Responses to Task-Irrelevant Visual Features by Primate Prefrontal Neurons

JOHAN LAUWEREYNS,[1] MASAMICHI SAKAGAMI,[1,2] KEN-ICHIRO TSUTSUI,[3] SHUNSUKE KOBAYASHI,[4] MASASHI KOIZUMI,[1] AND OKIHIDE HIKOSAKA[1]

[1]*Department of Physiology, Juntendo University, Tokyo 113-0033;* [2]*Brain Science Research Center, Tamagawa University, Tokyo 194-8610;* [3]*Department of Physiology, Nihon University, Tokyo 173-8610; and* [4]*Department of Neurology, University of Tokyo, Tokyo 113-8655, Japan*

Lauwereyns, Johan, Masamichi Sakagami, Ken-Ichiro Tsutsui, Shunsuke Kobayashi, Masashi Koizumi, and Okihide Hikosaka. Responses to task-irrelevant visual features by primate prefrontal neurons. *J Neurophysiol* 86: 2001–2010, 2001. The primate brain is equipped with prefrontal circuits for interpreting visual information, but how these circuits deal with competing stimulus-response (S-R) associations remains unknown. Here we show different types of responses to task-irrelevant visual features in three functionally dissociated groups of primate prefrontal neurons. Two Japanese macaques participated in a go/no-go task in which they had to discriminate either the color or the motion direction of a visual target to make a correct manual response. Prior to the experiment, the monkeys had been trained extensively so that they acquired fixed associations between visual features and required responses (e.g., "green = go"; "downward motion = no-go"). In this design, the monkey was confronted with a visual target from which it had to extract relevant information (e.g., color in the color-discrimination condition) while ignoring irrelevant information (e.g., motion direction in the color-discrimination condition). We recorded from 436 task-related prefrontal neurons while the monkey performed the multidimensional go/no-go task: 139 (32%) neurons showed go/no-go discrimination based on color as well as motion direction ("integration cells"); 192 neurons (44%) showed go/no-go discrimination only based on color ("color-feature cells"); and 105 neurons (24%) showed go/no-go discrimination only based on motion direction ("motion-feature cells"). Overall, however, 162 neurons (37%) were influenced by irrelevant information: 53 neurons (38%) among integration cells, 71 neurons (37%) among color-feature cells, and 38 neurons (36%) among motion-feature cells. Across all types of neurons, the response to an irrelevant feature was positively correlated with the response to the same feature when it was relevant, indicating that the influence from irrelevant information is a residual from S-R associations that are relevant in a different context. Temporal and anatomical differences among integration, color-feature and motion-feature cells suggested a sequential mode of information processing in prefrontal cortex, with integration cells situated toward the output of the decision-making process. In these cells, the response to irrelevant information appears as a congruency effect, with better go/no-go discrimination when both the relevant and irrelevant feature are associated with the same response than when they are associated with different responses. This congruency effect could be the result of the combined input from color- and motion-feature cells. Thus these data suggest that irrelevant features lead to partial activation of neurons even toward the output of the decision-making process in primate prefrontal cortex.

INTRODUCTION

There is no doubt that primate prefrontal cortex is involved in the executive control of behavior (Fuster 1997; Goldman-Rakic 1987; Passingham 1993). A large body of evidence from single-unit studies indicates that prefrontal neurons interpret visual information to determine the correct hand movement (Hoshi et al. 2000; Niki 1974; Rainer et al. 1998; Sakagami and Niki 1994a; Sakagami and Tsutsui 1999; Sakagami et al. 2001; Watanabe 1986; White and Wise 1999) or eye movement (Asaad et al. 1998, 2000; Funahashi et al. 1993; Hanes and Schall 1996; Kim and Shadlen 1999; Schall et al. 1995a). However, it remains unknown how such prefrontal neurons behave when they are presented with conflicting relevant and irrelevant visual information.

Recent evidence from an electrical-stimulation study in frontal eye field (Gold and Shadlen 2000) suggests that the brain's decision-making process to determine the required action consists of a gradual commitment toward a choice based on the accumulation of sensory evidence (see also Sakagami and Tsutsui 1999; Schall and Thompson 1999). In line with this view, it is possible that conflicting sensory evidence disturbs the development of a prefrontal neuronal code favoring one action over another. To investigate this prediction, we devised a conflict paradigm in which the monkey should discriminate one visual feature while ignoring another to make the appropriate behavioral response.

Depending on the discrimination condition, the monkey had to interpret either the color or the motion direction of the visual target. The monkey had been trained extensively prior to the experiment to acquire fixed associations between stimulus features and required responses (e.g., "purple = go"; "rightward motion = no-go"). In this design, as shown in Fig. 1*A*, the monkey was confronted with a visual target from which it had to extract relevant information (e.g., color in the color-discrimination condition, henceforth "color condition") while ignoring
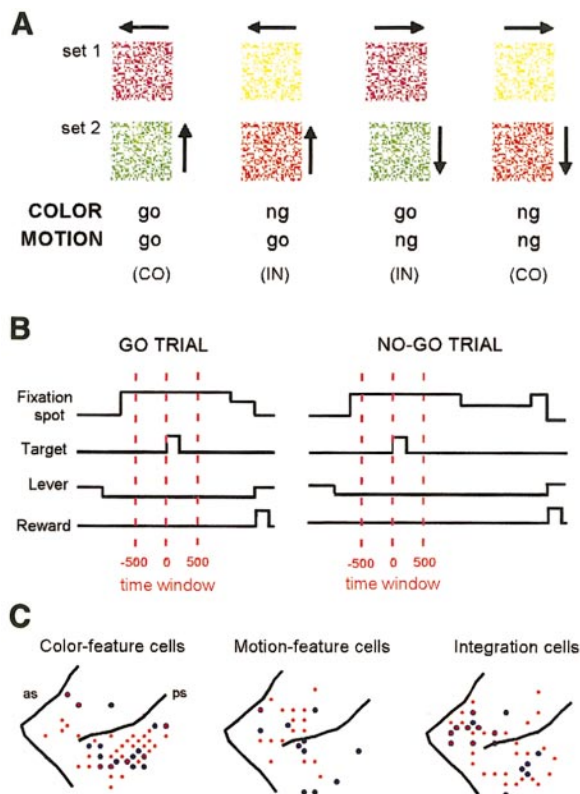
FIG. 1. *A*: schematic illustration of the experimental design. In *set 1,* the color purple and the leftward motion direction are associated with the go response (go), while the color yellow and the rightward motion direction are associated with the no-go response (ng). In *set 2,* the color green and the upward motion direction are associated with the go response (go), while the color red and the downward motion direction are associated with the no-go response (ng). The stimuli were exactly the same in both discrimination conditions, and so the irrelevant stimulus feature could be either congruent (CO) or incongruent (IN) with the required response. *B*: the sequence of events in a go trial (*left*) and in a no-go trial (*right*). The monkey initiated a trial by pressing the lever. In both types of trial, the monkey had to keep the lever pressed throughout target presentation and during a delay period of variable length. In a go trial, the monkey then had to release the lever as soon as possible on dimming of the fixation spot. In a no-go trial, the monkey had to refrain from releasing the lever until the fixation spot became bright again after the dim period. The vertical lines indicate the time windows used in the following figures of single-unit activity. *C*: the electrode penetrations in the right hemisphere of *monkey EC. Left, middle,* and *right*: in the same hemisphere, the distributions of, respectively, color-feature cells, motion-feature cells, and integration cells. The distributions are a good representation of the overall pattern of results with the 4 explored hemispheres. A red dot indicates a penetration in which we encountered at least 1 neuron that was unaffected by the irrelevant feature, whereas a blue circle represents a penetration in which we found at least 1 neuron that was influenced by the irrelevant feature; a blue circle superimposed on a red dot indicates a penetration in which we observed both types of neurons. PS, principal sulcus; AS, arcuate sulcus.

potentially confusing information (e.g., motion direction in the color condition).

In a previous behavioral study with a reaction-time version of this task, we found that the monkeys' manual responses were slower and less accurate when the irrelevant feature primed a different response than the relevant feature as compared with when both features primed the same response (Lauwereyns et al. 2000). The data suggested that irrelevant target information automatically activates hard-wired but presently inappropriate S-R associations. If this is true, such inappropriate S-R associations could be represented in prefrontal

neuronal activity. To test this hypothesis, we conducted a single-unit study in prefrontal cortex while the monkey performed the same type of visual multidimensional discrimination. We opted for a nonspeeded version of the task, introducing a delay between target presentation and manual response to prevent confounding with motor processes in the single-unit activity.

METHODS

Behavioral paradigm

The monkey was required to discriminate either the color or the motion direction of a visual target to make a correct go or no-go manual response. The behavioral meaning (go or no-go) of each target feature was fixed for each monkey during both training and experiments but could be either relevant or irrelevant depending on the discrimination condition. The color of the fixation spot indicated which of the target features the monkey should discriminate. Throughout a block of trials, the monkey had to select the appropriate behavior based on the same visual dimension (i.e., the color of the fixation spot remained constant). The correct response to a particular multidimensional target depended on only the relevant stimulus feature and so could vary across discrimination conditions (see Fig. 1*A*). Throughout an experimental block, we used one of two stimulus sets on which the monkey had been trained separately: *set 1,* the colors purple/yellow and the motion directions left/right; *set 2,* the colors red/green and the motion directions up/down.

With this design, in each trial irrelevant stimulus information could be either congruent or incongruent with the required response. Using a speeded version of the discrimination task, we confirmed that this paradigm leads to interference effects in the behavior of the two monkeys used in the present study with longer response times and increased error rates when the irrelevant feature was incongruent with the required response (Lauwereyns et al. 2000).

To prevent confounding between go/no-go discrimination and motor execution processes in the neuronal activity, we introduced a delay period between stimulus and response in the discrimination task for the neurophysiological recordings. In all other respects, the task and stimuli were exactly the same as for the behavioral test. In the paradigm with the delay period, the sequence of events in each trial was as follows (see Fig. 1*B*). The monkey initiated each trial with a lever press. The fixation spot (0.3° diam) appeared in the center of the CRT. After a variable period (1–2 s), the target stimulus was presented for 200 ms, the center of the target appearing at 4.1° either to the left or right of the center of the fixation spot. Following a variable delay (0.5–2 s), the fixation spot dimmed. The monkey then had to release the lever within 0.8 s (for a correct go response) or refrain from releasing the lever for at least 1.2 s (for a correct no-go response). In a no-go trial, the monkey could release the lever at any time after the 1.2-s no-go period. A drop of fruit juice was delivered on lever release as reward for every correct go or no-go response.

Eye movements were restricted to within 1° of the fixation spot by means of an infrared camera and associated equipment (R-21C-A, RMS) from 500 ms before until 500 ms after the onset of the target stimulus (with a sampling rate of 250 kHz). Trials in which an eye movement was detected outside the fixation window were aborted and counted as errors.

The monkey viewed a dynamic random dot pattern through a virtual square aperture (6.2 × 6.2°) as a target stimulus. All dots were of the same color and moved unidirectionally and coherently. Approximately 280 dots moving at 6°/s were used to cover 11% of the virtual aperture area. Apparent motion was produced by successive frame replacement (4 frames). All stimuli were presented on a 20-in CRT (HC39PEX, Mitsubishi) controlled by personal computers (PC386V, Epson). A lever consisting of a small plastic disk, with a

diameter of 2.0 cm, was used for the manual responses. The lever was attached to the monkey chair in front of the right hand at the height of the elbow in such a way that the monkey could reach it with the right hand only.

Electrophysiological recording

Recording was done in at least two blocks of 32 to 64 trials, one block using either stimulus *set 1* or *2* in the motion condition and one block using the same stimulus set in the color condition, in random order. The stimulus set was determined randomly; for some neurons, the entire experiment was repeated with the alternative stimulus set. Since many prefrontal cells show a spatial preference similar to the receptive fields found in visual cortices (Sakagami and Niki 1994b), we presented the target stimuli either ipsilaterally or contralaterally, where the cell showed the largest change in activity during preliminary investigation.

We recorded from four hemispheres in two Japanese monkeys (*Macaca fuscata*): *monkeys EC* and *FR*. The training history of the monkeys is described in Lauwereyns et al. (2000). After completion of the training, a head-holding device and a unit-recording chamber were implanted with standard surgical techniques under pentobarbital sodium anesthesia. During single-unit recording, the monkey's head was restrained, and a hydraulic microdrive (Narishige, MO-90) was attached to the chamber. A glass-coated elgiloy microelectrode with 10- to 15-μm tip exposure was used for unit recording. Action potentials were identified using a dual-voltage, time-window discriminator and were stored on computer at 1-kHz sampling rate. For detailed information on the histological procedures, see Sakagami and Tsutsui (1999).

All surgical and experimental protocols were approved by the Animal Care and Use Committees at Juntendo University and were in accordance with the Guidelines for the Care and Use of Laboratory Animals of the National Institutes of Health.

Data analysis

Trials in which the monkey made an incorrect manual go or no-go response were eliminated from analyses. To analyze cell activity, two-factor ANOVA (color $\times$ motion direction) was applied to the responses of each neuron (100- to 400-ms period from target onset) separately for each discrimination condition. This time window was chosen as it showed the highest discrimination between go and no-go trials for the entire population of recorded neurons; the limit of 400 ms also ensured that the neuronal responses were not confounded with eye movement or retinal eccentricity (as the monkey's gaze was restricted to the fixation point up to 500 ms after target onset).

Based on the ANOVA results, we selected cells that could discriminate between go and no-go targets based on the relevant feature in at least one discrimination condition. If the neuron was recorded with both stimulus sets, we used the set to which the neuron showed the largest differential response. Specifically, we selected cells that produced a statistically reliable main effect ($P < 0.05$) of color in the color condition (color-feature cells), of motion in the motion condition (motion-feature cells), or both (integration cells). Cells were considered to show interference from an irrelevant feature (e.g., motion direction in the color condition) if there was a significant main effect of the irrelevant feature or if there was a significant two-way interaction effect between the irrelevant and relevant feature. Post hoc tests consisted of two-tailed *t*-tests.

To characterize the direction of the neuronal responses, we compared the discrimination of one type of visual information (e.g., color) when the monkey was required to process this information (e.g., color in the color condition) versus when it was required to ignore this information (e.g., color in the motion condition). This comparison allows us to evaluate the similarity between the neuronal representation of relevant and irrelevant information. We calculated a relevant color index based on data from the color condition as follows

$$RC = \frac{C_{go} - C_{ng}}{C_{go} + C_{ng}}$$

RC refers to relevant color index; C_{go} refers to the average spike rate from 100 to 400 ms after target onset in case of a go-indicating color; and C_{ng} refers to the average in case of a no-go-indicating color. Similarly we computed an irrelevant color index (*IC*) based on data from the motion condition

$$IC = \frac{C_{go} - C_{ng}}{C_{go} + C_{ng}}$$

Note that the monkey should ignore the go or no-go meaning of the target's color in the motion condition. In the same way, we calculated a relevant and an irrelevant motion index (*RM* and *IM*, respectively) based on the spike rates in response to go- versus no-go-indicating motion directions. The relationship between the relevant and irrelevant indices was evaluated at the population level by computing Pearson correlation coefficients, which were then evaluated against zero by means of two-tailed *t*-test. To estimate the linear relation between the relevant and irrelevant indices, we used principal component analysis.

To examine the temporal properties of the neuronal responses to irrelevant information, we made sliding population histograms from 200 ms before to 400 ms after target onset. The sliding histograms were computed separately for color-feature, motion-feature, and integration cells. We calculated the three-point smoothed population average of the (go – ng) discrimination values in time epochs of 10 ms. To combine the data from cells showing a "go preference" with data from cells showing a "no-go preference," we reversed the sign of the index for the latter type of cells. From this population average, we subtracted the average (go – ng) value (per 10 ms) during the precue period, that is, at a time when differential values cannot reflect perceptual discrimination. This was done by way of control because there were different precue levels among the color-feature, motion-feature, and integration cells. For each population of cells, the precue level was estimated per 10 ms based on the data from −500 to 0 ms before target onset. To determine the onset latency of the irrelevant-feature discrimination, we used running one-tailed *t*-tests to check at which moment there were two consecutive time epochs in which the corrected (go – ng) value reliably exceeded zero.

RESULTS

Behavioral performance

Both monkeys performed the manual go/no-go task with very high accuracy, reaching a grand average of over 94% correct responses, with better performance ($P < 0.01$) in the color condition (94.9%) than in the motion condition (93.4%). With the delay between target presentation and cue to respond, and so without time pressure, the monkeys were able to optimize their decision strategy, making correct responses despite the irrelevant stimulus-response associations: The congruency effect between relevant and irrelevant features (i.e., error rates on incongruent versus congruent trials) was 0.2% in the color condition (not significant) and 1.2% in the motion condition ($P < 0.01$). As a consequence, for any given neuron, there were not enough error trials to compute a correlation between the neuron's response to irrelevant information and the probability of making an error. We therefore decided to exclude error trials from further analyses.

Database of neurons

We explored the lateral part of prefrontal cortex (see Fig. 1*C* for indications of electrode penetrations in one hemisphere). A total of 436 neurons discriminated reliably ($P < 0.05$, ANOVA) between go- and no-go-indicating targets based on color or motion direction or both. Among these, 162 neurons (37.2%) were influenced by irrelevant information, as indicated by a significant main effect of, or interaction with, the target feature that the monkey should ignore (see METHODS). That is to say, "irrelevant information" refers to the color feature in the motion condition, and the motion feature in the color condition.

Specifically, 139 neurons (31.9% of the total population) showed go/no-go discrimination in both discrimination conditions (integration cells); among these, 53 neurons (38.1%) were influenced by irrelevant information. There were 192 neurons (44.0% of the total population) that showed go/no-go discrimination only based on color (color-feature cells); among these, 71 neurons (37.0%) were influenced by irrelevant information. There were 105 neurons (24.1% of the total population) that showed go/no-go discrimination only based on motion direction (motion-feature cells); among these, 38 neurons (36.2%) were influenced by irrelevant information.

Thus irrelevant target information is processed quite extensively in lateral prefrontal cortex even though the monkey succeeds in making correct go or no-go responses. The effects of irrelevant target information, however, are different for the three functionally dissociated groups of prefrontal neurons. In this article, we aim to characterize these different neuronal responses to irrelevant visual features.

Congruency effect from irrelevant information in integration cells

Integration cells are able to interpret information from different visual dimensions. Figure 2 presents histograms and rasters of two prefrontal neurons that were classified as integration cells because they showed reliable differential responses to go- and no-go-indicating stimuli in both discrimination conditions. The cell shown at the top (Fig. 2*A*) is a good representation of the majority ($n = 86$) of integration cells, firing differentially for go- and no-go-indicating stimuli in the color condition (main effect of color, $P < 0.01$) as well as in the motion condition (main effect of motion, $P < 0.01$) without reliable effects from the irrelevant feature. Looking more closely at the firing rates especially in go trials, however, there did appear a tendency for stronger activity in congruent go trials than in incongruent go trials ($P < 0.10$).

Figure 2*B* presents an integration cell in which the influence from irrelevant information was much more pronounced ($n = 53$). Overall, this unit discriminated reliably between go and no-go trials in the color condition (main effect of color, $P < 0.05$) as well as in the motion condition (main effect of motion, $P < 0.05$), but the unit was influenced also by the congruency of the irrelevant feature. In the color condition, there was a significant main effect of the irrelevant motion feature ($P < 0.01$); in the motion condition, color caused a nonsignificant trend ($P < 0.10$).

The neuron's activity was observed while the monkey succeeded in making a correct manual response in each trial regardless of the irrelevant information. Yet, in both condi-
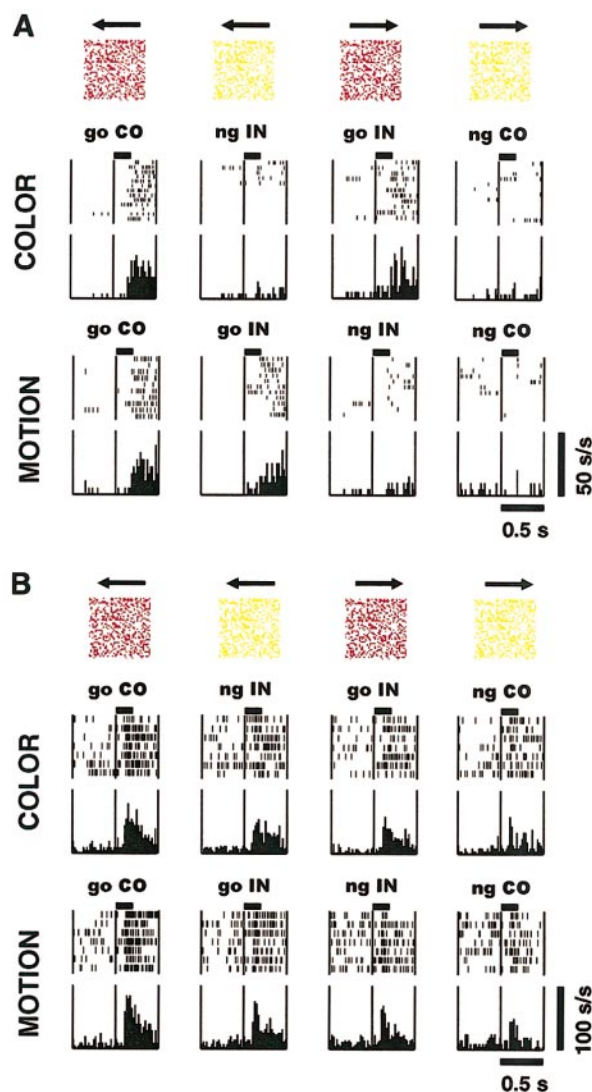

FIG. 2. *A*: the activity pattern of an integration cell unaffected by irrelevant information. Each pair of histogram and rasters illustrates the neuronal response in the type of trial indicated at the top (go CO, required go response with congruent irrelevant feature; go IN, required go response with incongruent irrelevant feature; ng IN, required no-go response with incongruent irrelevant feature; ng CO, required no-go response with congruent irrelevant feature). *Top*: the neuronal activity in the color condition; *bottom*: that in the motion condition. The vertical midline in each pair of histogram and rasters indicates target onset. The horizontal bar above the rasters indicates the target duration (200 ms). Bin width is 20 ms. *B*: the activity pattern of an integration cell that was influenced by irrelevant information.

tions, the neuronal go/no-go discrimination was degraded in case the two target features were incongruent. Thus the activity of this unit shows a blurred or suboptimal discrimination in case of conflict between the required response and the response primed by the irrelevant feature. As such, the direction of the influence from irrelevant information on the visual activity of this neuron reflects a general tendency of the population of neurons (see the population analyses in the following text).

Nonadaptive responses to one visual dimension in feature cells

In addition to integration cells, we found many neurons that were able to distinguish between go- and no-go-indicating

stimuli in only one discrimination condition (color- and motion-feature cells). As with integration cells, the majority of color- and motion-feature cells showed no significant response to irrelevant information. Figure 3A, for instance, shows a motion-feature cell that discriminated reliably between go and no-go trials in the motion condition (main effect of motion, $P < 0.01$) but not in the color condition (no effect of color; $F < 1$). The irrelevant target features did not lead to significant effects in the 2×2 ANOVA, neither in the color nor in the motion condition (we observed 67 cells of this type).

Although this motion-feature cell generally transmitted task-relevant information about motion, there was a small trend of influence from irrelevant information in its firing rate in the color condition (tendency toward a main effect of motion, $P < 0.10$). Specifically, the firing rate in congruent go trials was slightly higher than in incongruent no-go trials ($P < 0.05$). Thus the motion-feature cell did not remain entirely neutral in

the color condition, leaking some information pertaining to a task-irrelevant visual dimension. Such leakage of irrelevant information led to significant effects of irrelevant information in about one-third of the population of motion-feature cells.

Figure 3B presents an example of a motion-feature cell influenced by irrelevant information ($n = 38$). The cell fired differentially for go- and no-go targets in the motion condition (main effect of motion, $P < 0.01$), but not in the color condition (no effect of color; $F < 1$). In the color condition, however, the firing rate of this neuron was still determined by the motion direction of the target (main effect of motion, $P < 0.01$), even though the monkey successfully disregarded the motion direction to make the appropriate manual response. This neuron, then, seems to encode the target's motion direction regardless of the discrimination condition and so fails to adapt to the requirements of the task.

Similarly, Fig. 4, A and B, shows color-feature cells that are unable to fully adapt to the task requirements. Both cells reliably discriminated color (main effects of color, $P < 0.01$) in the color condition but also in the motion condition. We observed 71 cells of this type, whereas 121 color-feature cells responded to color information only when the monkey was required to interpret color. The cells illustrate very well the large range of irrelevant responses we found in color-feature cells.

Specifically, the cell in Fig. 4A showed some task-dependent modulation. In the color condition, this cell fired for yellow targets regardless of the motion direction (i.e., no reliable difference between ng IN and ng CO trials), whereas in the motion condition this cell fired more ($P < 0.05$) for yellow rightward moving targets (ng CO trials) than for yellow leftward moving targets (ng IN trials). Thus in the motion condition this cell suppressed its activity for yellow in case the irrelevant color-based S-R association (yellow = no-go) primed a different response than the relevant motion-based S-R association (leftward motion = go; ng IN trials in the motion condition). The fact that this cell changed its behavior depending on the task is further underscored by its activity in the waiting period right before the appearance of the target: the cell had a stronger anticipatory or background firing rate in the color condition than in the motion condition ($P < 0.01$). Subtle condition-dependent changes in background activity were quite common (25–30% of task-related neurons) (see also Sakagami and Niki 1994a). Typically, such cells changed their background activity slightly when the task required discriminating the preferred visual dimension. The cell shown in Fig. 4B, on the other hand, fired phasically for a yellow color, regardless of the discrimination condition, and without any changes in the background activity.
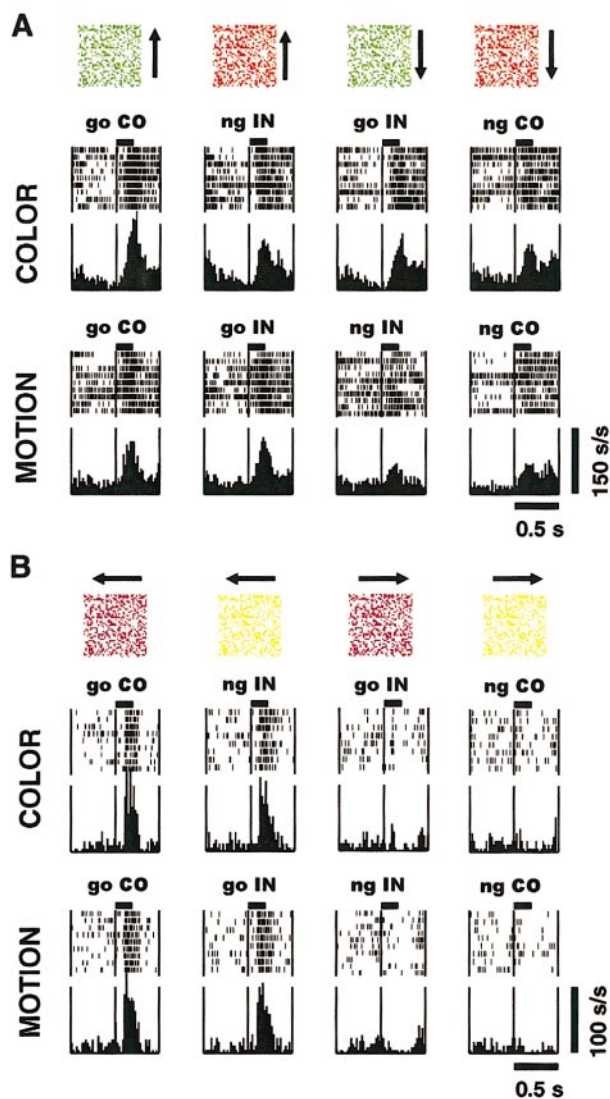

FIG. 3. A: the activity pattern of a motion-feature cell, showing reliable go/no-go discrimination when motion is relevant (in the motion condition) but not when motion is irrelevant (in the color condition). The presentation format is the same as in Fig. 2. B: the activity pattern of a motion-feature cell with a nonadaptive response to motion, both when motion is relevant (in the motion condition) and when motion is irrelevant (in the color condition).

Direction of the responses to irrelevant information

The effects of irrelevant information in the neurons shown in Figs. 2B, 3B, and 4, A and B, exhibit a common direction. The neurons' firing rate to a particular irrelevant feature shows the same go/no-go preference as when that feature is relevant to the monkey's task. Specifically, the neurons shown in Figs. 2B and 3B delivered more spikes in response to a go- than to a no-go-indicating motion direction (i.e., "go preference") when motion was relevant (in the motion condition) but also when motion was irrelevant (in the color condition). The neurons
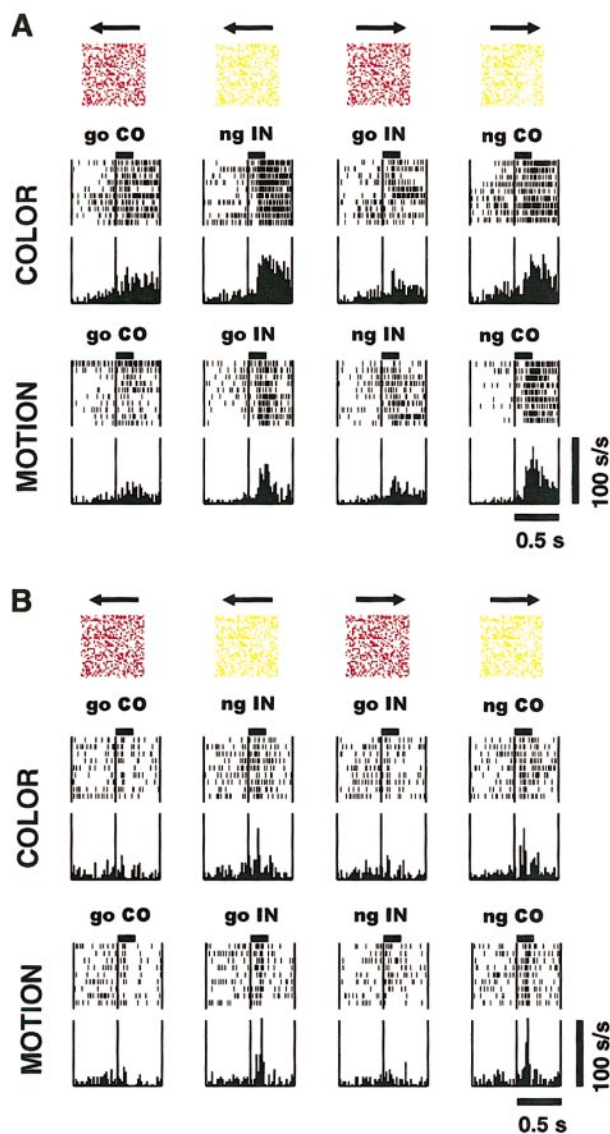
FIG. 4. *A*: the activity pattern of a color-feature cell, showing reliable go/no-go discrimination when color is relevant (in the color condition), but also, be it less clearly, when color is irrelevant (in the motion condition). The presentation format is the same as in Fig. 2. *B*: the activity pattern of a color-feature cell with a nonadaptive response to color, both when motion is relevant (in the motion condition) and when motion is irrelevant (in the color condition).

shown in Fig. 4, *A* and *B*, on the other hand, consistently preferred the no-go-indicating color across conditions.

To confirm this observation at the population level, we computed color and motion indices in both discrimination conditions (see METHODS). With these indices, we could examine the relation between a neuron's sensitivity to a particular visual dimension when this dimension was relevant versus irrelevant to the task.

Figure 5, *top,* presents the color indices (RC and IC) for three populations of neurons: color-feature cells (*left*), motion-feature cells (*middle*), and integration cells (*right*). Each point represents one neuron; the horizontal coordinate is determined by the color index in the color condition (RC, when color constitutes relevant information, indicated by a thick black bar), whereas the vertical coordinate is deter-

mined by the color index in the motion condition (IC, when color is irrelevant).

There was a positive correlation between RC and IC for color-feature cells ($r^2 = 0.244$; significantly different from 0, $P < 0.01$; Fig. 5, *top left*), confirming that the effect of irrelevant color on the neurons' activity showed the same direction as the effect of color when it was the relevant dimension. There was also a significant positive correlation between RC and IC with integration cells ($r^2 = 0.200$; significantly different from 0, $P < 0.01$; Fig. 5, *top right*) but not with motion-feature cells ($r^2 = 0.057$; not different from 0; Fig. 5, *top middle*).

In the same way as for color, Fig. 5, *bottom,* presents the motion indices (RM and IM) for the three populations of cells. There was a positive correlation between RM and IM for motion-feature cells ($r^2 = 0.337$; significantly different from 0, $P < 0.01$; Fig. 5, *bottom middle*), confirming once again that the effect of the irrelevant feature on the neurons' activity showed the same direction as the effect of the same feature when it was relevant. There was also a significant positive correlation between RM and IM with Integration cells ($r^2 = 0.348$; significantly different from 0, $P < 0.01$; Fig. 5, *bottom right*), and, be it less pronounced, with color-feature cells ($r^2 = 0.142$; significantly different from 0, $P < 0.01$; Fig. 5, *bottom left*).

Together, these positive correlations indicate that responses to irrelevant information can be characterized as a residual of S-R associations from a different context. In other words, the influence seems to be due to the neurons' inability to entirely suppress their go/no-go preference of features that are presently irrelevant to the monkey's task.

Temporal properties of responses to irrelevant information

To understand how the responses to irrelevant information develop over time, we made sliding population histograms, separately for different types of cells. The histograms are based on the running population average of irrelevant-discrimination indices (see METHODS).

Figure 6, *top,* shows the histograms of color discrimination when it is irrelevant, that is, in the motion condition. The histograms are shown for the two populations of cells that are responsive to color information, that is, color-feature cells and integration cells. Color-feature cells show a relatively sharp and fast discrimination, with a latency of 80 ms after target onset and a peak between 150 and 250 ms after target onset, followed by a gradual dissipation of the response. The curve of the irrelevant discrimination of integration cells, on the other hand, shows a slower and generally smaller response, with an onset latency of 90 ms after target onset and with no discernable peak in the response.

Similarly, Fig. 6, *bottom,* shows the histograms of responses to irrelevant motion in the color condition. The curve of the irrelevant responses of motion-feature cells appears relatively phasic, with an onset latency of 120 ms after target onset and a peak at 170 ms after target onset, followed by a gradual dissipation of the response. And again, the curve of the Integration cells shows a slower and smaller response, with an onset latency of 130 ms after target onset and with no discernable peak in the response.
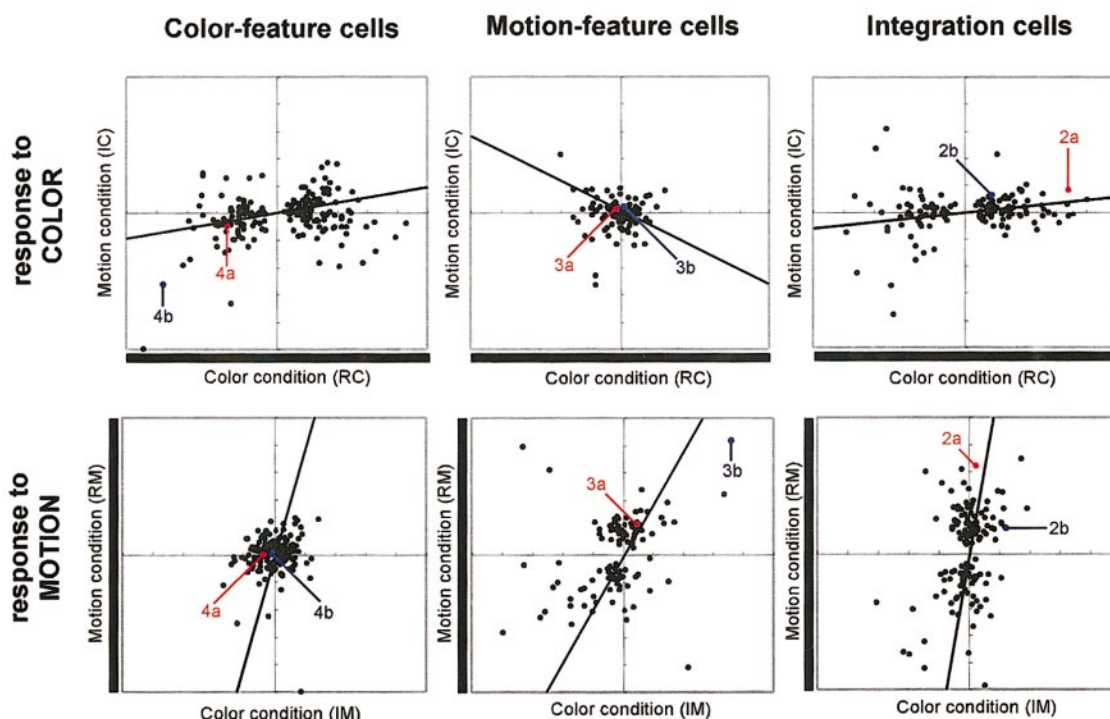
FIG. 5. The neurons' responsiveness to the same visual dimension in different discrimination conditions. In all box plots, each point represents 1 neuron; the horizontal coordinate is determined by the neuronal response in the color condition, whereas the vertical coordinate is determined by the response in the motion condition. *Top*: the neuronal discrimination of color [irrelevant vs. relevant color index (IC vs. RC)]; *bottom*: the neuronal discrimination of motion [relevant vs. irrelevant motion index (RM vs. IM)]. Thick black bars below or to the left of the plots indicate the axis of relevant discrimination. The data are shown separately for color-feature cells ($n = 192$; *left*), motion-feature cells ($n = 105$; *middle*), and integration cells ($n = 139$; *right*). The indices (ranging from -1 to $+1$) are based on the difference of the firing rate to a go-indicating feature minus the firing rate to a no-go-indicating feature in a time window from 100 to 400 ms after target onset, calibrated by the sum of these firing rates (see METHODS). The data points of cells shown in Figs. 2A–4B are indicated in red for cells unaffected by irrelevant information, and in blue for cells affected by irrelevant information. In each box plot, the diagonal line represents the orientation of the 1st component from principal component analysis.

Anatomical locations of different types of cells

Figure 1*C* indicates the locations of electrode penetrations in the primate prefrontal cortex (example from the left hemisphere of *monkey EC;* similar distributions were obtained in the other hemispheres, not shown here). Considering cells that are unaffected by irrelevant information (red dots), we found an anatomical segregation: color-feature cells appeared mainly ventral to the principal sulcus in areas 46 and the upper part of area 12, whereas motion-feature cells tended to be located dorsal to the principal sulcus in areas 46 and 8A. These results are consistent with anatomical data on the connections between color-sensitive areas in inferotemporal cortex and ventrolateral prefrontal cortex (Barbas 1988; Ungerleider et al. 1989) and the connections between motion-sensitive areas in parietal cortex and peri-arcuate prefrontal cortex (Andersen et al. 1990; Schall et al. 1995b). Integration cells appeared in area 8A, both in the dorsal and ventral sectors, as well as ventral to the principal sulcus.

Considering cells that are affected by irrelevant information (blue circles), it appears that the same anatomical segregation holds true for color-feature cells and for integration cells. With motion-feature cells, the segregation was less clear.

DISCUSSION

Our data are the first to show prefrontal representations of irrelevant information in a feature-discrimination task. Using a manual go/no-go task with two Japanese monkeys, we found many prefrontal neurons that were able to code the behavioral meaning of targets based on color and/or motion direction. More than one-third of these neurons, however, also coded S-R associations with presently irrelevant information even though the monkey successfully ignored the irrelevant information in its behavior.

Typically, the responses to irrelevant information appeared as the neuron's inability to completely adapt to the changed requirements of the task. For instance, color-feature cells (see Fig. 4, *A* and *B*) were able to discriminate between go and no-go targets in the color condition, but tended still to leak information about color even when this feature was irrelevant, that is, in the motion condition. Thus relevant and irrelevant S-R features appear to run in parallel in the brain even up to the stage of decision making, for which prefrontal cortex is presumed to be responsible (Kim and Shadlen 1999; Sakagami and Tsutsui 1999).

Among the neurons that were influenced by irrelevant information, the effects ranged from complete interaction between relevant and irrelevant information (such as with the neuron shown in Fig. 2*B*) to nonadaptive responses to only one visual dimension, regardless of whether this dimension is relevant to the monkey's discrimination task (such as with the neurons shown in Figs. 3*B* and 4*B*). These different types of neuronal activity may reflect distinct stages of visual interpretation.
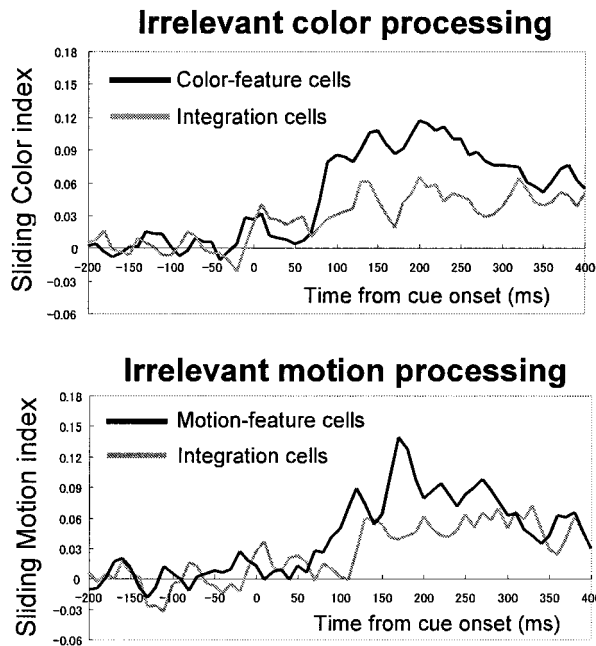
FIG. 6. Sliding population histograms of neuronal responses to irrelevant information. The top panel shows the population average of activity to irrelevant color information in the Motion condition. The histograms are shown separately for the two populations of cells that are responsive to color: Color-feature cells (black curve; $n = 192$), and Integration cells (gray curve; $n = 139$). Similarly, the bottom panel shows the activity to irrelevant motion information in the Color condition, separately for Motion-feature cells (black curve; $n = 105$) and for Integration cells (gray curve; $n = 139$). The population histograms were obtained after 3-point smoothing of the (go – ng) values of irrelevant information (Color index in the Motion condition; Motion index in the Color condition) in time epochs of 10 ms minus control [i.e., the average (go – ng) value during the precue period; see METHODS]. For cells with a higher firing rate to no-go-indicating stimuli we reversed the sign of the index.

Specifically, nonadaptive responsiveness to one visual dimension seems to be derived from a purely sensory neural code and so could reflect the input from extrastriate and/or association visual areas such as V4 and IT for color information and MT and MST for motion direction. The interaction between relevant and irrelevant information in integration cells, on the other hand, can be characterized as a congruency effect toward the output side of the decision-making process.

Sequential mode of information processing in prefrontal cortex

The conceptual scheme in Fig. 7 shows how responses to irrelevant features could be gated through prefrontal cortex. We propose that the conversion of visual information into appropriate behavior is a hierarchically organized decision-making process (Sakagami and Tsutsui 1999), in which feature-selective cells (i.e., color- and motion-feature cells) generate behavioral significance based on specific sensory properties and send their output to integration cells, which in turn encode the appropriate behavioral action. This process is represented in Fig. 7 with information flowing from independent sensory modules through feature units toward integration units, which in turn influence motor preparation.

In this scheme, it is presumed that long-term training enables fixed or automatic associations between representations in independent visual modules and feature-selective representations

of behavioral meaning. By this interpretation, experience with the task leads to automation of stimulus-response associations (Logan 1988). The notion that experience can alter visual processing in prefrontal cortex was established already in visual-interpretation tasks to guide manual behavior (Niki et al. 1990; Rainer and Miller 2000) or oculomotor behavior (Bichot and Schall 1999; Bichot et al. 1996). Such training effects can explain why feature-selective cells respond to one visual dimension (e.g., color) even when it is presently irrelevant to the task. As a consequence, these learning mechanisms would lead to a competition between relevant and irrelevant S-R associations.
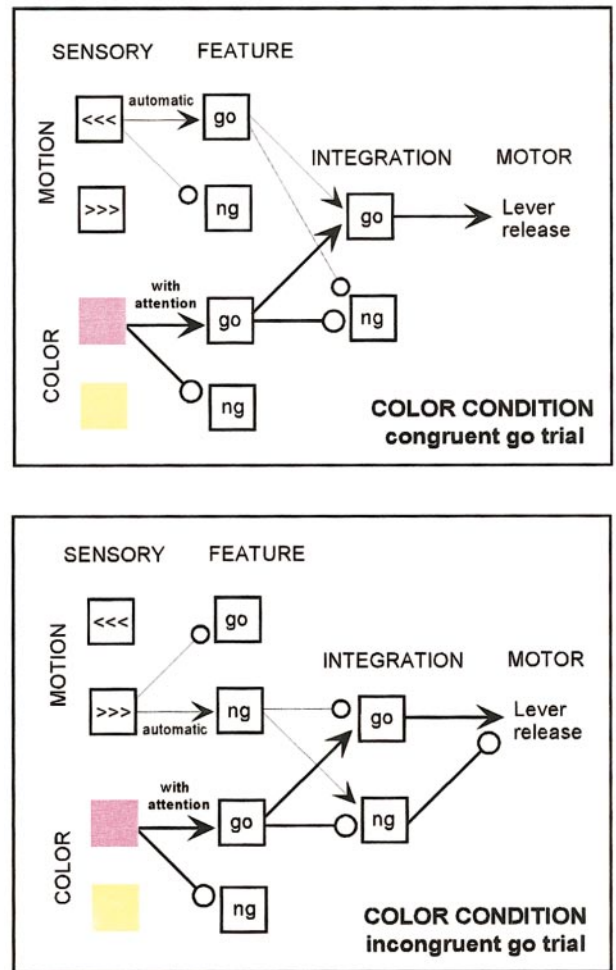

FIG. 7. Conceptual scheme of the responses to irrelevant visual information, exemplified by comparison of a congruent (*top*) and an incongruent (*bottom*) go trial in the color condition. The information flows from left to right, from independent sensory modules through feature-selective units toward integration units, which in turn control motor preparation. Excitatory connections are indicated with lines ending in arrowheads; inhibitory connections are indicated with lines ending in circles. Depending on the task demands, information processing is selectively enhanced for only the relevant visual pathway (in this example, the color pathway; indicated by thicker lines than in the motion pathway). Through the fixed associations, both relevant (thick lines) and irrelevant (thin lines) information arrives in the integration module. Because of the attentional enhancement, the relevant information (i.e., purple = go) outweighs the irrelevant information. However, the irrelevant information does influence the decision-making process to some extent, leading to less efficient discrimination when the irrelevant feature primes a different answer (rightward motion = no-go, *bottom*) as compared with when it primes the same answer (leftward motion = go, *top*).

Depending on the attentional demands, we assume that information processing is selectively enhanced for only the relevant visual pathway (e.g., in Fig. 7, the color pathway). Through the fixed associations, however, both relevant and irrelevant information travels from the feature modules to the integration module. Because of the attentional modulation (either enhancement or suppression), relevant information on average outweighs the irrelevant information. In this way, prefrontal circuits appear to be organized so that they can filter out irrelevant information during the decision-making process, leading to smaller responses to irrelevant features in the later than in the earlier stages of decision making (i.e., smaller irrelevant responses by integration cells, see Fig. 6). Yet the irrelevant information influences even integration cells to some extent, leading to less efficient discrimination when the irrelevant feature primes a different answer than the relevant feature (*bottom*) as compared with when both features prime the same answer (*top*).

The notion that decision-making in prefrontal cortex is indeed organized hierarchically finds support in the anatomical segregation of different types of cells (Fig. 1*C*) as well as the temporal differences between color- and motion-feature cells on the one hand and integration cells on the other hand (Fig. 6). Color- and motion-feature cells showed a more phasic and slightly faster response to the irrelevant feature, consistent with their presumed function toward the perceptual or input side of the decision-making process, whereas integration cells showed a more sustained and slower response, consistent with their presumed function toward the motor-preparation or output side of the decision-making process (see also Sakagami and Tsutsui 1999).

Another finding that supports the notion of a hierarchical organization is that integration cells show a high consistency in their preference for either go- or no-go-indicating features across both visual dimensions: Out of 139 integration cells, 75 cells (53.9%) consistently preferred go-indicating stimuli; 51 cells (36.7%) consistently preferred no-go-indicating stimuli; and only 13 cells (9.4%) showed a different preference in the color than in the motion dimension. This observation suggests that integration cells do indeed integrate behaviorally relevant information from multiple dimensions rather than showing a random combination of tuning curves to independent sensory features.

In sum, the activity of the neural population as a whole includes both sensory-derived input and behaviorally relevant output signals as should be expected from a neural substrate of decision making (Kim and Shadlen 1999; Leon and Shadlen 1998; Zhang et al. 1997). In this population, however, irrelevant stimulus representations compete with relevant representations throughout the entire decision-making process. Even toward the output side of this process, the irrelevant stimulus representations are still strong enough to influence the activity of Integration cells. Thus the cells showing partial responses to irrelevant features appear to be part of a network that takes sensory input and turns it into a decision output. As such, these cells could be merely intermediate in the computation process—that is, they could be hidden units in a multilayer neural network. Alternatively, the partial responses to irrelevant features could represent interference, which in turn may disturb the monkey's decision-making behavior.

In this regard, further research is needed to investigate to what extent the representations of irrelevant information in prefrontal neuronal codes cause interference on behavioral performance. The two monkeys in the present study showed clear interference effects in their behavioral reaction times in a speeded version of the task with otherwise exactly the same experimental set-up, whereas with the delayed version of the task we found interference effects in the error rates in the motion condition but not in the color condition. Reaction times are a more sensitive behavioral measure of interference effects than error rates because they can be related to the decision process rather than the decision outcome (MacLeod 1991). Given that the present study establishes that there exist responses to irrelevant features in prefrontal cortex at a cognitive stage dissociable from motor control, the next step in this research should be to estimate the influence of such irrelevant prefrontal neuronal codes on behavior. This can be done, for instance, by recording single-cell activity during a speeded discrimination task (Lauwereyns et al. 2000) so that trial-by-trial neuronal signals can be correlated to behavioral reaction times.

A related matter is the question of the relationship of the prefrontal neuronal code to different types of motor control. We used fixed one-to-one mapping between visual features and manual responses in the present task. Consequently, we cannot indicate whether the irrelevant neuronal activity pertains to the behavioral meaning of the visual features (go or no-go) or whether the irrelevant activity is more tightly linked to the response dimension (manual lever release). To tease apart these two possibilities, future research should examine the influence of irrelevant features on neuronal activity in situations with variable instead of fixed stimulus-response mapping. This issue can be resolved, for instance, by comparing the perceptual decision-making process of the same neurons in manual versus oculomotor tasks. In this respect, it is interesting to note that the present manual go/no-go task revealed neurons with multidimensional discriminative activity that could not be reduced to oculomotor activity in the frontal eye field. Our data, which were obtained while the monkey gazed at a fixation spot, suggest that the process of perceptual decision-making could be more independent from motor control than has been suggested recently (Gold and Shadlen 2000).

Partial prefrontal activation by irrelevant features

Partial activation by task-irrelevant information in prefrontal cortex might simply be regarded as the corollary of the notion that there are prefrontal representations of relevant information in tasks with complex stimuli (e.g., Asaad et al. 2000; Bichot and Schall 1999; Hoshi et al. 2000; Rainer et al. 1998, 1999; Sakagami and Niki 1994a; White and Wise 1999). However, in previous studies, which were not designed to study responses to task-irrelevant features, decision-making was performed in situations where there was no irrelevant visual dimension that could imply an alternative S-R association. For instance, the study by Bichot and Schall (1999) showed history effects from previous S-R associations in a situation where the monkey performed a conjunction task, for which it had to consider both the color and the shape of the target. The partial activation by a previous S-R association, then, was derived from a task-relevant visual dimension. In other studies, there were no competing S-R associations from different dimensions. For instance, in the study by Asaad et al. (2000), the monkey performed different tasks such as a spatial task and an object task, but the stimuli changed with the task as well so that the

monkey was presented with simple dots in the spatial task (with no possibility of interference from object information) or with a reference object at the center of the screen (with no possibility of interference from spatial information). Hoshi et al. (2000) used a similar design.

One intriguing study by White and Wise (1999) did show partial activation of irrelevant features in a conflict situation between two rules. Yet one of the two rules involved spatial information, which may not be ideal as a task-irrelevant dimension, especially when the monkey has to process visual information at the same position in space. Specifically, under the *conditional* rule in the White and Wise study, the monkey was required to identify an object at a particular position, while the same position implied an alternative behavioral meaning. It could be argued that the monkey had to allocate attention to the object's position to be able to identify the object, implying that spatial information was not entirely irrelevant to the monkey's task even under the *conditional* rule. The same argument can be applied to Sakagami and Niki (1994a) or to the interference effects observed with the anti-saccade paradigm (e.g., Funahashi et al. 1993).

In contrast, in the present study, the monkey was required to discriminate one of two visual dimensions, either color or motion, while ignoring the alternative visual dimension because it carried an alternative S-R association. Thus we could examine how prefrontal cortex responds to visual information that was clearly task-irrelevant but could prime the same or a different manual response than the task-relevant information. We found phenomenally different responses to irrelevant information in three functionally segregated groups of prefrontal neurons. These responses could all be characterized as residuals of S-R associations from the alternative discrimination condition (Fig. 5). Color-feature cells leaked information about color in the motion condition, whereas motion-feature cells still responded to motion in the color condition. In integration cells, relevant and irrelevant input from different visual dimensions appeared to be combined, leading to a congruency effect in the go/no-go discrimination.

REFERENCES

ANDERSEN RA, ASANUMA C, ESSICK G, AND SIEGEL RM. Corticocortical connections of anatomically and physiologically defined subdivisions within the inferior parietal lobule. *J Comp Neurol* 296: 65–113, 1990.

ASAAD WF, RAINER G, AND MILLER EK. Neural activity in the primate prefrontal cortex during associative learning. *Neuron* 21: 1399–1407, 1998.

ASAAD WF, RAINER G, AND MILLER EK. Task-specific neural activity in the primate prefrontal cortex. *J Neurophysiol* 84: 451–459, 2000.

BARBAS H. Anatomic organization of basoventral and mediodorsal visual recipient prefrontal regions in the rhesus monkey. *J Comp Neurol* 276: 313–342, 1988.

BICHOT NP AND SCHALL JD. Effects of similarity and history on neural mechanisms of visual selection. *Nat Neurosci* 2: 549–554, 1999.

BICHOT NP, SCHALL JD, AND THOMPSON KG. Visual feature selectivity in frontal eye fields induced by experience in mature macaques. *Nature* 381: 697–699, 1996.

FUNAHASHI S, CHAFEE MV, AND GOLDMAN-RAKIC P. Prefrontal neuronal activity in rhesus monkeys performing a delayed anti-saccade task. *Nature* 365: 753–756, 1993.

FUSTER JM. *The Prefrontal Cortex.* New York: Raven, 1997.

GOLD JI AND SHADLEN MN. Representation of a perceptual decision in developing oculomotor commands. *Nature* 404: 390–394, 2000.

GOLDMAN-RAKIC PS. Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In: *Handbook of Physiology. The Nervous System. Higher Functions of the Brain.* Bethesda, MD: Am. Physiol. Soc., 1987, sect. 1, vol. V, p. 373–417.

HANES DP AND SCHALL JD. Neural control of voluntary movement initiation. *Science* 274: 427–430, 1996.

HOSHI E, SHIMA K, AND TANJI J. Neuronal activity in the primate prefrontal cortex in the process of motor selection based on two behavioral rules. *J Neurophysiol* 83: 2355–2373, 2000.

KIM JN AND SHADLEN MN. Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nat Neurosci* 2: 176–185, 1999.

LAUWEREYNS J, KOIZUMI M, SAKAGAMI M, HIKOSAKA O, KOBAYASHI S, AND TSUTSUI K. Interference from irrelevant features on visual discrimination by macaques (*Macaca fuscata*): a behavioral analogue of the human Stroop effect. *J Exp Psychol Anim Behav Process* 26: 352–357, 2000.

LEON MI AND SHADLEN MN. Exploring the neurophysiology of decisions. *Neuron* 21: 669–672, 1998.

LOGAN GD. Toward an instance theory of automatization. *Psychol Rev* 95: 492–527, 1988.

MACLEOD CM. Half a century of research on the Stroop effect: an integrative review. *Psychol Bull* 109: 163–203, 1991.

NIKI H. Prefrontal activity during delayed alternation in the monkey. I. Relation to the direction of the response. *Brain Res* 68: 185–196, 1974.

NIKI H, SUGITA S, AND WATANABE M. Modification of the activity of primate frontal neurons during learning of a GO/NO-GO discrimination task and its reversal: a progress report. In: *Vision, Memory, and the Temporal Lobe,* edited by Iwai E and Mishkin M. New York: Elsevier, 1990, p. 295–304.

PASSINGHAM RE. *The Frontal Lobes and Voluntary Action.* New York: Oxford, 1993.

RAINER G, ASAAD WF, AND MILLER EK. Selective representation of relevant information by neurons in the primate prefrontal cortex. *Nature* 393: 577–579, 1998.

RAINER G AND MILLER EK. Effects of visual experience on the representation of objects in the prefrontal cortex. *Neuron* 27: 179–189, 2000.

RAINER G, RAO SC, AND MILLER EK. Prospective coding for objects in primate prefrontal cortex. *J Neurosci* 19: 5493–5505, 1999.

SAKAGAMI M AND NIKI H. Encoding of behavioral significance of visual stimuli by primate prefrontal neurons: relation to relevant task conditions. *Exp Brain Res* 97: 423–436, 1994a.

SAKAGAMI M AND NIKI H. Spatial selectivity of go/no-go neurons in monkey prefrontal cortex. *Exp Brain Res* 100: 165–169, 1994b.

SAKAGAMI M AND TSUTSUI K. The hierarchical organization of decision making in primate prefrontal cortex. *Neurosci Res* 34: 79–88, 1999.

SAKAGAMI M, TSUTSUI K, LAUWEREYNS J, KOIZUMI M, KOBAYASHI S, AND HIKOSAKA O. A code for behavioral inhibition on the basis of color, but not motion, in ventrolateral prefrontal cortex of macaque monkey. *J Neurosci* 21: 4801–4808, 2001.

SCHALL JD, HANES DP, THOMPSON KG, AND KING DJ. Saccade target selection in frontal eye field of macaque. I. Visual and premovement activation. *J Neurosci* 15: 6905–6918, 1995a.

SCHALL JD, MOREL A, KING DJ, AND BULLIER J. Topography of visual cortex connections with frontal eye field in macaque: convergence and segregation of processing streams. *J Neurosci* 15: 4464–4487, 1995b.

SCHALL JD AND THOMPSON KG. Neural selection and control of visually guided eye movements. *Annu Rev Neurosci* 22: 241–259, 1999.

UNGERLEIDER LG, GAFFAN D, AND PELAK VS. Projections from inferior temporal cortex to prefrontal cortex via the uncinate fascicle in rhesus monkeys. *Exp Brain Res* 76: 473–484, 1989.

WATANABE M. Prefrontal activity during delayed conditional go/no-go discrimination in the monkey. I. Relation to the stimulus. *Brain Res* 382: 1–14, 1986.

WHITE IM AND WISE SP. Rule-dependent neuronal activity in the prefrontal cortex. *Exp Brain Res* 126: 315–335, 1999.

ZHANG J, RIEHLE H, REQUIN J, AND KORNBLUM S. Dynamics of single neuron activity in monkey primary cortex related to sensorimotor transformation. *J Neurosci* 17: 2227–2246, 1997.