

Self-images in the video monitor coded by monkey intraparietal neurons

Atsushi Iriki^{a,b,c,*}, Michio Tanaka^{a,b}, Shigeru Obayashi^{a,b}, Yoshiaki Iwamura^b

^a *Section of Cognitive Neurobiology, Department of Maxillofacial Biology, Tokyo Medical and Dental University, Tokyo 113-8549, Japan*

^b *Department of Physiology, Toho University School of Medicine, Tokyo 143-8540, Japan*

^c *PRESTO, Japan Science and Technology Corporation, Tokyo 143-8540, Japan*

Received 5 December 2000; accepted 2 March 2001

Abstract

When playing a video game, or using a teleoperator system, we feel our self-image projected into the video monitor as a part of or an extension of ourselves. Here we show that such a self image is coded by bimodal (somatosensory and visual) neurons in the monkey intraparietal cortex, which have visual receptive fields (RFs) encompassing their somatosensory RFs. We earlier showed these neurons to code the schema of the hand which can be altered in accordance with psychological modification of the body image; that is, when the monkey used a rake as a tool to extend its reach, the visual RFs of these neurons elongated along the axis of the tool, as if the monkey's self image extended to the end of the tool. In the present experiment, we trained monkeys to recognize their image in a video monitor (despite the earlier general belief that monkeys are not capable of doing so), and demonstrated that the visual RF of these bimodal neurons was now projected onto the video screen so as to code the image of the hand as an extension of the self. Further, the coding of the imaged hand could intentionally be altered to match the image artificially modified in the monitor. © 2001 Elsevier Science Ireland Ltd and the Japan Neuroscience Society. All rights reserved.

Keywords: Body-image; Tool-use; Virtual reality; Conscious monkey; Single unit recordings; Bimodal neurons; Visuo-somatosensory integration

1. Introduction

When we use a virtual reality apparatus (such as teleoperator system) or when we play a video game, we feel our self-image projected onto the image of the hand or visual cues (i.e. a certain figure or an abstract point, such as cursors, which moves in accordance with the movement of our own hand) appearing in the video monitor, as a part of or an extension of our own hands (Harman et al., 1999). Essential elements of such an image of our own body should be comprised of neural representations about the dimension, posture and movement of the corresponding body parts in relation to the environmental space. Thus, its production requires integration of somatosensory (intrinsic) and

visual (extrinsic) information of our own body in space. Human patients with a lesion in the parietal cortex describe introspective experiences of distortion, neglect, or extinction of their body images (Holmes, 1918; Triggs et al., 1994; Berlucchi and Aglioti, 1997), and the monkey intraparietal cortex possesses multimodal representations of the body and the peripersonal space which are used for planning purposeful movements (Andersen et al., 1997). This intraparietal area of the cerebral cortex receives projections from the somatosensory cortex conveying intrinsic information about the current posture (Seltzer and Pandya, 1980; Iwamura, 1998) and is simultaneously fed (via a dorsal stream of visual information processing) with extrinsic information about spatial locations and movements of the objects and our own body parts (Ungerleider and Mishkin, 1982; Colby et al., 1988). Thus, it seems reasonable to assume that the body-image is created and stored as response properties of a group of neurons in this brain area.

* Corresponding author. Tel.: + 81-3-5803-5852; fax: + 81-3-5803-0186.

E-mail address: iriki.cnb@tmd.ac.jp (A. Iriki).

In our earlier report, we trained macaque monkeys to use rake-shaped tools to extend their reaching distance, and found in the intraparietal cortex a group of bimodal (somatosensory and visual) neurons which seemed to represent the image of the hand into which the tool was incorporated as its extension (Iriki et al., 1996). That is, around the somatosensory receptive field resided in the hand/forearm area was formed the visual receptive field defined as a territory in the space where a neuron responded to the moving visual stimuli. Tool-use induced an expansion of the visual receptive field only when monkeys intended to use tools to retrieve distant objects, but the modification was never induced when just holding it as an external object. Thus, this modification was not related to mere physical appearance of the tool simply held by the hand, but rather to psychological experience that the tool was assimilated to the hand. This mental process was recently confirmed to occur also in human patients' brain (Berti and Frassinetti, 2000; Farne and Ladavas, 2000; Marvita et al., 2001). We also demonstrated (Obayashi et al., 2000) that visual responses of these intraparietal bimodal neurons also represent mental (or introspective) processes to create and sustain 'subjective' body-image in the brain, not only when it is visible but also when it is invisible and creating such an image only in the brain. Thus, among other brain areas which contain neurons representing similar bimodal properties for the body parts in the direct environments (Leinonen et al., 1979; Leinonen and Nyman, 1979; Colby and Duhamel, 1991; Colby et al., 1993; Graziano and Gross, 1993, 1994; Fogassi et al., 1996; Graziano et al., 1997; Rizzolatti et al., 1997; Graziano, 1999), response property of bimodal neurons in intraparietal area is most likely to represent self-image in the monitor, which should necessitate complex mental processes for the self-images to be projected.

In the present experiment, we trained monkeys to recognize their image in a video monitor. Although it has been generally believed earlier that monkeys are not capable of doing so (Thompson and Boatright-Horowitz, 1994; Tomasello and Call, 1997), this experimental paradigm appears feasible because monkeys are known able to use the mirror to guide their movements (Itakura, 1987a,b). Using above paradigm, we attempted to prove above hypothesis by demonstrating that the visual receptive field (RF) of these bimodal neurons will be newly formed around the image of the hand projected onto the video screen so as to code the video-image of the hand as an extension of the self, and if represented images are subject to introspective manipulation independent of the state of the existing body parts in the actual space of reality. Preliminary results have been published in an abstract form (Iriki et al., 1997).

2. Materials and methods

2.1. Subjects and training

Four Japanese monkeys (*Macaca fuscata*; male, 4–7 kg) were used. After the monkeys were familiarized with the laboratory environment, the first surgery was performed (under Nembutal anesthesia; 30 mg/kg, intravenous (i.v.)) to install five screws on the skull to fix their head to the monkey chair. Then, the monkeys were trained to sit quietly on a primate chair with their head fixed in an upright position. Food pellets were placed on a table (75 × 85 cm) at the monkeys' waist height. Initially, they were trained to use a rake-shaped tool (Iriki et al., 1996) to take the food pellets when delivered on the table beyond their reaching distance. This was necessary to prevent the monkeys from just groping for the food blindly during the later training procedures. This first step of training took about 2 weeks in all four monkeys, which is common among other monkeys in which acquisition processes were closely analyzed (Ishibashi et al., 2000). Following this training, an opaque panel was installed under the monkey's eyes to mask their view, and instead of viewing the table directly, monkeys were trained to retrieve food relying entirely on the video monitor screen on which real time video image of their hand retrieving the bait was projected (Fig. 1A). This second training procedure took another 2 weeks. Neuronal recordings were conducted in one monkey after the first step of training was terminated, and in other three monkeys after the second training procedure was completed. This study was approved by the Animal Care and Use Committee in Tokyo Medical and Dental University and by Animal Research Committee in Toho University School of Medicine, and all husbandry and experimental procedures were in accordance with the Guide for the Care and Use of Laboratory Animals of the National Research Council (1996) and Guidelines for Animal Experimentation in Tokyo Medical and Dental University and Toho University School of Medicine.

2.2. Video equipments

Video equipments used in this study consisted of two video cameras, a visual effects generator, a chromakeyer, and a monitor. These pieces of equipment were connected as shown in Fig. 1B (with the video signal passing through a particular device if it was not being used at that time). The monkeys' hand/arm movements and the scanning with the visual probe on the table were captured by a small color CCD camera with fixed focal distance of 3.6 mm (Fig. 1B, 'Camera 1'; Sony, CCD-MC10) and the images were projected onto 10 inch color liquid crystal TV monitor (Panasonic, TH-10PC1) in real time.

In order to superimpose the visual stimuli in the monitor, images of the visual stimuli in front of a blue back panel were captured by the camera with variable focal distance (11–66 mm) placed outside the experimental chamber ('Camera 2'; Sony, CCD-G100ST), and then a portion of the image occupied by blue color (the key chroma) was electronically replaced, using digital chromakeyer ('Chromakeyer'; Sony, DCK-500), by the image captured by 'Camera 1' (an actual image of the monkey working on the table).

To erase the entire image except for the bright spot at the tip of the rake utilizing a 'luminancekey effect', the image captured by the 'Camera 1' was fed into digital multi-effects generator (Sony, DME-3000) and the part of the image except for the highlighted area (a white spot at the tip of the rake in this experiment) was replaced by uniform black mat screen, and then the visual stimulus was superimposed onto this altered image using the same chromakey method as described above. Other video effects such as enlargement, compression and displacement of the video image were done in real time also using the digital video effector.

2.3. Neuronal recordings

A recording chamber was installed after the first phase of training was completed. Chronic recordings were obtained from neurons in the banks of the intraparietal sulcus (IPS) of the contralateral hand/arm region. At the end of the experiments, electrode tracks and neuronal recordings sites were reconstructed from serial histological sections. Detailed techniques for in-

stalling chamber, chronic recordings, and histological reconstructions have been described elsewhere (Iwamura et al., 1993). A total of 203 electrode penetrations were attempted for four monkeys, during which 2560 single neurons were isolated. Only neurons having both somatosensory and visual RFs (bimodal neurons, $n = 261$) were analyzed.

2.4. Identification of somatosensory receptive fields

When clear single neuronal recordings were obtained, the body surface on which tactile stimulation induced spike discharges (tactile RF) was identified by palpation (light touch to) of the glabrous skin and/or hairy body surfaces using a hand-held probe, or a small-or-miniature paint brush. In a group of neurons, adequate somatosensory stimuli for activating them were passive manipulation (either extension, flexion or rotation) of joint(s) of the finger, wrist, elbow or shoulder. Some neurons could not be activated by passive somatic sensory stimulation, but were activated by active movements of the hand/arm or active touch.

2.5. Identification of visual receptive fields

The visual response properties were studied by scanning the peripersonal space by a hand held visual probe. The visual receptive field was defined as a territory in the space in which a hand held probe made neurons fire. The probe (typically sized 1 cm³, and the most effective one was a small piece of a bait) was moved freely by the experimenter's hand throughout

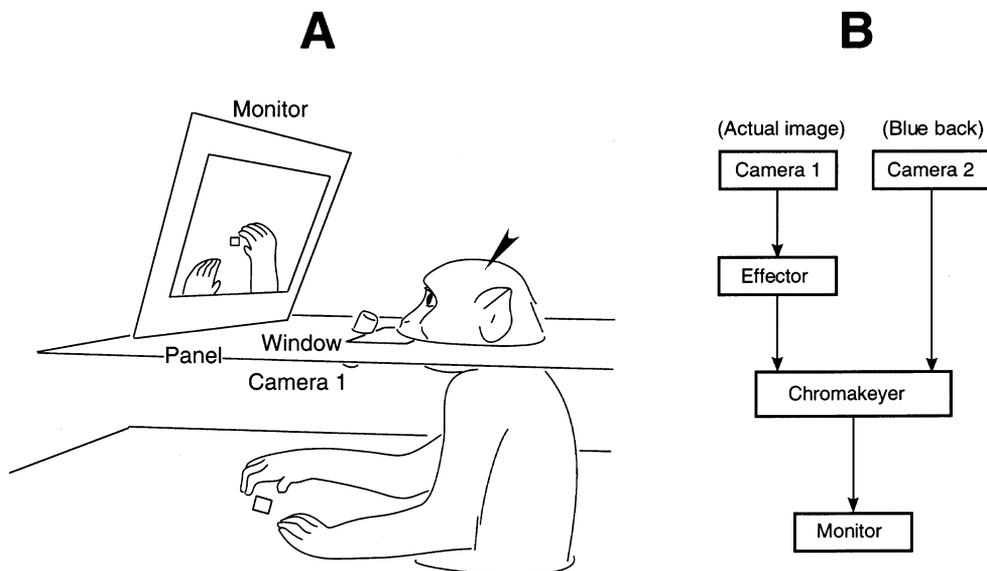


Fig. 1. Experimental setups. (A) Monkeys were trained to retrieve food by watching their hand/arm movements through a real-time video monitor or viewing directly through a window (at the near portion of the opaque panel between the monkey's eye and the table) when it is open. The arrowhead indicates that recording electrodes were penetrated to the hemisphere contralateral to the hand of which RFs were examined. (B) Connections of video equipments used (Section 2).

the space over the table in front of the monkey at a speed of about 0.5 m/s. Positions of the probe were measured (every 15–30 ms) by 3-dimensional electromagnetic tracking device (Polhemus, 3space Fastrak, spatial resolution, 0.8 mm). For the majority of bimodal neurons, the optimal movement of the probe for inducing visual responses was one that approached the somatosensory RF of the neuron. Other movements of the probe induced hardly any discharges. In order to avoid a recording bias, in which the density of scanning trajectory becomes uneven throughout space, points were placed in the scattergram only when the neurons fired at instantaneous firing frequency above twice their spontaneous firing rate. With this procedure, no RF-like crowding of data points were formed at the center of the scanning trajectories (the location of somatosensory RF) when there were no visually induced responses. During the recording of visual responses, the monkeys' hand was settled unmoved on the table. The gaze direction during scanning the space for the visual receptive field identification was determined (every 70–90 ms) by monitoring the center of gravity of the pupil image, through a video camera under infrared illumination (wavelength 840 nm), and analyzed by an image analyzer (Omron 3Z4SP-C22, spatial resolution of 10 pixels per mm).

3. Results

3.1. Utilization of self-image in the video monitor as a reference of hand movement

Monkeys sat in front of a table, and an opaque panel was installed under the eyes to mask their view (Fig. 1A) with a little window, which allowed direct viewing when it was open. Monkeys were trained to retrieve food (either with or without using a rake) relying entirely on a real-time video image shown on the monitor. For acquisition of this skill, monkey's hand-movement had to be displayed on the video monitor without any time delay. Thus, the coincidence of the movement of the real hand and the video-image of the hand seemed to be essential for training the monkey to use the video-image to guide their hand movements. After above training was completed, the monkeys exhibited the behavior suggesting that they became to recognize the image of their body parts in the video monitor as a part of their own body. That is, when they saw a spot of the laser pointer move onto their hand on the monitor, all the trained monkey tried to touch and pick up the spot with the other hand; a behavior that fulfills Gallup's criterion (Gallup, 1970, 1977) for mirror-self recognition, (i.e. an attempt to remove a spot marked on the face when that spot is first seen in a mirror). Also, when the experimenter displayed stimuli that the

monkeys dislike (such as a snake, spider or frogs) which could be seen only in the monitor on which the image of their hand is projected, monkeys expressed the facial expression of fear and threat (van Hooff, 1962; Redican, 1975), and tried to take their hands away from the disturbing stimuli. This would further support above idea that the monkeys have recognized the image projected in the video monitor as the image of themselves.

3.2. Parietal coding of self-image in the video monitor

Visual and somatosensory RF properties of 261 bimodal neurons recorded from intraparietal area of four trained monkeys were analyzed. Fig. 2 illustrates visual (A–C) and somatosensory (D) response properties of one representative bimodal neuron, recorded from the anterior bank of the left IPS. In the left column (A) are shown visual responses when the monkey could see his right hand directly through the open window. This neuron was activated when the scanning probe (a small visual target) was moved towards the monkey's hand but did not touch it (top panel, trajectories of the scanning probe with the image of monkey's hand are projected on a horizontal plane). Each data point in the middle panel (scatter map plotted on the same coordinates) represents the location of the probe when a spike discharged at an instantaneous firing rates > 3 Hz (twice the spontaneous firing rate of this cell). The densest area of excitation (namely, the visual RF) encompassed the area of the hand containing the somatosensory RF (Fig. 2D). As reported earlier (Iriki et al., 1996), this visual RF was formed around the hand regardless of the gaze direction (bottom panel); that is, the visual RF moved when the hand moved, but did not move when the gaze shifted. Thus, the visual RF is not retinotopic, but anchored to the somatosensory RF as if coding the image of the hand in space. Among 261 tested neurons, 227 neurons exhibited visual RF properties as described above coding the image of the hand in space.

In the next phase of the experiment, the window was closed to prevent direct vision, and the video image of the hand captured through Camera 1, whose view mimicked that of the monkey's eye, was projected onto the video monitor (Fig. 1A). Then, scanning with the visual probe was performed in the same manner as before (Fig. 2B). The area within the thick trapezoid (top and middle panel) represents the views within the video monitor (note that Camera 1 was placed oblique to the horizontal plane). In this condition, the monkey could see the scanning probe only through the video monitor. The bimodal neuron fired when the monkey saw the probe approaching the hand on the monitor. Here the monkey's gaze was restricted to the upper view field (bottom panel; the thick square represents the location of the video frame). Now the visual receptive

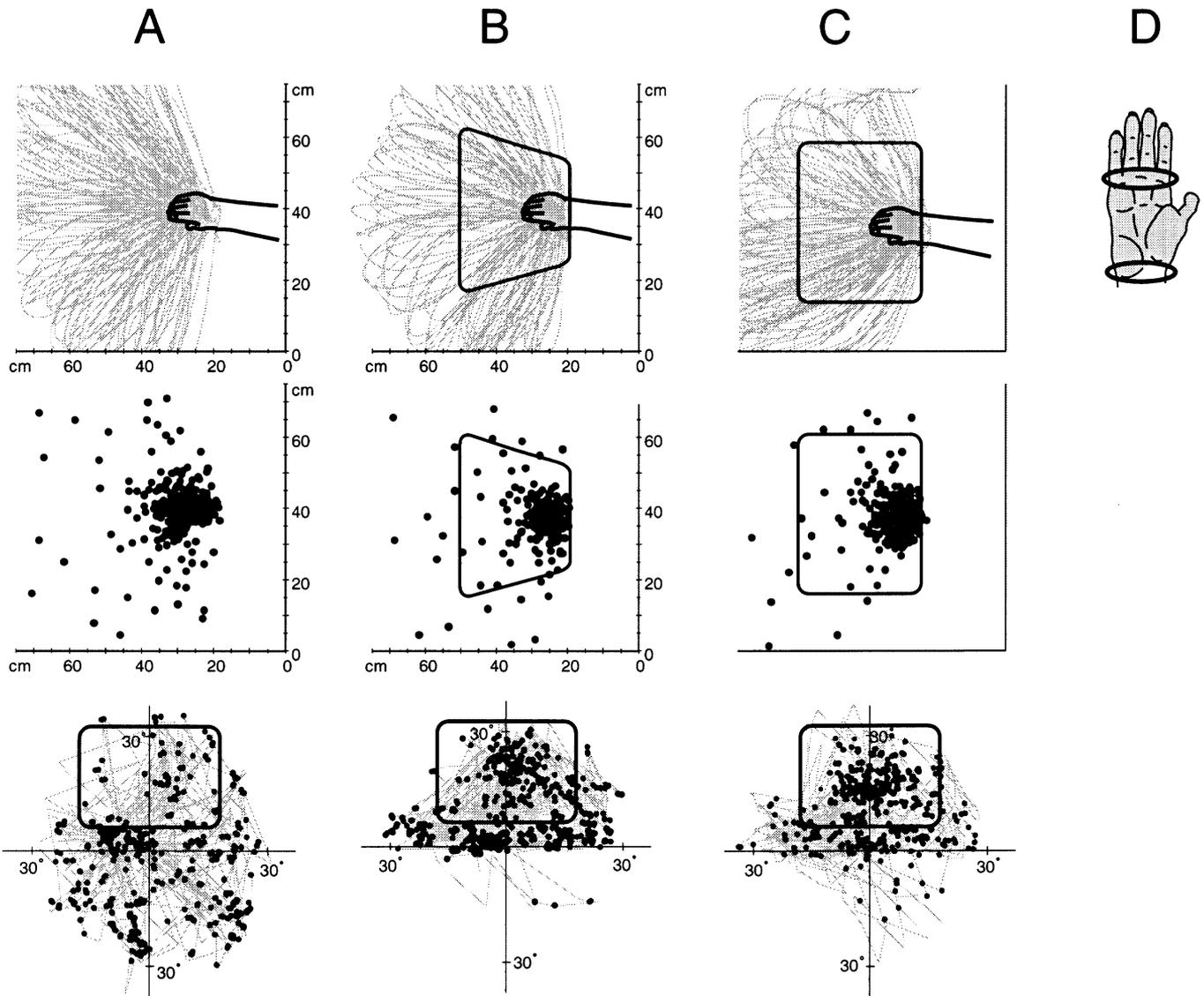


Fig. 2. Visual RF of a bimodal neuron, when visual target was presented under direct vision (A), through the video monitor (B) or superimposed in the monitor using chromakey-effect (C). Top panels; trajectories (shaded lines) of a scanning object projected to a horizontal plane. Middle panels; locations of the scanning object in the horizontal plane effective to make neurons fire (each dot represents 1 spike discharged at the instantaneous frequency higher than 3.0 Hz). Bottom panels; eye movements (shaded line) and gaze positions at which neuronal discharges occurred (dots). Thick trapezoids (B) show the boundary of the sight appearing in the monitor. Thick squares (C) indicate the monitor frame. D, somatosensory receptive field (tactile RF is shown by shaded area and site of joint manipulation is shown by circles around respective joints, which apply through Figs. 2–4).

field is located on the monitor, as if the monkey sees the image of its hand shown on the monitor as a part of or an extension of his own body. Of 227 bimodal neurons (coding the image of the hand under direct vision), 62 (27.3%) now showed visual RFs that coincided with the video image of the hand. In another experimental condition, the scanning with the visual probe was not performed in the actual field around the monkey's hand, but in neighboring room on an image captured by 'Camera 2' and superimposed on the monitor using a chromakey effect (see Section 2). This experiment was performed to exclude the possibility of

the monkey 'sensing' the approach of the probe to the hand by means of the events which the experimenter could not notice (for examples, by movement of the air, or occurrence of the very slight sound, or changes in reflection of noise, or whatever could happen accompanying probe manipulation). In this case, the visual probe was presented only on the monitor (Fig. 2C, squares indicate video frame), but again the visual RF of the neurons was located on the monitor. Among 62 neurons with visual RF on the monitor screen in the earlier viewing condition, the visual RF persisted in 46 (74.2%) neurons using this new method of stimulus

presentation. The reason why some neurons did respond to the condition depicted in Fig. 2B but not in Fig. 2C could not be identified presently. Nonetheless, we suspect that lack or suppression of the depth information by the chromakey effect, which could not be excluded, might be one of the potential reasons. Throughout these procedures, the somatosensory RF of this bimodal neuron (Fig. 2D) remained unchanged.

3.3. Intentional modification of the coding to match altered images in the monitor

We explored whether or not the projected self-image can be mentally manipulated to match modification of its visual appearance — independent of somatic sensation. The image of the hand captured by ‘Camera 1’ was modified using a video effects generator, and the scanning with the visual probe was superimposed on the modified hand image. As illustrated in Fig. 3, the visual RF of a bimodal neuron changed so as to accommodate the way in which the image of the hand was modified. That is, when the hand image was expanded (Fig. 3E) from the control image (D), the visual RF in the monitor was enlarged accordingly. When the image was compressed and displaced (A–C) the visual

RF followed it accordingly. The somatosensory RF of this neuron (F) was not modified during these procedures. Thus, the visual RF of this bimodal neuron coded the image of monkey’s own hand and changed as the image on the monitor was modified, even though, of course, the real hand did not change its size, shape, or location. This phenomenon was confirmed in all the 46 neurons that responded to chromakey-visual stimulation.

3.4. Tool-use in the monitor and modification of coding of the hand image along the tool

As earlier reported (Iriki et al., 1996), the visual RF of these bimodal neurons extend to include a hand-held tool used to retrieve distant food. Fig. 4A and B shows that this phenomenon also happens when the tool is viewed through the video monitor. Further, in the monitor, the image of the tool was gradually erased (during using the tool in the monitor) except for a bright spot at the end, using a luminance key effect. This makes the spot look like a cursor in the computer screen, in which situation we experience introspectively that the cursor has become our fingertip. Under this

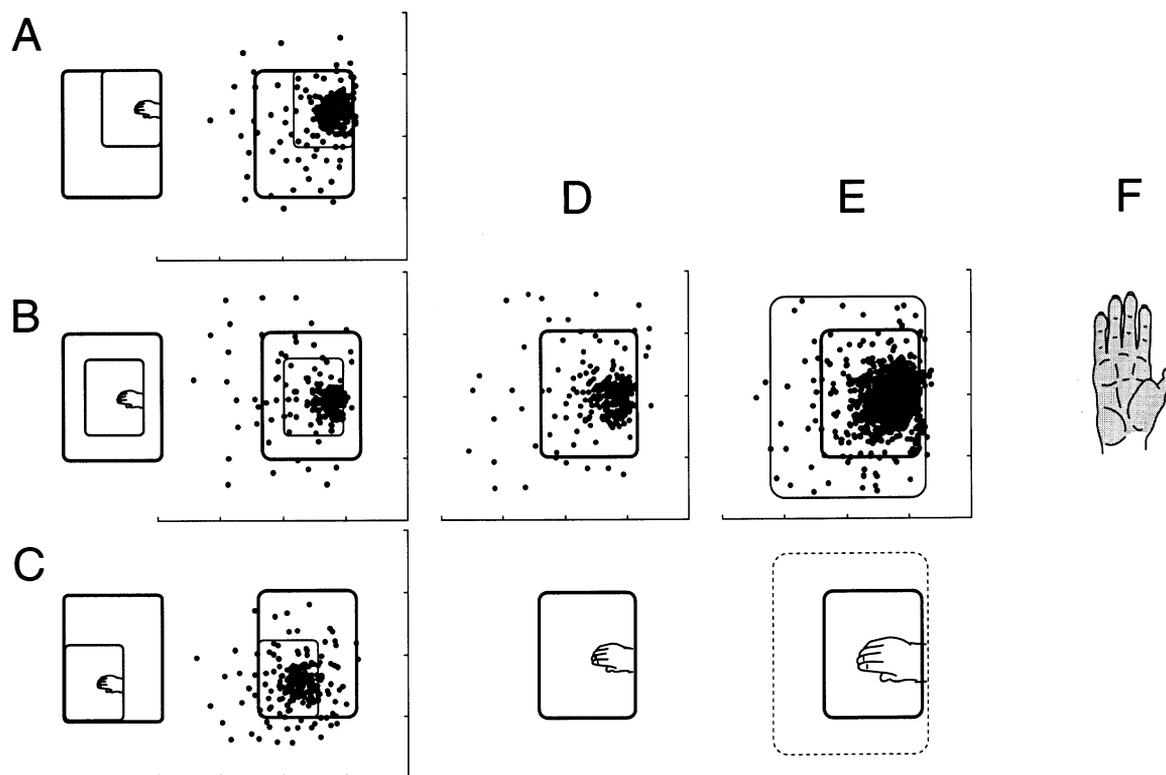


Fig. 3. Visual RF in the monitor, when the image of the hand was not altered (D), compressed and displaced (A–C), or expanded (E). Altered image of the hand in the monitor is shown in the inset attached to each graph. Coordinates of the graphs are relative to the monitor frame (thick squares). Locations of the scanning object (presented using a chromakey-effect) effective to make neurons fire are shown by dots. Each dot represents 1 spike discharged at the instantaneous frequency higher than 3.0 Hz. The posture or the position of the hand was not altered during modification of the screen image. F, somatosensory RF.

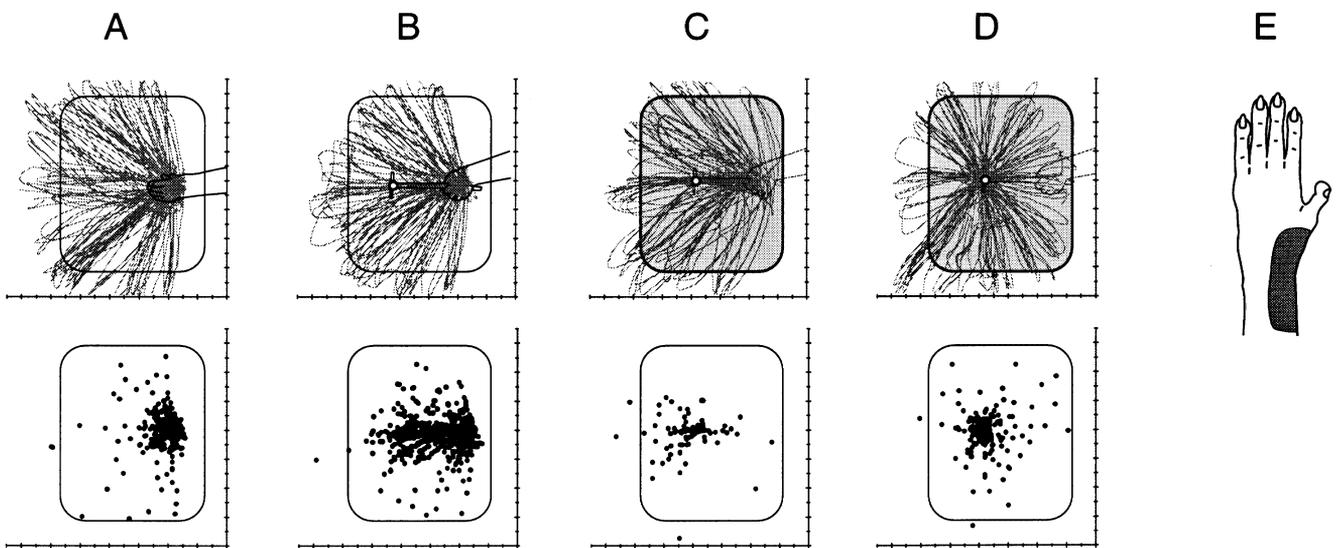


Fig. 4. Visual RF shifted in accordance with altered hand image. Top row; images taken by 'Camera 1' and presented in the Monitor with trajectory of a scanning probe superimposed by chromakeyer. Bottom row; positions of the probe (relative to the monitor frame) which drove the neuron to fire, with each dot representing one spike discharge. Thick squares indicate the monitor frame. Visual RF restricted to the hand (A) extended along the axis of the tool (B). C, D, Visual RF shrunk and limited to the area around the cue at the tool-tip, when the image except for the cue was concealed by luminance-key effect. Scanning was performed either hand-centripetally (C), or cue-centripetally (D). E, Somatosensory RF.

condition, scanning with the visual probe was less effective at inducing the neuron to fire when the probe was moved toward the hand (Fig. 4C), but was more effective when it approached the spot at the end of the tool (D). However, when the identical spot was shown in the monitor without any coincidence with the movement accompanying tool-use, this neuron never responded when the visual probe approached this 'independent' spot, indicating that these neurons were not responding to the appearance of the spot itself. This suggests that, in the present experimental condition, the tool was incorporated into the image of the hand, so that the one remaining spot of light came to represent the hand. Out of 18 bimodal neurons in which above procedure was examined, nine (50.0%) neurons exhibited this type of shift of hand image in the monitor.

3.5. Neuronal recording sites

Among 65 bimodal neurons, from four monkeys, in which new visual RFs were formed around the screen-image of the hand, most (53) neurons were found in the anterior bank of the IPSs contralateral to the hand shown in the video monitor, at the posterior extension of the forearm representation of the postcentral somatosensory cortex. But some were found in the posterior bank (7) or the fundus (4) of the intraparietal sulcus as well. There was no laminar bias in the distribution of these neurons. The remaining 196 neurons, which responded only to the visible hand, were distributed throughout the intraparietal cortex in a similar

manner. In other words, there was no evidence that the neurons that coded the image in the monitor were anatomically segregated from those that did not. In one monkey, recordings were attempted throughout the second stage of training with the monitor, (i.e. after the first stage of training to use the rake was completed). One electrode penetration was made on each day. Fig. 5 illustrates results obtained from this monkey. Before the monkey was able to manipulate his hand and the tool using the video monitor, no bimodal neurons responded to the image shown in the video monitor (Fig. 5A and B). Yet after the first day the monkey acquired the skill (at 13th day of the training for this monkey), 41.7% (20/48) of the bimodal neurons started to respond to the hand image in the video monitor. In short, the neurons depicting the present visual RF properties which seemed to code the image of the hand on the video monitor were created in this area of the cerebral cortex by training.

4. Discussion

4.1. Recognizing the self in the video monitor

It has been generally believed that macaque monkeys do not recognize their 'self image' in a mirror or video monitor (Thompson and Boatright-Horowitz, 1994; Tomasello and Call, 1997). Only primates of evolutionary higher level than the chimpanzee were thought to recognize the self in the mirror (Gallup, 1970, 1977,

1982). Although, macaque monkeys have been reported to be able to use the mirror to guide their movements which cannot be observed through direct vision (Itakura, 1987a,b). In the present study, we observed the behavior which would indicate that our monkeys have learned, by the training using combination of tool and the video monitor, to recognize the image in the video monitor as a part of their own body — the behavior which fulfills Gallup's criterion for mirror-self recognition (Gallup, 1970, 1977). This was achieved perhaps because our training procedures using tools helped the monkeys to realize their own body in the monitor. That is, without tools, monkeys hardly learned to recognize their images appeared in the monitor because it is easier for monkeys to just grope for the food blindly without watching the monitor (indeed they did prefer to grope blindly when not using the tool), but, with tools they cannot feel at the tip, and hence monkeys were obliged to rely solely on the self-images appeared in the monitor as a guidance to plan and execute the movements of their own body parts to retrieve the food.

However, the present behavioral result, by itself, might appear to support the nature of 'over-trained' conditional discrimination learning instead of an 'insightful' self-image formation, because it was attained only after extensive training. Notwithstanding, when combined with the present neurophysiological evidence,

interpretations would prefer the latter possibility — the identical neurons, which seemed to code the image of the hand in normal condition apparently responded in a similar manner to the image in the video monitor. An alternative interpretation for the present behavioral results may be that a novel visual coordinate system was created by the training to achieve efficient movement referring only to the image projected in the monitor. However, the presently observed phenomenon is difficult to interpret merely as a sort of coordinate transformations (Fogassi et al., 1996; Graziano et al., 1997; Rizzolatti et al., 1997), but rather more 'subjective' domain of the representations seemed to be involved. This is because the transformation was shown to occur without any changes in the actual hand position or movements (so no actual sensory input other than the image per se) but only by alterations of the appearance can induce the changes in the visual RF, and such a RF could be formed in the monitor placed anywhere in the space. Thus, although concrete neural mechanisms for its creation remain an open question at the moment, the concept of the subjective self-image was proposed. Now we assume that some sort of internal representation of the self image was formed elsewhere in the brain by the extensive training, and the presently observed neuronal properties reflect these mode of representations.

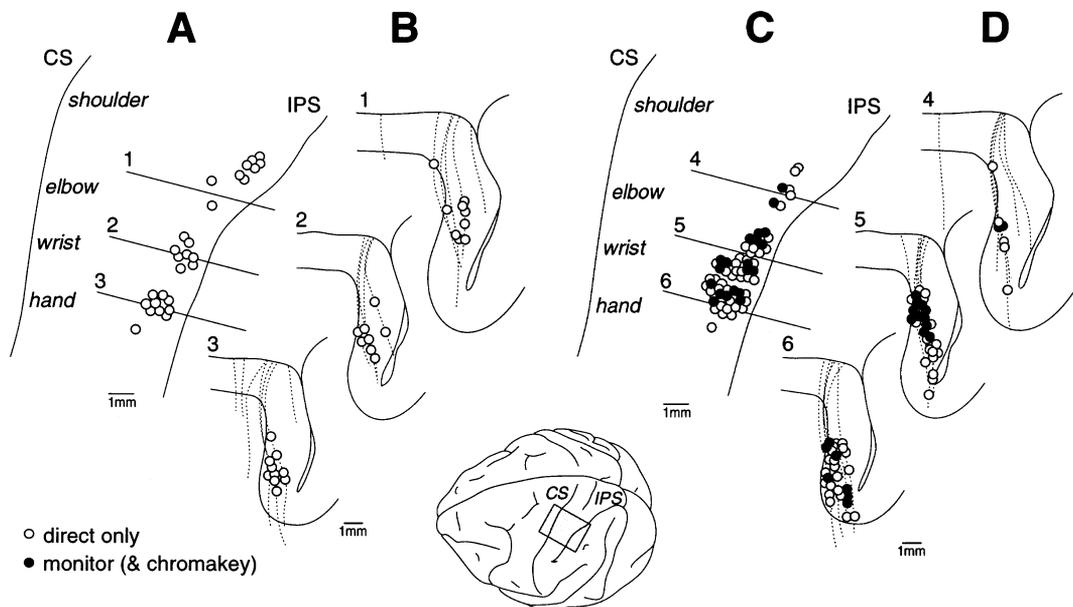


Fig. 5. Recording sites of bimodal neurons in one monkey, before (A, B) and after (C, D) the training. A, C, recording sites plotted on the dorsal surface of the left postcentral gyrus (area inside the square of). Open circles; neurons with visual RFs under direct vision only. Filled circles; neurons with visual RFs both under direct vision and in the video monitor (all these neurons responded also to the target presented using chromakey-effect). Approximate sites of representation of the body parts in area 3b are as indicated. B, D, off-sagittal sections orthogonal to the intraparietal sulcus (indicated by lines 1–3 in A and 4–6 in C), on which recording sites are projected. Dashed lines indicate electrode tracks. CS, central sulcus; IPS, intraparietal sulcus.

4.2. Possible neural mechanisms for creating representation of the body image in the parietal cortex

Bimodal (somatosensory and visual) properties in relation to the location of the body parts in relation to the environmental space, similar to the present findings, have been observed in other brain areas such as ventral premotor cortex (Fogassi et al., 1996; Graziano et al., 1997; Graziano, 1999), ventral (VIP) and medial (MIP) intra parietal areas (Colby and Duhamel, 1991; Colby et al., 1993; Colby and Duhamel, 1996; Rizzolatti et al., 1997; Graziano and Gross, 1998), area 7b (Hyvärinen and Poranen, 1974; Leinonen and Nyman, 1979; Leinonen et al., 1979), and putamen (Graziano and Gross, 1993). Unlike the present experiments, however, above neurons were ready to represent location of the body parts in untrained naive monkeys. In relation to above areas, the intraparietal area where present neurons were recorded from is located dorsomedial to VIP and lateral to MIP with more or less overlapping with their boundaries. One peculiar feature, which characterizes the present findings from above earlier reports, is that the visual response to code the self-image in the video monitor appeared only after the extensive training. Before training to use the monitor (but after the training to use the tool), no neuron responded to visual stimuli presented (using a chromakey-effect technique) around the image of the hand in the monitor screen. In the same monkey, immediately after the monkey learned to recognize the self-image in the monitor, a group of neurons with new visual RFs formed around the screen-image of the hand appeared. For acquisition of this skill, monkey's hand-movement had to be displayed on the video monitor without any time delay. That is, the coincidence of the movement of the real hand and the video-image of the hand seemed to be essential for training the monkey to use the video-image to guide their hand movements. Therefore, achievement of this skill might be subserved perhaps by matching visual input with the hand-image created in the brain referring to the efferent signals controlling the hand movement, to form a neural circuitry to subserve a novel mode of visuo-somatosensory integration necessary for representing newly formed body-image. Recent results from our laboratory show that during extensive training to use tools (although not using a video monitor like in this experiment), expression of immediate-early-genes (Ishibashi et al., 1999) as well as neurotrophic factors, such as brain-derived neurotrophic factor (BDNF) and its receptor trkB or NT3 (neurotrophic factor 3), are induced particularly in the intraparietal area where the present neurons were recorded (Ishibashi et al., 2001). After the training was completed these expressions no longer persisted. Behavioral analyses during this period (Ishibashi et al., 2000) suggested that reorganization of somatosensory-visual

integration seemed to be developed. Therefore, a novel mode of visual RF properties created during the training period may possibly be subserved by reformation of the neural circuitry in which molecular genetic processes in the cortical area described above are involved. Therefore, we suspect that these bimodal neurons may be 'reserved' in naive monkeys to take a role in further complicated higher functions which monkeys may potentially encounter eventually during the course of development or the evolutionary processes.

4.3. Functional considerations of the visual representation of the body in the brain

If the presently demonstrated phenomenon of the training-induced representation of the self image in the artificial conditions was readily induced by reinforced induction of genetic expressions as speculated above, it would provide clues to developmental as well as evolutionary pictures of higher cognitive functions. During early childhood of humans, their field of view is restricted to personal and immediately adjacent peripersonal space. An internal representation during this period should be a type of sensori-motor intelligence (Piaget, 1953) which is unconsciously acquired through experiencing various actions in the environmental space as accustomed patterns of action, and should correspond to the 'body schema' postulated by Head and Holmes (Head and Holmes, 1911). While becoming familiar with the surrounding space, children achieve (by the age of 9–10) abilities to handle action-free visual images of their own body, which dissociates an internal schema from existing actions. Having this mechanism in the brain, we might be able to feel 'reality' in the virtual reality or tele-existence apparatus. Further, by mentally manipulating this visual type of images independent of actual state of the existing body parts, we can mentally rotate objects and our own body parts independent of supporting actions (Bruner et al., 1966). Corresponding to this psychological experience, the size and position of visual RFs of the presently observed bimodal neurons were modified according to expansion, compression, or change of the position of the visual image in the video monitor. These changes were induced even if the posture or the position of the hand was not actually altered during modification of the screen image. Thus, the presently observed properties of the visual RFs would represent neural correlate for this sort of internal representations, indicating that macaque monkeys (along the course of primate evolution, akin to human children during development) attained a neural machinery which can become capable, when extensively forced, of representing an intentionally controllable visual representation of their own body. In addition, the neuron presented here respond not only to the natural image of the hand in the

monitor, but also to a sign which functionally substitutes for the actual hand in the monitor, whereas the sign with an identical appearance per se was not effective if it appeared in the monitor with no contextual relation to the function of the hand. Based on above coincidence between development of human mental representations and response properties of parietal neurons presently discovered in overtrained monkeys, we would like to postulate that evolutionary precursors for introspective manipulation of an abstract sign, or eventually a symbolic representation of the own body, might be already reserved as neural machinery in the monkey brain, but usually not in operation and is able to be recruited only when reinforced extensively. Therefore, we expect that by extending present experimental paradigms, monkey studies would potentially lead us to understanding of the neural mechanisms of our higher cognitive functions such as symbol manipulation, and perhaps eventually the language and metaphysical thoughts.

Acknowledgements

We thank Professor Melvyn A. Goodale for valuable comments in the earlier version of the manuscript. Supported by JSPS (Japan Society for the Promotion of Science) Research for the Future program.

References

- Andersen, R.A., Snyder, L.H., Bradley, D.C., Xing, J., 1997. Multimodal representation of space in the posterior parietal cortex and its use in planning movements. *Annu. Rev. Neurosci.* 20, 303–330.
- Berlucchi, G., Aglioti, S., 1997. The body in the brain: neural bases of corporeal awareness. *Trends Neurosci.* 20, 560–564.
- Berti, A., Frassinetti, F., 2000. When far becomes near: remapping of space by tool use. *J. Cogn. Neurosci.* 13, 415–420.
- Bruner, J.S., Olver, R.R., Greenfield, P.M., 1966. *Studies in Cognitive Growth*. Wiley, New York.
- Colby, C.L., Duhamel, J.-R., 1991. Heterogeneity of extrastriate visual areas and multiple parietal areas in the macaque monkey. *Neuropsychologia* 29, 517–537.
- Colby, C.L., Duhamel, J.-R., 1996. Spatial representations for action in parietal cortex. *Cogn. Brain Res.* 5, 105–115.
- Colby, C.L., Gattass, R., Olson, C.R., Gross, C.G., 1988. Topographic organization of cortical afferents to extrastriate visual area PO in the macaque: a dual tracer study. *J. Comp. Neurol.* 269, 392–413.
- Colby, C., Duhamel, J.-R., Goldberg, M.E., 1993. Ventral intraparietal area of the macaque: anatomic location and visual response properties. *J. Neurophysiol.* 69, 902–914.
- Farne, A., Ladavas, E., 2000. Dynamic size-change of hand peripersonal space following tool use. *Neuroreport* 11, 1645–1649.
- Fogassi, L., Gallese, V., Fadiga, L., Lupinno, G., Matelli, M., Rizzolatti, G., 1996. Coding of peripersonal space in inferior premotor cortex (Area F4). *J. Neurophysiol.* 76, 141–157.
- Gallup, G.G. Jr, 1970. Chimpanzees: self-recognition. *Science* 167, 86–87.
- Gallup, G.G. Jr, 1977. Self-recognition in primates: a comparative approach to the bidirectional properties of consciousness. *Am. Psychologist* 32, 329–338.
- Gallup, G.G. Jr, 1982. Self-awareness and the emergence of mind in primates. *Am. J. Primatol.* 2, 237–248.
- Graziano, M.S.A., 1999. Where is my arm? The relative role of vision and proprioception in the neuronal representation of limb position. *Proc. Natl. Acad. Sci. USA* 96, 10418–10421.
- Graziano, M.S.A., Gross, C.G., 1993. A bimodal map of space: somatosensory receptive fields in the macaque putamen with corresponding visual receptive fields. *Exp. Brain Res.* 97, 96–109.
- Graziano, M.S.A., Gross, C.G., 1994. The representation of extrapersonal space: a possible role for bimodal visual-tactile neurons. In: Gazzaniga, M.S. (Ed.), *The Cognitive Neurosciences*. MIT Press, Cambridge, pp. 1021–1034.
- Graziano, M.S.A., Gross, C.G., 1998. Spatial maps for the control of movement. *Curr. Opin. Neurobiol.* 8, 195–201.
- Graziano, M.S.A., Hu, X.T., Gross, C.G., 1997. Visuospatial properties of ventral premotor cortex. *J. Neurophysiol.* 77, 2268–2292.
- Harman, K.L., Humphrey, G.K., Goodale, M.A., 1999. Active manual control of object views facilitates visual recognition. *Curr. Biol.* 9, 1315–1318.
- Head, H., Holmes, G., 1911. Sensory disturbances from cerebral lesions. *Brain* 34, 102–154.
- Holmes, G., 1918. Disturbances of visual orientation. *Br. J. Ophthalmol.*, 2 449–468, 506–516.
- Hyvärinen, J., Poranen, A., 1974. Function of the parietal associative area 7 as revealed from cellular discharges in alert monkeys. *Brain* 97, 673–692.
- Iriki, A., Tanaka, M., Iwamura, Y., 1996. Coding of modified body schema during tool use by macaque postcentral neurons. *Neuroreport* 7, 2325–2330.
- Iriki, A., Tanaka, M., Iwamura, Y., 1997. Self image in the video monitor is coded by parietal neurons. *Soc. Neurosci. Abstr.* 23, 211.
- Ishibashi, H., Hihara, S., Takahashi, M., Iriki, A., 1999. Immediate-early-gene expression by the training of tool-use in the monkey intraparietal cortex. *Soc. Neurosci. Abstr.* 25, 889.
- Ishibashi, H., Hihara, S., Iriki, A., 2000. Acquisition and development of monkey tool-use: behavioral and kinematic analyses. *Can. J. Physiol. Pharmacol.* 78, 958–966.
- Ishibashi, H., Hihara, S., Takahashi, M., Heike, T., Yokota, T., Iriki, A., 2001. Tool-use learning selectively induces expression of brain-derived neurotrophic factor, its receptor trkB, and neurotrophin 3 in the intraparietal multisensory cortex of monkeys. *Cogn. Brain Res.* in press.
- Itakura, S., 1987a. Mirror guided behavior in Japanese monkeys (*Macaca fuscata fuscata*). *Primates* 28, 149–161.
- Itakura, S., 1987b. Use of mirror to direct their responses in Japanese monkeys (*Macaca fuscata fuscata*). *Primates* 28, 343–352.
- Iwamura, Y., 1998. Hierarchical somatosensory processing. *Curr. Opin. Neurobiol.* 8, 522–528.
- Iwamura, Y., Tanaka, M., Sakamoto, M., Hikosaka, O., 1993. Rostrocaudal gradients in the neuronal receptive field complexity in the finger region of the alert monkey's postcentral gyrus. *Exp. Brain Res.* 92, 360–368.
- Leinonen, L., Nyman, G., 1979. II. Functional properties of cells in anterolateral part of area 7 associative face area of awake monkeys. *Exp. Brain Res.* 34, 321–333.
- Leinonen, L., Hyvärinen, J., Nyman, G., Linnankoski, I., 1979. I. Functional properties of neurons in lateral part of associative area 7 in awake monkey. *Exp. Brain Res.* 34, 299–320.
- Marvita, A., Husain, M., Clarke, K., Driver, J., 2001. Reaching with a tool extends visual-tactile interactions into far space: Evidence from cross-modal extinction. *Neuropsychologia*, in press.

- Obayashi, S., Tanaka, M., Iriki, A., 2000. Subjective image of invisible hand coded by monkey intraparietal neurons. *Neuroreport* 11, 3499–3505.
- Piaget, J., 1953. *The Origin of Intelligence in the Child*. Routledge & Kegan Paul, London.
- Redican, W.K., 1975. Facial expressions in nonhuman primates. In: Rosenbloom, L.A. (Ed.), *Primate Behavior*. Academic, pp. 103–194.
- Rizzolatti, G., Fogassi, L., Gallese, V., 1997. Parietal cortex: from sight to action. *Curr. Opin. Neurobiol.* 7, 562–567.
- Seltzer, B., Pandya, D.N., 1980. Converging visual and somatic sensory cortical input to the intraparietal sulcus of the rhesus monkey. *Brain Res.* 192, 339–351.
- Thompson, R.L., Boatright-Horowitz, S.L., 1994. The question of mirror-mediated self-recognition in apes and monkeys: some new results and reservations. In: Taylor Parker, S., Mitchell, R.W., Boccia, M.L. (Eds.), *Self-Awareness in Animals and Humans*. Cambridge University Press, New York, pp. 330–349.
- Tomasello, M., Call, J., 1997. *Primate Cognition*. Oxford University Press, New York.
- Triggs, W.J., Gold, M., Gerstle, G., Adair, J., Heilman, K.M., 1994. Motor neglect associated with a discrete parietal lesion. *Neurology* 44, 1164–1166.
- Ungerleider, L.G., Mishkin, M., 1982. Two cortical visual systems. In: Ingle, D., Goodale, M.A., Mansfield, R.J.W. (Eds.), *Analysis of Visual Behavior*. MIT Press, Cambridge.
- van Hooff, J.A.R.A.M., 1962. Facial expression in higher primates. *Symp. Zool. Soc. Lond.* 8, 97–125.