RESEARCH ARTICLE

Acquisition of an externalized eye by Japanese monkeys

Yumiko Yamazaki · Hiromi Namba · Atsushi Iriki

Abstract Many animals use tools to augment motor function ("motor tools", like rake), while the use of tools to acquire sensory information or to augment sensory function ("sensory tools", like endoscope) has been reported only in humans. In the present study, we trained Japanese monkeys to acquire the sensory tool use to re-construct a possible developmental course of the human-specific tool use via motor tool use training. After they mastered the rake use, we systematically introduced a series of external mirror and video arrangements, so as to separate visual cues from their actual origins in visuomotor space. Finally, the monkeys could acquire the use of sensory tool—a sort of endoscope attached to a rake—to explore the experimental space to find and retrieve the food. The results indicated a critical role of environmental control to develop even higher order behavioral sequences like human-specific sensory tool use in nonhuman primates.

Y. Yamazaki · H. Namba · A. Iriki (✉)
Laboratory for Symbolic Cognitive Development,
RIKEN Brain Science Institute, 2-1 Hirosawa,
Wako, 351-0198 Saitama, Japan
e-mail: iriki@brain.riken.jp

Y. Yamazaki
The Graduate School of Human Relations,
Keio University, Tokyo, Japan
e-mail: yumyam@brain.riken.jp

Introduction

Perhaps, one of the most outstanding behavioral differences between nonhuman animals and humans is the degree to which we utilize and modify our environment. Rather than developing the ability to see in darkness or cut through tough skin with specialized teeth, humans began tending fire and fashioning knives. These cultural, rather than biological, methods of modifying the environment could be not only enjoyed but also further improved upon by each new generation. The present study, establishes a novel experimental paradigm whose aim is to hypothesize possible stages of epigenetic evolution in humans by inducing tool use in a nonhuman primate—the Japanese monkey, which shows no natural inclination for tool use in wild—through intensive training.

Tools used by nonhuman animals (e.g., Beck 1980) have been found in broad range of species. For example, New Caledonian crows make hooked twig tools to stimulate and then extract their prey (e.g., Hunt 1996). Some primate species use different stones combinatorially as anvil and hammer for nut cracking (Fragaszy et al. 2004; Matsuzawa 1994). In our lab, Japanese macaques rapidly learned to use rakes and spontaneously used rakes of different length combinatorially (Hihara et al. 2003), even though these monkeys show no inclination to use tools either in the wild or in captivity. In the above examples, tools are typically sticks or stones, which simply extend or strengthen the functions of an animal's effectors (Asano 1994), such as the hand or beak. Although existing research on tool use, and even the basic definition of a tool (Beck 1980), has exclusively

focused on these "motor tools", humans are known to employ a separate class of tool that enhance or externalize the sense organs (Goldenberg and Iriki 2007). An endoscope, binoculars, or a stethoscope represent such "sensory tools" (Asano 1994). We define a sensory tool as a tool used for adjusting, aiding, and processing sensory information so as to enhance the original function of the sensory organ. This is still compatible with previously accepted definitions of tools (Beck 1980; Visalberghi and Fragaszy 2006), which in our terminology are all motor tools. Based on our criterion of the definition, it is clear that no animals but humans have been reported to use sensory tools to obtain information that otherwise cannot be available by unaided sensory organs. However, there is no empirical framework for understanding how motor and sensory tools have developed, what the relation between them is, and which cognitive functions are recruited by each—all questions we aim to address in the present study.

These two classes of tool can be distinguished by purely functional aspects. While motor tool use directly connects to specific behavioral consequences, sensory tools provide discriminative cues to guide many types of behavior to follow (Asano 1994). Because sensory tool users can broaden their own environment by dealing with stimuli that are not available without the tools, they apparently possess additional cognitive abilities for handling information in a flexible and abstract manner. Previous work (Iriki et al. 1996) from our lab has demonstrated the change in neural activity in Japanese monkeys after motor tool use training. A group of bimodal (visual–tactile) neurons in the intraparietal area that originally coded the area around the hand and arm extended their receptive fields to include the tool (a rake) during active use. That is, the body image (body schema) encoded by these neurons expanded as if the tool, were part of the body.

The present experiment builds on this result (Iriki et al. 1996) and achieves a complementary process, abstracting sensory information away from its body-derived, eye-centric origin to an externalized sensory tool. Through systematic training, two Japanese monkeys became fully adept at using an externalized eye—a sort of endoscope made of a rake with a tiny forward-facing video camera mounted on the shaft ("camera rake", bottom of Fig. 1). We invented this device as a visual sensory tool to enable subjects to make purposive, active search of the environment. This induction of the novel tool use in the monkeys suggests the following recursive relationship leading to the acquisition of a human-specific sensory tool: becoming adept with a motor tool changes the user's relationship with his environment, which would induce changes in relevant cognitive functions (Iriki et al. 1996), which leads to the ability to use a sensory tool to utilize the novel information which can be available only after the training.
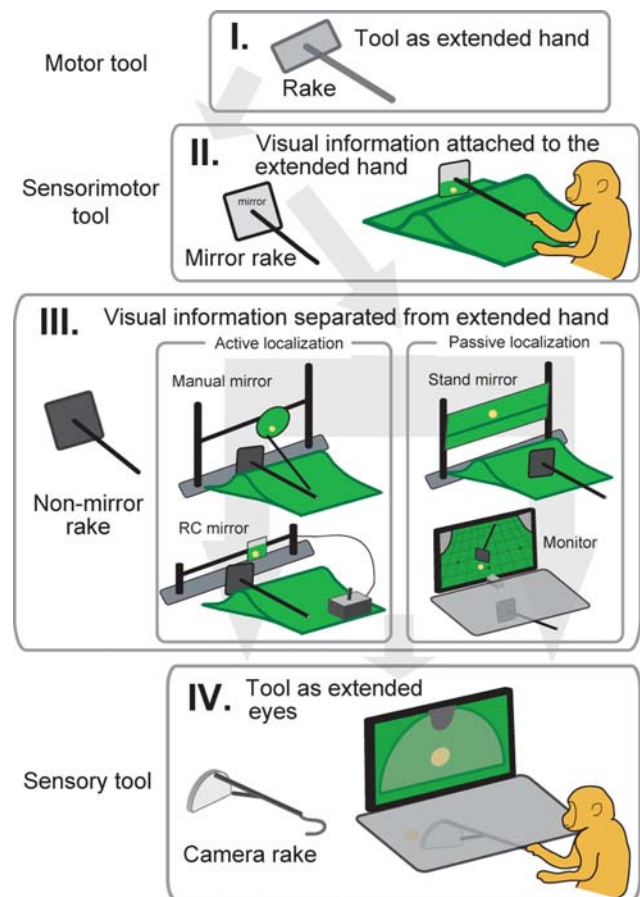
Fig. 1 The training structure for acquisition of a sensory tool. The training was divided into four Stages (I–IV), and the tool used in each stage is depicted on the left. (I) Rake, as a motor tool, to retrieve food on a flat table. Platform was used to train the monkey to raise the rake. (II) Mirror rake, as a sensorimotor tool, to retrieve food hidden behind a hump. To locate the food, the monkey had to use the food's reflection in the mirrored tip of the rake. (III) Non-mirror rake, the same shape as the mirror rake, on a humped table. The rake could be used to retrieve food, but could no longer be used to locate it. There were two lines of training for mastering the use of visually displaced cues. In case of the manual mirror and remote-controlled (*RC*) mirror, the animal actively operated the tools to search for the food. In case of the stand mirror and the monitor, he passively observed the captured image of the hidden food. (IV) Camera rake, as a sensory tool, with a transparent tip and a forward-facing camera mounted in the handle. Because of the opaque screen, the monkey had only the captured image on the monitor to guide its searching and retrieval

Pilot study

Failure of M1 to localize food using the monitor

Before starting the training sequence described below, one of the monkeys (M1) had been trained in a three choice discrimination task. In this task a tunnel-like apparatus with three semicircular openings (7.0 cm diameter each) was used, and he was required to insert the rake to get the

food at the end of the tunnel. The food was not directly visible to the monkey; he had to watch the image captured by a CCD camera located at the back of the apparatus and presented on the 32″ TV monitor in front of him. In this experiment, a rake with semicircular tip (6 cm diameter, 41 cm long) was used. For habituation to the apparatus, at first we placed the food at one of the tunnel's entrances where the subject could see it directly. Then we placed the food inside the tunnel, where he could not see it directly, and would have had to use the monitor to locate it. Although we continued the training more than 3 months, using various additional cues—such as pointing with other tools, moving and showing the food before placing it, and releasing his arm after requiring him to gaze at the correct position—his performance never rose to the planned acquisition criterion of 80% correct [though it was above the level of statistical significance (41.9%, $\chi^2(1) = 13.69$, $P < 0.01$)]. Thus, we changed the training plan entirely, developed the one described below with smaller steps, and introduced various kinds of mirror-based tools for increasing the degree of visual displacement involved in monitor use.

Materials and methods

Subjects

Two male Japanese monkeys (*Macaca fuscata*, M1 and M2) participated in the experiment. M1 was adult, 8 kg, and had 8 months of experience with the rake before the present study. M2 was juvenile, but he had complete permanent dentition when we started the experiments with him. His weight was 4 kg, and he was experimentally naive at the beginning of the experiments. They were individually housed in cages, where water was freely available. During the training period, they were fed daily in their cages with monkey chow. Apples and sweet potatoes were reserved as rewards in the training sessions. The study was approved (H16-2B033) by the Animal Experiment Committee and was conducted in accordance with the guidelines for conducting animal experiments of the RIKEN Brain Science Institute.

Apparatus and procedure

During the training sessions, each monkey sat on the monkey chair and could freely move his arms. The chair was fixed to the table [110 (*w*) × 150 (*d*) × 71 (*h*) cm, 77 cm (*w*) at the monkey's side]. For clarity, we separately describe the apparatus and procedures for each stage after "General experimental design" below.

General experimental design

Only M2 was given "pre-tool use training" and stepwise training with the camera rake. Otherwise, training conditions, were basically the same for both subjects. The training procedure of the monkeys was divided into four major stages, each of which we used particular apparatus and training procedure. Except for the camera rake training in Stage IV, our acquisition criterion was five successful food retrievals in five consecutive trials. For Stage IV, another index was employed to evaluate the acquired performance (see "Data analysis"). Sometimes, after reaching criterion we overtrained the monkeys to ensure that they could perform each task stably and skillfully.

The training consisted of four major Stages (I–IV, Fig. 1), each with several minor steps (different tasks) and phases (time courses within a step). Tools unique to each step characterized the function to be acquired in each respective stage, as follows. Stage I: rake as a motor tool on a flat table—a prerequisite for "extending" the body image. The monkey grasped the visuomotor relationship between food and rake using direct, unmediated vision. Stage II: mirror-tipped rake as a sensorimotor tool on a table with a hump. The food was placed behind the hump and could not be seen directly, but could be seen displaced as a reflection in the mirror. This stage was the first step toward recognizing and manipulating the displaced visual cues in an extended eyes (i.e., in the tip of the rake). Stage III: combination of two different types of tool on a humped table—one to explore for food either actively [manual mirror and remote-control (RC) mirror] or passively (stand mirror and monitor), and the other to retrieve food (non-mirror rake). Through these steps, visual cues were gradually separated from the area of the "extended" hand. Stage IV: camera rake as a sensory tool on a flat table under an opaque screen. Exploration and retrieval were possible only using the real-time camera view displayed on a monitor—the goal of the training. Except for the camera rake training in Stage IV, our acquisition criterion was five successful food retrievals in five consecutive trials (for details, see "Data analysis").

Stage I

Apparatus

Training materials only for M1

The rake used for initial training had a semicircular tip (6 cm diameter) attached to an aluminum bar (40 cm long, 1 cm in diameter). In the platform step, an acrylic pane

[30(*w*) × 15 (*d*) × 10 cm (*h*)] was installed parallel to and above the table that forced him to lift the rake in order to reach for the food.

Training materials only for M2

Before beginning the rake use training, we gave M2 "pre-tool use training," using two types of tool. The first one (35.5 cm long) had a spoon-shaped pink tip (5 cm in diameter). The second (44.5 cm long) had a ring-shaped blue tip (5 cm in diameter). In each case, food was placed on or in the tool's tip and then the monkey was allowed to grab the handle. The purpose of this was to train him to pull food toward himself using tools. The only difference between the two was that the spoon-tipped tool could be used to pick up or drag the food across the table while the ring-tipped tool could be used only for dragging.

The rake used to train M2 had a transparent, acrylic, rectangular tip (6 cm × 12 cm), attached to an aluminum handle (25 cm long, 1 cm in diameter). In the platform step we used two platforms [26.5 cm (*w*) × 15.0 cm (*d*)] of different heights: 1.0 and 3.0 cm. Both were made of black acrylic. As M2 practiced with the platform, the rake was exchanged for one with a smaller tip (10 cm × 6 cm) and a longer handle (30.5 cm).

Procedure

Pre-tool use training

As mentioned earlier, M2 was first trained to pull on a spoon-tipped tool to get food, and then was switched to the ring-tipped tool. It took six and 16 trials within a day to reach the criterion of five consecutive successful retrievals in spoon- and ring-tipped tools, respectively.

Rake use

M1 was trained to use a rake to retrieve food placed beyond arm's reach. Initially, the rake was placed near the food so that the subject only had to pull the rake back to his side (rake I). The distance between the food and the rake tip was gradually increased (II). Then he was trained to move the rake not only vertically, but also horizontally (III and IV, again with increasing distances). We continued the rake use training until he could make smooth, circular movements with the rake, without any pause or interruption within a trial (V). In the final phases he was trained to retrieve food from a position beyond the rake (VI and VII, again with increasing distances). Daily training sessions consisted of 100–500 trials.

Next we introduced the platform, so that the subject had to raise the rake, place the rake tip beyond the food, and

pull it back to his side. We knew (through earlier training experiences) that the platform step enhances a monkey's skill with the rake. In this step, an acrylic pane [30(*w*) × 15(*d*) × 10 cm (*h*) for M1 and 26 (*w*) × 15 (*d*) × 3 cm (*h*) for M2, placed 30 cm in front of each monkey] was mounted above the table so the monkey had to raise the rake to get the food. Firstly, the monkey was trained to raise the rake, and then to raise and place the rake to retrieve the food on it.

Transitional training I

Before introducing the mirror rake, we inserted a training phase in which visual information could be obtained both from direct sight and the tip of the mirror rake. The only difference from the previous task was the mirror rake itself.

Stage II

Apparatus

In the mirror rake step (Stage II), each monkey was trained to use a rake with a square, reflective, acrylic tip (10 cm × 10 cm) attached to a 40 cm aluminum handle. For hiding the food from the monkey's sight, we created a speed bump-shaped barrier by laying a green vinyl mat [71.5 cm (*w*) × 67.0 cm (*d*)] over a cylindrical hump (5.9 cm at its highest point). To ensure that food placed behind the hump could not be seen directly, the hump's slope was less steep on the side facing the monkey, and steeper on the far side.

Procedure

Mirror rake use

A critical and difficult aspect of mastering a visual sensory tool is learning to abstract away from direct-line vision with one's own eyes and using visually displaced information in its stead. We introduced a mirror tool to aid the transition between the tasks with and without visual displacement. The mirror was selected because it has been reported that some species of macaque monkeys (Anderson 1986) are capable of using a (fixed position, non-manipulated) mirror for finding food. We define the mirror rake as a sensorimotor tool because of its dual function in searching for and retrieving food. Additionally, the hump was introduced. The subject could no longer locate food rewards directly, but could only find them by placing and sliding the tip of the mirror rake on the far side of the hump and searching for the food's reflection. At first, an experimenter placed the mirror rake just behind the food, so as to let the subject

see it reflected in the mirror and trigger him to pull the rake handle. Then food was placed only a little away from the mirror, and the experimenter showed the food explicitly in front of him for triggering active search (mirror rake I). The distance between the food and the initial mirror position was gradually increased. Finally, we released the monkey's arm anywhere on the table, so he had to search for the food using the mirror rake without ever having been shown the food explicitly (II). During training sessions, we used a manual or electrical shutter hung in front of the monkey's face so he could not observe the placement of the food behind the hump. In order to avoid giving the monkey non-verbal cues about the food's location, the experimenter operating the shutter averted his eyes while a second experimenter placed the food. After placement, the first experimenter opened the shutter and released the monkey's arm, allowing him to start exploring with the mirror rake.

Transitional training II

Before proceeding to Stage III, we trained the monkey in a situation, where a stand mirror was placed behind the hump and the monkey used the mirror rake to retrieve the food. That is, redundant visual information from both the stand mirror and the mirror rake, were available to the monkey.

Stage III

Apparatus

In Stage III the monkey was further trained to separate and abstract visual information using four different tools. We differentiated the tools into two categories, as depicted in Fig. 1. One (in the left column: manual mirror, RC mirror) required active search to locate the food. The other (in the right column: stand mirror, monitor) presented visually displaced images of the food that could be seen through passive observation.

Manual mirror

In the manual mirror training step, each monkey was trained to move a circular mirror (20 cm diameter) using an attached aluminum handle (55 cm long). The mirror could slide left and right along a 70 cm horizontal rail mounted 21 cm above the table.

Remote-control (RC) mirror

In the RC mirror step, we used a mirror that was electronically connected to a joystick. The mirror was initially rectangular [14.5 cm (h) × 19.4 cm (w)], and was later exchanged for a square one (10 cm × 10 cm). Both mirrors were hung on a 64-cm-long horizontal bar mounted 8 cm above the table. Left and right on the joystick corresponded, respectively to left and right movements by the mirror.

Stand and semi-transparent mirrors

The stand mirror and transparent mirror were of the same size [70.5 cm (w) × 30.0 cm (h)] and were mounted at an angle of 60°–75° and at a height of 28–48 cm (the mirror's horizontal midline) relative to the surface of the table. The semi-transparent mirror had a reflectivity of 40%, so that when it was placed just in front of a TV monitor (see below), the food was more easily viewed onscreen than in the mirror.

Monitor

In the monitor step, a 32″ liquid TV monitor [70.5 cm (w) × 39.5 cm (h)] was placed at visual angle of 38° relative to the subject. In these phases, an opaque horizontal screen [75.5 cm (w) × 33.0 cm (d)] was installed 10.5 cm above the table. A small aperture in the screen allowed the monkey to move his arms and bring successfully retrieved food items up to his mouth, but otherwise the screen blocked his view of the table. The tabletop beneath the screen was illuminated by LED lights installed along the screen's far edge. On the tabletop, a checkerboard grid pattern (5 cm^2) was drawn in the area, where the rake was moved around. Three black dots were painted at three particular intersections: the farthest point was centered on the table, 56 cm from the monkey, and the other two points were 28 cm from each side of the table, 51 cm from the monkey. These points were used to train the monkey to retrieve food from fixed positions. Initially the food was put only on the center dot. Once he was retrieving food from the center dot within five pulling movements in consecutive five trials, we enlarged the area, he had to search to encompass all three dots.

Procedure

The training sequence of Stage III was: stand mirror, manual mirror, RC mirror, monitor. This order was determined so as to train the animal with each type of visual displacement (active, passive) alternately, building generalizations from each preceding step.

Stand mirror use

This step began identically to transitional training II, except now the rake was non-mirrored. At first the stand mirror

was placed just behind the hump, and then was raised to a height of 28 cm from, and was tilted 75° relative to, the surface of the table.

Manual mirror use

The monkey was trained to use a manual mirror, which as described earlier could be slid back and forth along a horizontal rail using an attached handle. First, we started training the monkey to move the mirror. If he moved the mirror and it reflected the food, one of the experimenters stopped him moving the mirror, and then gave him the non-mirror rake to retrieve it. At this point, the mirror could be moved through a restricted range of 40 cm (manual mirror I). Gradually he learned to stop the mirror on his own when he spotted the food's reflection. To avoid the chance that the experimenter might inadvertently cue him to stop the mirror, we waited for him to ask us to pass him the non-mirror rake. To ask for the rake, M1 was required to pat the experimenter's hand and M2 was required to grab the experimenter's arm. At this point the mirror was fully movable along the bar's entire 70 cm length (II).

Remote-controlled (RC) mirror use

After mastering the manual mirror, each monkey was trained to use a joystick to control the position of a mirror that could slide along another horizontal bar behind the hump. The training steps were the same as in the manual mirror step. In the first phase a larger rectangular mirror was used, and one of the experimenters stopped the controller when the mirror came to the position that reflected the food (RC mirror I). In the second phase the prompt was omitted, but the same large mirror was used (II). Finally a smaller square mirror was used without the prompt (III).

Transparent mirror and monitor use

A transparent mirror just like the stand mirror was used. The mirror functioned the same as the stand mirror when not backlit, but became transparent when the monitor behind it was turned on. We used this transparent mirror to bridge stand mirror use and monitor use. The transparent mirror was placed in front of the monitor at increasing heights, first at 28 cm (transparent mirror I), then at 38 cm (II, III), and finally at 48 cm (IV). The distance between the subject and the mirror was also increased, from 65 (I, II) to 80 cm (III and onward). On the monitor the back of the hump was shown, as captured by a fixed camera located at the center of the table just in front of the monitor. Left and right were inverted on the monitor, so that the monitor would behave similarly to a mirror. The luminance of the monitor was gradually increased. In phase V, the monkey

could use the visual information from both the transparent mirror and the monitor. Finally, in transparent mirror VI, we removed the transparent mirror, leaving only the monitor available for finding the food. In the monitor use training, the hump was removed and the monkey had to search for the food on the flat surface of the table.

Stage IV

Apparatus

The camera rake was acquired as a sensory tool. The camera rake had a transparent, acrylic, semicircular tip (10 cm diameter) attached to a round aluminum handle (45 cm). A small CCD camera (0.6 cm diameter, AS-807-SP-10, ARS) was mounted inside the handle, facing forward, 9 cm from the tip of the rake. The camera's video feed was presented on the same monitor used earlier under the same viewing conditions. For food rewards, pieces of apple, sweet potato, raisins, and peanut were alternated. We introduced a rubber, green, toy frog [10 (d) × 17 (w) × 65 (h)] in later phase of the training.

Procedure

Camera rake training for M1

M1 was trained to use a camera rake without any transitional steps. We randomly placed the food on the table under the opaque screen and released his arm at either edge of the table. All the visual information used for searching was displayed on the monitor, captured by the small camera mounted inside the rake. The training was continued until the monkey could make smooth, circular arcs to find and retrieve the food.

Camera rake training for M2

M2 was trained with the camera rake using a sequential training procedure in which the level of difficulty was gradually increased. For the first 92 trials, one experimenter placed the camera rake on the table and another experimenter placed a piece of food between the rake tip and the camera. Thus, the monkey could get food simply by pulling the rake back to his side. For the next 64 trials, the camera rake and the subject's arm were released at the edge of the table and the food was always in the same location, on the central dot on the table. Then, for the next 183 trials, the food's position varied among all three dots. Next, we began placing the food randomly on table. Up through these preliminary phases, the area for moving the camera rake was limited to within a 45 cm wide range, limited by barriers.

M2 was trained in this procedure for 296 trials. Finally, the area was fully enlarged (77 cm) to the same range that had been used for M1.

To confirm that the monkeys used the camera rake to explore on the table, not just to pull the rake accidentally on viewing the food, we introduced trials in which we did not placed the food at all, but pretend to place it on the table, and released the monkeys to see how they moved the rake, in later phase of the training. These trials were ended when the monkeys spontaneously ceased to move the camera rake. In addition, to see whether the monkey used the camera rake to see objects in front of it in general, we placed a toy frog, which we knew that it induced scarring response in M1, on the table, and let him to explore with the rake.

Data analysis

We defined success by, whether the monkey could get the food with the rake in the first trial in Stages I, II, and III. Latency to retrieve the food would be another index for evaluating the change of performance, but the time was very much variable in each trial depending on the places of the food. Because we wanted to evaluate the functional use of the rake, rather than skillful use, we selected the number of successful trial as a measure of acquisition.

The same criterion, however, could not be applied in Stage IV because the animals had to move the rake several times, both for seeking and retrieving the food. Thus, we analyzed the trajectories of the camera rake to see development of the skillfulness at different training points. Additionally, we counted the number of "pulling-back of the camera rake" to see if the number would decrease after repeated training trials, suggesting that the monkeys used the camera rake for obtaining sensory inputs, not for retrieving the food by accident. "Pulling-back" of the rake was defined as moving the rake tip more than 25 cm vertically to the monkey's side, after he started searching for food at the beginning of each trial.

Results

We trained the two Japanese monkeys, M1 and M2. M1 was trained first. We exploratively trained him using various kinds of tool and apparatuses, and analyzed what types of behavioral component were essential for training aimed at the acquisition of sensory tools. Based on our insights from training M1, a reorganized training protocol was applied to M2 to prove the effectiveness and replicability of the result, though the training process for both monkeys was nearly identical. Thus, most of the quantitative analysis presented here comes from M2. The schematic protocol of

M2's training is depicted in Fig. 1. The essence of the process was to train rake use first (motor tool), then gradually separate the visual cues from their actual origins in visuo-motor space through a series of external mirror and video display devices until exploration, reaching and food retrieval were completely guided by the camera rake.

Table 1 shows the total number of trials required for meeting the criterion in each phase in pre-tool-use, Stages I, II, and III. Fig. 2 shows the learning curve for the first phase of each training step in M2, indexed by the cumulative number of successful trials against the total number of trials. All but two of the learning curves followed a simple increasing function; the exceptions, the platform and mirror rake steps, are discussed below.

Performance of Stage I

During the first few phases of the rake use training (pre-tool-use, rake I and II of Stage I), the monkey smoothly acquired the ability to wield the rake ("initial rake" in Fig. 2), as has been repeatedly described earlier (Ishibashi et al. 2000), taking 38 trials to meet criterion in total

Table 1 The number of the total trials required for meeting criterion in each phase of pre-tool training and Stages I–III

Stage	Phase	Total trials
Pre-tool use	Spoon-tipped tool	6
	Ring-tipped tool	16
Stage I	Rake I	10
	Rake II	6
	Rake III	50
	Rake IV	124
	Rake V	184
	Rake VI	107
	Rake VII	32
	Platform	138
Stage II	Mirror rake I	163
	Mirror rake II	55
Stage III	Stand mirror	39
	Manual mirror I	145
	Manual mirror II	16
	RC mirror I	61
	RC mirror II	38
	RC mirror III	8
	Transparent mirror I	5
	Transparent mirror II	12
	Transparent mirror III	5
	Transparent mirror IV	6
	Transparent mirror V	5
	Transparent mirror VI	5
	Monitor	96
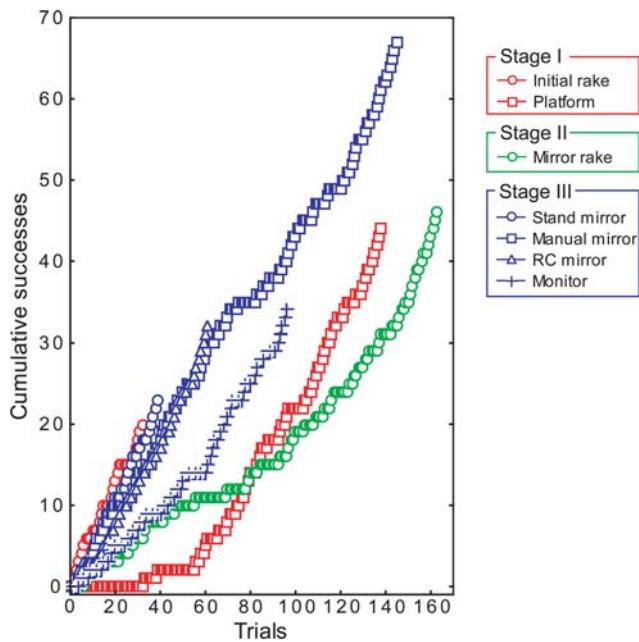
Fig. 2 Cumulative number of correct trials until reaching our criterion of successful retrieval in five consecutive trials, calculated for the first phase of each training step for M2

(Table 1). As the training required more complex movements of the rake in phase III–VII, the number of the required trials was increased, especially in phase V which required them to make circular arc of the rake tip. During the last phase of Stage I the "platform" was introduced ("platform" in Fig. 2). Prior to the platform's introduction, the monkey simply slides the rake in nearly straight lines across the tabletop to the food. Apparently the platform forces him to see and act in three dimensional space, not just in the two dimensions of the tabletop. Failure during this stage continued for the first 60 trials, then was followed by steady improvement until meeting criterion by trial 138. In total, Stage I took 651 trials over 3 days to complete.

Performance of Stage II

Stage II was the most crucial stage of the whole training procedure, as it comprised first step toward externalizing the eye via the hand. Food items placed behind the hump could not be seen directly, so the monkey had to actively explore the area behind the hump using reflected visual images in the tip of the rake (Supplementary video S1). Performance improved, though much more slowly than in other tasks, as clearly shown by the most delayed acceleration of the curve, until around trial 120. Prior to this, the monkey swept the mirror rake back and forth until it reflected the food by chance, which triggered pulling (Pepperberg et al. 1995) just as in the initial rake training stage. After trial 120, he started active search along the entire range of the hump. The learning process through this step

was relatively slow, taking 163 trials before meeting criterion. It took 218 trials within a day to finish Stage II in total.

Performance of Stage III

As shown in Fig. 2, the RC mirror was mastered faster than the manual mirror (61 and 145 trials, respectively), suggesting that competence with the manual mirror transferred effectively to the novel apparatus. The monitor took longer to master than the stand mirror (96 vs. 39 trials, respectively). In sum, in this stage the monkey used a variety of mirroring tools to locate and guide its reaching for food (Supplementary videos S2–S5). It took 9 days in total to finish all steps in Stage III.

Performance of Stage IV: acquisition of sensory tool use

After Stage III, all the essential cognitive and motor components were in place to acquire the use of an externalized eye. In Stage IV each animal mastered food retrieval using the camera rake. We analyzed the trajectories of the camera rake in different learning periods (i.e., first 100 and 400 trials after starting the training) in Stage IV to evaluate its functional use as an externalized eye. During the first 100 trials, M1 often pulled and pushed the rake back and forth in random directions until the rake accidentally caught the food (Fig. 3a). However, he occasionally stared at the monitor as if he noticed the correspondence between his own arm movements and the moving image onscreen. After 400 trials, the trajectory gradually changed into smooth, circular arcs (Supplementary videos S6 for M1, S7 for M2). At this point, we began to compare the trajectory of the camera rake's tip between two conditions: when the experimenter actually put food on the table (Fig. 3b) or only pretended to do so (Fig. 3c). The monkey immediately moved the camera rake in a circular trajectory from right to left. If food was present, he smoothly caught it with the rake tip and pulled it in; whereas if food was absent, he searched the entire table with multiple swoops of the rake, and then ceased to work. The difference clearly indicates that the monkey did not simply wait for the image of food to appear on the monitor independently of the tool's motion, but used the correspondence between the motion of the extended hand (which includes the tool) and the optic flow of the image on the monitor for functional exploration. Further evidence for functional exploration came when the experimenter moved the food horizontally on the table (Supplementary video S8) and M1 followed the moving food with the camera rake (upper-left block in the video). It took 600 trials over 5 days for M1 and 1,100 trials over 8 days for M2 to come to stably make circular movements to retrieve food. We consider these points of training as acquisition of the camera rake use.
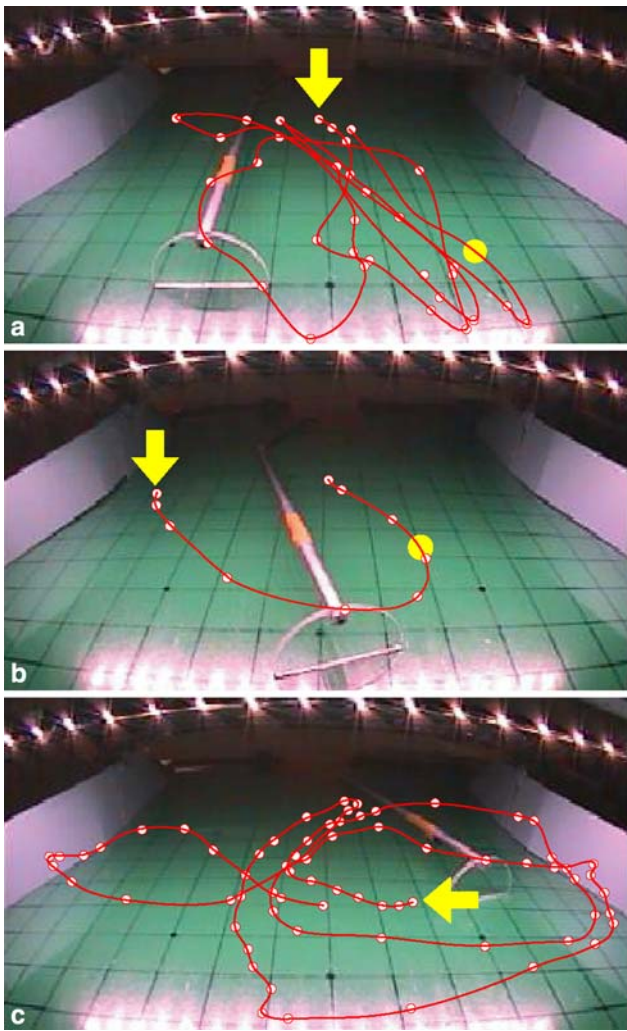
Fig. 3 Representative trajectories of M1's camera rake after food has been placed (**a**, **b**) and after sham placement (**c**), during early (**a**) and late (**b**, **c**) phases of Stage IV. The camera was fixed at the center of the table and faced the monkey from a distance of 70 cm. The video capture rate was 30 frames/s. *Yellow arrows* starting point of rake. *Yellow circles* (**a**, **c**) food position. *Red circles* center of camera rake tip every 1/3 s (every 10th frame). **a** Early in the training, the monkey just pushed and pulled the rake back and forth repeatedly until it retrieved the food by chance. The whole process took 5.6 s. **b** Later in the training, M1 began to make arcing trajectories with the rake. When the food was put on the table, he searched actively with the camera rake and retrieved the food very efficiently, in just 1.5 s. **c** When no food was present, the monkey searched the entire surface of the table repeatedly with the rake before giving up. The amount of time spent searching before giving up varied, but usually lasted longer than 10 s. In this instance it lasted 7.5 s

The number of pulling-back the camera rake also changed in accordance with the change in trajectories in Fig. 3. When the number of pulling-back in the final training condition, where the food was randomly placed on the table, was analyzed by averaging in 100 trials to the point of acquisition, it was linearly decreased ($y = -0.574x + 6.79$, $R^2 = 0.92$) from the first (6.09) to the last, sixth blocks

(3.66) in M1. On the other hand, M2 showed a bit non-linear decreasing function. In the first block of 100 trials, he pulled the camera rake 3.28 times in average, but the number was suddenly increased to 8.44 at the second block. However, he showed monotonically decreasing function afterwards ($y = -0.49x + 8.21$, $R^2 = 0.86$). Figure 4 depicted the number of pulling-back in the early, (where maximum number of pulling-back was observed: 1st block in M1, and 2nd block in M2) and the last blocks (6th block in M1, and 11th block in M2) of the training. The difference between the blocks was significant in both subjects [t (99) = 3.39, $P < 0.01$ in M1, t (99) = 5.88, $P < 0.001$ in M2].

Another evidence of functional use of the camera rake was observed, when we did not placed the food on the table, but just pretended to do so, both subjects started searching for the food with it by making circular arcs. After a while, however, they stopped working or repeatedly fumbled the rake with the direction upside down. In case of M1, who hated the toy frog, when we placed it, instead of food, on the table without showing it to him, he started searching, but immediately after catching the sight of the toy by the camera rake on the monitor, he threw the rake on the table. These cases were observed spontaneously after acquisition of the camera rake.

Discussion

Through our trial and error training of M1, we developed a training procedure that successfully and efficiently replicated M1's sensory tool use learning in the naive M2. Step-by-step transition from motor tool to sensory tool learning was crucial, as indicated by the failure of M1 to localize the food in the monitor before the present training. In our interpretation, extension of the hand (Iriki et al. 1996) through motor tool training induced a gradual objectification of the hand until hand and tool were perceived as equivalent (Iriki 2006). Then, through staged manipulation of visual cues and feedback, vision was separated from the eye and transferred to the mirror or monitor. By the end of training, the sensory organ was functionally externalized to an artificial device, the camera rake. Thus, our monkeys acquired the ability to make active use of a sensory tool, an "externalized eye". This is the first demonstration of this ability in nonhuman animals.

During Stage II, M2 started active search in the middle of the training, indicating a novel type of visual recognition—namely, mirror-mediated object discrimination and spatial localization (Pepperberg et al. 1995). This step was critical because the animal no longer used directly obtainable visual information, but used the displaced cues on the mirror. Without this step, the animal could not
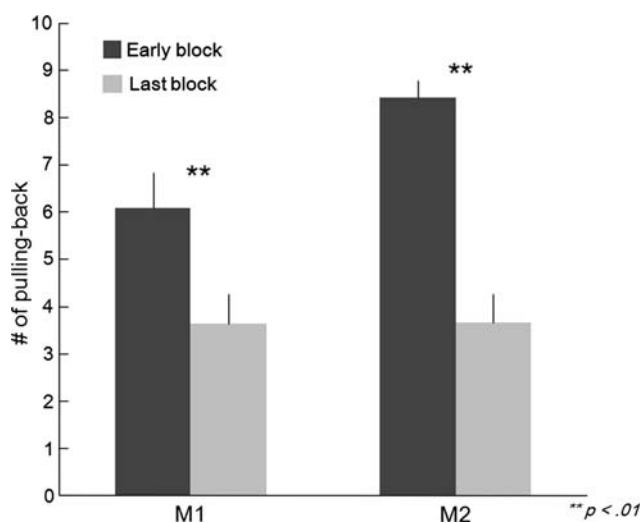
Fig. 4 The number of "pulling-back" the camera rake in the early (*dark grey bar*) and the last (*light grey bar*) blocks in Stage IV. The scores are averaged for 100 trials, depicted with standard errors. "Pulling-back" the rake was defined as moving the rake tip more than 25 cm vertically to the monkey's side, after he started searching for food at the beginning of each trial. The numbers were counted until 600 trials in M1 and 1,100 trials in M2, at which they showed smooth and circular movement of the rake to retrieve the food. The data of the early and the last blocks were obtained from the first and sixth blocks in M1, whereas they were in the second and eleventh blocks from M2

learn about the visual displacement reliably, as was justified by M1's failure of localizing food using the monitor (see "Results").

As shown in Fig. 2, there are some training steps in which the animal clearly had difficulty to reach the acquisition criterion other than mirror rake, such as platform and monitor. The apparent difficulty presented by the platform may be attributable to the novel motor and cognitive challenges this step introduced: to lift the rake to a higher position and to recognize and target the food in three dimensional space. Though seemingly small change from the preceding condition, manipulation of the tool in three dimensional space should have involved precise localization of the target and the tool, relying on the newly developed body image in the motor tool training. One plausible reason why monitor use seemed to be more difficult than stand mirror use could be loss of generalization: the size, color, shadow and shape of the food and other materials (such as apparatus and rake) was more distorted in the captured image on the monitor than in the reflected image on the stand mirror. Nevertheless, the animal could learn monitor use—unlike in the case of the three choice discrimination task tested earlier on M1—because it had already learned about visual displacement in the prior training steps.

Although both the monitor and camera rake steps involve locating and retrieving food using a captured

image, the cognitive abilities required for each task differ markedly. In the monitor task, the camera's position is fixed. The animal simply picks upon the geometric correspondence between the positions of the rake and the food on the monitor (Menzel et al. 1985), which might be achieved through simple sensorimotor learning. But in the camera rake task, each arm movement leads to a new camera view, with the position of and distance to the food changing continuously. This requires a much more sophisticated construction of the space in front of him, requiring something more than remapping of the sensorimotor coordination.

In camera rake training, M1 acquired the task much faster than M2, as shown in Fig. 4. This would reflect the difference in tool use experience between the two. Before being trained in the procedures described above, M1 was trained to use rake for more than 3 months, though he did not show any improvement in using captured image on the monitor. However, at the time when they mastered the camera rake, both monkeys took only one or two pulling movement to retrieve the food when they could use the rake with their own eyes, so one could easily understand how costly to use the camera rake to guide their behavior for reaching the same goal.

One could argue that manipulation of the camera rake could be achieved simply by remapping sensorimotor coordination—much as one adjusts to wearing deviating prism glasses and quickly resumes using correct reaching angles, as the monkeys were already probed to be capable of remapping their body parts (hand) and the target object (food) on the artificial device (monitor) (Iriki et al. 2001). In fact, the time courses of the number of pulling-back the camera rake (Fig. 4) in both monkeys suggested that they did not spontaneously combine the trained elements for camera rake use in the novel situation. If remapping were the explanation, however, the monkeys should have learned each task at a monotonic rate. But the facts that, first, some training phases displayed delayed acquisition as shown in Fig. 2, and second, that M1 failed to learn to use the motor tool with displaced visual cues without any step-by-step training, invalidate this argument, and suggest that some novel combinations of cognitive functions must have been employed to manage the task.

Other criticisms might be raised that evidence of the sensory tool use in nonhuman has already been appeared previously, by arguing mirror use in macaques (Anderson 1986) and monitor use in macaques (Washburn and Rumbaugh 1992) and chimpanzees (Menzel et al. 1985) as representative examples or even spiders utilizing the vibration on trap web would be an example of the sensory tool. However, we have clear reason to argue that these examples are not the evidence of sensory tool use in nonhuman animals. They do involve subjects gathering sensory inputs via external

objects, but this is a superficial similarity. If one refers to a detailed definition of tool use (Beck 1980), none of these examples qualify, because the objects are not carried or manipulated by the users, and the users are not responsible for the effective orientation between tools and targets (Thompson and Boatright-Horowitz 1994). Most critically, actions performed by the subjects do not produce any sensory feedback stimuli for guiding further action. Only our study cleared the criterion of tool use, and had specific function of the sensory tool.

The neural mechanisms subserving the acquisition of sensory tool use are unknown. However, we have shown elsewhere that motor tool training induces a novel extension of the neural connections between the parietal area, (where the modified body image that accompanies tool use is coded, Iriki et al. 1996) and higher visual cortices, (where information about visual flow is coded (Hihara et al. 2006; Siegel and Read 1997). In the present study, the training protocol required the monkeys to employ at least three different cognitive abilities for understanding (1) the relationships between the visual inputs directly from their original eyes and the one indirectly from the artificial objects (mirror and monitor), (2) the relationships between the proprioceptive, motion inputs from hand and tool and the visual input on the artificial object, and (3) spatial relationships of the objects on the table and the visual input. That is, the monkeys were required to integrate several different sources of inputs, such as visual, proprioceptive, spatial, and motion. Because some combinations of these abilities were involved in each task and therefore, they must be more demanded than those required for motor tool training, these multimodal integrations would have needed advanced, novel mode of neural coding in the monkey cortex (cf., Knudsen 2002).

Our monkeys highlight the potent significance of external variables—as opposed to genetically determined tool use granting factors—for constructing neatly connected reorganization of cognitive functions in the primate brain. Human cognitive evolution was probably not driven purely "genes up" but also "culture down," involving continual circular interaction between individuals, with modified neural connectivity, and environments, modified through tool use, over generations (cf., "niche construction", Laland et al. 2000). Although monkeys may lack certain neural adaptations that give humans an unparalleled facility for cultural transmission, they share many advanced sensorimotor and cognitive functions with us that, as the present study has attempted to show, can be induced into novel, higher order functional relationships through a properly structured training environment. Thus, the present study proposes an empirical framework for studying the evolution of higher human cognitive abilities by recapitulating

the route to them through epigenetic manipulation (Iriki 2006).

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

Anderson JR (1986) Mirror-mediated finding of hidden food by monkeys (*Macaca tonkeana* and *M. fascicularis*). J Comp Psychol 100:237–242

Asano T (1994) Tool using behavior and language in primates. In: Hayes SC, Hayes LJ, Sato M, Ono K (eds) Behavior analysis of language and cognition. Context Press, Reno, pp 145–148

Beck BB (1980) Animal tool behavior. Garland, New York

Fragaszy D, Izar P, Visalberghi E, Ottoni EB, de Oliveira MG (2004) Wild capuchin monkeys (*Cebus libidinosus*) use anvils and stone pounding tools. Am J Primatol 64:359–366

Goldenberg G, Iriki A (2007) From sticks to coffee-maker: mastery of tools and technology by human and non-human primates. Cortex 43:285–288

Hihara S, Obayashi S, Tanaka M, Iriki A (2003) Rapid learning of sequential tool use by macaque monkeys. Physiol Behav 78:427–434

Hihara S, Notoya T, Tanaka M, Ichinose S, Ojima H, Obayashi S, Fujii N, Iriki A (2006) Extension of corticocortical afferents into the anterior bank of the intraparietal sulcus by tool-use training in adult monkeys. Neuropsychologia 44:2636–2646

Hunt GR (1996) Manufacture and use of hook-tools by New Caledonian crows. Nature 379:249–251

Iriki A (2006) The neural origins and implications of imitation, mirror neurons and tool use. Curr Opin Neurobiol 16:660–667

Iriki A, Tanaka M, Iwamura Y (1996) Coding of modified body schema during tool use by macaque postcentral neurones. Neuroreport 7:2325–2330

Iriki A, Tanaka M, Obayashi S, Iwamura Y (2001) Self-images in the video monitor coded by monkey intraparietal neurons. Neurosci Res 40:163–173

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

Knudsen EI (2002) Instructed learning in the auditory localization pathway of the barn owl. Nature 417:322–328

Laland KN, Odling-Smee J, Feldman MW (2000) Niche construction, biological evolution, and cultural change. Behav Brain Sci 23:131–175

Matsuzawa T (1994) Field experiments on the use of stone tools by chimpanzees in the wild. In: Wrangham RW, MacGrew WC, de Waal FBM, Heltne PG (eds) Chimpanzee cultures. Harvard University Press, Cambridge, pp 351–370

Menzel EW Jr, Savage-Rumbough ES, Lawson J (1985) Chimpanzee (*Pan troglodytes*) spatial problem solving with the use of mirrors and televised equivalents of mirrors. J Comp Psychol 99:211–217

Pepperberg IM, Garcia SE, Jackson EC, Marconi S (1995) Mirror use by African grey parrots (*Psittacus erithacus*). J Comp Psychol 109:182–195

Siegel RM, Read HL (1997) Construction and representation in the visual space in the inferior parietal lobule. In: Rockland KS, Kaas JH, Peters A (eds) Cerebral cortex. Plenum, New York, pp 499–526

Thompson RL, Boatright-Horowitz SL (1994) The question of mirror-mediated recognition in apes and monkeys: some new results and reservations. In: Parker ST, Mitchell RW, Boccia ML (eds) Self-awareness in animals and humans: developmental perspectives. Cambridge University Press, New York, pp 330–349

Visalberghi E, Fragaszy D (2006) What is challenging about tool use? The capchin's perspective. In: Wasserman EA, Zentall TR (eds)

Comparative cognition. Oxford University Press, New York, pp 529–552

Washburn DA, Rumbaugh DM (1992) Comparative assessment of psychomotor performance: target prediction by humans and macaques (*Macaca mulatta*). J Exp Psychol General 121:305–312