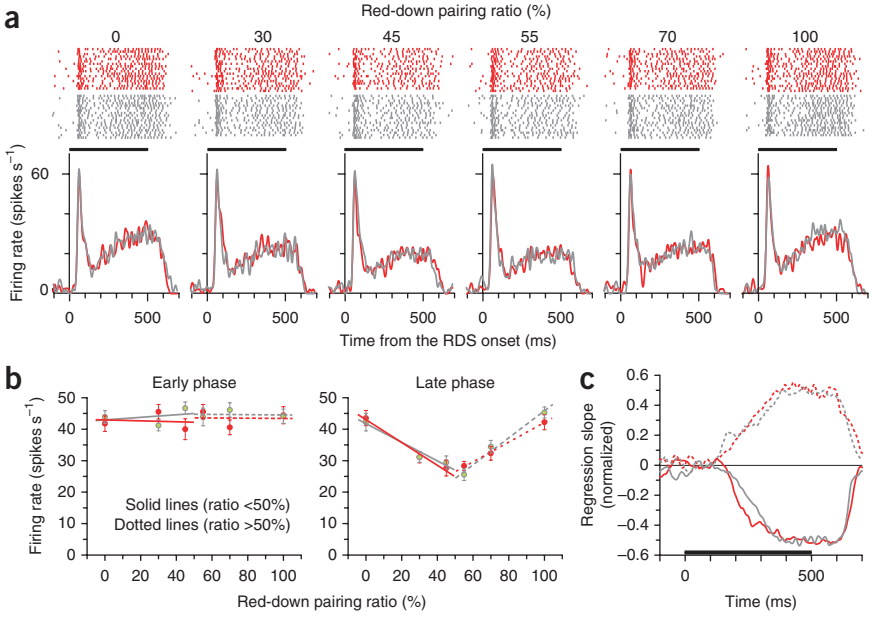


Figure 2 Graded modulation of pulvinar responses with stimulus ambiguity. (**a**) A representative response of a pulvinar neuron during the perceptual categorization task. Rastergrams and spike-density functions show the neuronal response to each RDS. Red and gray indicate data in the red and green task–relevant trials, respectively. The thick horizontal bar indicates the period during which the RDS was presented. (**b**) Regression slopes for the early and late responses of a pulvinar neuron (**a**) versus the lower ratio (<50%) and higher ratio (>50%) ratios of red-down pairing (error bars, s.e.m.; $n = 211$ trials). (**c**) Time course of the average regression slope for normalized visual response (normalized to maximal response magnitude for each neuron) versus the lower (solid line) and higher (dotted line) ratios of red-down pairing, calculated from the population data ($n = 163$ neurons).

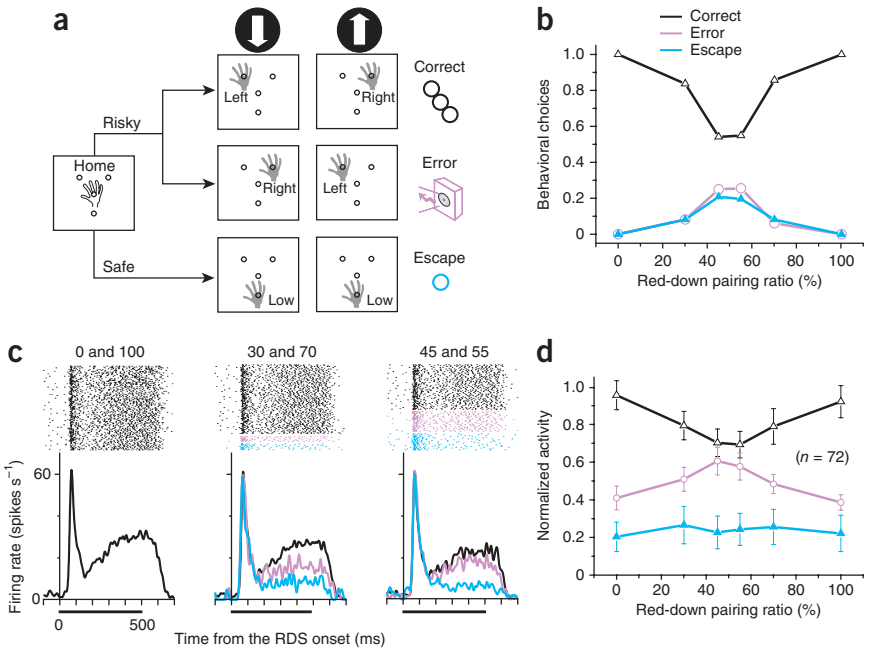


Figure 3 Relationships between opt-out task performances and pulvinar activities. (**a**) Outline of an opt-out task. The reward size was larger in correct risky trials than in safe trials. A 'beep' sound signaled when the outcome was an error. (**b**,**c**) Opt-out task performance (**b**) and pulvinar activities (**c**) as a function of stimulus parameter ($n = 298$ trials). In **c**, six red-down pairing ratios (0%, 30%, 45%, 55%, 70% and 100%) were combined into three groups (0% and 100%, 30% and 70%, and 45% and 55%) based on stimulus ambiguity. (**d**) Normalized firing rate (normalized to maximal response magnitude for each neuron; mean ± s.d.) as a function of stimulus parameter and the monkeys' choices for all of the tested neurons ($n = 72$ neurons).

for the lower and higher ratios, respectively. The results indicate that the response modulations with stimulus ambiguity appeared gradually after the onset of RDS.

Pulvinar responses during an opt-out task

Next, we examined whether the V-shaped response modulations merely reflected physical ambiguity in the stimulus or reflected perceptual ambiguity in the subject. Using the opt-out task[15–17], we reevaluated the pulvinar activities (**Fig. 3**). In this task, we gave the monkeys the option of abandoning the perceptual discrimination and choosing instead a third bar, touching which always resulted in a small reward (**Fig. 3a**). We hypothesized that the monkeys would accept the discrimination (risky option with a big reward only for correct trials) when they were confident in their perceptual categorization but would avoid the discrimination and choose a third bar (safe 'escape' option with a small reward in all trials) when they were less confident. The behavioral performance of the monkeys was consistent with the hypothesis because they chose the third bar more frequently as the stimulus ambiguity increased (**Fig. 3b**). We then investigated the relationship between the pulvinar activities and the monkey's choice in each trial (**Fig. 3c** and **Supplementary Fig. 2**). The magnitude of the pulvinar post-stimulus activities in the later epoch decreased (Kruskal-Wallis test, $H = 890.4$, $P < 0.001$; with post-*hoc* Steel-Dwass test, $P < 0.001$) in the order of correct risky choices, erroneous risky choices and safe escape choices. For all of the tested neurons ($n = 72$, **Fig. 3d**), the activity modulation as a function of the red-down pairing ratio displayed V-shaped and inverted V–shaped patterns for correct (permutation test, $P < 0.001$) and erroneous risky choices ($P < 0.001$), respectively, but was not apparent when the monkey chose the safe escape option ($P > 0.28$).

Computational basis of the pulvinar responses

To assess the response patterns of pulvinar neurons, we used a theoretical model of confidence that has been established in a previous study[18] of a binary categorization task. In the model (**Fig. 4a**), a binary choice was determined by comparing the stimulus value and a category boundary ($s < b$ or $s > b$), and a confidence value was calculated from the distance between the stimulus value and the category boundary ($|s - b|$). In the opt-out task, the subject chose the escape or risky response when the confidence value was smaller or larger than a threshold value, respectively. This algorithm generated a psychometric function in the model, which replicated the V-shaped pattern for correct risky choices and the inverted V–shaped patterns for both erroneous risky choices and safe escape choices (**Fig. 4b**). Based on this algorithm, the confidence values (d) as a function of stimulus type exhibited a V-shaped pattern for correct risky choices, an inverted V–shaped pattern for erroneous risky choices and a flat pattern for safe escape choices (**Fig. 4c**). The modeling results closely matched the characteristic patterns of the pulvinar activities (**Fig. 3d**) and suggest that pulvinar responses reflect a subject's confidence derived from a categorical decision process.

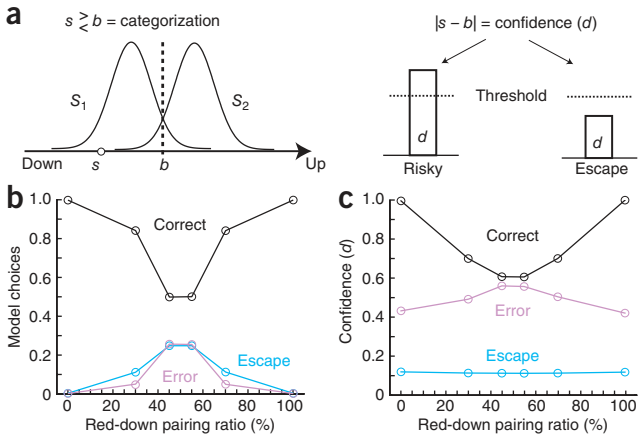


Figure 4 A model of categorization and confidence. (**a**) Schematic of the calculation of categorization and confidence during a motion-categorization task with an opt-out option. Each RDS was encoded as a distribution of values (S_1 or S_2). In each trial, a stimulus (s) was drawn from the respective distribution of the values. A categorical choice depended on comparing the stimulus value and a boundary ($s > b$ or $s < b$), and a confidence value (d) was determined by calculating the distance between the two ($|s - b|$). A risky or escape choice was defined by a confidence value above or below a threshold value, respectively. (**b**) Model choices as a function of stimulus parameter, with proportions of correct, error and escape choices plotted. (**c**) Average confidence values in correct, error, and escape trials as a function of stimulus parameter.

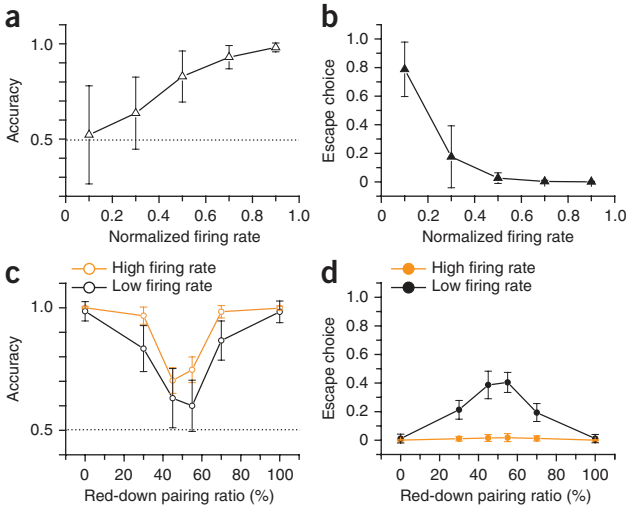
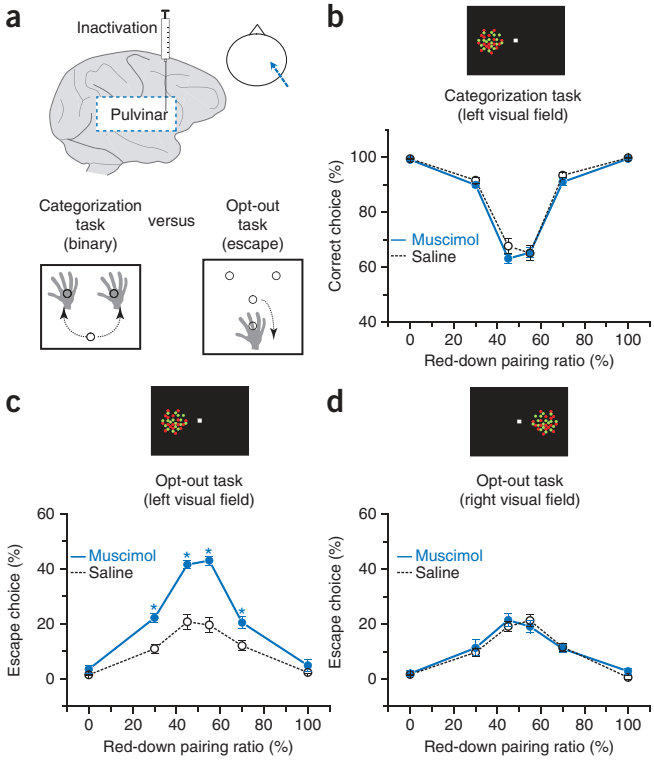


Figure 5 Pulvinar activity predicts the monkeys' behaviors in the opt-out task. (**a**,**b**) Proportion (mean ± s.d.) of accurate (**a**) and escape (**b**) choices as a function of the normalized firing rates (binned in five ranges) for all neurons tested during the opt-out task ($n = 72$ neurons). Note that the accuracy rates (**a**) were significantly above chance (0.5, dotted line; one-sample sign test, $P < 0.01$), except that for the bin with the lowest firing rate ($P = 0.61$). (**c**,**d**) Proportion (mean ± s.d.) of accurate (**c**) and escape (**d**) choices as a function of the stimulus parameter, conditioned of high or low firing rates ($n = 72$ neurons). Note that the accuracy (**c**) under any conditions of firing rates and stimulus types was significantly above chance (0.5, dotted line; one-sample sign test, $P < 0.01$).

Predicting the monkeys' behaviors from pulvinar activities

We also analyzed whether the pulvinar responses predicted the monkeys' upcoming behaviors. In **Figure 5a**,**b** we plotted the accuracy (proportion of correct choices among risky choices) or likelihood of escape choices as a function of the firing rates for all the neurons ($n = 72$, results from individual neurons are available in **Supplementary Fig. 3**). When the neurons fired at higher rates, the accuracy increased (**Fig. 5a**, Kruskal Wallis test, $H = 235.3$, $P < 0.001$; with post-*hoc* Scheffe test, $P < 0.001$), whereas the monkeys more frequently chose the escape options when the neurons fired at lower rates (**Fig. 5b**, Friedman test, $\chi^2 = 271.6$, d.f. = 4, $P < 0.001$; with post-*hoc* Scheffe test, $P < 0.001$). We then divided the trials into two groups characterized by high or low firing rates based on values that were higher or lower than the median spike count. We plotted accuracy (**Fig. 5c**) and escape (**Fig. 5d**) rates as a function of stimulus type for trials characterized by high or low firing rates. For any stimulus types, the accuracy and escape rates increased and decreased, respectively, when the neurons fired at high compared to low rates (Wilcoxon signed-rank test, $P < 0.01$). The results demonstrate that knowing whether firing rates were high or low improved the predictions of the monkeys' accurate and escape choices across any stimulus conditions. Such predictability also indicates that the trial-by-trial fluctuations in pulvinar responses correlated with changes in the subject's confidence.



Functional silencing effect of the visual thalamus

Next, we examined whether the neural responses in the pulvinar and LGN were functionally linked to the behaviors in the categorization and opt-out tasks (**Fig. 6a** and **Fig. 7a**, see **Supplementary Figs. 4** and **5** for each monkey's performance). We inactivated the unilateral visual thalamus (the right side) by injecting the GABA receptor agonist muscimol and compared the behavioral data with those obtained after saline injections. Inactivation of the pulvinar did not significantly affect the monkeys' binary choices in the catgorization task (**Fig. 6b** and **Supplementary Fig. 6a**; Wilcoxon signed-rank test corrected by Holm-Bonferroni method, $P > 0.34$) but increased the monkeys' choice of the escape option in the opt-out task (**Fig. 6c**; $P < 0.017$). We observed the effects of pulvinar inactivation when the RDS was presented to the monkey in the visual field that was contralateral (left side) to the injection site but not when the RDS was presented in the ipsilateral visual field (right side; **Fig. 6d**; $P > 0.77$). In contrast, inactivation of the LGN impaired performance on the categorization task (**Fig. 7b**; $P < 0.039$; but see **Supplementary Fig. 6b** for the RDS presented in the right hemifield, $P > 0.57$) and increased the frequency of the escape choice in the opt-out task (**Fig. 7c**; $P < 0.002$). These results indicate that the pulvinar is crucial for perceptual confidence but is not necessary for perceptual categorization, whereas the LGN underpins both perceptual categorization and confidence.

Functional topography of the pulvinar

Finally, we investigated the functional map of the pulvinar (**Fig. 8**). Based on the anatomical scheme in some previous studies[20–22], we divided the pulvinar into dorsal and ventral parts, which were separated at the level of the brachium of the superior colliculus (**Supplementary Fig. 7**). We reconstructed the recording sites using

Figure 6 Selective effects of inactivating the pulvinar on the tasks. (**a**) Reversible inactivation of the right pulvinar ($n = 22$ sites, two monkeys) during the perceptual categorization task and the opt-out task. (**b**–**d**) Behavioral performance (mean ± s.e.m.) as a function of stimulus ambiguity with muscimol versus saline injections. Images show where RDS was presented on the monitor. Inactivation effects on binary choices in the categorization task did not appear regardless of whether RDS was presented on the left side (**b**) or right side. Significant effects of inactivation (*$P < 0.05$, Wilcoxon signed-rank test) on escape choices in the opt-out task were observed when RDS was presented in the left visual field (**c**) but not when RDS was presented in the right visual field (**d**).

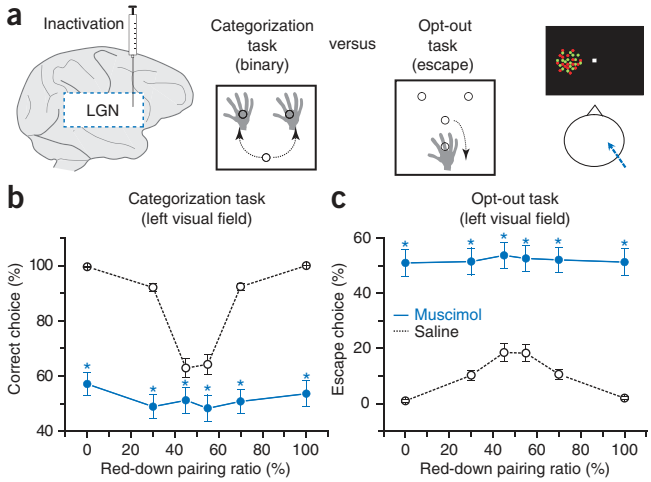


Figure 7 Effects of inactivating the LGN on the task performance. (**a**) Reversible inactivation of the right LGN ($n = 16$ sites, two monkeys) during the two types of task when RDS was presented in the left hemifield. (**b**,**c**) Proportion (mean ± s.e.m.) of correct choices in the binary categorization task (**b**) and that of escape choices in the opt-out task (**c**) as a function of stimulus ambiguity. Significant effects of inactivation ($^*P < 0.05$, Wilcoxon signed-rank test) were observed in both tasks.

MRI (magnetic resonance imaging) scans (**Fig. 8a**) and located the 247 neurons (164 in the dorsal pulvinar and 83 in the ventral pulvinar) with the types of response modulation identified during the categorization task. The majority of dorsal pulvinar neurons ($n = 141$ neurons, 86.0%) exhibited V-shaped response modulations, whereas we observed this type of response modulation less frequently for ventral pulvinar neurons ($n = 22$, 26.5%). Moreover, all muscimol-treated sites that resulted in increasing escape rates during the opt-out task were located in the dorsal pulvinar (**Fig. 8b**).

DISCUSSION

Using a combination of psychophysics, neural recordings and theoretical modeling, we identified the neural correlates of confidence in the primate pulvinar. Previous studies have shown that the pulvinar is involved in visual attention[23–28]. These studies raise a possibility that our experimental data may be explained solely by an 'attention factor'. First, difficulty of the task owing to the stimulus could change demands on attention, which potentially caused V-shaped modulations of the pulvinar responses in correct trials. Second, lapses in attention could decrease the firing rate and increase the escape rates in the opt-out task. The attention factor by itself, however, cannot account for the peculiar tuning curves of the pulvinar responses showing the V-shaped and inverted V–shaped modulations in correct and error trials, respectively (**Fig. 3d**). Other studies have shown that the pulvinar responses are modulated by eye position[29–31], although changes in eye position do not explain the pulvinar responses described in the current study: pulvinar responses markedly changed as a function of the stimulus and the monkey's choice, whereas eye position during the task was not significantly affected by the different stimuli and monkeys' choices (Kruskal-Wallis test, $P > 0.35$; **Supplementary Fig. 8**). In contrast, all the patterns of response modulations during correct, error and escape trials were systematically explained by a confidence model (**Figs. 3d** and **4a**,**c**). Additionally, the monkeys' psychometric function in the opt-out task was also reproduced by this model (**Fig. 4b**). Finally, the pulvinar activities were predictive of the monkeys' performances beyond the stimulus information (**Fig. 5**). These results led us to conclude that the observed response modulations of the pulvinar neurons most likely reflect the subject's confidence.

The experiment of pharmacological inactivation revealed a specific role of the pulvinar in the tasks: pulvinar inactivation affected opt-out choices but not visual categorization (**Fig. 6**). In contrast, inactivation of the LGN affected both categorization and opt-out choices (**Fig. 7**). From a modeling perspective, we considered inactivation of the pulvinar to represent conditions in which the distance between the

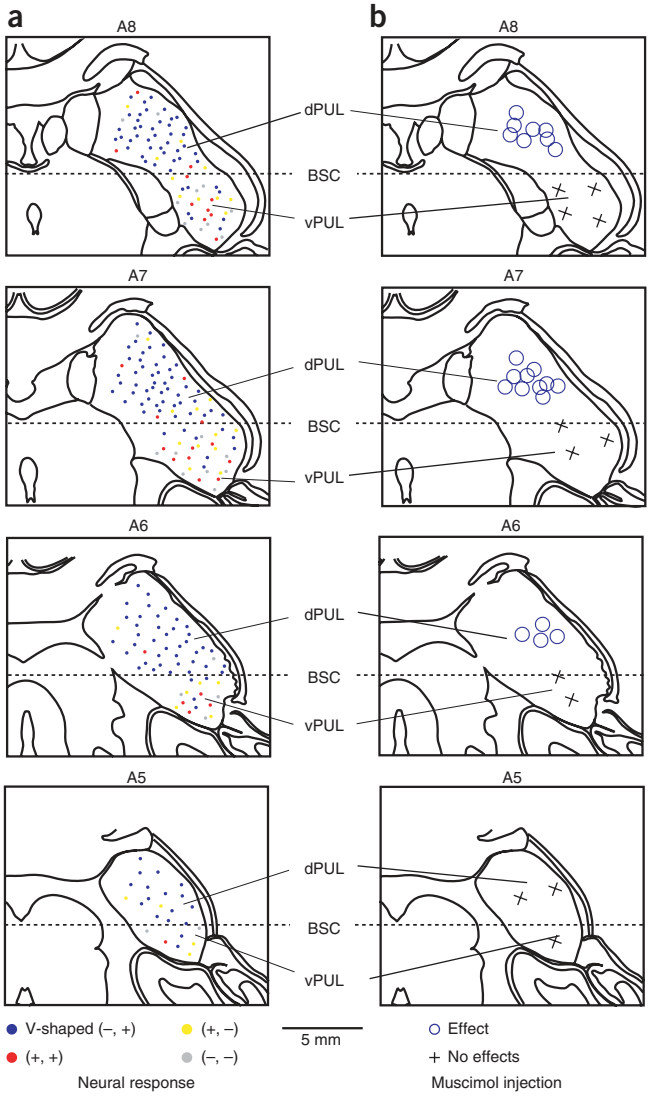


Figure 8 Functional topography of the pulvinar. (**a**) Locations of neurons showing the four types of response modulation during the categorization task are plotted on representative coronal sections of the right thalamus. Blue and yellow dots indicate neurons with negative-positive (V-shaped) and positive-negative slopes for the response modulations, respectively. Red and gray dots indicate neurons with positive-positive and negative-negative slopes, respectively. (**b**) Sites of muscimol injection in the pulvinar. Circles indicate sites that, when inactivated by muscimol, affected the behavioral results of the opt-out task. Crosses indicate sites that did not affect the results of the task after muscimol injection. The dorsal pulvinar (dPUL) and pulvinar ventral (vPUL) were defined as pulvinar regions dorsal and ventral to the brachium of superior colliculus (BSC), respectively. The value above each section indicates the anterior distance from ear bar zero.

stimulus and boundary (the confidence value) was underestimated (**Supplementary Fig. 9**). Although we could make similar predictions by elevating the threshold for opt-out choices without changing confidence values, the model of reduced confidence provided a more plausible explanation for the inactivation results (**Fig. 6c**), particularly because the majority of pulvinar neurons showed visual responses that were modulated by confidence levels (**Figs. 2** and **3**). In contrast, inactivation of the LGN was thought to devastate the stimulus value itself. These behavioral and modeling results of inactivation support the view on perceptual process[13,14,32,33] that the LGN (first-order visual thalamus) acts upstream of perceptual categorization whereas the pulvinar (higher-order visual thalamus) acts downstream of visual categorization and has an essential role in the subject's confidence about the visual categorization.

Neurons with V-shaped response modulation were located primarily in the dorsal pulvinar rather than the ventral pulvinar (**Fig. 8a**). The dorsal region corresponds to the classically defined medial pulvinar and dorsal portion of the lateral pulvinar[12,21,34–36]. This region receives input from multiple cortical areas, including the frontal and parietal cortices[12–14,33–36], which have been shown to contain neural correlates of confidence in previous studies[18,19,37]. The response profiles of the cortical neurons, however, differed from those of the pulvinar neurons. The frontal cortical neurons[18,19] represented uncertainty (the opposite of confidence) more frequently than confidence and were activated during the outcome-anticipation period (mainly between the behavioral choice and reward delivery). The parietal cortical neurons[37] represented the decision about where to direct a saccade, and their responses showed monotonic modulations in relation to confidence levels. Thus, it is unlikely that the pulvinar responses passively reflected these cortical responses. The pulvinar neurons exhibited undifferentiated early responses and differentiated later responses, with V-shaped modulation during presentation of the visual stimulus. Based on anatomical connectivity, we presume that the early responses likely reflect bottom-up sensory inputs via visual cortical areas, whereas the late responses reflect recurrent activity arising from interactions with multiple cortical areas. Additional experiments, such as simultaneous neural recordings from the different brain areas during the same task, are necessary to study the neural circuits that generate these responses.

Our results can offer a new perspective on previous studies of the pulvinar. First, our findings do not deny a hypothesis about the pulvinar contributing to attention, which has been suggested in many studies[23–28]. Rather, a possible interpretation is that confidence-related signals in the pulvinar can be used to regulate attention, because recent studies indicated that estimates of uncertainty or confidence influence how a subject will explore the world and allocate attention resources[19,38–41]. Second, clinical studies in humans[5,42,43] have shown that pulvinar damage results in hemineglect, that is, failure to notice a stimulus that is present in one hemifield. A study of monkey neurophysiology also has suggested that pulvinar responses correlate with the visibility of a stimulus[22]. Visibility has been often measured by rating the confidence in whether the stimulus can be perceived and is intimately tied to stimulus awareness[6,7,44–46]. In this study, the pulvinar neurons representing confidence were activated only during visual events (**Figs. 2** and **3**) and effects of pulvinar inactivation during the opt-out task were specific to contralateral visual targets (**Fig. 6**). These observations are consistent with the view that confidence signals in the pulvinar are closely related to the extent of perceptual visibility in the corresponding visual field. Taken together with these previous results, our data indicate that the pulvinar houses a confidence map of the visual world, which potentially serves for a subject's visual awareness 'here and now' and explorations of the environment.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Supplementary information is available in the online version of the paper.

ACKNOWLEDGMENTS

We thank K. Kawano, S. Dehaene, R. Kanai and S. Phillips for valuable comments and critical discussions; R. Tamura and K. Numata for continuous encouragement; T. Mega and A. Muramatsu for technical help and animal care. This work was supported in part by Precursory Research for Embryonic Science and Technology program from Japan Science and Technology Agency, Grant-in-Aids for Young Scientists (A) and Scientific Research on Innovative Areas from Ministry of Education, Culture, Sports, Science and Technology, Japan (to Y.K.).

AUTHOR CONTRIBUTIONS

Y.K. and N.H. designed the experiment. Y.K., A.N., N.H. and A.M. collected data. Y.K., A.N., T.U. and N.H. contributed to data analysis. All authors discussed the results and wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Peirce, C.S. & Jastrow, J. On small differences of sensation. *Mem. Nat. Acad. Sci.* **3**, 73–83 (1884).
2. Seth, A.K., Baars, B.J. & Edelman, D.B. Criteria for consciousness in humans and other mammals. *Conscious. Cogn.* **14**, 119–139 (2005).
3. Frith, C., Perry, R. & Lumer, E. The neural correlates of conscious experience: an experimental framework. *Trends Cogn. Sci.* **3**, 105–114 (1999).
4. Sahraie, A., Weiskrantz, L. & Barbur, J.L. Awareness and confidence ratings in motion perception without geniculo-striate projection. *Behav. Brain Res.* **96**, 71–77 (1998).
5. Driver, J. & Vuilleumier, P. Perceptual awareness and its loss in unilateral neglect and extinction. *Cognition* **79**, 39–88 (2001).
6. Lau, H.C. A higher order Bayesian decision theory of consciousness. *Prog. Brain Res.* **168**, 35–48 (2008).
7. Dehaene, S. & Changeux, J.P. Experimental and theoretical approaches to conscious processing. *Neuron* **70**, 200–227 (2011).
8. Haynes, J.D., Deichmann, R. & Rees, G. Eye-specific effects of binocular rivalry in the human lateral geniculate nucleus. *Nature* **438**, 496–499 (2005).
9. Tononi, G. An information integration theory of consciousness. *BMC Neurosci.* **5**, 42 (2004).
10. Crick, F. & Koch, C. Constraints on cortical and thalamic projections: the no-strong-loops hypothesis. *Nature* **391**, 245–250 (1998).
11. Berman, R.A. & Wurtz, R.H. Exploring the pulvinar path to visual cortex. *Prog. Brain Res.* **171**, 467–473 (2008).
12. Kaas, J.H. & Lyon, D.C. Pulvinar contributions to the dorsal and ventral streams of visual processing in primates. *Brain Res. Rev.* **55**, 285–296 (2007).
13. Guillery, R.W. & Sherman, S.M. Thalamic relay functions and their role in corticocortical communication: generalizations from the visual system. *Neuron* **33**, 163–175 (2002).
14. Casanova, C. The visual functions of pulvinar. in *The Visual Neurosciences* (eds. Chalupa, L.M. & Werner, J.S.) 592–608 (MIT Press, 2003).
15. Shields, W.E., Smith, J.D., Guttmannova, K. & Washburn, D.A. Confidence judgments by humans and rhesus monkeys. *J. Gen. Psychol.* **132**, 165–186 (2005).
16. Kornell, N., Son, L.K. & Terrace, H. Transfer of metacognitive skills and hint seeking in monkeys. *Psychol. Sci.* **18**, 64–71 (2007).
17. Persaud, N., McLeod, P. & Cowey, A. Post-decision wagering objectively measures awareness. *Nat. Neurosci.* **10**, 257–261 (2007).
18. Kepecs, A., Uchida, N., Zariwala, H.A. & Mainen, Z.F. Neural correlates, computation and behavioural impact of decision confidence. *Nature* **455**, 227–231 (2008).
19. Kepecs, A. & Mainen, Z.F. A computational framework for the study of confidence in humans and animals. *Phil. Trans. R. Soc. Lond. B* **367**, 1322–1337 (2012).
20. Olszewski, J. *The Thalamus of the Macaca Mulatta, An Atlas for Use with the Stereotaxic Instrument* (Karger Press, Basel, 1952).
21. Gutierrez, C., Cola, M.G., Seltzer, B. & Cusick, C. Neurochemical and connectional organization of the dorsal pulvinar complex in monkeys. *J. Comp. Neurol.* **419**, 61–86 (2000).
22. Wilke, M., Mueller, K.M. & Leopold, D.A. Neural activity in the visual thalamus reflects perceptual suppression. *Proc. Natl. Acad. Sci. USA* **106**, 9465–9470 (2009).

23. Petersen, S.E., Robinson, D.L. & Morris, J.D. Contributions of the pulvinar to visual spatial attention. *Neuropsychologia* **25**, 97–105 (1987).
24. Rafal, R.D. & Posner, M.I. Deficits in human visual spatial attention following thalamic lesions. *Proc. Natl. Acad. Sci. USA* **84**, 7349–7353 (1987).
25. Kastner, S. & Pinsk, M.A. Visual attention as a multilevel selection process. *Cogn. Affect. Behav. Neurosci.* **4**, 483–500 (2004).
26. LaBerge, D. & Buchsbaum, M.S. Positron emission tomographic measurements of pulvinar activity during an attention task. *J. Neurosci.* **10**, 613–619 (1990).
27. Saalmann, Y.B., Pinsk, M.A., Wang, L., Li, X. & Kastner, S. The pulvinar regulates information transmission between cortical areas based on attention demands. *Science* **337**, 753–756 (2012).
28. Desimone, R., Wessinger, M., Thomas, L. & Schneider, W. Attentional control of visual perception: cortical and subcortical mechanisms. *Cold Spring Harb. Symp. Quant. Biol.* **55**, 963–971 (1990).
29. Petersen, S.E., Robinson, D.L. & Keys, W. Pulvinar nuclei of the behaving rhesus monkey: visual responses and their modulation. *J. Neurophysiol.* **54**, 867–886 (1985).
30. Benevento, L.A. & Port, J.D. Single neurons with both form/color differential responses and saccade-related responses in the nonretinotopic pulvinar of the behaving macaque monkey. *Vis. Neurosci.* **12**, 523–544 (1995).
31. Robinson, D.L., McClurkin, J.W. & Kertzman, C. Orbital position and eye movement influences on visual responses in the pulvinar nuclei of the behaving macaque. *Exp. Brain Res.* **82**, 235–246 (1990).
32. Malpeli, J.G. & Baker, F.H. The representation of the visual field in the lateral geniculate nucleus of Macaca mulatta. *J. Comp. Neurol.* **161**, 569–594 (1975).
33. Purushothaman, G., Marion, R., Li, K. & Casagrande, V.A. Gating and control of primary visual cortex by pulvinar. *Nat. Neurosci.* **15**, 905–912 (2012).
34. Bender, D.B. & Youakim, M. Effects of attentive fixation in macaque thalamus and cortex. *J. Neurophysiol.* **85**, 219–234 (2001).
35. Imura, K. & Rockland, K.S. Long-range interneurons within the medial pulvinar nucleus of macaque monkeys. *J. Comp. Neurol.* **498**, 649–666 (2006).
36. Romanski, L.M., Giguere, M., Bates, J.F. & Goldman-Rakic, P.S. Topographic organization of medial pulvinar connections with the prefrontal cortex in the rhesus monkey. *J. Comp. Neurol.* **379**, 313–332 (1997).
37. Kiani, R. & Shadlen, M.N. Representation of confidence associated with a decision by neurons in the parietal cortex. *Science* **324**, 759–764 (2009).
38. Yu, A.J. & Dayan, P. Uncertainty, neuromodulation, and attention. *Neuron* **46**, 681–692 (2005).
39. Luck, S.J., Hillyard, S.A., Mouloua, M. & Hawkins, H.L. Mechanisms of visual-spatial attention: resource allocation or uncertainty reduction? *J. Exp. Psychol. Hum. Percept. Perform.* **22**, 725–737 (1996).
40. Itti, L. & Baldi, P. Bayesian surprise attracts human attention. *Vision Res.* **49**, 1295–1306 (2009).
41. Dayan, P., Kakade, S. & Montague, R.P. Learning and selective attention. *Nat. Neurosci.* **3** (suppl.), 1218–1223 (2000).
42. Karnath, H.O., Himmelbach, M. & Rorden, C. The subcortical anatomy of human spatial neglect: putamen, caudate nucleus and pulvinar. *Brain* **125**, 350–360 (2002).
43. Zihl, J. & von Cramon, D. The contribution of the 'second' visual system to directed visual attention in man. *Brain* **102**, 835–856 (1979).
44. Kunimoto, C., Miller, J. & Pashler, H. Confidence and accuracy of near-threshold discrimination responses. *Conscious. Cogn.* **10**, 294–340 (2001).
45. Koch, C. & Preuschoff, K. Betting the house on consciousness. *Nat. Neurosci.* **10**, 140–141 (2007).
46. Kolb, F.C. & Braun, J. Blindsight in normal observers. *Nature* **377**, 336–338 (1995).

ONLINE METHODS

Experiments with monkeys. All surgical and experimental procedures were approved by the Animal Care and Use Committee of the National Institute of Advanced Industrial Science and Technology (AIST, Japan), and performed in accordance with the Guidelines for the Care and Use of Animals of AIST. Details of procedures used in the present study are similar to those previously described[47–49]. Two male macaque monkeys (*Macaca fuscata*, weighing 5–9 kg, 3–8 years of age) participated in the experiments. We minimized the number of monkeys used by considering the ethics and data similarity. Each monkey had his own cage with environmental enrichment. The housing area was maintained on a 12-h light–12-h dark cycle. The experiments were conducted during the light cycle. We used the naive monkeys without a history of any other experiments. For this study, a head holder, recording chamber and search coils were implanted into the monkeys under anaesthesia and sterile surgical conditions. The head holder, the recording chamber and the eye coil connectors were all embedded in dental acrylic that covered the top of the skull and were connected to the skull using titanium screws. The search coils were surgically placed under the conjunctiva of the eyes for recording of eye movements. The recording chamber was placed over the parietal cortex, tilted laterally by 20° and aimed at the visual thalamus.

Visual stimulus. Visual stimuli were generated using VSG 2/5 (Cambridge Research Systems). A CRT color display (Sony, GDM-F500) was placed 57 cm in front of the monkey. The display subtended a visual angle of 40° × 30° with a resolution of 688 × 516 pixels, and refreshed at 100 Hz. Random dots (0.165° in diameter each) were drawn with densities of 16.8 per degree^2 inside circle apertures with diameters of 4.7°–8.8° and moved at the speed of 15.8 degrees s^{-1}. In the task, the random-dots stimulus covered the neuron's receptive field, mapped initially with a moving bar. The dot color was red (CIE chromaticity, $x = 0.623$, $y = 0.34$) or green ($x = 0.288$, $y = 0.6$) at the same luminance (10 cd m^{-2}). Apart from the color cue, the fixation point (0.33°) was 10 cd m^{-2} white ($x = 0.255$, $y = 0.219$), then dimmed to 2 cd m^{-2} white.

Behavioral tasks. The behavioral task was controlled by REX, a QNX-based real-time experimentation data-acquisition system. All trials started with the presentation of a fixation point at the center of the display to attract the focus of the monkey. In the perceptual categorization task, the color of the fixation point (red or green) indicated which colored dots the monkey should judge (**Fig. 1a**). If the monkey gazed at it for 600 ms, the color of the fixation point turned white. After the monkey fixated on it for 300–600 ms, the RDS appeared for 500 ms. Between 400 ms and 900 ms after the RDS was turned off, the fixation point dimmed, signaling that the monkey could release the home bar and was required to touch the left or right bar within 1 s. Correct responses were rewarded with a drop of water or juice. The two cue colors were randomly interleaved from trial to trial. In any tasks, the monkey was required to maintain fixation within a 1.4-degree square window throughout all the visual events. The eye position was monitored with a search coil system (Datel) and stored at 1 kHz. Based on the psychophysical assessment (**Fig. 1c**), we selected RDSs with six kinds of the pairing ratios (0%, 30%, 45%, 55%, 70% and 100%) in the electrophysiological experiments. Each RDS was presented randomly from trial to trial at nearly equal frequency.

Electrophysiology. Single units were recorded with flexible tungsten microelectrodes (Microprobe). The recording chambers were implanted in accordance with coordinates determined by MRI. Before the start of the recording experiments, we first searched for auditory neural responses by clapping and thereby localized the margins of the medial geniculate body, serving as a physiological landmark. We then determined the posterior pole of the lateral geniculate nucleus (LGN), which demarcated the anterior border of the pulvinar. Based on the coordinates derived from these structural and physiological localizing procedures, the microelectrode was then inserted stereotaxically stepwise with a micromanipulator (MO-97A, Narishige) into various parts of the visual thalamus. At the beginning of each recording session, visual receptive-field mapping was done while the monkey looked at the central fixation spot. All the LGN neurons showed the visual responses with the contralateral small receptive fields. Although some pulvinar neurons had large response fields extending into the ipsilateral hemifields, most pulvinar neurons also had the contralateral receptive fields. Thus we selected the neurons whose receptive fields were restricted in the contralateral hemifield. Spiking data were collected with a MAP recording system (Plexon).

Recording sites. To position the electrode at the target site, we advanced it through a guide tube that placed in a Crist grid and transversed the dura matter of the monkey. The Crist grid with holes separated by 1 mm was attached to the inside of the recording chamber[50]. During the recording experiments, we visualized electrode tracks in MRI scans eight times in total. After each scan, we confirmed by comparing the MRI image to coronal sections in a histological atlas[51] of *M. fuscata* that the tip of the electrode was located in the LGN or pulvinar (**Supplementary Fig. 7**). These procedures enabled us to reconstruct the recording sites: locations of the recorded neurons were plotted onto monkey brain atlas, based on MRI images, electrode positions in the chamber and recording depth (**Fig. 8a**). In the present study, we divided the pulvinar into the dorsal and ventral parts, separated at the level of the brachium of superior colliculus; similar approaches have been used in previous studies[20–22]. The dorsal pulvinar contains sites previously referred to as the dorsal portion of the lateral pulvinar and medial pulvinar regions, whereas the ventral pulvinar contains sites previously referred to as the ventral portion of the lateral pulvinar and inferior pulvinar.

Data analysis. We analyzed behavioral and spiking data using custom software written in Matlab (Mathworks). Spike trains were smoothed by convolution with a Gaussian kernel ($\sigma = 10$ ms) to obtain spike-density functions. The early response of the pulvinar neurons was defined as the number of spikes in the first 100 ms from 50 ms to 150 ms after the onset of RDS. The early response of the LGN neurons was defined as the number of spikes in the first 100 ms after the onset of RDS because the response latencies of the LGN neurons were shorter than those of the pulvinar neurons. The late response was defined as the number of spikes from 400 ms to 500 ms after the onset of RDS. A neuron was classified as 'visual responsive' if the early or late response differed from the number of spikes in the 100-ms control period just before the onset of RDS (Wilcoxon signed-rank test, $P < 0.01$).

To statistically assess how the visual responses of the single neurons were related to the red-down pairing ratios, we first performed the following linear regression analysis, in which the firing rate of the neurons was given by

$$f = \begin{cases} f_1 + k_1 x & (0 \le x < 50) \\ f_2 + k_2 x & (50 < x \le 100) \end{cases} \quad (1)$$

where x is the ratio of red-down pairing, and k_1 and k_2 indicate the regression coefficient (slope) in the lower and higher ratios of red-down pairing, respectively, and f_1 and f_2 indicate the intercept in the lower and higher ratios of red-down pairing, respectively. We evaluated the statistical significance of the coefficient using a permutation test by randomly reassigning each firing rate to the red-down pairing ratio. This was repeated 1,000 times to calculate P value for each coefficient. If $P < 0.01$ in both k_1 and k_2, neural activity was defined as modulated by the RDS of different red-down pairing ratios. Second, for cells satisfying this criterion, the response modulations were classified into four types, depending on whether the value of k_1 or k_2 was positive or negative: in other words, (k_1, k_2) was (positive, positive), (negative, positive), (negative, positive) or (positive, negative). If (k_1, k_2) was (negative, positive), the response modulation was classified as V-shaped.

To normalize the firing rates across neurons, we first averaged neural responses to the RDS with any color-motion pairings for different tasks and then used the maximum among these as a reference value for each neuron. Next, we divided the response of each neuron to the different condition with respect to the reference value, and this quotient was defined as a normalized activity in each condition. To track the time course of V-shaped response modulations at the population level (**Fig. 2c**), we calculated the slope of the linear regression of the normalized firing rates in the 50-ms sliding windows with 1-ms moving steps on the lower and higher pairing ratios.

Each statistical test was performed, following checks on normality and equal variances of the samples by the Kolmogorov-Smirnov test and Levene test, respectively. Throughout this study, two-sided tests were used when we compared the data between the two groups. We used the Holm-Bonfferoni method to adjust for multiple comparisons, if necessary (**Figs. 6** and **7** and **Supplementary Figs. 4–6**). The data were analyzed by the persons with and without knowing relations between the experimental procedures and data. These blinded and unblended analyses had no effects on the results. We found similar tendencies in behavioral performances across the different monkeys (**Supplementary Figs. 1**, **4** and **5**) and

doi:10.1038/nn.3393

in the neural responses across the different cells (**Figs. 2** and **3** and **Supplementary Figs. 2** and **3**), which indicated the biological replicates.

Characterizing neural responses. We encountered 618 neurons showing visual responses to the RDS during the categorization task. Of these neurons, 247 showed response modulations during the perceptual categorization task. The response modulations of the neurons were classified into four types: negative-positive ($n = 163$ neurons), positive-positive ($n = 28$), negative-negative ($n = 25$) and positive-negative ($n = 31$) slopes in the lower-higher pairing ratio, respectively. These neurons were located in the pulvinar. The responses of all the LGN neurons were not modulated during the categorization task, like the neuron for which data are shown in **Supplementary Figure 10**. The LGN neurons exhibited the early transient responses to the onset of RDS, but, unlike the pulvinar neurons, did not show the vigorous responses during the later phase of the RDS presentation period. To understand the major role of the visual thalamus during our task, we focused on the vast proportion of the neurons that produced negative-positive slopes ($n = 163$, 66% of 247 neurons, chi-squared test, $P < 0.01$), resulting in V-shaped response modulations like those shown in **Figure 2b**.

We recorded 133 neurons during the opt-out task. Out of the 163 neurons with V-shaped response modulation during the categorization task, we consecutively tested 38 in the opt-out task. Ninety-five cells were recorded only during the opt-out task. Among these 95 neurons, 34 showed V-shaped response modulation in correct trials. Thus, we analyzed 72 neurons in total for the opt-out task. Of these, 65 neurons and 7 neurons were recorded from the dorsal and ventral pulvinar, respectively.

Modeling. We used a confidence model in a previous study[18]. Our task can be thought of as a binary motion-categorization task (up or down) after the target color is determined. In the model, each stimulus (s) with the red-down pairing ratio (x) was encoded as a distribution with Gaussian noise (n):

$$s = m(x - 50)/50 + n, n \in N(0,\sigma),$$

where σ^2 was the noise variance, and m determined mean values of the extreme distributions for stimulus type with 0-100 pairing ratio. In a given trial, a stimulus sample was drawn from the respective distribution.

The categorical choice was determined by comparing stimulus and boundary:

$$\text{Categorization} = \{\text{up} \mid s > b; \text{down} \mid s < b\}.$$

The confidence value (d) was calculated from the distance between stimulus and boundary ($|s - b|$). In the opt-out task, a risky or escape decision was determined by comparing the confidence value and a threshold value (c):

$$\text{Decision} = \{\text{risky} \mid d > c; \text{escape} \mid d < c\}.$$

For fitting the model, m, σ and c were adjusted using the least-square method. The yielded confidence values as a function of stimulus type were normalized to the maximum value of 1. The qualitative patterns of the simulated data for correct, error and escape trials were robust to different parameters, which affected the slope and level of each function. In the present model, we did not calibrate the confidence value. For simplicity, the category boundary ($b = 0$) and threshold (c) were fixed without variance.

Reversible inactivation. After the recording experiment, the inactivation experiments were conducted. A small amount of a GABA agonist (muscimol hydrobromide; 5 μg μl^{-1} in saline; volume = 1.0 μl) was pressure-injected at a rate of about 0.1 μl per 30 s using a 32-gauge injection needle that was connected to a microsyringe and was penetrated through a 23-gauge guide tube. Site and depth of injection were adjusted by using an x-y stage attached on top of the cylinder and a micromanipulator that was also used for the unit recording experiments. We injected muscimol in the pulvinar ($n = 34$ sites) and LGN ($n = 16$ sites). Sites of inactivation in the pulvinar were within the area containing neurons showing V-shaped response modulations during the categorization task. To rule out confounds from physical effects resulting from needle penetration, the effect of inactivation for each site was assessed across any stimulus types by comparing behavioral data obtained after injection with muscimol and saline (volume = 1.0 μl). If there were any differences in the proportions of correct or escape choices between the two injection conditions (two-sample proportion test, $P < 0.05$), the sites were classified as effective. All the LGN sites with muscimol injection significantly affected the performances during the categorization and opt-out tasks (**Fig. 7**). Among the 34 muscimol-treated sites in the pulvinar (**Fig. 8b**), 22 affected the opt-out task performances whereas the other 12 sites did not exhibit any effects on the task performance. In **Figure 6**, the population data based on muscimol injections into these effective sites of the pulvinar is shown. The muscimol or saline injection was conducted from day to day in random order.

We did not measure the spread of muscimol directly. However, a previous study[52] showed using glucose autoradiography that muscimol spread ~1.7 mm from the injection site when 1.0 μl of muscimol solution was injected during a 4-min period. In the presesnt study, we injected 1.0 μl of muscimol solution during a 5-min period. The minimum distances between the injection sites in the pulvinar and the LGN were more than 3.0 mm, which was beyond the range that muscimol would spread. Thus, it is unlikely that the injected muscimol inactivated both of the two regions of the visual thalamus. The effect of inactivation persisted for the period of the experiments but disappeared the next day.

47. Inaba, N., Shinomoto, S., Yamane, S., Takemura, A. & Kawano, K. MST neurons code for visual motion in space independent of pursuit eye movements. *J. Neurophysiol.* **97**, 3473–3483 (2007).
48. Komura, Y. *et al.* Retrospective and prospective coding for predicted reward in the sensory thalamus. *Nature* **412**, 546–549 (2001).
49. Komura, Y. *et al.* Auditory thalamus integrates visual inputs into behavioral gains. *Nat. Neurosci.* **8**, 1203–1209 (2005).
50. Crist, C.F., Yamasaki, D.S., Komatsu, H. & Wurtz, R.H. A grid system and a microsyringe for single cell recording. *J. Neurosci. Methods* **26**, 117–122 (1988).
51. Kusama, T. & Mabuchi, M. *Stereotaxic atlas of the brain of macaca fuscata* (University of Tokyo Press, Tokyo; University Park Press, 1970).
52. Martin, J.H. Autoradiographic estimation of the extent of reversible inactivation produced by microinjection of lidocaine and muscimol in the rat. *Neurosci. Lett.* **127**, 160–164 (1991).