Sensorimotor imitation experience enhances automatic imitation of robotic action

Clare Press*, Helge Gillmeister and Cecilia Heyes

Department of Psychology, University College London, 26 Bedford Way, London WC1H 0AP, UK

Recent research in cognitive neuroscience has found that observation of human actions activates the ‘mirror system’ and provokes automatic imitation to a greater extent than observation of non-biological movements. The present study investigated whether this human bias depends primarily on phylogenetic or ontogenetic factors by examining the effects of sensorimotor experience on automatic imitation of non-biological robotic stimuli. Automatic imitation of human and robotic action stimuli was assessed before and after training. During these test sessions, participants were required to execute a pre-specified response (e.g. to open their hand) while observing a human or robotic hand making a compatible (opening) or incompatible (closing) movement. During training, participants executed opening and closing hand actions while observing compatible (group CT) or incompatible movements (group IT) of a robotic hand. Compatible, but not incompatible, training increased automatic imitation of robotic stimuli (speed of responding on compatible trials, compared with incompatible trials) and abolished the human bias observed at pre-test. These findings suggest that the development of the mirror system depends on sensorimotor experience, and that, in our species, it is biased in favour of human action stimuli because these are more abundant than non-biological action stimuli in typical developmental environments.

Keywords: mirror system; automatic imitation; ASL; action observation; animacy; expertise

1. INTRODUCTION

The ‘mirror system’ consists of a network of areas in human ventral premotor and parietal cortices which is active, not only when actions are executed but also when the same actions are passively observed (e.g. Buccino et al. 2001; Gangitano et al. 2004). Behavioural and neuroimaging studies have found that this system shows a human bias; it is activated more by observation of human action than by observation of physically similar non-biological movement (Tai et al. 2004). Behavioural studies have used the inadvertent tendency to copy observed body movements as an index of mirror system functioning. For example, in a simple reaction time (RT) study, Brass et al. (2001) found that index finger movements (e.g. lifting) were executed faster in response to observed compatible movements (lifting) than in response to observed incompatible movements (tapping). This compatibility effect on RT is called ‘automatic imitation’ (e.g. Heyes et al. 2005) because it reflects facilitation of matching responses relative to non-matching responses, and this bias is not intended by the participant. A recent study of this kind has found that human models elicit substantially more automatic imitation than robotic models (Press et al. 2005).

Little is known about the origins of the mirror system. One hypothesis suggests that the mirror system’s capacity to match observed with executed actions is a product of phylogenetic evolution, and that it is an adaptation with respect to higher sociocognitive functions, such as understanding the mental states of others (e.g. Gallese & Goldman 1998). In contrast, the associative sequence learning model (ASL; e.g. Heyes 2001, 2005) suggests that the mirror system acquires its mirror properties through sensorimotor learning. Experience in which observation of an action is correlated with its execution establishes excitatory links between the sensory and motor representations of the same action, and these mediate mirror system activation. Both of these hypotheses are consistent with the finding that observation of human actions will activate the mirror system, and generate automatic imitation, to a greater extent than observation of non-biological movements. If the mirror system evolved through natural selection to support inferences about mental states, it should not be tuned to the movements of non-biological systems which lack mental states. Similarly, if the mirror system emerges through correlated sensorimotor experience, one would expect a human bias because self-observation, mirrors and synchronous social activities ensure that there are many more opportunities in the course of human development to execute actions while observing the same human actions than while observing the same non-biological movements.

The mirror system’s human bias could thus be due primarily to phylogenetic or ontogenetic factors (Heyes 2003). The present study sought to distinguish between these two hypotheses by investigating the influence of correlated sensorimotor training with robotic stimuli on automatic imitation of these stimuli. Automatic imitation of human and robotic movements was assessed before (pre-test) and after (post-test) training in which participants executed actions that matched, or were ‘compatible’ with (group CT), those of a robotic hand. A second group was included to control for the effects of unimodal sensory, and unimodal motor, experience on automatic imitation. During training, the participants in this control group (TT) observed and executed the movements with the same frequency as group CT, but instead of experiencing a match between observed and
executed actions, they experienced a non-matching or 'incompatible' sensorimotor contingency.

If the mirror system’s human bias is a product of natural selection, it is unlikely to be modified by a relatively brief period of sensorimotor training. Although the operation of an 'innate' system could, in principle, be modified by experience, it has been argued that experience-based alteration of an innate cognitive system would usually be maladaptive, and therefore that natural selection is likely to have acted to prevent such modification (Pinker 1997). Therefore, the phylogenetic hypothesis would predict that, compared with group IT, automatic imitation of robotic stimuli in group CT should not differ systematically between pre- and post-test. In contrast, if the mirror system develops through correlated experience of observing and executing actions, then training which involves the execution of actions that are compatible with those of observed robotic stimuli should promote automatic imitation of those stimuli. Therefore, compared with group IT, group CT should show a smaller human bias at post-test than at pre-test.

2. MATERIAL AND METHODS
Twenty healthy participants (eight male, mean age = 24.4 years) gave informed consent to take part in this study. All were right-handed, had normal or corrected-to-normal vision, and were naive with respect to the purpose of the experiment. Four participants were excluded from training and post-test sessions because they did not demonstrate numerically larger automatic imitation effects with human stimuli than with robotic stimuli at pre-test. The remaining 16 participants were randomly assigned in equal numbers to groups CT and IT. The study was approved by the University College London ethics committee and performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki.

(a) Pre- and post-test
The participant’s right forearm lay in a horizontal position across his/her body, parallel with the stimulus monitor, and was supported from elbow to wrist by an armrest. In each block of the simple RT task, participants were required to make a pre-specified response (to open or to close their right hand) as soon as the stimulus hand moved (either opened or closed). After making each response, participants were required to return their hand to a neutral starting position. They were instructed to refrain from moving their hand in catch trials, in which the stimulus hand did not move. Each stimulus was a naturalistic or a schematic representation of a human or a robotic hand. Naturalistic stimuli were used to establish ecological validity, and schematic stimuli were used to ensure that any human bias was not due to stimulus salience, e.g. to larger or brighter human stimuli. The four stimulus formats (human naturalistic, robotic naturalistic, human schematic and robotic schematic) are shown in figure 1. The naturalistic human and robotic stimuli differed in shape, colour palette (flesh versus metallic tones), luminance and surface area. The human stimuli were slightly brighter and occupied a larger area of the screen. Although not identical, the sizes of the naturalistic human and robotic stimuli were similar. The schematic human and robotic stimuli differed in shape but were controlled for colour (all were blue), size, luminance and surface area (see Press et al. (2005) for full details of stimulus control).

All trials began with the presentation of the warning stimulus (fingers closed and pointing upwards in parallel with the thumb). In most trials, the warning stimulus was replaced 800–1500 ms later by an imperative stimulus, an opened or closed hand, which was of 480 ms duration. Replacement of the warning stimulus by the imperative stimulus gave rise to apparent motion. The stimulus onset asynchrony (SOA) varied randomly between 800 and 1500 ms in 50 ms steps. The variable SOA, along with presence of catch trials, ensured the processing of stimulus movement before response initiation. After the presentation of the imperative stimulus, the screen went black for 3000 ms before the warning stimulus for the next trial appeared. In catch trials, the warning stimulus remained on the screen for 1980 ms before the 3000 ms inter-trial interval. In each session (pre- and post-test), the participants completed eight blocks of 36 trials. In each block, imperative stimuli showing an opened posture and those showing a closed posture were equiprobable and were randomly intermixed with six catch trials. There were two blocks with each of the four stimulus formats, one in which closing the hand was the required response and one in which opening the hand was the required response. Participants were instructed before, for example, opening response blocks: ‘when you see the hand move, regardless of the movement type, you should open your hand’.

For both the opening and closing hand movement responses, response onset was measured by recording the
electromyogram (EMG) from the first dorsal interosseous muscle of the right hand using disposable AG/AgCl surface electrodes. Signals were amplified, mains hum filtered at 50 Hz and digitized at 2.5 kHz. They were rectified and smoothed using a dual-pass Butterworth filter, with cut-off frequencies of 20 and 1000 Hz. To define a baseline, EMG activity was registered for 100 ms when the participant was not moving at the beginning of each trial. A window of 20 ms was then shifted progressively over the raw data in 1 ms steps. Response onset was defined by the beginning of the first 20 ms window after the onset of the imperative stimulus in which the standard deviation for that window, and for the following 20 ms epoch, was greater than 2.75 times the standard deviation of the baseline. This criterion was chosen during the 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 participant. The RT interval began with the onset of the imperative stimulus and ended with EMG onset. Errors were recorded manually.

(b) Training
During training, robotic naturalistic and robotic schematic stimuli were presented in a choice RT task. Participants were required to respond to opening and closing movements of the robotic hand by opening or closing their own hand in a compatible (group CT) or incompatible fashion (group IT). Group CT were instructed: ‘when you see the robotic hand open, open your hand, and when you see the robotic hand close, close your hand’. Group IT were told: ‘when you see the robotic hand open, close your hand, and when you see the robotic hand close, open your hand’.

All trials began with the presentation of the warning stimulus, which was replaced 1000 ms later by the imperative stimulus (480 ms duration). After the presentation of the imperative stimulus, the screen went black for 3000 ms before the next trial. Robotic naturalistic and robotic schematic stimuli were presented in separate blocks. Participants completed six blocks of 36 trials with each of these two stimulus formats. Naturalistic and schematic stimulus formats were presented in alternating blocks. Impressive stimuli consisting of opened and closed postures were equiprobable and presented in random order in each block.

3. RESULTS
(a) Preliminary analysis
During pre- and post-test sessions, participants initiated movement in 3.8% of catch trials. This low rate implies that, in standard trials, participants obeyed the task instructions by using stimulus movement as the imperative stimulus. Catch trials, practice trials, incorrect responses (0.09% in pre- and post-test, 1.59% during training) and response omissions (0.14% in pre- and post-test, 0.09% during training) were excluded from the analysis, as were RTs smaller than 100 ms and greater than 1000 ms (0.01% in pre- and post-test, 0.09% during training).

Figure 2 shows mean RTs in groups CT and IT during training. Analysis of these RTs confirmed that responding in group IT was slower than in group CT ($F_{1, 14} = 12.6, p = 0.005$) and that in both groups, there was a linear decline in RT over blocks ($F_{1, 14} = 14.6, p = 0.002$), indicative of learning.

(b) Pre- and post-test
Figure 3 shows the mean values of the compatibility effects induced by human and robotic stimuli at pre- and post-test for training groups CT (figure 3a) and IT (figure 3b). A human bias was clearly evident before training: in both groups, human stimuli elicited more automatic imitation than robotic stimuli. However, as predicted by the ontogenetic hypothesis, after training the human bias was preserved in the control group (IT), but was abolished in the group that had experienced matching sensorimotor contingency (CT). At post-test in group CT, the robotic movement stimuli elicited as much automatic imitation as the human movement stimuli.
These impressions were confirmed by ANOVA in which animacy (human and robotic) and test session (pre- and post-test) were within-subjects variables, and training type (CT or IT) was a between-subjects variable. This analysis indicated a significant three-way animacy × session × training type interaction ($F_{1,14} = 15.7, p = 0.001$). This effect was similar for naturalistic and schematic stimuli (animacy × session × training type × stimulus type, $F < 1$; animacy × session × training type × stimulus type for naturalistic stimuli, $F_{1,15} = 6.7, p < 0.03$, and for schematic stimuli, $F_{1,15} = 5.1, p < 0.05$). Simple effects analyses comparing effects of compatibility with human and robotic stimuli showed that there was a significant effect of animacy in both training groups at post-test (group CT: $F_{1,7} = 41.4, p < 0.001$; group IT: $F_{1,7} = 18.8, p < 0.005$). At post-test, the effect of animacy was preserved in group IT ($F_{1,7} = 17.3, p < 0.005$), but not in group CT ($F < 1$). Thus, these analyses confirm that compatible, but not incompatible, training with robotic stimuli increased automatic imitation of robotic stimuli to the level of that elicited by human stimuli. (When ANOVA was applied to the RT data from compatible and incompatible trials separately, in a compatibility × animacy × session × training-type design, no additional main effects or interactions were significant.)

In group CT, the magnitude of the compatibility effect for robotic stimuli increased between pre- and post-test, but in all other cases (group CT human stimuli, group IT human and robotic stimuli), the compatibility effect was smaller after training than before training (main effect of session: $F_{1,11} = 5.2, p < 0.04$). Increasing familiarity with general task demands is likely to have resulted in RT reduction between pre- and post-test, and previous studies have indicated that the magnitude of movement compatibility effects declines with RT (Brass et al. 2001; Press et al. 2005). Therefore, it is probable that the post-test reduction in compatibility effects was due to a reduction in RT. Additional analyses supported this interpretation by showing that RTs were shorter at post-test (mean = 251.1 ms, s.e. = 11.9 ms) than at pre-test (mean = 290.2 ms, s.e. = 12.6 ms; $F_{1,14} = 40.5, p < 0.001$), and, via quintile analyses (Ratcliff 1979), that the magnitude of the compatibility effects decreased as RT decreased ($F_{4,56} = 24.0, p < 0.001$, Greenhouse Geisser corrected).

4. DISCUSSION

The results of this study show that automatic imitation of robotic stimuli can be enhanced, and the human bias eliminated, by sensorimotor experience in which hand movements are executed during the observation of matching movements of a robotic hand. They are therefore consistent with the ontogenetic hypothesis, which suggests that the development of the mirror system depends on correlated experience of observing and executing the same actions (e.g. Heyes & Ray 2000; Heyes 2001, 2005; Brass & Heyes 2005; Heyes et al. 2005). The ASL model proposes that such experience establishes bidirectional excitatory links between visual and motor representations of action, and that these are responsible for the ‘mirror’ properties of the mirror system. This account implies that human actions usually promote more mirror system activation, and thereby more automatic imitation, than non-biological movements, because human developmental environments typically provide more experience in which action execution is paired with the observation of the same human action, than with the observation of the same non-biological action. The present study shows that when people are given experience of the latter kind, the human bias disappears.

Our findings are consistent with those of a number of recent studies indicating that the activity of the mirror system covaries with expertise (Jarvelainen et al. 2004; Calvo-Merino et al. 2005, 2006; Ferrari et al. 2005; Haslinger et al. 2005; Cross et al. 2006; Vogt & Thomaschke 2007). For example, in a functional magnetic resonance imaging (fMRI) study of expert ballet and capoeira dancers, Calvo-Merino et al. (2005) found that when participants (e.g. ballet dancers) observed those actions that they had been trained to perform (ballet movements), there was greater activation in premotor and parietal cortices than when they observed actions they had not been trained to perform (capoeira movements). Cross et al. (2006) showed similar effects of expertise in dancers; left premotor and parietal activation was correlated with the dancers’ own ratings of their competence in performing the observed movements.

These studies implicate experience in the development of the mirror system, but, unlike the present experiment, they do not tell us whether it is sensory experience, motor experience or, as the ASL model predicts, sensorimotor experience, which is critical. Calvo-Merino et al. (2006) found that premotor and parietal cortices were activated to a greater extent when ballet dancers (e.g. female) observed movements performed by their own gender (female), compared with the opposite gender (male). This implies that sensory (visual) experience alone is insufficient for the development of the mirror system, but does not isolate the effects of motor experience from those of sensorimotor experience. Since ballet dancers make extensive use of mirrors while training, female dancers have not only more motor experience of female movements but also more sensorimotor experience of female movements, than male dancers. In contrast, the results of the present experiment clearly and specifically implicate sensorimotor experience as the engine of mirror system development. During training, the compatible and incompatible groups had equal amounts of visual experience of the robotic stimuli, and equal amounts of motor experience of the hand actions—they observed and performed the opening and closing responses equally often. However, it was only the group that experienced a compatible sensorimotor contingency, or matching relationship, between the robotic movements and their own movements, that showed an increase in automatic imitation of the robotic actions. These findings are consistent with those of a recent fMRI study, which used a similar training regime with human stimuli, and found an effect of training type on premotor and parietal cortex activation during action observation (Bird et al. submitted).

The results of the present study do not support the hypothesis that the functional characteristics of the mirror system evolved through natural selection to support mental state understanding (e.g. Gallese & Goldman 1998; Melzoff & Decety 2003; Kilner et al. 2003).
If phylogenetic factors determined the preferential mirroring of human action over non-biological movement, it is unlikely that this bias would be affected by a relatively brief period of training (Pinker 1997). Although inconsistent with a purely phylogenetic account of the origins of the mirror system, the findings of the present study leave open the possibility that the mirror system provides input to higher-level sociocognitive functions, such as action understanding (Jacoboni et al. 2005), empathy (Carr et al. 2003; Gallese 2003) and theory of mind (Gallese & Goldman 1998). The pre-test results presented here are in line with previous evidence that human stimuli generate more activation in the mirror system, and elicit more automatic imitation, than non-biological stimuli (Stevens et al. 2000; Kilner et al. 2003; Tai et al. 2004; Ottop et al. 2005; Press et al. 2005, 2006). Therefore, assuming that humans have mental states and that non-biological systems do not, their differential activation of the mirror system, in the context of normal human development.

In a recent fMRI study, Gazzola et al. (2007) failed to find more mirror system activation during the observation of human rather than robotic action stimuli and suggested that, when it is observed, the human bias is due to the more repetitive nature of robotic movements. It is certainly true that responses to repetitive motion are more likely to habituate and that some studies reporting a human bias have confounded animacy with movement invariance. Therefore, it is plausible that some reports of a human bias in the mirror system are unreliable. However, in the present study, and in previous studies using automatic imitation as an index of mirror system functioning (Press et al. 2005, 2006), the robotic action stimuli were no more, or less, repetitive than the human action stimuli. Therefore, our results indicate that movement invariance is not the only factor that diminishes the power of robotic stimuli, and that the human mirror system is genuinely biased in favour of human action stimuli. Whether or not this bias is detected in any given study is likely to depend on the extent to which the context, procedure and task instructions focus attention on the kinematic and morphological variables that distinguish human from robotic action stimuli.

5. CONCLUSION
The present study demonstrates that correlated sensorimotor training, in which participants perform actions while observing compatible movements of a robotic model, can increase automatic imitation of robotic movements and eliminate the human bias. These findings are consistent with an ontogenetic account of the origins of the mirror system; they suggest that the system’s functional capacity to match observed with executed actions, and its human bias, are products of sensorimotor experience in the course of normal human development.

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 C.P. by the Biotechnology and Biological Sciences Research Council (BBSRC).

REFERENCES


