Visual Adaptation to Goal-directed Hand Actions

Nick E. Barraclough[1,2], Rebecca H. Keith[1], Dengke Xiao[2],
Mike W. Oram[2], and David I. Perrett[2]

Abstract

■ Prolonged exposure to visual stimuli, or adaptation, often results in an adaptation ''aftereffect'' which can profoundly distort our perception of subsequent visual stimuli. This technique has been commonly used to investigate mechanisms underlying our perception of simple visual stimuli, and more recently, of static faces. We tested whether humans would adapt to movies of hands grasping and placing different weight objects. After adapting to hands grasping light or heavy objects, subsequently perceived objects appeared relatively heavier, or lighter, respectively. The aftereffects increased logarithmically with adaptation action repetition and decayed logarithmically with time. Adaptation aftereffects also indicated that perception of actions relies predominantly on view-dependent mechanisms. Adapting to one action significantly influenced the perception of the opposite action. These aftereffects can only

be explained by adaptation of mechanisms that take into account the presence/absence of the object in the hand. We tested if evidence on action processing mechanisms obtained using visual adaptation techniques confirms underlying neural processing. We recorded monkey superior temporal sulcus (STS) single-cell responses to hand actions. Cells sensitive to grasping or placing typically responded well to the opposite action; cells also responded during different phases of the actions. Cell responses were sensitive to the view of the action and were dependent upon the presence of the object in the scene. We show here that action processing mechanisms established using visual adaptation parallel the neural mechanisms revealed during recording from monkey STS. Visual adaptation techniques can thus be usefully employed to investigate brain mechanisms underlying action perception. ■

INTRODUCTION

The application of visual adaptation to understanding brain mechanisms underlying visual perception has a long history. In psychophysical experiments, adaptation consists of prolonged exposure to a stimulus with closely defined visual parameters, and this affects a suppression of the neural mechanisms that underlie the coding of those precise visual characteristics. Perceptual judgments of test stimuli after the adapting period are often biased, and the character of the ''aftereffects'' can illustrate the function of the adapted neural mechanisms underlying the perception of the stimuli. Indeed, single-unit recording in animals has shown that cell responses are significantly reduced after adaptation to preferred stimuli (stimuli to which they are ''tuned''), but are little affected by adaptation to nonpreferred stimuli. Stimulus-specific reduction in cell responses after adaptation has been seen at many levels in the visual system including, for example, in V1 cells after spatial frequency adaptation (Saul & Cynader, 1989), in V5 cells after motion adaptation (van Wezel & Britten, 2002), and in the infero-temporal cortex and superior temporal sulcus (STS) cells after adaptation to objects and complex images (Baylis &

Rolls, 1987). In the last few years, several research groups have demonstrated that it is possible to selectively adapt mechanisms in the visual system that code social stimuli (e.g., Leopold, O'Toole, Vetter, & Blanz, 2001).

Experiments using static faces as adapting stimuli have helped elucidate how they are coded in the human visual system. Benton et al. (2007) demonstrated that adapting to facial expressions seen from one viewpoint has a different influence on subsequently presented test faces depending upon the relative viewing angle of the adapting and test faces. Strongest aftereffects were observed with test faces seen from the same viewpoint as the adapting face; as the difference in viewing angle was increased, aftereffects decreased, although some aftereffects were still evident when the adapting and test faces were seen from views 90° apart. Benton et al. argue that this demonstrates that the coding of facial expressions relies on mechanisms that are both viewpoint dependent and viewpoint independent. Indeed, facial identity also appears to rely on a combination of both viewpoint-dependent and viewpoint-independent mechanisms (Benton, Jennings, & Chatting, 2006).

Often, the mechanisms underlying human perception, established using psychophysical adaptation paradigms, show remarkable parallels to the neural mechanisms revealed during monkey neurophysiological studies. Earlier studies investigating monkey STS cells preferentially

[1]University of Hull, Hull, East Yorkshire, UK, [2]University of St Andrews, Scotland, UK

sensitivity to faces showed that separate populations of cells code faces with and without respect to the viewpoint of the observer (Perrett et al., 1991; Hasselmo, Rolls, Baylis, & Nalwa, 1989; Perrett, Smith, Potter, et al., 1985). It is likely that the monkey's assessment of faces engages both of these populations of cells.

Other characteristics of adaptation aftereffects seen in psychophysical studies with humans are also mirrored in the cellular coding of facial stimuli in monkey single cells. In humans, facial identity aftereffects are relatively insensitive to small (6°) changes in the position of the stimulus (Leopold et al., 2001), and face shape aftereffects, although maximal when test faces are the same size as the adapting stimuli, are also substantial when the sizes of adapting and test faces are different (Zhao & Chubb, 2001). Many cells in the monkey temporal cortex that are sensitive to faces often have large receptive fields (Bruce, Desimone, & Gross, 1981) and are relatively insensitive to the position of the face within the receptive field (Tovee, Rolls, & Azzopardi, 1994). Monkey STS cells are also broadly tuned to the size of the facial stimulus and can therefore be relatively insensitive to stimulus size (Perrett, Oram, & Ashbridge, 1998; Rolls & Baylis, 1986; Perrett, Rolls, & Caan, 1982).

In humans, adaptation aftereffects are further observed after viewing complex motion stimuli including expanding fields of random dots (Meng, Mazzoni, & Qian, 2006) and "biological motion" stimuli (Troje, Sadr, Geyer, & Nakayama, 2006). "Biological motion" stimuli contain information about the local movement trajectories of, usually, points of articulation in walking human figures; information, however, about the form or shape of the walker can be unavailable when static. Cells within a posterior region of the monkey STS, the medial superior temporal area (MST) respond to expanding random dot fields (Saito et al., 1986). In the more anterior STS, many cells respond selectively to walking human figures (Oram & Perrett, 1996), and a third of STS cells sensitive to walking humans will respond to "biological motion" stimuli (Oram & Perrett, 1994). In summary, these studies have shown separately that the human visual system can adapt to faces and complex motion stimuli. In addition, the perceptual mechanisms revealed using adaptation paradigms in humans closely parallel the cellular coding of similar stimuli seen during monkey neurophysiological studies.

We wanted to investigate if it was possible to use the technique of visual adaptation to examine the mechanisms underlying the perception of hand actions in human observers. Visual adaptation typically shows a characteristic build up in strength with increasing exposure to the adapting stimulus, and exponential decay with time (e.g., Rhodes, Jeffery, Clifford, & Leopold, 2007). To test if adaptation to hand actions would show similar dynamics to those seen previously with simple visual stimuli, we measured adaptation aftereffects in human observers after adapting to a grasping action and a placing action. We varied both the number of times the

adapting action was repeated and the duration of the interstimulus interval (ISI) between the adapting and test stimuli. In order to test whether hand actions are coded by view-dependent or view-independent mechanisms, we measured adaptation aftereffects when the adapting and test actions were seen from similar or different perspectives. To investigate if different hand actions with different goals are coded by the same or different neural mechanisms, we tested if adapting to one action, with a specific goal, influences the subsequent perception of a different action with a different goal.

Psychophysical adaptation experiments are used to infer the tuning properties of cells underlying the perception of individual stimuli, as the responses of neurons tuned to the adapting stimulus are selectively reduced by repeated exposure (e.g., see Kohn, 2007 for a review). Such neuronal properties have been demonstrated in many visual areas, including the STS (Baylis & Rolls, 1987). We wanted to know if the response properties of monkey STS cells sensitive to hand action stimuli confirmed the properties determined during our adaptation experiments. Rather than investigate the effect of repeated exposure to our hand action stimuli, we made the assumption that STS cells would show reduced responses to repeated instances of hand action stimuli as has been demonstrated in STS cells with other stimuli (Baylis & Rolls, 1987); we instead tested the sensitivity of STS cells that responded preferentially to either grasping or placing actions to the opposite action. To test whether viewpoint was important for STS cell coding of hand actions, we measured responses to grasping and placing actions seen from different perspectives. To test whether the interaction between hand and object was critical for STS cell responses to the hand actions, we measured the responses to the actions with and without the presence of the object.

METHODS

Stimuli

Stimulus Set 1

A female hand reaching out and lifting from a table an abstract black object of 500 g weight was filmed with a 3CCD digital video camera (Canon XL1s) at 25 frames/sec progressive scan, 720 × 576 pixels, 16-bit color depth, and then digitized. The action consisted of a hand reaching toward the object (reach phase), grasping the object and lifting it away again (withdraw phase; Figure 1), and the complete action lasted 1520 msec (38 frames, at 40 msec/frame).

Previous research has showed that subjects use the various kinematics of the grasping action in order to judge the weight of lifted objects (e.g., Hamilton, Joyce, Flanagan, Frith, & Wolpert, 2005). When the withdraw phase of the action is prolonged compared to the reach phase of the action, objects are perceived as being
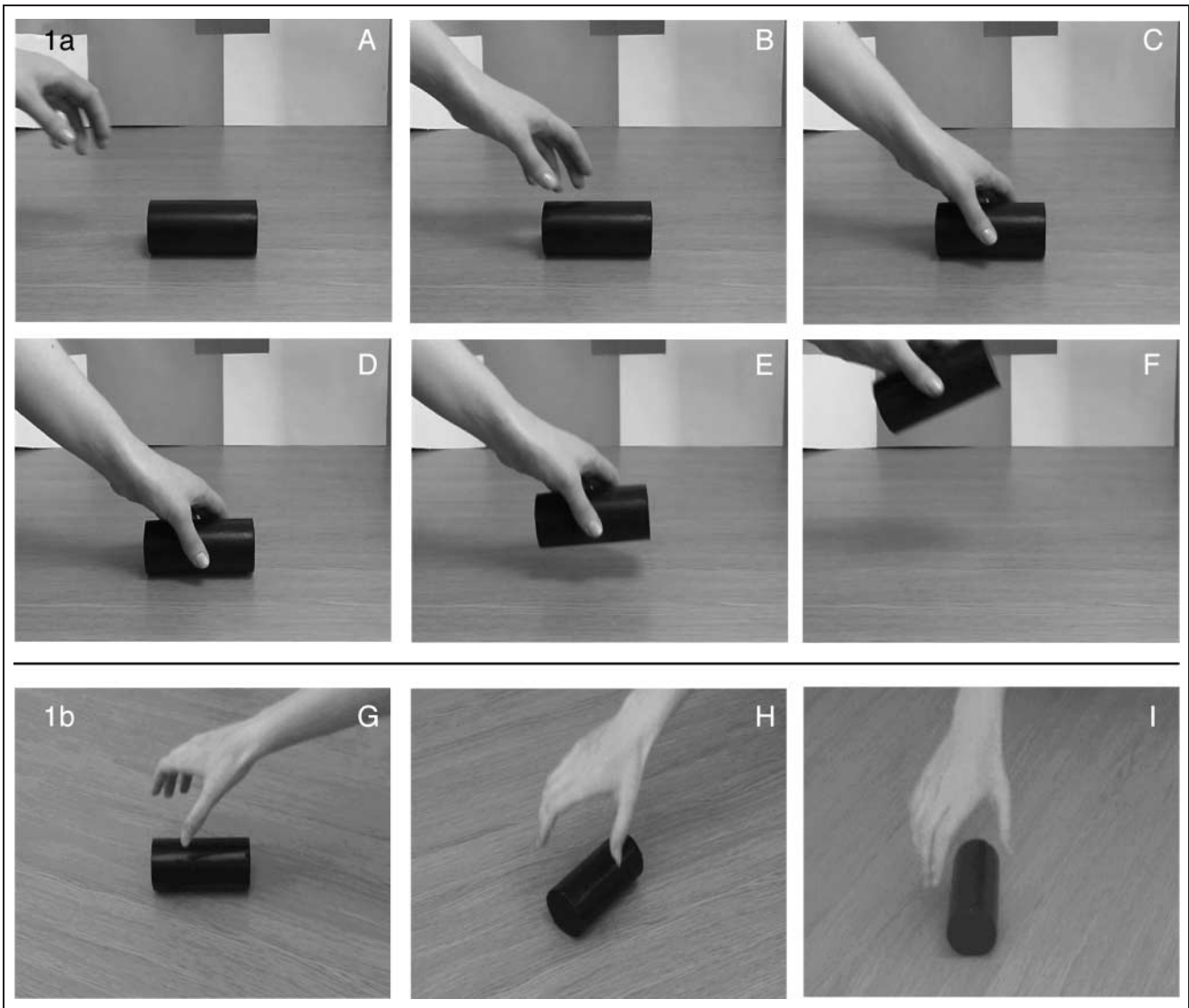
Figure 1. Example images from Stimulus Sets 1 and 2. (1a) A–F: Example images from Stimulus Set 1 used in experiments to test the dependence of action adaptation on action repetition and the duration of action adaptation. Images 2 (A), 8 (B), 15 (C), 22 (D), 29 (E), and 36 (F) taken from the grasping action with ratio −0.053. (1b) G–I: Example images from Stimulus Set 2 used to test the dependence of action adaptation on viewing angle and action goal. Image 9 from grasping actions with ratio 0.142, viewed from 0° (G), 45° (H), and 90° (I).

heavier (Hamilton et al., 2005). We edited the original action stimulus in order to generate a set of new action stimuli where the hand appeared to grasp different weight objects. Video frames were removed from or duplicated at regular intervals within the two phases of the original action in order to respectively shorten or extend the reach or withdraw phase durations. Ten actions in total were generated, ranging from an action with an extended reach phase (1200 msec) and a brief withdraw phase (320 msec), to an action with a brief reach phase (120 msec) and an extended withdraw phase (1400 msec).

Consequently, each of the 10 generated grasping actions lasted 1520 msec and showed the same hand shape and the same movement trajectory, the difference between the actions was the relative speed of movement of the hand in the different phases, or "action ratio." The action ratio was calculated as: (duration reach phase − duration withdraw phase)/(duration reach phase + duration withdraw phase). Positive action ratios indicate that the reach phase of the action was longer than the withdraw phase of the action; negative action ratios indicate that the withdraw phase of the action was longer than the reach phase; ratios ranged from 0.58 to −0.84. Preliminary informal testing showed that subjects consistently judged the object lifted in grasping actions with positive action ratios as light, and judged the object lifted in grasping actions with negative action ratios as heavy.

Stimulus Set 2

As for Set 1, a similar action was filmed simultaneously from three different viewpoints with three cameras

(35 images/frames, duration 1400 msec). The principal filming angle (Canon XL1s) was orthogonal to the trajectory of movement of the hand grasping the object. A second camera (Sony DV: DSR-PD100AP) filmed the action from $45°$ to the left of the principal camera; a third camera (JVC, GR-D720 DVCAM) filmed the action from $90°$ to the left of the principal camera. All films were digitized and synchronized to each other by matching the exact kinematics of the action. As for Stimulus Set 1, the three action films were edited simultaneously to generate 27 action movies (3 views, 9 different weight objects). The luminance of each action frame was shifted to the mean luminance of the display monitor, the range of contrasts in each frame was made equal. Still frames of the grasping action filmed from these three different views are illustrated in Figure 1.

Movies from Stimulus Set 2 were displayed in forward sequence to generate grasping actions and in reverse order to generate placing actions. Thus, for any one view, there is a grasping and placing action that contain the same images, and therefore, the same hand and object identity, shape, and movement over the same region of visual space. Reversing the movie also reverses the perceived weight: Reversing grasping a light object (slow reach, fast withdraw: positive action ratio) generates placing a heavy object (fast reach, slow withdraw: negative action ratio). Equivalently reversing grasping a heavy object (negative action ratio) generates placing a light object (positive action ratio). Irrespective of whether the action is grasping or placing, the action ratio has the same meaning: More positive action ratios indicate lighter objects and more negative ratios indicate more heavy objects.

Stimulus Set 3

A third stimulus set was made to test STS cells' responses to actions similar to those used in the human adaptation experiments. Stimuli consisted of 24-bit color movies of either a human hand or a monkey hand grasping either an abstract pink ball or a raisin, respectively. Actions were filmed with a 3CCD digital video camera (Panasonic, NV-DX110), lasted 800 msec (20 frames), and were cropped to 256×256 pixels. Each individual frame of each movie was also flipped horizontally to create further movies with the hand acting from the opposite side ($180°$ away). An additional set of movies was prepared where the object or hand was edited out of each movie so either the hand appeared to mime the action, or the object appeared to move alone.

Human Subjects

Subjects consisted of students and staff from the University of Hull, students ether received course credit or were paid for participating. All subjects had normal or corrected-to-normal vision; subjects were naïve to the purpose of the experiments (except authors N. B. and R. K., who took part in all experiments). Fifteen subjects took part in Experiment 1 (11 women, $\mu = 21.3$ years, $SD = \pm3.8$). Eleven subjects took part in Experiment 2 (8 women, $\mu = 21.4$ years, $SD = \pm4.5$), all except one took part in Experiment 1. Seventeen subjects took part in Experiment 3 (12 women, $\mu = 21.7$ years, $SD = \pm5.1$), of these subjects, seven had also taken part in Experiments 1 and 2. Twelve subjects took part in Experiment 4 (7 women, $\mu = 24.9$ years, $SD = \pm10.5$), of these subjects, six had also taken part in Experiments 1, 2, and 3. The ethics committee of the Department of Psychology, University of Hull, approved all experiments.

Human Psychophysical Experiments

A PC running MATLAB 2006a and the Cogent toolbox was used to control the experiment, display the stimuli, and record subject responses. Subjects sat in a darkened room approximately 57 cm from a 22-in. flat screen CRT monitor (Phillips 202P40, 1280×1024 pixels, 100 Hz refresh rate) on which all visual stimuli were presented. Action movies were shown in the middle of a mid-gray (luminance $= 9.7$ cd m^{-1}) background at full resolution (720×576 pixels) and subtended $22.3° \times 16.6°$ at the eye. This was achieved by rendering on-screen in sequence each image/bitmap from the action movie at 25 frames/sec.

Prior to the first two adaptation experiments, each subject took part in a preadaptation test phase. Subjects were told that they would be shown movies of the act of a hand lifting different objects weighing anything between 100 and 900 g. Each of the 10 actions from Stimulus Set 1 was presented pseudorandomly 10 times (100 trials), where no stimulus was presented for the $n + 1$th time before all had been presented n times. On each trial, subjects were required to rate the weight of the object in the test movie and then indicate their response using the keyboard number pad (from $1 = 100$ g to $9 = 900$ g). After the subject had indicated their response, the screen remained blank (gray) for a period of 1500 msec before the start of the next trial.

Experiment 1: Influence of Action Repetition on the Action Adaptation Aftereffect

In the first experiment, we tested how the adaptation aftereffect varied with repetition of adapting action. The task was similar to the preadaptation test phase, except that on each trial the test movie was preceded by an adapting stimulus. The adapting stimuli consisted of a hand grasping a light object (action ratio 0.579) or grasping a heavy object (action ratio −0.842). On each trial, the adapting movie was shown a variable number of times in immediate succession (1, 2, 4, or 8 times; total adapting

stimulus durations 1520, 3040, 6080, or 12160 msec), followed by an ISI of 520 msec where the screen remained blank (gray). During the final 200 msec of the ISI, a small yellow fixation cross was presented in the center of the screen to inform the subject that the next movie was to be the test stimulus. Test stimuli were presented once and consisted of one of four movies of grasping intermediate weight objects. The 16 different conditions (4 × adaptation durations, 4 × test movies) were presented pseudorandomly five times (total 80 trials). Subjects adapted to one action on the first day (half: lightweight object) and returned on a subsequent day to perform the same procedure adapting to the other action.

Experiment 2: Duration of the Action Adaptation Aftereffect

During the second experiment, a similar procedure to the action repetition experiment was performed, except that the adapting movie was always shown four times in immediate succession (6080 msec) and the ISI was varied (200 msec, 4000 msec, or 8000 msec).

Experiments 3 and 4: Dependence of Hand Action Aftereffect on View and Adapting Action Goal

Before any adapting experiments were performed, subjects were initially tested in a preadaptation test phase using the actions taken from Stimulus Set 2. During this phase, subjects rated the weight of objects in four test movies covering a range of different action ratios, viewed from three different angles (0°, 45°, and 90°), displayed in a forward or backward sequence so that the actions appeared as both grasping and placing. All conditions (4 action ratios × 3 views × 2 action types) occurred 10 times each (total 240 trials).

The adaptation experiment was similar to the preadaptation test phase, except that on each trial, the test movie was preceded by the adapting stimulus (hand grasping a heavy object: action ratio −0.714, viewed from 0°, repeated 4 times: duration 5600 msec). The ISI lasted 150 msec and contained a yellow fixation cross throughout; finally the test stimulus was presented (the same stimuli as for the preadaptation test phase). Subjects took a 5 minute break at the midpoint of the experiment in order reduce eyestrain and help maintain subject concentration throughout testing.

In order to confirm our results, we performed a very similar experiment where all parameters were identical but the adapting stimulus was a hand placing a heavy object (action ratio −0.657). Of those subjects that performed both experiments, they adapted to grasping and placing actions on different days, half adapted to the grasping action first.

For every preadaptation and adaptation experiment, subjects' mean responses to each of the test movies were plotted against the test movie's action ratio (see Figure 2). For each separate experiment, a linear function was fitted to the data. Where this function crossed a threshold value of 5 (the midpoint of the 1–9 stimulus rating scale) was recorded for each experiment. The difference between action ratios, at this threshold, before and after adaptation, indicated the nature of the adaptation aftereffect. A positive value indicated that adaptation caused the object to appear heavier and a negative value indicated that adaptation caused the object to appear lighter. Analysis using the intercept the $x = 0$ produced equivalent results.

Monkey Physiological Subjects, Recording and Reconstruction Techniques

One rhesus macaque, aged 9 years, was trained to sit in a primate chair with head restraint. Using standard techniques (Perrett, Smith, Mistlin, et al., 1985) and carried out in accordance with the UK Animals (Scientific Procedures) Act 1986, recording chambers were implanted over both hemispheres to enable electrode penetrations to reach the STS. Single neurons were recorded using tungsten microelectrodes inserted through the dura. The subject's eye position (accuracy ±1°) was monitored (IView, SMI, Germany). A Pentium IV PC with a Cambridge electronics CED 1401 interface running Spike 2 recorded eye position, spike arrival, and stimulus on/offset times.

After each electrode penetration, x-ray photographs were taken coronally and parasagitally. The positions of the tip of each electrode and its trajectory were measured with respect to the intra-aural plane and the skull's midline. Using the distance of each recorded neuron along the penetration, a three-dimensional map of the position of the recorded cells was calculated. Coronal sections were taken at 1-mm intervals over the anterior–posterior extent of the recorded neurons. Alignment of sections with the x-ray coordinates of the recording sites was achieved using the location of microlesions and injection markers on the sections (see Harries & Perrett, 1991, for full details).

Monkey Physiological Experiments

All visual stimuli were stored on an Indigo2 Silicon Graphics workstation hard disk and presented centrally on a black monitor screen (Sony GDM-20D11, resolution 25.7 pixels/degree, refresh rate 72 Hz), 57 cm from the subject. Cell responses were isolated using standard techniques, and visualized using an oscilloscope. Systematic screening was performed with a search set of (on average 55) images and movies of different objects, body parts, and actions previously shown to activate neurons in the STS (Barraclough, Xiao, Oram, & Perrett, 2005; Foldiak, Xiao, Keysers, Edwards, & Perrett, 2003). Static images (subtending 19° × 19°, duration = 125 msec) and actions (subtending up to 25° × 20.5°, frame rate =

42 msec/bitmap) were presented in a random sequence with a 500-msec ISI. Presentation of this screening set commenced when the subject fixated (±3°) a yellow dot presented centrally on the screen for 500 msec (to allow for blinking, deviations outside the fixation window lasting <100 msec were ignored). Fixation was rewarded with the delivery of fruit juice. Spikes were recorded during the period of fixation; if the subject looked away for longer than 100 msec, spike recording and presentation of stimuli stopped until the subject resumed fixation for >500 msec. Responses to each stimulus in the screening set were displayed as on-line rastergrams and poststimulus time histograms aligned to stimulus onset.

Within the screening set were the movies taken from Stimulus Set 3, played forward (grasping action) and played backward (placing action). Occasionally, shortened versions of each action (every third frame, 7 frames in total) were included instead. A subset of those cells that showed a maximal response to either the grasping or placing action during the screening process was subsequently tested with an extended stimulus set. The extended stimulus set contained the same grasping and placing actions as used in the screening set, in addition, those same actions were presented from a different view where the movie frames were left–right flipped; all these actions were additionally presented without the presence of the object. Neural responses to all stimuli were recorded to hard disk for off-line filtering and analysis.

Single-cell Analysis

Off-line isolation of single cells was performed using a template matching procedure and principal components analysis (Spike 2). Each cell's response to a stimulus was calculated by aligning segments in the continuous recording, on each occurrence of that particular stimulus (trials). In order to account for blinking by the animal, eye movement information was used to include only those trials where the subject was fixating for over 80% of the first 300 msec of stimulus presentation (Barraclough et al., 2005).

For each stimulus, a poststimulus time histogram was generated and a spike density function (SDF) was calculated by summing across trials (bin size = 1 msec) and smoothing (Gaussian, σ = 10 msec). Background firing rate was measured in the 100-msec period prior to stimulus onset. Response latencies to each stimulus were measured as the first 1-msec time bin where the SDF exceeded 3 standard deviations above the background firing rate for over 15 msec in the period 0–400 msec following stimulus onset (Edwards, Xiao, Keysers, Foldiak, & Perrett, 2003; Oram & Perrett, 1992, 1996).

Responses to the grasping action and placing action were compared within a 100-msec window starting at each stimulus response latency. If no response latency could be obtained, then a default latency of 100 msec was used. For cells that were tested only with the screen-

ing set, data from this set were analyzed if the cell showed the biggest response to either the grasping action or placing action when compared to all other stimuli in the screening set. For cells that had been additionally tested with the extended stimulus set, data from this experiment were used. For each cell, the responses to the grasping and placing action were entered into a one-way ANOVA [action (n = 2) with trials as replicates], if there was a significant visual response to either of the actions (response > background firing rate, t test p < .05) and all the conditions contained more than five trials.

Responses to grasping and placing actions seen from the two different views were compared in a similar manner; if no action response latency could be found at the nonpreferred view, then the latency to action from the preferred view was used. For each cell, responses were entered into a two-way ANOVA [action (n = 2) by view (n = 2) with trials as replicates]. Responses to grasping and placing actions with and without the presence of the object were compared similarly; if no response latency could be calculated to the action without the object, the latency to the action with object was used. For each cell, responses were also entered into a two-way ANOVA [action (n = 2) by presence of object (n = 2) with trials as replicates].

Population Analysis

Cell responses to both grasping and placing actions seen from both views were combined within condition to create an average cell response to preferred and nonpreferred actions seen from the most effective and least effective views. In addition, cell responses to the preferred action with and without objects present, and the object presented without the action, were combined within condition to create an average cell response to the three different stimulus conditions. First, each contributing cell's SDFs to the conditions were normalized with respect to the peak response to the action and view that generated the largest response (or action with object). Second, each cell's SDFs to the preferred and nonpreferred actions seen from the preferred view (or preferred action with object) were shifted in time such that the cell's visual response latencies were aligned at 100 msec and the SDFs of the respective actions seen from the opposite view (or action without object and object without action) were shifted equivalent amounts.

RESULTS

Psychophysical Studies in Humans

One subject's responses to different test action movies (Figure 1) before and after adaptation are shown in Figure 2 to illustrate the hand adaptation aftereffect and our analysis. The kinematics of the action in the movie
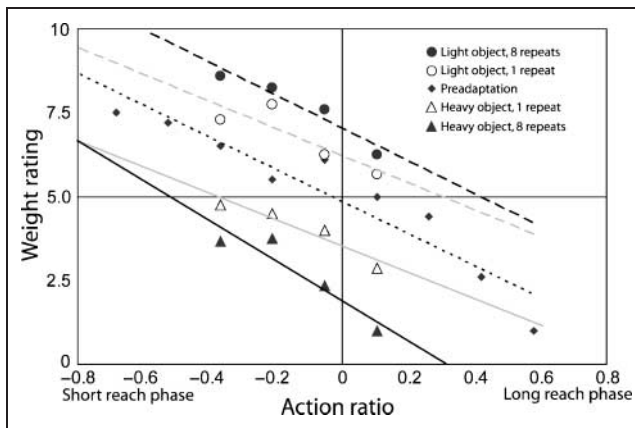
Figure 2. Single subject ratings of hand actions before and after adaptation. Mean ratings of the weight of a grasped object in each test movie, before and after adaptation, are plotted against each test movie's "action ratio." For the grasping action, the first phase was the reach toward the object and the second phase was withdraw. Positive action ratios: longer reach phase; negative action ratios: longer withdraw phase. Best fitting linear functions are plotted. Preadaptation test phase: black diamonds, dotted black line, $r = .961$; action ratio for weight rating of 5 is -0.031. After adapting to a movie of a hand grasping a light object repeated once: open circles, dashed gray line, $r = .864$; action ratio for weight rating of 5 is 0.301; and repeated eight times: black circles, dashed black line $r = .959$; action ratio for weight rating of 5 is 0.417. After adapting to a hand grasping a heavy object repeated once: open triangles, solid gray line, $r = .950$; action ratio for weight rating of 5 is -0.380; and repeated eight times: black triangles, solid black line, $r = .936$, action ratio for weight rating of 5 is -0.520.

have a significant effect on how heavy the weight appears [Figure 2, black diamonds; one-way ANOVA: $F(9, 90) = 34.48, p < .001$]; actions with a short reach phase and long withdraw phase appeared to grasp heavy weights, actions with a long reach phase and short withdraw phase appeared to grasp light objects. Prior adaptation to a hand grasping a light object resulted in the weights of the objects in the subsequently presented test action movies appearing relatively more heavy. The kinematics of the test actions still influence the perception of the weight of the objects, but the weight of the objects grasped during all actions tested appeared heavier. Adapting to a hand grasping a heavy object has the opposite effect: a general decrease in the perceived weight of the object across all test actions. When the adapting stimulus is repeated eight times (black circles and triangles), the perceived weight of the objects in the test actions appears more profoundly influenced than when the adapting stimulus is repeated once (open circles and triangles).

Influence of Action Repetition on the Action Adaptation Aftereffect

We compared the effect of repeating the adapting action (1, 2, 4, 8 times) and the adapting action object weight (light, heavy) on the aftereffect using ANOVA (while in-

verting the sign of the aftereffect induced by the adapting action grasping the heavy weight). Increasing the number of times the adaptation stimulus is repeated increases the adaptation aftereffect [ANOVA, main effect of adapting action repetition: $F(1.4, 19.6) = 4.75, p < .05$, Greenhouse–Geisser correction applied], which is illustrated clearly in Figure 3. Adapting to a hand grasping a light object appeared to generate a more pronounced aftereffect than adapting to a hand grasping a heavy object, although this was not significant [ANOVA, main effect of adapting action object weight: $F(1, 14) = 2.162, p = .128$]. There was no interaction between action repetition and adapting action object weight.

We tested the difference in the aftereffects induced after adapting to a hand grasping light and heavy objects when at different levels of adapting action repetition using Bonferonni-corrected t tests. With only one presentation of the adapting action, there was no significant differences in aftereffects [paired-sample t test: $t(14) = 1.32, p = .209$, two-tailed]. There was, however, a significant difference in aftereffects after presenting the adapting action more than once [2, 4, 8 times, paired-sample t tests: $t(14) > 3.44, p < .005$, two-tailed].

Duration of the Action Adaptation Aftereffect

We compared the effect of increasing the ISI and the adapting action object weight on the aftereffect using ANOVA (the sign of the aftereffect induced by adapting action grasping the heavy weight was inverted). The adaptation aftereffect was greatest with short ISIs [ANOVA, main effect of ISI: $F(1.2, 11.5) = 4.81, p < .05$, Greenhouse–Geisser correction applied, illustrated in
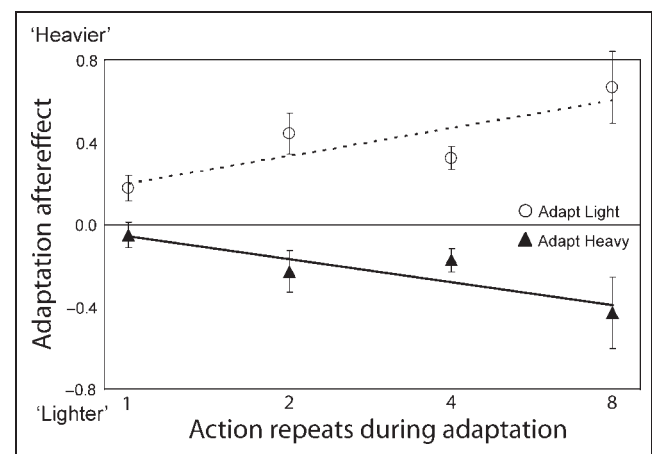

Figure 3. Effect of action repetition on the action aftereffect. Adaptation aftereffects for the influence of hands grasping light objects (circles, dashed line) or heavy objects (triangles, solid line) are plotted on a semi-log scale against the number of times the adapting action is repeated. Positive aftereffects indicate that subsequent test actions appear heavier; negative aftereffects indicate that subsequent test actions appear lighter. Logarithmic functions are fitted to the data shown: hands grasping light objects ($R^2 = .712$), hands grasping heavy objects ($R^2 = .782$). Error bars indicate *SEM*.

Figure 4]. There also appeared to be a larger difference in the aftereffect after adapting to a hand grasping a light object, but this was not significant [ANOVA, main effect of adapting action object weight: $F(1, 10) = 0.793, p = .39$]. There was no interaction between ISI and adapting action object weight.

Dependence of Hand Action Aftereffect on View and Adapting Action Goal

In a factorial design, we measured the effect of adapting to a hand grasping a heavy object on the perception of both grasping actions and placing actions (grasping action movies played in reverse) seen from the same and different views. As expected, after adapting to hands grasping heavy objects, subsequently viewed test grasping actions appeared to grasp lighter objects (see Figure 5). Importantly, the effect of adapting to a hand grasping a heavy object caused subsequently viewed placing actions to appear to place even heavier objects. We compared these effects using ANOVA [angle of separation between views of adapting and test actions (0, 45, 90) vs. type of test action (same as adapting action, opposite to adapting action)], while inverting the sign of the aftereffects measured with test actions opposite to the adapting action.

The adaptation aftereffects were greatest when test actions were seen from the same viewpoint as the adapting action [ANOVA, main effect difference in viewpoint: $F(2, 32) = 4.51, p < .05$]. There was no significant difference between the sizes of the aftereffects generated when viewing the same (grasping) or different (placing) test actions; there was no significant interaction between the aftereffects generated with different test
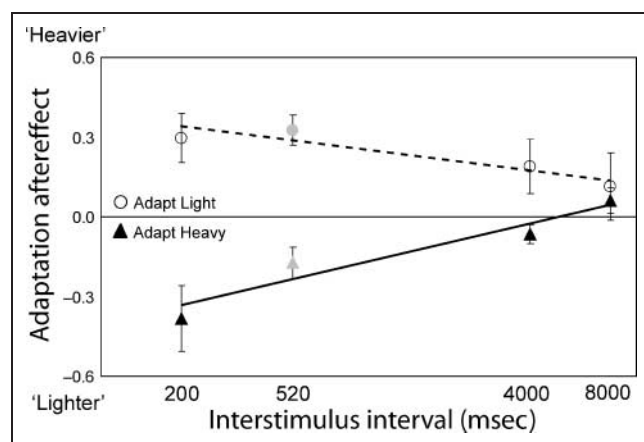

Figure 5. Grasping adapting action: viewpoint dependence. Adapting stimulus is hand grasping heavy object, aftereffects are plotted for test actions of hands grasping objects (same action type, triangles) and hands placing objects (different action type, circles) seen from three different viewpoints. Difference in viewpoint is the angular difference between adapting and test action perspectives. Linear functions are fitted; error bars indicate *SEM*.

actions and when they were observed from different views. We compared the difference in aftereffects induced in the same and different actions at different levels of viewpoint using Bonferonni-corrected *t* tests. When test actions were seen from 0° or 45° away from the adapting action, there was a significant difference in the aftereffect [paired-sample *t* tests: $t(16) > 4.71, p < .001$, two-tailed]. When test actions were seen from a viewpoint 90° away from the adapting action, the aftereffects were not significantly different [paired-sample *t* test: $t(16) = 1.42, p < .173$, two-tailed]. These results indicate that the perception of grasping actions relies primarily on neural mechanisms that are view dependent, but also to some extent on view-independent neural mechanisms. The influence of the adapting grasping action had a similar effect on the test placing actions.

We confirmed these results by testing the perception of the same test actions after adapting to a hand placing a heavy object (see Figure 6); results were analyzed as for above. Adapting to a hand placing a heavy object made subsequent placing actions appear to be placing lighter objects, subsequent grasp actions appeared to be grasping even heavier objects. There was a significant effect of varying the degree of separation between the view of the adapting placing action and the view of the test actions [ANOVA, main effect difference in viewpoint: $F(2, 22) = 10.16, p < .001$]. There was no significant difference between the sizes of the aftereffects in the same (placing) or different (grasping) test actions, nor was there a significant interaction between the aftereffects generated with different test actions and when they were observed from different views. We compared the difference in aftereffects induced in the same and different actions at different levels of viewpoint using

Figure 4. Effect of ISI on the action aftereffect. Conventions as for Figure 3 except: the aftereffect is plotted on a semi-log scale against the interstimulus interval (ISI, 200, 4000, 8000 msec). Each adapting stimulus was repeated four times. Data from Experiment 1 are included where adapting stimuli were repeated four times, with 520 msec ISI (gray symbols). Logarithmic functions are fitted to all four data points separately for both adapting stimuli: hands grasping light objects ($R^2 = .864$), hands grasping heavy objects ($R^2 = .924$).
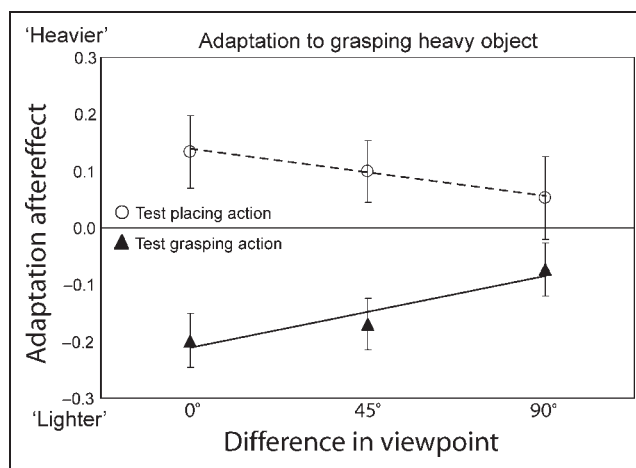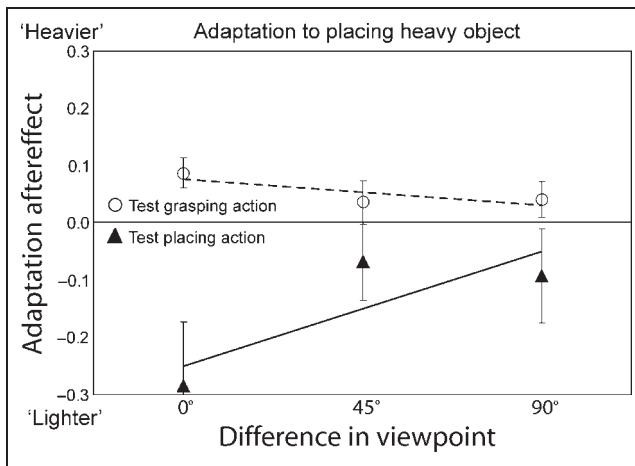
Figure 6. Place adapting action: viewpoint dependence. Conventions as for Figure 5 except: adapting stimulus is hand placing heavy object, aftereffects are plotted for test actions of hands placing objects (same action type, triangles) and hands grasping objects (different action type, circles) seen from three different viewpoints.

Bonferonni-corrected t tests. When test actions were seen from the same viewpoint as the adapting action, there was a significant difference in the aftereffect [paired-sample t tests: $t(11) = 3.53$, $p < .005$, two-tailed]. When test actions were seen from viewpoints $45°$ and $90°$ away from the adapting action, the aftereffects were not significantly different [paired-sample t tests: $t(11) < 1.37$, $p > .199$, two-tailed], indicating that the mechanisms

underlying the perception of the actions are largely view dependent.

Physiological Responses to Grasping and Placing Actions

We recorded the responses of single units and multiple units from 643 recording sites in both hemispheres of the temporal lobe (upper and lower banks of the STS) from one monkey. At 301/643 sites (47%), we found single or multiunit responses that were visually responsive. At 63/643 recording sites (10%), we found single or multiunits that were responsive to hand actions (grasping, placing, grooming, manipulating, tearing, etc.). At each recording site, we could isolate between one and five units; in total, we recorded 95 units that showed a significant response ($p < .05$) to a hand action. For 58/95 units (61%), the grasping and placing actions produced greater responses than other actions. Eighteen out of 58 units (31%) showed maximal responses to the grasping action, and 40/58 (69%) showed maximal responses to the placing action. For all 58 cells, we compared the mean responses to the grasping and placing actions; 20/58 cells (34%) showed significantly (ANOVA, $p < .05$) different responses to the two actions. Of the 18 cells that preferred grasping actions, 6 (33%) showed significantly ($p < .05$) greater responses to the grasping than the placing action. For these cells, the average response to the grasping action was 30.8 spikes/sec

Figure 7. Responses of single cells to grasping and placing actions. Responses to grasping actions (closed circles) and responses to placing actions (open circles) are plotted as spike density functions (SDFs, gray = *SEM*) for each cell. (A) Single cell that responds preferentially to a grasping action, when the hand is empty. (C) Single cell that responds preferentially to a grasping action, when the hand has reached out and touched the object. The individual frames that make up the movie of the grasping action are illustrated with left-hand edge of frame aligned at time of onset. (B) Single cell that responds preferentially to a placing action, when the hand is holding the object. (D) Single cell that responds preferentially to a placing action, after the hand has placed the object. Placing action frames are aligned as for grasping action.
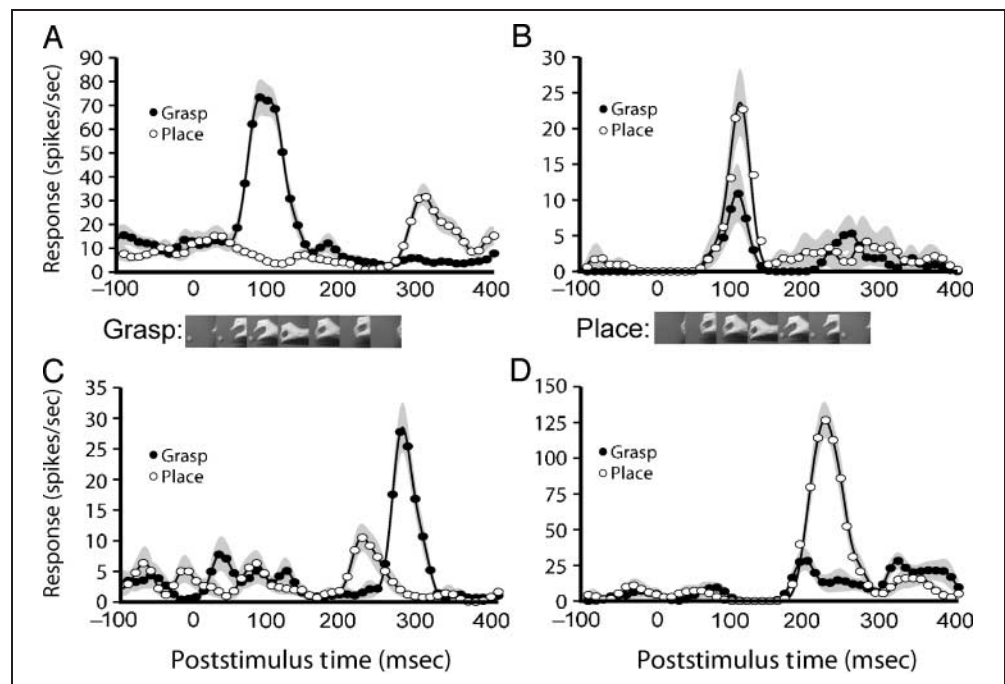
(*SEM* = ±10.5 spikes/sec), there was also a substantial average response to the placing action at 12.7 spikes/sec (*SEM* = ±5.7 spikes/sec), 41% of the size of the grasping action response. Of the 40 cells that preferred placing actions, 14 (35%) showed significantly (*p* < .05) greater responses to the placing than the grasping action. For these cells, the average response to the placing action was 51.6 spikes/sec (*SEM* = ±12.7 spikes/sec), and there was also a substantial average response to the grasping action at 31.1 spikes/sec (*SEM* = ±12.7 spikes/sec), 60% of the size of placing action response. For the 38/58 (66%) cells that showed no significant difference between the responses to the two actions, the average response to the grasping action was 32.0 spikes/sec (*SEM* = ±4.3 spikes/sec), and the average response to the placing action was 34.9 spikes/sec (*SEM* = ±4.7 spikes/sec). Figure 7A and C illustrates the responses of two single cells that responded preferentially to grasping actions and Figure 7B and D responses of two single cells that responded preferentially to placing actions. For all 58 cells that showed a significant response to either action, the distribution of the ratios between the action responses is plotted in Figure 8A.

Across the population of cells recorded, there were considerable differences in the latencies of the responses to the two actions. Some cells had early response latencies to grasping actions (e.g., Figure 7A) when the hand and object were not touching, and others had late responses (e.g., Figure 7C) when the hand and object were touching. Other cells had early response latencies to placing actions (e.g., Figure 7B) when the hand and object were touching, and others had late response latencies (e.g., Figure 7D) when the hand and object were not touching. For all cells that responded preferentially to grasping actions, the average grasping action response latency was 159 msec, and placing action response latency was 137 msec. For all cells that responded preferentially to the placing action, the average placing action response latency was 126 msec, and grasping action response latency was 128 msec. Although the latencies of the responses to both actions for cells that responded preferentially to grasping actions appear later than those for the cells that responded preferentially to placing actions, this was not significant [two-tailed independent-samples *t* test: grasping response latencies, $t(55) = 1.62$, $p = .11$; placing response latencies, $t(55) = 0.64$, $p = .53$]. Figure 8B illustrates the distribution of latencies of the responses to the action that produces the maximal response for all cells that were tested with stimuli that lasted for seven frames ($n = 42$).

Cell Sensitivity to View

We measured the responses of 23 cells that responded preferentially to either the grasping ($n = 8$) or placing ($n = 15$) actions to the grasping and placing movies and when the movies were flipped horizontally by 180°, thus
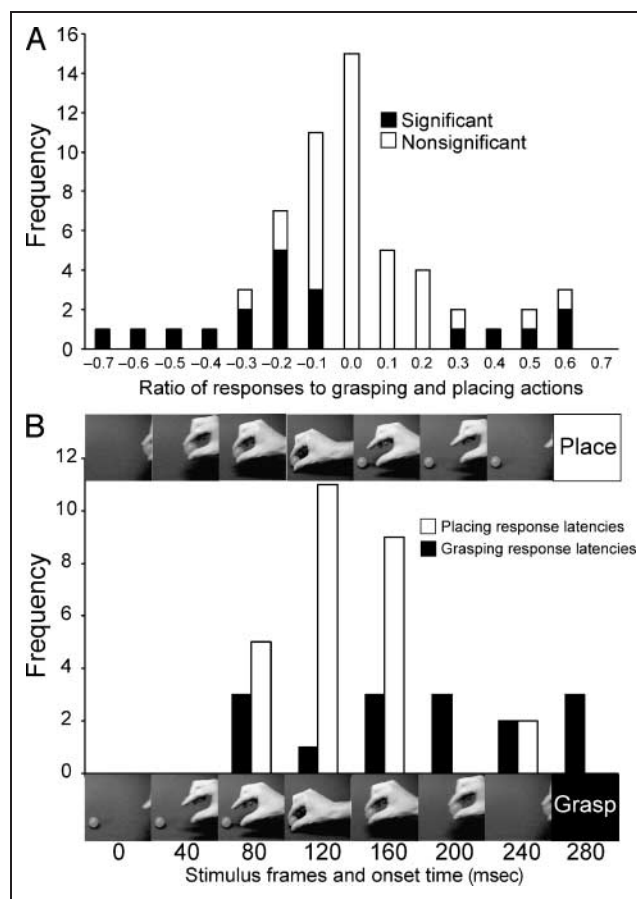

Figure 8. Distribution of cell response ratios and response latencies. (A) Distribution of the ratios ([response to grasp − response to place]/ [response to grasp + response to place]) for all cells tested. Dark bars illustrate ratios calculated from cells that showed significantly different responses to grasping ($n = 6$) and placing actions ($n = 14$); light bars: no significant difference in action responses ($n = 38$). (B) Distribution of the latencies of the responses to grasping actions for cells that responded preferentially to grasping actions (black bars) and latencies of responses to placing actions for cells that responded preferentially to placing actions (white bars), for all cells tested with stimuli lasting 7 frames.

viewed from a different perspective. Of the 23 cells tested, 16 (70%) showed a significant influence of the view from which the actions were seen (ANOVA, main effect view or interaction Action × View, $p < .05$).

Figure 9 shows the average cell responses (see Methods) to the four different actions calculated from the 23 cells. The view from which the actions were seen had a significant effect on the average responses [ANOVA, main effect of view: $F(1, 22) = 18.18$, $p < .0001$; interaction Action × View: $F(1, 22) = 4.59$, $p < .05$]. The average response to the preferred action seen from the preferred view was 34.4 spikes/sec; when the preferred action was seen from a view 180° away, this was significantly reduced to 21.2 spikes/sec [paired *t* test: $t(22) = 6.19$, $p < .0001$, two-tailed], a 38% reduction. The average response to the nonpreferred action was 25.8 spikes/sec; when seen from a view 180° away, this was significantly

reduced to 16.9 spikes/sec [paired t test: $t(22) = 2.70, p < .005$, two-tailed], a 34% reduction.

Cell Sensitivity to Presence of Object

In order to affirm whether the presence of the object was critical for cell responses to the actions, we tested the responses of 15 STS cells that preferred either grasping ($n = 4$) or placing ($n = 11$) actions to the actions performed with and without the presence of the object, and to the object without the action. Of the 15 cells tested, 10 (66%) showed significantly different responses to the different stimuli (ANOVA, $p < .05$).

Figure 10 shows the average responses of the 15 cells to the three tested stimuli; responses are aligned similarly to Figure 9. There was a significant difference in the average responses to the different stimuli [ANOVA: $F(1.07, 15.0) = 6.92, p < .05$, Greenhouse–Geisser correction applied]. Planned contrasts indicated that, on average, cells showed significantly [$F(1, 14) = 6.38, p < .05$] bigger responses to the preferred action performed with the object present 43.2 spikes/sec ($SEM = \pm11.2$ spikes/sec) than to the preferred action performed without the object 28.0 spikes/sec ($SEM = \pm7.2$ spikes/sec); and also significantly [$F(1, 14) = 7.17, p < .05$] bigger responses than when the object was present without the action 14.9 spikes/sec ($SEM = \pm3.8$ spikes/sec).

The anterior–posterior extent of the recorded cells was from 6 to 10 mm anterior of the interaural plane consistent with previous studies showing visual responses to hand actions in this region (Barraclough et al., 2005; Perrett et al., 1989). There appeared to be a largely similar distribution of cells showing all types of visual sensitivity over both the upper and lower banks of the STS. Most cells (44/58), however, were recorded from the lower bank, and these cells tended to show greater
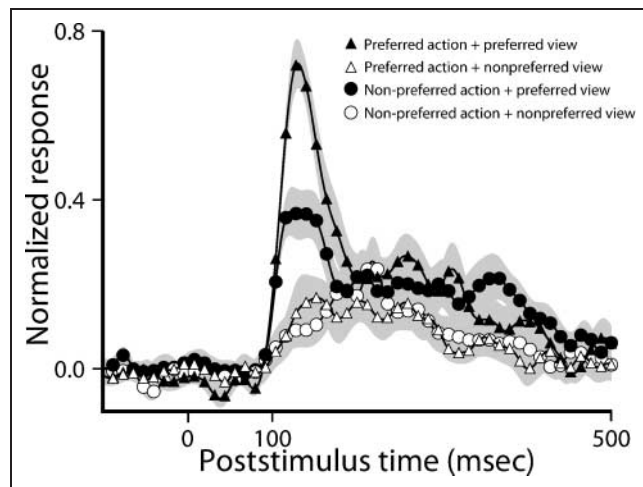

Figure 9. Responses of the average cell to the preferred and nonpreferred actions seen from preferred and nonpreferred views averaged across cells. SDFs (gray = *SEM*) from responses of 23 cells, of which 8 preferred grasping actions and 15 preferred placing actions.
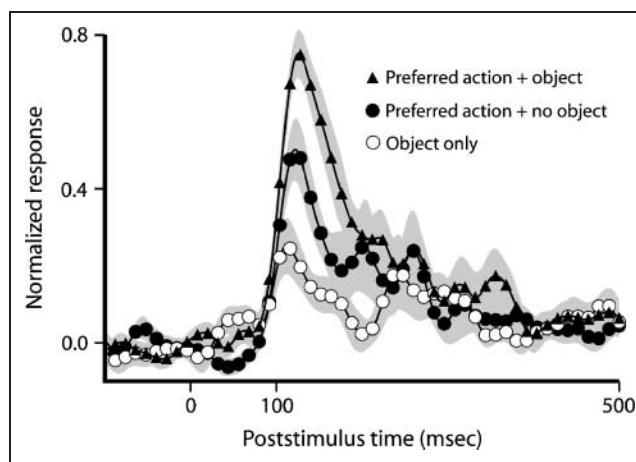

Figure 10. Responses of the average cell to the preferred action with and without the presence of the object, and when the object is presented alone averaged across cells. SDFs (gray = *SEM*) from responses of 15 cells, of which 4 preferred grasping actions and 11 preferred placing actions.

selectivity (38% of tested cells) for the preferred hand action than those cells recorded in the upper bank (21% of tested cells). Cells that were significantly selective for the view of the action were found in both banks of the STS and cells that showed significantly bigger responses to the action when the object was present were all found in the lower bank of the STS.

DISCUSSION

We have shown that judgments about the interaction between a human hand and an object are susceptible to visual adaptation. Prior observation of hand–object interactions influences our subsequent perception of other hand actions. Visual adaptation can occur after observing a hand action just once, seeing grasping of a light object biases subsequent grasped objects to appear heavier. Adaptation increases with repetition of the adapting stimulus and decreases with time. In our final two adaptation experiments, we found that as the difference in the viewpoint of the adapting action and the test action was increased, the effect of adaptation decreased. In addition, adapting to one hand action (grasping) influenced the perception of an opposite hand action (placing) and vice versa. STS cells that responded to the grasping or placing actions, often responded to the opposite action, were sensitive to the view of the action, and showed a reduced response in the absence of the object.

Hand Action Adaptation Dynamics

The adaptation aftereffects we see here using natural goal-directed hand actions have much in common with recent demonstrations of adaptation aftereffects with other visual stimuli. We see both a logarithmic increase

in adaptation with action repetition and a logarithmic decrease with ISI, inconsistent with a simple priming effect, but consistent with studies investigating the dynamics of tilt (Magnussen & Johnsen, 1986), motion (Hershenson, 1989), face identity (Leopold, Rhodes, Muller, & Jeffery, 2005), face configuration (Rhodes et al., 2007), and biological motion aftereffects (Troje et al., 2006). As the dynamics of the hand action aftereffect follows this classic time course, it suggests that the adapted mechanism is perceptual in nature and neither an artifact of subject behavior during the experimental task, nor perhaps due to other postperceptual mechanisms.

The hand lifting the light object appeared to generate bigger aftereffects than the hand lifting the heavy object as illustrated by the larger offset in the slopes plotted for the aftereffects of the hand lifting light objects (Figures 3 and 4). This larger offset with the hand grasping the light object can largely explain why only this action generates a significant aftereffect after one repetition of the adapting action. The slopes in Figures 3 and 4 for actions grasping both the light and heavy objects appear largely similar (albeit with different signs). This suggests that the effect of adapting to a hand grasping a heavy object is the same as adapting to a hand grasping a light object, although the adapting effect does not develop as quickly when observing a hand lifting a heavy object. During the period when the hand is holding the object, there is more motion energy generated by the action of the hand grasping the light object (it moves faster). This may explain this stimulus' greater effect on adaptation.

The hand action aftereffect appears to be very strong, occurring after adaptation to just two repetitions of a hand grasping. Over the duration of the experiment, the subjects are exposed to many presentations of the adapting stimulus, and the observed aftereffect could be due to an overall build up of adaptation in the visual system. When we varied the repetition of adapting stimuli, the approximate time interval between the adapting stimuli on separate trials was 4040 msec (assuming a conservative 500-msec reaction time of subjects to test stimuli). The effect of adaptation was nonsignificant after 4000 msec (see Figure 4) and so we believe that build up of adaptation would be unlikely.

Sensitivity to View

The strongest adaptation aftereffects occur typically when the adapting and test actions are seen from the same view. After adapting to a hand grasping a heavy object, the greatest effect is on grasping actions seen from the same view. There is still a significant effect on actions seen from 45° rotated away; this suggests that a mixture of view-dependent and view-independent neural mechanisms underlie the perception of the hand actions. A related effect was seen by Benton et al. (2006, 2007) when investigating facial identity and expression aftereffects, although they observed aftereffects when

adapting and when test faces were viewed from angles 90° apart. Our results suggest that grasping action perception is relatively more reliant on view-dependent mechanisms than facial identity or expression. Indeed, we found that STS cells in the monkey showed significantly bigger responses to grasping and placing actions when seen from one view. Jellema, Baker, Wicker, and Perrett (2000) also observed STS cells that responded to reaching actions that were sensitive to the direction of the reach. These cells, however, were "nontransitive" cells, showing equally sized responses whether the reach was directed toward an object or not. The cells we report here are "transitive"; they show significantly increased responses when the action is directed toward an object. Thus, this is the first demonstration of view-dependent cellular coding of transitive hand actions in the STS, which complements those cells that code transitive hand actions in a view-independent manner found by Perrett et al. (1989).

Action coding in monkey STS neurons typically shows a predominance of view-dependent mechanisms (Jellema & Perrett, 2006) and perhaps a similar proportion of action-sensitive cells exist in humans. Human neuroimaging studies reveal that the homologous brain region, the posterior STS (pSTS), and a network of other brain regions (including the inferior parietal cortex, ventral premotor and inferior frontal cortex) are involved in the perception of hand actions (Thompson, Hardee, Panayiotou, Crewther, & Puce, 2007; Grèzes, Frith, & Passingham, 2004; Wheaton, Thompson, Syngeniotis, Abbott, & Puce, 2004; Buccino et al., 2001; Rizzolatti et al., 1996). It is not yet clear, however, to what extent these brain regions process hand actions in a view-dependent or -independent manner.

A further possibility is that the balance between engagement of view-dependent and view-independent action adaptation we see here is influenced by the nature of the task itself. We asked subjects to make a judgment of the weight of the object being grasped (or placed) and indicated they could use any cues available on screen. Jellema and Perrett (2006) have argued that view-dependent mechanisms are best suited for interacting with objects under visual guidance and during visuomotor tasks (e.g., see Craighero, Belloa, Fadiga, & Rizzolatti, 2002). Conversely, view-independent mechanisms would be best suited to recognizing objects and scenes. By varying the task of the subject in future hand action adaptation experiments, it might be possible to increase the relative contribution of the view-independent mechanisms.

Cross-action Adaptation

Intriguingly, we see that human adaptation to one action (grasping or placing) influences the subsequent perception of the opposite hand action. When viewed from the same angle, the size of the aftereffect is considerable

irrespective of the adapting action. This suggests that the perception of one action relies on mechanisms common to both actions. Indeed, most STS cells we report here responded to both grasping and placing actions; the 34% of cells that showed significantly larger responses to one of the actions also showed substantial responses to the opposite action. Thus, it is likely that the perception of each action relies both on cells that respond preferentially to that action and cells that preferentially respond to the other action.

Possible Mechanisms Underlying Action Adaptation

Retinotopic Adaptation

First, there is a possibility that there is some influence of low-level retinotopic adaptation during the task. If eye movements during the adapting and test stimuli were identical, then the same retinotopic mechanisms will be affected by the adapting and test stimuli, leading to the aftereffects we observe. For example, as the visual angle between adapting action and test action is increased, the overlap in the low-level features between the two stimuli decreases. This reduction in retinotopic overlap might explain the reduction in observed aftereffects with angular separation. Although we cannot rule out the presence of these aftereffects, we believe they are unlikely to dominate as they cannot explain the influence of adapting to one action on the perception of the opposite action (see below). In addition, during measurement of adaptation to similarly "high-level" stimuli, different identity faces, Rhodes et al. (2007) found no change in aftereffect magnitude with and without controls for low-level retinotopic adaptation. Our adapting stimuli typically lasted several seconds and test actions lasted 1520 msec (or 1400 msec). Subjects freely viewed both the adapting and test actions and were able to make many eye movements during the period of stimulus presentation.

Low-level Adaptation

A second possibility is that the action adaptation we observe relies on a simple low-level adaptation to the speed of movement during each action. Adaptation to speed of movement, slow or fast, can make subsequent intermediate movement speeds appear respectively faster or slower (Hammett, Champion, Morland, & Thompson, 2005); often, the effect of adaptation is described as repelling subsequent responses (Clifford, 2002). The adapting stimulus of grasping a heavy object consists of a fast initial reach movement (fast in) followed by a slow withdraw (slow out). If the effect of adaptation is at a level where local movement vectors are coded (with little regard for the presence/absence of the object in the hand or for the goal of the action), then the output of

motion detectors signaling fast movement in would be suppressed and the output of motion detectors signaling slow movement out would also be suppressed. The effect on subsequent actions performed at intermediate speeds would be to make the movement in appear slower and the movement out appear faster. This would make grasping test actions appear to be grasping lighter objects (as indeed we observe), but also to make placing test actions appear to be placing lighter objects. In fact, we see that placing test actions appear to be placing even heavier objects. A similar logic can be used to understand the influence of adapting to a hand placing a heavy object on the subsequent perception of a hand grasping an object. Indeed, we also see that after adapting to a hand placing a heavy object, subsequent grasping actions appear to be grasping even heavier objects. Thus, irrespective of the adapting action in our experiments, we find that an explanation based upon speed adaptation alone cannot account for all the effects we observe.

Weight Judgment Adaptation

Adaptation of a "cognitive" weight judgment mechanism would result in the same effects as if adaptation of a low-level motion adaptation mechanism dominated. After adapting to grasping or placing a heavy object, all subsequent weights, irrespective of action, should be judged as being lighter. In reality, we see that adapting to one action, lifting a heavy weight, influences subsequent opposite actions to appear to be interacting with even heavier weights.

High-level Adaptation

Although we cannot rule out any of the adaptation effects just described, we believe the dominating effect is the adaptation of a high-level mechanism that codes the interaction between the hand and the object. Adapting to the phase of the action where the hand touches the object would influence the phase of the test action where the hand touches the object, whereas adapting to the phase of the action where the hand is not touching the object would influence the equivalent phase of the test action. For example, the adapting action grasping a heavy object consists of a *fast movement in with no object* followed by a *slow movement out with object*. The phases of the subsequent test actions without object would appear slower, and phases with object would appear faster. For test grasping actions, the *movement in without object* would appear slower and the *movement out with object* would appear faster, and thus, the action would appear to be grasping a lighter object; for test placing actions, the *movement in with object* would appear faster, the *movement out without object* would appear slower, and thus, the action would appear to be placing a heavier object. A similar logic can be used to

describe the influence of adapting to a hand placing a heavy object on the subsequent perception of a hand grasping an object. This explanation is consistent with the results we see in both Figures 5 and 6.

Supporting this argument, we see that STS cells that respond to grasping and placing actions take into account whether the hand is touching the object or not. Many STS cells respond to grasping and placing actions when the hand is touching the object. Other cells, however, respond to grasping and placing actions when the hand is not touching but is near the object. In either case, the presence of the object in the scene is important, and the cells are ''transitive'' as cell responses are significantly reduced if the action is mimed.

The adaptation aftereffects we observe here could be explained by a combination of the outputs from such STS cells that take into account the presence of the carried object. The cells that respond when the hand is touching the object (with object) might represent one population of cells. If speed and direction are also taken into account, then the responses of this cell population could represent a continuum of hand movements from fast out with object through slow out with object, slow in with object, and finally, to fast in with object. These movements are components of, respectively, a hand grasping a light object, a hand grasping a heavy object, a hand placing a light object, and a hand placing a heavy object. Here, adapting to a hand grasping a heavy object (slow out with object) would repel the cell population response along the continuum in both directions. The cell population response would be weighted toward both fast out with object movements (grasping light object) and fast in with object movements (placing heavy object). As the slow movement out with object during the adapting grasping action is last, this phase would have the most influence on subsequent actions; hence, changes in the responses of the population of cells responding when hand and object touch would dominate.

The cells that respond when the object is not touching the hand (no object) might represent a separate continuum from fast out no object through slow out no object, slow in no object, and finally, to fast in no object. These movements are components of, respectively, a hand placing a light object, a hand placing a heavy object, a hand grasping a light object, and a hand grasping a heavy object. When adapting to a hand placing a heavy object, the last, and thus, more influential component of the adapting action, would be the slow movement out without object; here the change in the responses of the population of cells responding when the hand and object do not touch would dominate. Thus, the effect would be to repel the cell population response along the continuum in both directions: The cell population response would be weighted toward both fast out no object movements (placing light object) and fast in no object movements (grasping heavy object). We might also expect that as the kinematics of the component of

the action without the presence of the object are less constrained, then there might be a reduced reliance on populations of cells responding when hand and object are not touching. Indeed, we see that the influence of adapting to the placing action on the opposite action is smaller than the effect of adapting to the grasping action on the opposite action.

Conclusion

In conclusion, we see strong adaptation aftereffects after observing hand actions. The dynamics of these effects are similar to those seen in previous studies using more simple stimuli and are likely to be due to adaptation of neurons with properties similar to those STS neurons we describe here in the monkey. These results indicate that our perception of human acts is strongly influenced by our immediate prior experience, and that adaptation techniques can be usefully employed to investigate the brain mechanisms underlying the perception of human action.

REFERENCES

Barraclough, N. E., Xiao, D.-K., Oram, M. W., & Perrett, D. I. (2005). Integration of visual and auditory information by STS neurons responsive to the sight of actions. *Journal of Cognitive Neuroscience, 17,* 377–391.

Baylis, G. C., & Rolls, E. T. (1987). Responses of neurons in the inferior temporal cortex in short term and serial recognition memory tasks. *Experimental Brain Research, 65,* 614–622.

Benton, C. P., Etchells, P. J., Porter, G., Clark, A. P., Penton-Voak, I. S., & Nikolov, S. G. (2007). Turning the other cheek: The viewpoint dependence of facial expression after-effects. *Proceedings of the Royal Society of London, Series B, Biological Sciences, 274,* 2131–2137.

Benton, C. P., Jennings, S. J., & Chatting, D. J. (2006). Viewpoint dependence in adaptation to facial identity. *Vision Research, 46,* 3313–3325.

Bruce, C. J., Desimone, R., & Gross, C. G. (1981). Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *Journal of Neurophysiology, 46,* 369–384.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., et al. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: An fMRI study. *European Journal of Neuroscience, 13,* 400–404.

Clifford, C. W. G. (2002). Perceptual adaptation: Motion parallels orientation. *Trends in Cognitive Sciences, 6,* 136–143.

Craighero, L., Belloa, A., Fadiga, L., & Rizzolatti, G. (2002). Hand action preparation influences the responses to hand pictures. *Neuropsychologia, 40,* 492–502.

Edwards, R., Xiao, D.-K., Keysers, C., Foldiak, P., & Perrett, D. I. (2003). Color sensitivity of cells responsive to complex stimuli in the temporal cortex. *Journal of Neurophysiology, 90,* 1245–1256.

Foldiak, P., Xiao, D.-K., Keysers, C., Edwards, R., & Perrett, D. I. (2003). Rapid serial visual presentation for the determination of neural selectivity in area STSa. *Progress in Brain Research, 144,* 107–116.

Grèzes, J., Frith, C. D., & Passingham, R. E. (2004). Inferring false beliefs from the actions of oneself and others: An fMRI study. *Neuroimage, 21,* 744–750.

Hamilton, A. F., Joyce, D. W., Flanagan, J. R., Frith, C. D., & Wolpert, D. M. (2005). Kinematic cues in perceptual weight judgement and their origins in box lifting. *Psychological Research, 71,* 13–21.

Hammett, S. T., Champion, R. A., Morland, A. B., & Thompson, P. G. (2005). A ratio model of perceived speed in the human visual system. *Proceedings of the Royal Society of London, Series B, Biological Sciences, 272,* 2351–2356.

Harries, M. H., & Perrett, D. I. (1991). Modular organisation of face processing in temporal cortex: Physiological evidence and possible anatomical correlates. *Journal of Cognitive Neuroscience, 3,* 9–24.

Hasselmo, M. E., Rolls, E. T., Baylis, G. C., & Nalwa, V. (1989). Object-centred encoding by face selective neurons in the cortex in the superior temporal sulcus of the monkey. *Experimental Brain Research, 75,* 417–429.

Hershenson, M. (1989). Duration, time constant, and decay of the linear motion aftereffect as a function of inspection duration. *Perception & Psychophysics, 45,* 251–257.

Jellema, T., Baker, C. I., Wicker, B., & Perrett, D. I. (2000). Neural representation for the perception of the intentionality of actions. *Brain and Cognition, 44,* 280–302.

Jellema, T., & Perrett, D. I. (2006). Neural representations of perceived bodily actions using a categorical frame of reference. *Neuropsychologia, 44,* 1535–1546.

Kohn, A. (2007). Visual adaptation: Physiology, mechanisms, and functional benefits. *Journal of Neurophysiology, 97,* 3155–3164.

Leopold, D. A., O'Toole, A. J., Vetter, T., & Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level aftereffects. *Nature Neuroscience, 4,* 89–94.

Leopold, D. A., Rhodes, G., Muller, K.-M., & Jeffery, L. (2005). The dynamics of visual adaptation to faces. *Proceedings of the Royal Society of London, Series B, Biological Sciences, 272,* 897–904.

Magnussen, S., & Johnsen, T. (1986). Temporal aspects of spatial adaptation. A study of the tilt aftereffect. *Vision Research, 26,* 661–672.

Meng, X., Mazzoni, P., & Qian, N. (2006). Cross-fixation transfer of motion aftereffects with expansion motion. *Vision Research, 46,* 3681–3689.

Oram, M. W., & Perrett, D. I. (1992). Time course of neural responses discriminating views of the face and head. *Journal of Neurophysiology, 68,* 70–84.

Oram, M. W., & Perrett, D. I. (1994). Responses of anterior superior temporal polysensory (STPa) neurons to "Biological Motion" stimuli. *Journal of Cognitive Neuroscience, 6,* 99–116.

Oram, M. W., & Perrett, D. I. (1996). Integration of form and motion in the anterior superior temporal polysensory area (STPa) of the macaque monkey. *Journal of Neurophysiology, 76,* 109–129.

Perrett, D. I., Harries, M. H., Bevan, R., Thomas, S., Benson, P. J., Mistlin, A. J., et al. (1989). Frameworks of analysis for the neural representation of animate objects and actions. *Journal of Experimental Biology, 146,* 87–113.

Perrett, D. I., Oram, M. W., & Ashbridge, E. (1998). Evidence accumulation in cell populations responsive to faces: An account of generalisation of recognition without mental transformations. *Cognition, 67,* 111–145.

Perrett, D. I., Oram, M. W., Harries, M. H., Bevan, R., Hietanen, J. K., Benson, P. J., et al. (1991). Viewer-centered and object-centered coding of heads in the macaque temporal cortex. *Experimental Brain Research, 86,* 159–173.

Perrett, D. I., Rolls, E. T., & Caan, W. (1982). Visual neurons sensitive to faces in the monkey temporal cortex. *Experimental Brain Research, 47,* 329–342.

Perrett, D. I., Smith, P. A. J., Mistlin, A. J., Chitty, A. J., Head, A. S., Potter, D. D., et al. (1985). Visual analysis of body movements by neurons in the temporal cortex of the macaque monkey: A preliminary report. *Behavioural Brain Research, 16,* 153–170.

Perrett, D. I., Smith, P. A. J., Potter, D. D., Mistlin, A. J., Head, A. S., Milner, A. D., et al. (1985). Visual cells in the temporal cortex sensitive to face view and gaze direction. *Proceedings of the Royal Society of London, Series B, Biological Sciences, 223,* 293–317.

Rhodes, G., Jeffery, L., Clifford, C. W. G., & Leopold, D. A. (2007). The timecourse of higher-level face aftereffects. *Vision Research, 47,* 2291–2296.

Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu, E., Perani, D., et al. (1996). Localization of grasp representations in humans by PET: 1. Observation versus execution. *Experimental Brain Research, 111,* 246–252.

Rolls, E. T., & Baylis, G. C. (1986). Size and contrast have only small effects on the responses to faces of neurons in the cortex of the superior temporal sulcus of the monkey. *Experimental Brain Research, 65,* 38–48.

Saito, H., Yukie, M., Tanaka, K., Hikosaka, K., Fukada, Y., & Iwai, E. (1986). Integration of direction signals of image motion in the superior temporal sulcus of the macaque monkey. *Journal of Neuroscience, 6,* 145–157.

Saul, A. B., & Cynader, M. S. (1989). Adaptation in single units in visual cortex: The tuning of aftereffects in the spatial domain. *Visual Neuroscience, 2,* 593–607.

Thompson, J. C., Hardee, J. E., Panayiotou, A., Crewther, D., & Puce, A. (2007). Common and distinct brain activation to viewing dynamic sequences of face and hand movements. *Neuroimage, 37,* 966–973.

Tovee, M. J., Rolls, E. T., & Azzopardi, P. (1994). Translation invariance in the responses to faces of single neurons in temporal visual cortical areas of the alert macaque. *Journal of Neurophysiology, 72,* 1049–1060.

Troje, N. F., Sadr, J., Geyer, H., & Nakayama, K. (2006). Adaptation aftereffects in the perception of gender from biological motion. *Journal of Vision, 6,* 850–857.

van Wezel, R. J. A., & Britten, K. H. (2002). Motion adaptation in area MT. *Journal of Neurophysiology, 88,* 3469–3476.

Wheaton, K. J., Thompson, J. C., Syngeniotis, A., Abbott, D. F., & Puce, A. (2004). Viewing the motion of human body parts activates different regions of premotor, temporal, and parietal cortex. *Neuroimage, 22,* 277–288.

Zhao, L., & Chubb, C. (2001). The size-tuning of the face-distortion after-effect. *Vision Research, 41,* 2979–2994.