BOTTOM-UP, NOT TOP-DOWN, MODULATION OF IMITATION
BY HUMAN AND ROBOTIC MODELS

Clare Press*, Helge Gillmeister & Cecilia Heyes
*Corresponding author
Department of Psychology
University College London
Gower Street, London. WC1H OAP
United Kingdom

Phone: 44 (0)20 7679 5379
Fax: 44 (0)20 7436 4276
Email: c.press@ucl.ac.uk
URL: http://www.psychol.ucl.ac.uk/celia.heyes/netintro.htm

Number of text pages: 18
Number of figures and tables: 4
Word count overall: 3833 (excl. refs)
Word count abstract (<250): 242
Word count introduction (<500): 500

Running title (21 characters): IMITATION AND ANIMACY

Keywords: Action observation, visuomotor priming, animacy, mirror system.
Visual observation of human actions provokes more motor activation than observation of robotic actions. We investigated the extent to which this visuomotor priming effect is mediated by bottom-up or top-down processing. The bottom-up hypothesis suggests that robotic movements are less effective in activating the ‘mirror system’ via pathways from visual areas via the superior temporal sulcus to parietal and premotor cortices. The top-down hypothesis postulates that beliefs about the animacy of a movement stimulus modulate mirror system activity via descending pathways from areas such as the temporal pole and prefrontal cortex. In an automatic imitation task, subjects performed a prespecified movement (e.g. hand opening) on presentation of a human or robotic hand making a compatible (opening) or incompatible (closing) movement. Speed of responding on compatible trials, compared with incompatible trials, indexed visuomotor priming. In the first experiment, robotic stimuli were constructed by adding a metal and wire ‘wrist’ to a human hand. Questionnaire data indicated that subjects believed these movements to be less animate than those of the human stimuli, but the visuomotor priming effects of the human and robotic stimuli did not differ. In the second experiment, when the robotic stimuli were more angular and symmetrical than the human stimuli, human movements elicited more visuomotor priming than the robotic movements. However, subjects’ beliefs about the animacy of the stimuli did not affect their performance. These results suggest that bottom-up processing is primarily responsible for the visuomotor priming advantage of human stimuli.
INTRODUCTION

Single-cell recording has identified neurons in the premotor and parietal cortices of the monkey that discharge, not only when the animal performs an action, but also when it observes the same action being performed by another agent (e.g. Fogassi et al. 2005; Rizzolatti et al. 1996; Ferrari et al. 2003). Studies using functional magnetic resonance imaging (e.g. Gangitano et al. 2004; Buccino et al. 2001), transcranial magnetic stimulation (Gangitano et al. 2004), and behavioural measures (e.g. Heyes et al. 2005; Press et al. 2005), indicate that an homologous ‘mirror system’ is present in humans.

In addition to mediating visuomotor priming - the generation of motor activation by visual stimuli - the mirror system has been implicated in a variety of higher sociocognitive functions such as action understanding (Iacoboni et al. 2005), empathy (Gallese 2003; Carr et al. 2003) and theory of mind (Gallese & Goldman 1998). These functions may depend, not only on the transmission of information from the mirror system to more anterior regions, but also on top-down modulation of the mirror system by areas that specialise in inferential processing of social stimuli (Nishitani et al. 2004; Oberman et al. 2005). For example, it has been suggested that the temporal pole and/or medial prefrontal cortex enhance attention to social stimuli by modulating activity in the superior temporal sulcus (STS) (Castielli et al. 2002; Allison et al. 2000). This could affect activity in the mirror system through the connections between STS and premotor cortex via the rostral part of the inferior parietal lobule (Rizzolatti 2005).
Several recent studies have reported that observation of human movements elicits more mirror system activity, and more automatic imitation, than observation of robotic movements (e.g. Tai et al. 2004; Press et al. 2005). If the mirror system participates in higher sociocognitive functions, this visuomotor priming effect could be due to bottom-up and/or top-down processing. The mirror system may be better tuned to visual input from human movements (bottom-up), and/or its activity may be enhanced by knowledge that the movement stimulus is human (top-down).

To investigate the bottom-up and top-down contributions to visuomotor priming we assessed the impact of stimulus variables and of beliefs on automatic imitation of human and robotic hand movements (e.g. Heyes et al. 2005). In Experiment 1, the moving parts of human and robotic stimuli were identical, while subjects’ beliefs about their identity (human vs robotic) varied as a function of information supplied by the experimenter. If beliefs about animacy exert a top-down influence on functioning of the mirror system, one would expect more effective priming by the human than by the robotic stimuli. In Experiment 2, one group of subjects observed human movements but, half the time, believed them to be of robotic origin. The other group observed more angular and symmetrical movements of a robotic hand but, half the time, believed them to be of human origin. In this factorial design, an effect of stimulus type would indicate bottom-up processing, and an effect of belief would indicate top-down processing.
EXPERIMENT 1

In an automatic imitation task, Experiment 1 required subjects to perform a prespecified response (e.g. hand opening) on presentation of a stimulus hand making a compatible (opening) or incompatible (closing) movement. Speed of responding on compatible trials, compared with incompatible trials, indexed automatic imitation. All subjects participated in two sessions of testing; in one session they were informed that the stimuli were human, and in the other that they were robotic. In reality, the moving parts of all stimuli were human. If beliefs about animacy exert a top-down influence on visuomotor priming, stimuli believed to be human should evoke greater automatic imitation than stimuli believed to be robotic.

Materials and methods

Twelve consenting healthy subjects with an average age of 23 years, 4 male, took part in the study with local ethical committee approval. They were naïve with respect to the purpose of the experiment.

The subject’s right forearm lay in a horizontal position across his/her body, parallel with the stimulus monitor. It was supported from elbow to wrist by an armrest, and therefore the subject’s hand was free to move. In each block of the simple reaction time (RT) task, subjects were required to make a prespecified response (to open or to close their right hand) as soon as the stimulus hand started to move (either open or close). After making each response, subjects were required to return their hand to a neutral starting position. They were instructed to refrain from
moving their hand in catch trials, when the stimulus hand did not move. The stimulus was a blue life-sized silhouette of a human right hand on a black background, with either a naturalistic, flesh-coloured wrist (stimulus which subjects were informed was human) or a wrist made from metal and wires (stimulus which subjects were informed was robotic). The two stimulus formats (‘human’ and ‘robotic’) are shown in Figure 1. All subjects viewed ‘human’ and ‘robotic’ stimuli in separate sessions of testing, and were informed of the stimulus type that they would view at the beginning of each session.

INSERT FIGURE 1 ABOUT HERE

All trials began with presentation of the warning stimulus (fingers closed and pointing upwards in parallel with the thumb). In stimulus trials, this was replaced 800, 1600, or 2400ms later by onset of the opening or closing movement, which was of 480ms duration. After the imperative stimulus, the screen went black for 3000ms before the warning stimulus for the next trial appeared. In catch trials, the warning stimulus remained on the screen for 2880ms before the 3000ms inter-trial interval. In each session, 240 stimulus trials and 48 catch trials were delivered. These trials were presented within 4 separate blocks; in half of these blocks subjects were instructed to open their hand whenever the stimulus hand moved and in the other half they were instructed to close their hand. This resulted in 20 trials of each type, defined by combination of the response movement (opening and closing), stimulus movement (opening and closing) and stimulus onset asynchrony (800, 1600, 2400ms) variables.
For both open and close responses, response onset was measured by recording the electromyogram (EMG) from the first dorsal interosseus (FDI) muscle using disposable Ag/AgCl surface electrodes (Unomedical, Redditch, Worcestershire, UK). Signals were amplified, high-pass filtered at 20Hz, mains-hum filtered at 50Hz and digitised at 2.5kHz (1902 amplifier, Cambridge Electronic Design, Cambridge, Cambridgeshire, UK). They were rectified and smoothed using a dual-pass Butterworth filter, with a cut-off frequency of 50Hz. Signals were not low-pass filtered. To define a baseline, EMG activity was registered for 100ms when the subject was not moving at the beginning of each trial. A window of 20ms was then shifted progressively over the raw data in 1ms steps. Response onset was defined by the beginning of the first 20ms window after the onset of the imperative stimulus in which the standard deviation for that window, and for the following 20ms epoch, was greater than 2.75 times the standard deviation of the baseline. This criterion was chosen during initial calibration of the equipment as the most effective in discriminating false positives from misses. Whether the criterion correctly defined movement onset in the present experiment was verified by sight for every trial performed by each subject. Stimulus onset marked the beginning, and EMG onset marked the end, of the response time (RT) interval. Errors were recorded manually.

Following completion of the simple reaction time task in each session, subjects were given a 14-item questionnaire assessing their beliefs about the animacy of the stimulus observed during testing. The first nine questions assessed beliefs about animacy directly, for example, ‘Did the movement seem to be active or passive?’ Questions 10-14 assessed beliefs about animacy indirectly by probing subjects’ feelings about interaction with the stimulus, for example, ‘How would you
feel about the hand assisting in surgery on your body?’ Responses were measured by movement of a scroll-bar for each item. The lower end of the bar (scored as 0) indicated minimal animacy and the higher end (scored as 50) indicated maximal animacy.

Results and discussion

ANOVA applied to the questionnaire data indicated that animacy ratings (an average of questions 1-14) were higher for stimuli which subjects had been told were human (M = 28.3/50, SEM = 1.9) than for stimuli which they had been told were robotic (M = 23.5/50, SEM = 1.5), F(1,11)= 12.1, p<0.01. This suggests that the instructions were effective in influencing subjects’ beliefs about the objects depicted in the stimulus images. The effect of instruction on animacy ratings was small, but highly reliable. It is possible that the effect was small because subjects were reluctant to use the extremes of the rating scale.

Participants initiated movement in 10.1% of catch trials. These data were not analysed further. Practice trials and all RTs smaller than 100ms and greater than 1000ms (0.24%) were excluded from the analysis. There were no incorrect responses or response omissions. The reaction time (RT) data are shown in Figure 2. On each trial, the stimulus movement was either the same as (compatible) or different from (incompatible) the prespecified response. On average, responses were initiated 20ms faster when the stimulus was response compatible (M = 361.8ms, SEM = 11.1ms) than when it was response incompatible (M = 382.5ms, SEM = 13.4ms). This observation was confirmed by ANOVA in which the within-Ss variables were
stimulus-response compatibility (compatible and incompatible) and belief about stimulus type (human and robotic), F(1, 11) = 32.0, p<0.001. This RT difference was not greater for stimuli believed to be human than for stimuli believed to be robotic. In fact, the RT difference was numerically, but not statistically, greater for stimuli believed to be robotic than for stimuli believed to be human, F(1,11) = 1.8, p = 0.2. No other effects or interactions were significant.

Thus, Experiment 1 did not detect any top-down effect on visuomotor priming of beliefs about stimulus animacy.

EXPERIMENT 2

The results of Experiment 1 suggested that visuomotor priming is not modulated by beliefs about animacy. To provide a further test of this hypothesis, and to investigate any bottom-up contributions to visuomotor priming, in Experiment 2 we manipulated genuine stimulus animacy in addition to beliefs about stimulus animacy. One group of subjects always observed human movements and in one session they were informed that the movements were human and in the other session that they were robotic. The other group observed more angular and symmetrical movements of a robotic hand and in one session they were informed that the movements were generated by human movement and in the other session that they were generated by robotic movement. In this factorial design, an effect of genuine
stimulus type would indicate bottom-up processing, and an effect of belief about stimulus type would indicate top-down processing.

Materials and methods

Twenty-four new consenting healthy subjects with an average age of 25 years, 7 male, took part in the study with local ethical committee approval. They were naïve with respect to the purpose of the experiment. Subjects were randomly assigned, in equal numbers to groups that viewed a genuinely human stimulus, or a genuinely robotic stimulus, across both days of testing.

The two stimulus formats (human and robotic) are shown in Figure 3. To create a factorial design in which genuine stimulus animacy and beliefs about stimulus animacy varied orthogonally, the genuinely human stimuli were those from Experiment 1 but were now identical irrespective of whether subjects were told that they were human or robotic, that is, they did not have discriminating wrists. The genuinely robotic stimuli were opening and closing movements made by a blue silhouette of an angular, symmetrical robotic hand. Each frame of the opening and closing robotic movements approximately matched each frame of the opening and closing human movements in surface area, luminance, horizontal and vertical visual angles, and aperture between closest effectors. Subjects presented with the genuinely human stimulus were told that the hand was either human or robotic in the two sessions of testing. Subjects presented with the genuinely robotic stimulus were told that the movement was generated by either human or robotic movement. In the latter case, they were told that the stimuli had been generated by mapping important
features of human or robotic motion onto an inanimate image “in a similar way to that in which they make characters move in films like Shrek©”. All other aspects of the method were the same as Experiment 1.

**Results and discussion**

ANOVA applied to the questionnaire data indicated that animacy ratings (an average of questions 1-14) were higher for stimuli which subjects had been told were human (M = 24.1/50, SEM = 1.0) than for stimuli which they had been told were robotic (M = 21.2/50, SEM = 0.9), F(1,22) = 17.0, p<0.001. This confirms that, as in Experiment 1, our instructions were effective in influencing subjects’ beliefs about stimulus animacy. Indeed, the ratings for genuinely human stimuli and genuinely robotic stimuli did not differ, F(1,22) = 2.3, p = 0.14, suggesting that the instructions about stimulus origin were so effective that they blocked any effects of stimulus properties on beliefs.

Participants initiated movement in 4.8% of catch trials. These data were not analysed further. Practice trials, incorrect responses (0.11%), response omissions (0.05%) and all RTs smaller than 100ms and greater than 1000ms (0.03%) were excluded from the analysis. The RT data are shown in Figure 4. RTs were 10ms faster when the stimulus movement was response compatible (M = 283.5ms, SEM = 8.8ms) than when it was response incompatible (M = 293.1ms, SEM = 9.8ms). This difference between compatible and incompatible trials was larger when subjects were
responding to genuinely human stimuli (14.2ms) than when they were responding to
genuinely robotic stimuli (5.2ms), but it was not influenced by whether subjects
believed the stimulus they were observing was human (10.0ms) or robotic (9.3ms).

INSERT FIGURE 4 ABOUT HERE

These observations were confirmed by ANOVA in which the within-Ss
variables were stimulus-response compatibility (compatible and incompatible) and
belief about stimulus type (human and robotic), and the between-Ss variable was
genuine stimulus type (human and robotic). There was a main effect of compatibility,
\[ F(1, 22) = 20.7, p < 0.001, \] and a genuine stimulus type x compatibility interaction,
\[ F(1, 22) = 4.5, p < 0.05. \] Simple effects analyses indicated that the difference in RT
between compatible and incompatible trials was significant, not only for genuinely
human stimuli, \[ F(1,11) = 13.7, p<0.005, \] but also for genuinely robotic stimuli,
\[ F(1,11) = 8.0, p < 0.02. \] The belief x compatibility interaction was not significant
\( (F<1). \)

The results of Experiment 2 indicate that stimulus variables can have a
bottom-up effect on visuomotor priming. However, like those of Experiment 1, they
provide no evidence in support of the view that beliefs about stimulus animacy play a
significant role in modulating automatic imitation of hand movements.
GENERAL DISCUSSION

To investigate bottom-up and top-down contributions to visuomotor priming we assessed the impact of stimulus variables and of beliefs about animacy on automatic imitation of human and robotic hand movements. In Experiment 1, the moving parts of the stimuli were human, but subjects’ beliefs about their identity (human vs robotic) varied as a function of information supplied by the experimenter. Experiment 1 found no evidence that beliefs about animacy can exert a top-down influence on visuomotor priming; the human movements were no more effective as visuomotor primes when subjects believed that they were human than when they believed that they were robotic. Experiment 2 indicated that, rather than beliefs about stimulus origin, properties of the movement stimuli influence visuomotor priming potential. In a factorial design, subjects were presented with stimuli that were genuinely human or genuinely robotic, and were informed in one session that the stimuli were human and in another session that the stimuli were robotic. This experiment demonstrated that automatic imitation of genuinely human stimuli was greater than automatic imitation of genuinely robotic stimuli, and that this difference was not modulated by beliefs about the human or robotic origin of the stimuli.

Several previous behavioural studies have indicated that observation of human actions elicits more automatic imitation than observation of robotic actions (Press et al. 2005; Kilner et al. 2003; Oztop et al. 2004). Neurological studies have also suggested a human visuomotor priming advantage, by showing that observation of human actions gives rise to more activity in premotor cortex than observation of robotic actions (Gallese et al. 1996; Tai et al. 2004). For example, Tai et al. (2004)
required subjects to observe a human hand and a robotic hand performing grasping actions, and found that observing human actions activated premotor cortex to a greater degree than observing a static image of a human hand, but observing robotic actions did not activate premotor cortex more than observing a static image of a robotic hand. However, as far as we are aware, no previous study has attempted to isolate and distinguish top-down and bottom-up contributions to the human bias in visuomotor priming.

The finding that visuomotor priming depends on stimulus properties rather than beliefs about stimulus origin suggests that robotic movements are less effective in activating the ‘mirror system’ through pathways from visual areas via the superior temporal sulcus to parietal and premotor cortices. This is consistent with the suggestion that the mirror system influences higher sociocognitive functions via ascending connections with systems involved in drawing mental state inferences (Kilner et al. 2003; Gallese & Goldman 1998). It is generally assumed that humans and other animals have mental states whereas robotic systems do not. Therefore, it would be adaptive for inputs from biological stimuli to have privileged access to processes that generate inferences about mental states.

The present results are consistent with the hypothesis that mirror systems provide input to higher-level sociocognitive functions, but not with the converse hypothesis, that higher-level functions modulate processing in the mirror system. We found no evidence of top-down modulation by beliefs about stimulus animacy. In our experiments, the movement stimuli were in full view, presentation of the human and robotic stimuli was blocked, and subjects were not explicitly encouraged to think
about the meaning of the observed actions or the intentions of the actor. These conditions resemble those in which a human visuomotor priming advantage has been demonstrated previously (Press et al. 2005; Tai et al. 2004; Gallese et al. 1996; Kilner et al. 2003; Oztop et al. 2004), and are not unusual in everyday life. However, our results do not exclude the possibility that top-down modulation occurs under contrasting conditions (Ramnani & Miall 2004).

We used a questionnaire to assess subjects’ beliefs about stimulus animacy, and questionnaire responses are susceptible to demand effects. In principle, subjects may have responded in accordance, not with their beliefs about the objects depicted in the stimulus images, but with inferences about the experimenter’s expectations. However, it is unlikely that our questionnaire measure was contaminated by demand effects for two reasons. First, the experimenter conspicuously did not observe subjects when they were completing the questionnaire, and they were assured that all data would be stored anonymously. Second, questionnaire responses were made using a scroll-bar that was not marked with numbers. Therefore, it would have been difficult for subjects to remember when completing the questionnaire after their second session (e.g. robot instructions) how they had responded after their first (e.g. human instructions), at least 24 hours earlier.

The results of Experiments 1 and 2 therefore suggest that the mirror system is better tuned to movement stimuli with human perceptual properties than to movement stimuli with robotic perceptual properties. This tuning may have occurred on an evolutionary or a developmental timescale. It has been proposed that the mirror system evolved to support higher sociocognitive functions (Kilner et al. 2003; Gallese
& Goldman 1998). This hypothesis implies that the tuning was done by natural selection, and therefore that the human advantage in visuomotor priming is an adaptation. Alternatively, the associative sequence learning model (Heyes et al. 2005; Heyes & Ray 2000; Heyes 2001) suggests that the mirror system is tuned, and the potential for visuomotor priming established, through experience in which the individual contiguously observes and executes the same actions. Our environment provides us with many opportunities to form associative cortical links between visual representations of human actions and congruent motor representations. For example, visuomotor priming of hand actions may depend on cortical links established during visual observation of one’s own hand while performing actions. In comparison, there would be relatively little opportunity to form similar associations between motor and visual representations of robotic actions. This hypothesis that self-observation can result in the formation of associations between visual and motor representations of action has also been advanced in ideomotor theories of action (e.g. Hommel et al. 2001; James 1890; Lotze 1852).

In conclusion, the results of the present study suggest that observation of human actions provokes more visuomotor priming than observation of robotic actions, and that this bias is not modulated by beliefs about whether the observed stimuli are human or robotic. Instead, the present findings indicate that the greater potency of human stimuli in eliciting visuomotor priming depends on visual properties of the stimuli; that it is predominantly a bottom-up, rather than a top-down, effect.
Acknowledgements

This research was supported by the Economic and Social Research Council (ESRC) research centre for Economic Learning and Social Evolution and by a PhD studentship awarded to CP by the Biotechnology and Biological Sciences Research Council (BBSRC). We are grateful to Geoffrey Bird, Rüdiger Flach and Patrick Haggard for their advice and assistance.
References


Lotze R.H. (1852) *Medizinische Psychologie oder die Physiologie der Seele [Medical psychology or physiology of the soul]*. Weidmann'sche Buchhandlung, Leipzig, Germany.


Figure legends

Figure 1: Stimuli in Experiment 1: A ‘human’ hand, B ‘robotic’ hand. The moving parts of both stimuli were genuinely human, but the ‘human’ hand had a naturalistic, flesh-coloured wrist and the ‘robotic’ hand had a wrist made from metal and wires. Within each panel, the image on the left is the last frame of the hand opening video and the image on the right is the last frame of the hand closing video.

Figure 2: Experiment 1. Mean RT on compatible and incompatible trials for stimuli believed to be human and stimuli believed to be robotic. Vertical bars indicate the standard error of the mean.

Figure 3: Experimental stimuli in Experiment 2: A genuinely human hand, B genuinely robotic hand. Within each panel, the image on the left is the last frame of the hand opening video and the image on the right is the last frame of the hand closing video.

Figure 4: Experiment 2. Mean RT on compatible and incompatible trials for genuinely human (left) and robotic (right) stimuli, both when believed to be human and when believed to be robotic. Vertical bars indicate the standard error of the mean.