Journal of Experimental Psychology: General
Running head: Contextual modulation of counter-mirror responding
Word count: 8,786

Contextual modulation of mirror and counter-mirror sensorimotor associations

Richard Cook*[1], Anthony Dickinson[2], and Cecilia Heyes[3]

*Corresponding Author: email r.cook@ucl.ac.uk

[1]Cognitive, Perceptual and Brain Sciences Research Department, University College London, London WC1H 0AP, UK.

[2] Department of Experimental Psychology, University of Cambridge, Cambridge CB2 3EB, UK.

[3]All Souls College & Department of Experimental Psychology, University of Oxford Oxford, OX1 4AL, UK.

Key words: Automatic imitation, associative sequence learning, mirror neuron system, context, counter-mirror training

ABSTRACT

Automatic imitation – the unintended copying of observed actions - is thought to be a behavioural product of the mirror neuron system (MNS). Evidence that the MNS develops through associative learning comes from previous research showing that automatic imitation is attenuated by counter-mirror training, in which the observation of one action is paired contingently with the execution of a different action. If the associative account of the MNS is correct, counter-mirror training should show context-specificity, because counter-mirror associations render action stimuli ambiguous, and ambiguity promotes contextual control. Two experiments are reported which confirm this prediction. In Experiment 1 we found less residual automatic imitation when human participants were tested in their counter-mirror training context. In Experiment 2, sensorimotor training where participants made action responses to novel abstract stimuli was insensitive to the same context manipulation, confirming that the former result was not a procedural artefact. Contextual modulation may enable the MNS to function effectively in spite of the fact that action observation often excites multiple conflicting MNS responses.

INTRODUCTION

Mirror neurons (MNs) are commonly characterized as single neurons that respond selectively during both the observation and execution of the same action - for example, to the observation and performance of a precision grip, but not during the observation or performance of a power grip. However, a close reading of the results from single-cell recording indicates that substantial proportions of macaque MNs (di Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992; Ferrari, Gallese, Rizzolatti, & Fogassi, 2003; Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Rozzi, Ferrari, Bonini, Rizzolatti, & Fogassi, 2008; Umilta et al., 2001) and human sensorimotor units (Mukamel, Ekstrom, Kaplan, Iacoboni, & Fried, 2010) discharge indiscriminately during the observation of multiple actions, or selectively during the observation and execution of different actions. These findings suggest that the sight of an action may often excite several different congruent and non-congruent motor representations. Such conflicting MN responses are potentially problematic for accounts which propose that the MNS mediates action understanding (Rizzolatti & Craighero, 2004; Rizzolatti, Fadiga, Gallese, & Fogassi, 1996), imitation (Heyes, 2001, 2011; Iacoboni, 2009), or generic action selection (Hickok & Hauser, 2010). However, through its appeal to the principles of associative learning to explain the development of MNs, the associative sequence learning (ASL) model (Heyes, 2001, 2010a, 2010b; Ray & Heyes, 2011) suggests that contextual modulation may enable the mirror neuron system (MNS) to function effectively in spite of the conflicting responses of individual MNs. To investigate whether the MNS is subject to contextual modulation, the present study examined the effects of context on automatic imitation – a behavioral effect widely thought to be mediated by the human MNS.

Automatic imitation and the mirror neuron system

Automatic imitation is a robust behavioral effect in which the topographical features of task-irrelevant action stimuli facilitate similar, and interfere with dissimilar, motor responses (Brass, Bekkering, & Prinz, 2001; Heyes, Bird, Johnson, & Haggard, 2005; Sturmer, Aschersleben, & Prinz, 2000). For example, participants are faster to make hand-open responses to the onset of hand opening stimuli than to the onset of hand-close stimuli (Heyes et al., 2005). Similarly, participants execute finger lift responses faster to the onset

of finger lifting stimuli than finger tapping stimuli (Brass et al., 2001). The finding that participants make faster imitative responses than non-imitative responses is highly robust (Heyes, 2011), having been reported for several effector systems (Gillmeister, Catmur, Liepelt, Brass, & Heyes, 2008; Leighton & Heyes, 2010), and for both transitive (Craighero, Bello, Fadiga, & Rizzolatti, 2002) and intransitive actions (Press, Bird, Walsh, & Heyes, 2008). Moreover, studies have confirmed that this effect is truly imitative; it depends on the topography of observed actions - on how body parts move relative to one another; not merely on spatial compatibility - the position of the action relative to an external frame of reference (Catmur & Heyes, 2010; Cooper, Catmur, & Heyes, under review).

It is likely that automatic imitation is mediated by the same processes as motor mimicry and mirror effects (Heyes, 2011; Wang, Newport, & Hamilton, 2011). The term 'motor mimicry' is used to describe the occurrence of spontaneous, unconscious imitation in naturalistic social settings (Chartrand & Bargh, 1999; Cook, Bird, Lunser, Huck, & Heyes, 2011). For example, participants are more likely to engage in foot-tapping than face-touching behaviours in the presence of a foot-tapping confederate, while the opposite pattern is observed in the presence of a confederate prone to touching their face (Chartrand & Bargh, 1999). 'Mirror effects' provide evidence of covert imitation within the human motor system (Dimberg, Thunberg, & Elmehed, 2000; Fadiga, Fogassi, Pavesi, & Rizzolatti, 1995; Strafella & Paus, 2000). For example, during passive observation of actions, motor evoked potential (MEPs) recorded from the muscles involved in performing the observed action are greater than MEPs recorded from task-irrelevant muscles (Fadiga et al., 1995; Strafella & Paus, 2000).

Automatic imitation is widely thought to be a product of a human MNS (Bien, Roebroeck, Goebel, & Sack, 2009; Blakemore & Frith, 2005; Catmur, Walsh, & Heyes, 2009; Ferrari, Bonini, & Fogassi, 2009; Heyes, 2011; Iacoboni, 2009; Longo, Kosobud, & Bertenthal, 2008; Rizzolatti, Fogassi, & Gallese, 2001). MNs are single units identified in the ventral premotor (area F5) and inferior parietal cortices (area PF) of the macaque (di Pellegrino et al., 1992; Fogassi et al., 2005; Gallese et al., 1996; Rizzolatti et al., 1996) which respond to both the sight and execution of transitive and communicative actions. Since the discovery

of MNs in monkeys, considerable indirect evidence has accumulated suggesting that humans also have a MNS. Numerous neuroimaging studies have revealed premotor and parietal areas of the human brain which respond to the observation and execution of transitive and intransitive actions (Buccino et al., 2001; Chong, Cunnington, Williams, Kanwisher, & Mattingley, 2008; Fadiga et al., 1995; Gazzola, Rizzolatti, Wicker, & Keysers, 2007; Grezes, Armony, Rowe, & Passingham, 2003; Iacoboni et al., 1999; Kilner, Neal, Weiskopf, Friston, & Frith, 2009). Single-cell recording in patients with intractable epilepsy has also identified neurons with mirror properties in the medial wall (supplementary motor area, cingulate cortex) and the medial temporal lobe (hippocampus, parahippocampal gyrus, entorhinal cortex and amygdala) of the human brain (Mukamel et al., 2010). Although some studies have failed to find evidence of a human MNS (Dinstein, Gardner, Jazayeri, & Heeger, 2008; Lingnau, Gesierich, & Caramazza, 2009), the balance of evidence provides clear support for the existence of such a network (Grezes & Morin, 2008; Molenberghs, Cunnington, & Mattingley, 2009, 2011).

Evidence from functional magnetic resonance imaging (fMRI) paradigms suggests that the MNS mediates automatic imitation. For example, sensorimotor training that modulates the magnitude of participants' automatic imitation effects also modulates the magnitude of the blood oxygen level dependent (BOLD) response in premotor areas thought to be part of the MNS (Catmur et al., 2008). Convergent evidence has also been derived using transcranial magnetic stimulation (TMS). Specifically, the application of disruptive TMS to the inferior frontal gyrus (IFG) has been shown to abolish the automatic imitation effect (Catmur et al., 2009; Heiser, Iacoboni, Maeda, Marcus, & Mazziotta, 2003; Newman-Norlund, Ondobaka, van Schie, van Elswijk, & Bekkering, 2010). The IFG is thought to be the human homologue of the macaque premotor region F5 and is widely regarded as a key component of the human MNS (Kilner et al., 2009). Reports that virtual lesions to this area abolish automatic imitation are therefore important because they suggest that the MNS, and the IFG in particular, makes a necessary causal contribution to this effect.

MNs are frequently characterised as neurons that respond selectively to the observation and execution of the same action (Arnstein, Cui, Keysers, Maurits, & Gazzola, 2011; Chong et al., 2008; Dinstein, Hasson, Rubin, & Heeger, 2007; Dinstein, Thomas, Behrmann, &

Heeger, 2008; Keysers & Gazzola, 2010). However, a surprising number of MNs do not exhibit sensorimotor congruency. The few studies of macaque MNs which report detailed congruency analyses all describe substantial proportions (6-30%) of so-called 'non-congruent' (or 'logically-related') MNs which respond to the observation and execution of different actions (di Pellegrino et al., 1992; Gallese et al., 1996; Rozzi et al., 2008). In addition, they indicate that a significant proportion of macaque MNs (15-65%) are 'ambiguous' in that they respond to the observation or execution of multiple actions (di Pellegrino et al., 1992; Ferrari et al., 2003; Gallese et al., 1996; Rozzi et al., 2008; Umilta et al., 2001). Similarly, the study of sensorimotor units in human participants has revealed large proportions of neurons with both non-congruent and ambiguous sensorimotor properties (Mukamel et al., 2010).

Ontogenetic origins of the mirror neuron system

Considerable evidence has accumulated indicating that learning plays a crucial role in the development of the human MNS. Neuroimaging studies of sensorimotor expertise have taken advantage of naturally occurring variation in participants' sensorimotor experience to better understand how learning shapes the motor responses elicited by action observation. For example, in expert pianists but not in non-pianist controls, the fMRI BOLD response in premotor areas is stronger during the observation of piano-related than of arbitrary finger movements (Haslinger et al., 2005). Similarly, capoeira and ballet dancers show stronger motor responses when viewing sequences of dance movement from their own genre (Calvo-Merino, Glaser, Grezes, Passingham, & Haggard, 2005). Using a similar design, it has also been reported that male and female ballet dancers show greater motor activation when viewing dance sequences unique to their gender (Calvo-Merino, Grezes, Glaser, Passingham, & Haggard, 2006). Because dancers had equivalent visual exposure to the movements of both sexes, this finding cannot be attributed to visual experience alone.

'Counter-mirror' training studies suggest, not only that learning plays a crucial role in the ontogeny of the MNS, but more specifically that its development is driven by associative learning based on sensorimotor experience. The logic behind these training studies is simple: If the MNS develops through associative learning, then markers of MNS activity should be reduced or even reversed by training in which the execution of one action is

contingent upon the observation of a different action. The first counter-mirror training study was reported by Heyes et al. (2005). Participants in the counter-mirror group first completed a training session during which they repeatedly made hand-open responses to hand-close stimuli and hand-close responses to hand-open stimuli. When tested 24-hours later, the counter-mirror group showed much smaller residual automatic imitation effects than a 'mirror' control group trained to execute imitative responses. Counter-mirror training has since been shown to reverse the mirror pattern of MEPs seen during action observation (Catmur, Mars, Rushworth, & Heyes, 2010; Catmur, Walsh, & Heyes, 2007), and to modulate the BOLD response in fMRI studies of the human MNS (Catmur et al., 2008). By showing that counter-mirror training has parallel effects on behavioral responses (automatic imitation) and on electrophysical and neuroimaging markers of MNS function, these studies provide convergent evidence that the MNS mediates automatic imitation, and that counter-mirror learning serves to modulate the activity of MNs.

The counter-mirror training paradigm was developed to test the ASL model of the origins of the MNS. This model proposes that MNs are a product of the same domain-general mechanisms of associative learning that produce Pavlovian and instrumental conditioning phenomena in humans and animals (Brass & Heyes, 2005; Heyes, 2001, 2010a, 2010b; Ray & Heyes, 2011). Where the observation of an action is contingent on the execution of an action, or vice versa, ASL posits that the respective visual and motor representations will become associated. It is these associations which are thought to endow MNs with their sensorimotor properties. Many of the sensorimotor contingencies we experience are matching, where observed and executed actions correspond, and therefore give rise to congruent MNs and automatic imitation. Experience of this kind is provided by, for example, visual monitoring of one's own actions, synchronous activity in response to a common stimulus, and being imitated by others (Ray & Heyes, 2011). However, exposure to non-matching sensorimotor contingencies may cause the emergence of non-congruent MNs through the same learning mechanisms (Cook, Press, Dickinson, & Heyes, 2010). Similarly, where the sight of an action predicts the execution of more than one action, or where action execution predicts the observation of several actions, ambiguous MNs may emerge with multiple effective actions.

Ambiguity and context

The sensorimotor properties of ambiguous and non-congruent MNs suggest that observing a given action often excites several different motor representations. This ambiguity is problematic for accounts which propose that the MNS makes a functional contribution to action understanding (Rizzolatti & Craighero, 2004; Rizzolatti et al., 1996), imitation (Heyes, 2001, 2011; Iacoboni, 2009), or generic action selection (Hickok & Hauser, 2010), because it implies uncertainty about how observed actions should be 'understood' or which responses 'selected'. The conflicting motor representations excited by congruent and non-congruent MNs makes it unclear which interpretation or response is appropriate in a given situation. The associative account of MNS development, embodied in the ASL model, suggests a potential mechanism for resolving the ambiguity caused by conflicting MN responses. Associative learning theory suggests that where a stimulus is rendered ambiguous by virtue of being associated with multiple outcomes, the ambiguity is resolved through contextual modulation. While the first associations formed with a novel stimulus generalize well to other contexts, subsequently acquired associations, which give rise to ambiguity, demonstrate greater contextual specificity (Bouton, 1993, 1994; Nelson, 2002).

Our understanding of contextual modulation has been advanced by the study of two related conditioning effects, renewal and counter-conditioning. In renewal designs, participants are first placed in Context A where they learn that a stimulus predicts a certain outcome. Once the initial association has been acquired, participants are transferred to Context B, where they learn that the same stimulus no longer predicts the outcome. However, the extinction learning that takes place during this second phase is subject to contextual control: When transferred back to Context A (ABA renewal), or placed in a novel Context C (ABC renewal), a renewal of responding occurs, i.e. participants exhibit the conditioned response acquired during the first phase (Bouton & King, 1983; Nelson, Sanjuan Mdel, Vadillo-Ruiz, Perez, & Leon, 2011). In these studies 'context' typically refers to the physical environment provided by the conditioning chamber or testing cubicle. However the definition of 'context' may be extended to include internal states, such as those induced by the presence of alcohol or a tranquilizer (Bouton, 1993, 1994).

Of more direct relevance to counter-mirror training are demonstrations that stimulus ambiguity plays a similar role in the contextual control of counter-conditioning. Counter-conditioning is an associative learning paradigm in which the conditioned responses to a stimulus vary across successive phases of a training procedure. Typically participants are first placed on a training schedule where they learn that a stimulus signals one outcome, before being placed on a second conditioning schedule in which the same stimulus predicts a different outcome. Importantly, the asymmetric learning seen in renewal is also observed in counter-conditioning. While the conditioned response acquired in the first phase is relatively insensitive to changes in context, the learning that occurs in the second phase is often far more context-specific. For example, Peck and Bouton (1990) initially trained rats to expect a mild electric shock following a tone in Context A, before transferring them to a second schedule, where the tone signalled the delivery of food, in Context B. Although the original conditioned response was reduced during training in Context B, it re-emerged when the rats were returned to Context A or placed in a novel Context C.

The present study

The aim of the present study was to test the hypothesis, advanced by the ASL model, that when the sight of an action is associated with rival mirror and counter-mirror responses, the resulting ambiguity is resolved through contextual modulation. Applying the terminology used in studies of counter-conditioning and renewal, ASL implies that the congruent MNs responsible for automatic imitation are a product of first-learned 'mirror' associations. Each of these associations connects a sensory representation of an action with a motor representation of the same action, and is acquired through everyday experience of a predictive relationship between the observation and execution of the same actions. The ASL model further implies that, during counter-mirror training, participants acquire a set of second-learned 'counter-mirror' associations, connecting sensory representations of actions with motor representation of different actions. Therefore, if the ASL model is correct in suggesting that MNS development and modification depend on associative learning, then counter-mirror learning, like counter-conditioning, should be subject to contextual control. This prediction was tested in Experiment 1, where automatic imitation was measured in red and blue contexts, both before and after two sessions of counter-mirror training completed in either red or blue contexts. The purpose of Experiment 2 was to confirm that any

context-specificity observed in Experiment 1 was due, as the ASL model suggests, to conflict between pre-experimental learning and counter-mirror training, and not to an artefact of procedure.

EXPERIMENT 1

Our first experiment sought to test whether counter-mirror learning is context-specific. In the first of four sessions, two separate pre-tests were conducted to establish participants' baseline automatic imitation effects in distinctive red and blue contexts. During the second and third sessions, participants received counter-mirror training in either the red or the blue contexts. In the final session two separate post-tests were conducted to establish participants' residual automatic imitation effects in both the red and blue contexts. Context-specificity was indexed by comparing the magnitude of the pre- to post-test reduction in automatic imitation when participants were tested in the context in which they received counter-mirror training and in their untrained context. If the MNs responsible for automatic imitation acquire their properties through associative learning, counter-mirror learning should behave like counter-conditioning - it should show context-specificity due to the conflict between the first-learned mirror associations and the second-learned counter-mirror associations.

Method

Participants

Sixteen healthy adults (five males) with a mean age of 22.4 years served as participants in the experiment in return for a small honorarium. All were right handed, had normal or corrected-to-normal vision, and were naive to the purpose of the experiment. Participants were assigned randomly to either the train-red or train-blue groups in equal numbers. The study was approved by the University College London ethics committee and performed in accordance with the ethical standards set out in the 1964 Declaration of Helsinki.

Apparatus and stimuli

The stimuli used in Experiment 1 consisted of 12 digital images of a model's right hand in naturalistic tones (Figure 1). Six of the images were presented against a red background, and six against a blue background. Each colour set comprised images of male and female

neutral hands; male and female closed hands; and male and female open hands. The open stimulus showed fingers and thumb splayed. The male stimulus subtended approximately 17° of visual angle horizontally and 20° vertically, whereas the female stimulus subtended approximately 16° of visual angle both horizontally and vertically. The closed stimulus depicted a fist. The male stimulus subtended approximately 11° horizontally and 14° vertically, and the female stimulus subtended approximately 8° horizontally and 12° vertically. In the neutral hand stimulus fingers were shown together, pointing upwards in parallel with the thumb. Both training and test trials were presented in colour on a laptop PC with a 38 cm screen (resolution 1,024 x 678 pixels).

Contexts

Four elements were varied to provide two distinctive 'red' and 'blue' contexts. As described above, hand stimuli were presented against either red or blue backgrounds. In addition, the testing cubicle was lit by either red or blue light, provided by a Eurolite PAR-38 RGB LED spotlight. The experimental procedure required the use of an armrest to support the participants' hands and an occluder to prevent them observing their own responses. Distinctive armrests and occluders were constructed for use in the red and blue contexts. The occluders were identical in all features except their colour (either red or blue). The armrests were identical in size (15cm x 35cm x 25cm) but differed in both colour and surface texture. The red armrest was covered in a red coarse woollen fabric. The blue armrest was covered with a fine-grain plastic material. Display backgrounds (Cook et al., 2010; Nelson et al., 2011) and lighting changes (Van Gucht, Vansteenwegen, Beckers, & Van den Bergh, 2008; Vansteenwegen et al., 2005) have been shown to be effective context manipulations in single-room procedures.

Data recording and analysis

Both the training and test procedures took the form of reaction time (RT) tasks in which electromyography (EMG) was used to establish response onset. Recordings were taken from the first dorsal interosseus (FDI) muscle using disposable surface electrodes manufactured by Unomedical Limited, UK. The EMG signal was 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 Hz and 1,000 Hz. The EMG signal

was registered for 100 ms before the onset of the imperative stimulus (see below) to define the baseline. A window of 20 ms was then shifted incrementally over the raw data in 1 ms steps. Response onset was defined as the start of the earliest 20 ms window following presentation of the imperative stimulus in which the standard deviation for that window, and for the following 20ms interval, was greater than 2.75 times the standard deviation of the baseline. That this criterion reasonably defined response onset was verified by sight for every training and test trial.

Procedure

The experiment was conducted over four sessions each completed approximately 24 hr apart. During the first session participants completed two test procedures to establish their baseline automatic imitation effects in the red and blue contexts. In sessions two and three, participants completed six blocks of counter-mirror training spread evenly over the two days, in either the red or blue context. In the final session, participants' residual automatic imitation effects were again measured in the red and blue contexts. The order in which participants completed the red and blue tests was counterbalanced, but was held constant across sessions one and four.

In all four sessions, participants sat at a viewing distance of approximately 60 cm with their shoulders parallel to the stimulus display. Participants' arms were bent at the elbow, with their forearm positioned to face downwards, also parallel to the display. The elbow and forearm were supported by an armrest, with both the hand and forearm occluded from view. Each participant's hand and wrist were positioned such that their fingers moved upwards during open responses and downwards during close responses. Stimulus postures were presented in the lateral plane (left–right), thus ensuring that response movements were orthogonal to stimulus postures throughout. This feature of the design allows automatic imitation to be isolated from left–right spatial compatibility.

The test procedure was a simple RT task (see Figure 1), in which EMG recording was used to establish response onset. Tests comprised two blocks of 80 trials during which participants made speeded pre-specified movements (hand-open responses in one block, hand-close in the other) in response to the onset of open and closed stimuli. This factorial

manipulation constitutes a stimulus-response compatibility (SRC) design whereby responses can be either compatible with the observed stimulus (e.g. hand-open response to hand-open stimulus) or incompatible (e.g. hand-open response to hand-close stimulus). Each test trial started with the participant's hand in the neutral starting position, and with a neutral hand warning stimulus on the screen. Thereafter participants were required to make the pre-specified response as soon as an action stimulus appeared before returning to the neutral position ready for the next trial. The action stimulus was present for 480 ms until replaced by a blank display for 3,000 ms, prior to the warning stimulus for the next trial. Stimulus onset asynchrony (SOA) was varied randomly between 800 ms and 1,500 ms in 50 ms increments. The order in which participants completed the open and close blocks was counterbalanced across groups.

Twenty catch trials were included in each test block, in which the warning stimulus was displayed throughout the trial to which the participants had been instructed to make no response whatsoever. Because participants were making speeded pre-specified responses, catch trials were included to prevent habitual, anticipatory responding and to encourage participants to continue to monitor what the stimulus was doing and to remain engaged with the task. On catch trials, the warning stimulus was presented for 1,980 ms before the 3,000 ms inter-trial interval.

The training procedure took the form of a choice RT task (see Figure 1). Each trial required the participant to make either an open response to a close-hand stimulus or a close response to an open-hand stimulus. Having made each response participants returned to the starting position ready for the next trial. The order of the action stimuli presentation was randomized during the training trials. The beginning of each trial was indicated by the appearance of the warning stimulus, which was then replaced by one of the action stimuli, presented for 480 ms. SOA was varied randomly between 800 ms and 1,500 ms in 50 ms increments. The hand depicted in the warning stimulus was identical to that in the subsequent action stimulus, giving rise to apparent motion. Following the offset of the action stimulus, the screen went blank for 3,000 ms, until the warning stimulus for the subsequent trial was presented. Each training block comprised 144 counter-mirror training

trials. Half of the participants completed the counter-mirror training task in the red context and half in the blue context.

Figure 1

Results and Discussion

Training and test trials in which participants made incorrect responses, no response, or where the point of movement onset was equivocal were excluded from all further analyses (2.8% of training trials and 4.1% of test trials). Thereafter, any remaining data points beyond 2.5 standard deviations of a participant's mean response latency for a given block were also excluded (2.0% of training trials and 2.9% of test trials). EMG signals on catch trials were examined to ensure that the participants obeyed task instructions, but were excluded from all further analyses. During the test sessions, participants initiated movements on only 5.4% of catch trials, indicating that they were not making anticipatory responses on the stimulus trials.

Training

Figure 2 shows mean RTs for the groups trained in red and blue contexts. These training data were analysed using a mixed-model ANOVA with training block (B1-B6) as a within-subjects factor and training group (train-red, train-blue) as a between-subjects factor. Trend analysis revealed a highly significant linear decline across the six training blocks [$F(1,14) = 19.83$; $p < .001$; $\eta^2 = .59$] suggestive of learning. This trend did not vary as a function of group [$F(1, 14) = .02$; $p > .80$; $\eta^2 = .00$] indicating that the improvement in training task performance was comparable. Although Figure 2 suggests that responses were faster in the red context than in the blue, no main effect of group was present [$F(1, 14) = 2.33$; $p > .14$; $\eta^2 = .14$], implying that overall RTs during the training task were broadly equivalent. Simple effects analysis revealed that the difference between the groups at Day 1, Block 1 was not significant [$t(14) = .87$; $p > .40$] indicating no difference in baseline performance. No higher order trends or other trend x group interactions were observed ($p > .07$).

Figure 2

Tests

Figure 3a shows the mean RTs observed at pre- and post-test, in the trained and untrained contexts, on compatible and incompatible trials. Figure 3b depicts the mean automatic imitation effects observed on the four tests (pre-test in the trained context; pre-test in the untrained context; post-test in the trained context; post-test in the untrained context). Automatic imitation reflects the tendency to execute imitative responses faster than non-imitative responses. The magnitude of the automatic imitation effect was therefore calculated by subtracting the mean RT on imitative trials (open responses to open stimuli; close responses to close stimuli) from the mean RT on non-imitative trials (open responses to close stimuli; close responses to open stimuli) for each participant. The greater this RT difference, the stronger the tendency to imitate. In Table 1, performance is further broken down for the groups trained in the red and blue contexts.

Table 1

The RT data were analyzed using a mixed-model ANOVA with test (pre-test, post-test), context (trained, untrained) and stimulus-response compatibility (compatible, incompatible) as within-subjects factors and group (train-red, train-blue) as a between-subjects factor. The analysis revealed a highly significant main effect of stimulus-response compatibility [$F(1,14) = 98.23$; $p < .001$; $\eta^2 = .88$] whereby participants were slower to make incompatible responses (M = 303.4ms; SD = 71.6ms) than compatible responses (M = 282.2ms; SD = 68.9ms) indicative of automatic imitation. In addition, a marginally significant test x compatibility interaction was observed [$F(1,14) = 3.94$; $p = .067$; $\eta^2 = .22$] indicating that automatic imitation effects were generally smaller at post-test (M = 16.3ms; SD = 15.1ms) than at pre-test (M = 26.2ms; SD = 16.4ms). Crucially, this test x compatibility interaction varied as a function of context [$F(1,14) = 7.46$; $p < .025$; $\eta^2 = .35$]. Simple effects analysis indicated that there was a significant reduction in automatic imitation between the pre-test (M = 26.9 ms; SD = 14.5 ms) and post-test (M = 12.6 ms; SD = 15.3 ms) when tested in the trained context [$t(15) = 2.89$; $p < .025$]. In contrast, the reduction in automatic imitation between the pre-test (M = 25.5 ms; SD = 18.6 ms) and post-test (M = 20.1 ms; SD = 14.5 ms) when tested in the untrained context was not

significant [$t(15) = 1.05$; $p > .30$]. This test x context x compatibility interaction did not vary as a function of group [$F(1, 14) = .40$; $p > .5$; $\eta^2 = .03$].

In addition to the effects of principal interest described above, a significant group x compatibility interaction was observed [$F(1,14) = 10.25$; $p < .025$ $\eta^2 = .42$]. Those participants given counter-mirror training in the red context generally showed greater automatic imitation (M = 28.1 ms; SD = 14.5 ms) than those trained in the blue context (M = 14.4 ms; SD = 15.5 ms). A marginally significant main effect of test was also observed [$F(1,14) = 3.94$; $p = .067$; $\eta^2 = .22$] indicating that participants generally responded faster at post-test (M = 277.5 ms; SD = 60.5 ms) than at pre-test (M = 308.2 ms; SD = 77.3 ms). None of the other main effects or interactions approached significance ($p > .10$).

Figure 3

The results of Experiment 1 indicate that greater reductions in automatic imitation were seen when participants were tested in their trained context than in their untrained context. This finding suggests that counter-mirror learning behaves like counter-conditioning, in that both show context-specificity (Bouton & Peck, 1992). Associative accounts argue that counter-conditioning comes under contextual control in order to resolve ambiguity (Bouton, 1993, 1994; Nelson, 2002). Rather than simply overwriting the original learning, there is considerable evidence that subsequent conditioning to the same stimulus sets up parallel, second-learned associations (Bouton & Peck, 1992; Brooks, Hale, Nelson, & Bouton, 1995). However, the conflict between the first- and second-learned associations renders the stimulus ambiguous: it is not clear what the appropriate behaviour is in the presence of the stimulus. To resolve this ambiguity, the excitability of both sets of associations is modulated by the context. Second-learned associations become active only in the trained context, whereas first-learned associations are inhibited by the cues present in the trained context, but remain active in all other contexts.

EXPERIMENT 2

The ASL model (Heyes, 2001, 2010a, 2010b; Ray & Heyes, 2011) proposes that automatic imitation, and the congruent MNs that are thought to mediate automatic imitation, are a

product of sensorimotor associations acquired during development, e.g. through self-observation, synchronous activity and while being imitated. This model therefore implies that the context-specificity observed in Experiment 1 was due to conflict between the second-learned counter-mirror associations established by experimental training, and first-learned mirror associations acquired long before the experiment began.

Experiment 2 sought to confirm this interpretation, by testing whether first-learned sensorimotor associations generalize across contexts within our procedure. Novel S-R learning should generalize well to untrained contexts because there is no conflict with pre-experimental learning, and consequently the stimuli remain unambiguous. Evidence of context-specific learning in Experiment 2 would therefore indicate that the context-specificity observed in Experiment 1 was not due to conflict with pre-experimental learning, and may instead reflect an artefact of procedure.

We used a four session design, identical to that employed in Experiment 1, to train arbitrary stimulus-response (S-R) mappings with abstract geometric shapes. Participants again made open- and close-hand responses, but this time to the onsets of geometric trapezoid forms. In the first session they completed simple RT pre-test procedures in the red and blue contexts to confirm the absence of any pre-existing compatibility effects. They then completed two training sessions with a choice RT task, either in the red or blue contexts, where they learned to make open and close responses to the onset of top- and bottom-heavy trapezoids, respectively. In the final session they were tested again in both the red and blue contexts to determine the magnitude of their newly acquired SRC effect.

Method

Participants

A further 16 healthy adults (4 males) with a mean age of 21.9 years served as participants in the experiment in return for a small honorarium. All were right handed, had normal or corrected-to-normal vision, and were naive to the purpose of the experiment. Participants were assigned randomly to either the train-red or train-blue training groups in equal numbers.

Apparatus and stimuli

The stimuli used in Experiment 2 consisted of 12 digital images of grey quadrilaterals (Figure 4). Two sets of lighter and darker stimuli, each with different aspect ratios were employed to parallel the pigmentation and scale variation in the male and female hand stimuli. Six of the images were presented against a red background, and six against a blue background. Each colour set included two rectangles, one lighter and one darker, and four Isosceles trapezoids bisected by a horizontal black line, one top-heavy in lighter tones, one top-heavy in darker tones, one bottom-heavy in lighter tones and one bottom-heavy in darker tones. The lighter grey rectangle (160 on a 0-255 scale) had an aspect ratio of 1:1.23 and subtended approximately 12° horizontally. The darker rectangle (140 on a 0-255 scale) had an aspect ratio of 1:1.57 and subtended approximately 9° horizontally. The lighter grey trapezoids subtended 14° at the wider end (170 on a 0-255 scale) and 10° at the narrower end on a (150 on a 0-255 scale). The darker grey trapezoids subtended 11° at the wider end (150 on a 0-255 scale) and 7° at the narrower end (130 on a 0-255 scale).

Procedure and design

As in Experiment 1, the order in which participants completed the red and blue tests was counterbalanced, but held constant across pre- and post-test. The test procedure employed in both contexts again took the form of a simple RT task. Participants made speeded pre-specified movements (open responses in one block, close in the other) in response to the onset of top-heavy and bottom-heavy trapezoid stimuli (see Figure 4). Each test trial started with the participant's hand in the neutral starting position, and with a rectangular warning stimulus on the screen. Thereafter participants were required to make the pre-specified response as soon as a trapezoid stimulus appeared and then return to the neutral position ready for the next trial. The rectangle-trapezoid display sequence gave rise to apparent motion, either of the top or bottom of the rectangle moving forwards, depending on the trapezoid presented. The trapezoid stimulus was present for 480 ms until replaced by a blank display for 3,000 ms, prior to the warning stimulus for the next trial. SOA was varied randomly between 800 ms and 1,500 ms in 50 ms increments. Twenty catch trials were included in each test block. On catch trials, the warning stimulus was presented for 1,980 ms before the 3,000 ms inter-trial interval. The order in which participants completed the

open and close blocks was counterbalanced across groups. The test procedure was therefore identical to that employed in Experiment 1.

Figure 4

The training procedure again took the form of a choice RT task (see Figure 4). Each trial required the participant to make either an open response to a top-forwards stimulus or a close response to a bottom forwards stimulus. To ensure that the design and analyses were comparable with those of Experiment 1, this arbitrary S-R mapping was not counterbalanced. The order of stimulus presentation during the training trials was randomized. The beginning of each trial was indicated by the appearance of the rectangle warning stimulus, which was then replaced by one of the trapezoid stimuli, presented for 480ms. SOA was varied randomly between 800 ms and 1,500 ms in 50 ms increments. Following the offset of the trapezoid stimulus, the screen went blank for 3,000 ms, until the warning stimulus for the subsequent trial was presented. Each training block comprised 144 training trials. Half of the participants completed the training task in the red context and half in the blue context.

Results and Discussion

Training and test trials in which participants made incorrect responses, no response, or where the point of movement onset was equivocal were excluded from all further analyses (3.4% of training trials and 4.5% of test trials). Thereafter, any remaining data points beyond 2.5 standard deviations of a participant's mean response latency for a given block were also excluded (3.0% of training trials and 3.3% of test trials). EMG signals on catch trials were examined to ensure that participants obeyed task instructions, but were excluded from all further analyses. During the test sessions, participants initiated movements on only 5.7% of catch trials, indicating that they were not making anticipatory responses on the stimulus trials.

Training

Figure 5 shows mean RTs for the groups trained in red and blue contexts. A mixed-model ANOVA with training block (B1-B6) as a within-subjects factor and training group (train-

red, train-blue) as a between-subjects factor, revealed a highly significant linear decline across the six training blocks [$F(1,14) = 14.19$; $p < .01$; $\eta^2 = .50$] suggestive of learning. The group x linear trend interaction failed to reach significance [$F(1, 14) = 3.47$; $p > .08$; $\eta^2 = .19$] indicating that the linear decline was broadly comparable across groups. Trend analysis also revealed a significant quadratic trend [$F(1,14) = 14.07$; $p < .01$; $\eta^2 = .50$] however, this did not vary as a function of group [$F(1, 14) = .66$; $p > .40$; $\eta^2 = .05$]. No main effect of group was present [$F(1, 14) = .03$; $p > .80$; $\eta^2 = .00$], implying that overall RTs during the training task were also broadly equivalent. Simple effects analysis revealed that the difference between the groups at Day 1, Block 1 was not significant [$t(14) = .49$; $p > .60$] indicating that baseline performance was comparable.

Figure 5

Tests

Figure 6 shows the mean RTs observed at pre- and post-test, in the trained and untrained contexts, on compatible and incompatible trials. Figure 6 also depicts the mean compatibility effects seen on the four tests (pre-test in the trained context; pre-test in the untrained context; post-test in the trained-context; post-test in the untrained context). Because the trained S-R mappings were open to top forwards and close to bottom forwards, the size of each SRC effect was calculated by subtracting the mean RT on compatible trials (open to top forwards; close to bottom forwards) from the mean RT on incompatible trials (open to bottom forwards; close to top-forwards). The greater the RT difference, the stronger the participant's tendency to execute the trained response in the presence of the test stimuli. These SRC effects are conceptually equivalent to the automatic imitation effects calculated in Experiment 1. Performance is further broken down for the groups trained in the red and blue contexts in Table 2.

Table 2

The RT data were analyzed using a mixed-model ANOVA with test (pre-test, post-test), context (trained, untrained) and stimulus-response compatibility (compatible, incompatible) as within-subjects factors and group (train-red, train-blue) as a between-subjects factor. The

analysis revealed a significant main effect of stimulus-response compatibility [$F(1,14) = 12.60$; $p < .01$; $\eta^2 = .47$] whereby participants were faster to make responses compatible with the trained mapping (M = 280.8 ms; SD = 52.3 ms) than incompatible responses (M = 284.1 ms; SD = 51.9 ms) when collapsed across test. In addition, a significant test x compatibility interaction was observed [$F(1,14) = 9.37$; $p < .01$; $\eta^2 = .40$] indicating that stimulus-response compatibility effects were larger at post-test (M = 7.08 ms; SD = 6.62 ms) than at pre-test (M = -0.61 ms; SD = 5.43 ms). Crucially, this test x compatibility interaction did not vary as a function of context [$F(1, 14) = .27$; $p > .60$; $\eta^2 = .02$] indicating that the change in stimulus-response compatibility was equivalent in trained and untrained contexts. Moreover, the test x compatibility x context interaction was comparable in both groups [$F(1, 14) = .18$; $p > .65$; $\eta^2 = .01$]. No other main effects or interactions approached significance ($p > .25$).

Figure 6

In Experiment 1 counter-mirror learning was context-specific; the associations acquired manifested more strongly in the trained than in the untrained context. In contrast, the results from Experiment 2 show that novel S-R learning generalizes to untrained contexts. Given that identical designs were used in the two experiments, this contrast indicates that the context-specificity observed in Experiment 1 was not an artefact of procedure. Rather, the results of Experiment 2 accord with the claim that, in Experiment 1, context-specificity was generated by a conflict between second-learned counter-mirror associations and first-learned associations established during ontogeny.

GENERAL DISCUSSION:

The present study tested the hypothesis, advanced by the ASL model, that when the sight of an action is associated with rival mirror and counter-mirror responses, this ambiguity is resolved through contextual modulation – the same process thought to resolve ambiguity resulting from conflicting associations in counter-conditioning experiments (Bouton, 1993, 1994; Nelson, 2002). The results were affirmative. When open and close hand stimuli were rendered ambiguous through periods of counter-mirror training, the second-learned counter-mirror associations manifested more strongly in the trained context, resulting in

less automatic imitation. Our second experiment confirmed that this context-specificity was due to conflict between second-learned counter-mirror associations and first-learned mirror associations established during the course of normal, pre-experimental development by demonstrating that analogous first-learned arbitrary sensorimotor associations do generalize across our context manipulation.

Contextual modulation of automatic imitation

Several previous findings indicate that automatic imitation may be modified through sensorimotor training (Cook et al., 2010; Gillmeister et al., 2008; Heyes et al., 2005; Press, Gillmeister, & Heyes, 2007). Crucially, however, the present data provide further evidence that this learning is associative; that it is mediated by the same processes responsible for Pavlovian and instrumental conditioning in humans and animals. Counter-conditioning and renewal effects demonstrate that second-learned associations show greater context specificity than the first associations formed with a novel stimulus (Bouton, 1994; Bouton & King, 1983; Nelson, 2002; Nelson et al., 2011; Peck & Bouton, 1990). That the sensorimotor associations established during counter-mirror learning are modulated by context argues that counter-mirror associations are equivalent to second-learned associations acquired during counter-conditioning. This finding suggests a clear parallel between mirror associations acquired during development and first-learned associations established during conditioning. In this respect the present data accord with evidence that counter-mirror learning is sensitive to the contingency between the sensory and motor representations; the extent to which action observation predicts the incompatible response (Cook et al., 2010). Insofar as contextual modulation and contingency sensitivity are both predictions derived from conditioning experiments, these effects support the view that the acquisition of conceptually equivalent mirror and counter-mirror sensorimotor links is mediated by phylogenetically ancient mechanisms of associative learning (Heyes, 2001, 2010a, 2010b; Ray & Heyes, 2011).

It might be suggested that counter-mirror training reduces automatic imitation, not by establishing new sensorimotor associations, but by encouraging participants to use an intentional, ‘do opposite’ response strategy (Heyes, 2011). The context-specificity observed in Experiment 1 suggests that this is not the case. One of the defining characteristics of

intentional response strategies is their flexibility. For example, that incompatible stimulus-response training transfers from one spatial task to another, has been cited as evidence that learning is strategic (Vu, 2007). In contrast, the context-specificity of counter-mirror learning shows that it is relatively inflexible.

Associative accounts argue that the context specificity of second-learned associations emerges as a means to resolve ambiguity (Bouton, 1993, 1994; Nelson, 2002). Rather than over-write or unlearn existing associations, there is considerable evidence that extinction and counter-conditioning establish additional parallel associations (Bouton & Peck, 1992; Brooks et al., 1995). However, being associated with two opposing responses renders the stimulus ambiguous. Because of the conflict between the first- and second-learned associations, it is not clear what the appropriate behaviour is in the presence of the stimulus. To resolve this ambiguity, the excitability of the first- and second-learned associations may be modulated by the context. Specifically, counter-conditioning may create associations that inhibit the original response, as well as excite the second-learned response, both of which may be selectively activated by the counter-conditioning context. When applied to the present paradigm this account argues that counter-mirror training sets-up second-learned associations which inhibit the mirror response and excite the appropriate counter-mirror response. Both inhibitory and excitatory counter-mirror associations are selectively 'turned-on' by the contextual cues present in the training environment. When participants leave the training context the newly acquired associations become less excitable.

The contrasting behaviour of first- and second-learned associations suggests that the temporal order in which we encounter sensorimotor contingencies is an important factor in determining which responses generalise to novel contexts. In particular, the nature of early sensorimotor experience may be crucial. Our results suggest that whichever associations are acquired first, be they mirror or counter-mirror, will generalise readily across contexts, whereas associations acquired subsequently will show greater context-specificity. Parents and caregivers imitate newborns (Malatesta & Haviland, 1982; Pawlby, 1977), and these interactions are likely to provide some of the earliest experience of contingencies between action execution (e.g. neonate executes smile) and action observation (e.g. neonate

observes caregiver smiling) (Ray & Heyes, 2011). The early acquisition of mirror associations may cause 'mirroring' to become the default, context-general response which manifests in novel situations. If, however, newborns were first exposed to predominantly non-matching sensorimotor contingencies, automatic counter-mirror or complementary responses might become prepotent in novel contexts.

In the present study context was manipulated by changing visual (lighting, background colours) and tactile variables (texture of the arm-rest). However, in addition to physical variables 'context' may also encompass internal states. If extinction of a conditioned response is conducted while participants are receiving tranquilisers (Bouton, Kenney, & Rosengard, 1990) or alcohol (Cunningham, 1979), learning is found to be state-dependent. A renewal effect is observed whereby the conditioned response is renewed when the effects of the drug manipulations wear off. Several non-matching sensorimotor contingencies may also be modulated by the context provided by internal state. For example, sports provide a wealth of non-matching contingencies. In tennis, the execution of a serve is predictive of the sight of a return, and in boxing the execution of jab predicts the sight of an opponent raising his guard. However, while sportsmen and -women experience these contingencies, levels of cortisol and adrenalin, heart-rate and body temperature are frequently elevated. These internal contexts may serve to modulate conflicting responses, by changing the relative excitability of mirror and counter-mirror associations in different situations.

Insofar as the present study did not include neurophysiological methods, we cannot specify precisely where in the brain these learning effects occurred. One possibility is that the learning observed reflects the acquisition of arbitrary stimulus-response associations thought to be represented within dorsal premotor cortex (Passingham, 1993; Wise & Murray, 2000). However, this account is not compelling because it implies that the action stimuli were perceived, not as actions, but as arbitrary visual stimuli. The robust automatic imitation effects seen at pre-test indicate that the action stimuli were sufficiently life-like to excite previously learned sensorimotor associations. Rather than reflecting associations acquired with arbitrary stimuli, these data demonstrate that sensory representations of actions came to excite non-congruent motor representations, as a result of counter-mirror training. Based on a purely functional definition, these counter-mirror training effects are

therefore likely to be mediated by MNs (i.e. units that respond to both the sight and execution of an action). Any alternative account would appear to require the existence of a new class of motor neuron, one which also responds to action observation, but is somehow qualitatively distinct from MNs.

Implications for the functioning of the 'mirror neuron system'

Assuming (i) that automatic imitation is a behavioural product of congruent MNs; and (ii) that counter-mirror learning occurs within the MNS, our findings may have important implications for the functioning of this system. In the present study, the open and close hand stimuli were rendered ambiguous by periods of counter-mirror training. However, there is evidence that sensorimotor ambiguity may be a typical feature of the MNS, and not merely a product of laboratory S-R training. Substantial proportions of macaque MNs (di Pellegrino et al., 1992; Ferrari et al., 2003; Gallese et al., 1996; Rozzi et al., 2008; Umilta et al., 2001) and human sensorimotor units (Mukamel et al., 2010) discharge indiscriminately during the observation of multiple actions, or selectively during the observation and execution of different actions. The resulting sensorimotor ambiguity is challenging for theories which posit that MNs make a functional contribution to action understanding (Rizzolatti & Craighero, 2004; Rizzolatti et al., 1996), imitation (Heyes, 2011; Iacoboni, 2009), or generic action selection (Hickok & Hauser, 2010). Because the sight of an action excites multiple motor representations, it is unclear which action the observer should 'understand' or 'select'. However, modulation of MN responses by context might help to explain how the MNS can contribute to coherent behaviour in spite of the conflicting responses of single MNs; for example, how observation of the same action might trigger an imitative response in one context, but a complementary or counter-mirror response in another.

Much of the contingent sensorimotor experience we receive in everyday life involves matching; the execution of an action predicts the observation of the same action. However, we are also exposed to non-matching sensorimotor contingencies, where the execution of one action predicts the observation a different action. During coordinated instrumental action, the sight of another agent releasing an object reliably predicts the performance of a grasping action (Newman-Norlund, van Schie, van Zuijlen, & Bekkering, 2007; van Schie,

van Waterschoot, & Bekkering, 2008). Similarly, in social control situations, involving dominance and submission, expansive gestures by one interactant predict contraction movements by the other, and vice versa (Tiedens & Fragale, 2003). The contexts in which fMRI and single-cell recording takes place are necessarily very different from the contexts in which we receive non-matching experience. However, where non-matching contingencies represent second-learned associations, neural populations coding such relationships may only be evident in the presence of specific contextual cues. Taking steps to invoke learning contexts during neuroimaging and single-cell procedures may therefore reveal greater numbers of non-congruent MNs.

Contextual modulation may also help explain why MNs appear to code abstract 'motor goals'. MNs respond differently to the sight of an experimenter grasping a food item depending on whether the goal of the action is to eat the food or to place it in a shoulder-mounted cup (Bonini et al., 2010; Fogassi et al., 2005). That MN responses differed before the observed action trajectories could be differentiated is taken as evidence that MNs help us understand the motor intentions of others. However, the present data demonstrate that a simple associative account can be advanced to explain these data, without postulating any role of MNs in higher-order action understanding. Prior to the experiment, it is reasonable to assume that the macaques had ample opportunity to acquire S-R associations between the sight of grasping and grasp-to-eat motor programs. During the experiment the animals acquired further S-R associations such that the sight of grasping could also excite grasp-to-place motor programs. Because these second-learned associations served to render the sight of grasping ambiguous they became subject to contextual control. Crucially, whenever the monkeys were required to grasp-to-place, the target cup was always present, either in front of the macaque or on their shoulder. The second-learned grasp-to-place associations are likely to have been modulated by the presence of this salient contextual cue. When the macaques subsequently observed grasping, either in the presence or absence of the cup, MNs were either excited or inhibited by this context, thus giving the impression that they were coding action intention.

In summary, while counter-mirror learning shows context-specificity, the acquisition of arbitrary S-R associations generalizes well to untrained contexts. Associative accounts

argue that counter-mirror learning comes under contextual control because it establishes second-learned associations which render the sight of actions ambiguous. In contrast, novel S-R associations generalise to untrained contexts because there is no conflict with previous learning. That the sensorimotor associations established during counter-mirror learning are modulated by context indicates that counter-mirror associations are equivalent to second-learned associations acquired during counter-conditioning experiments. Our results therefore suggest a parallel between mirror associations acquired during development and first-learned excitatory associations established during conditioning. We have argued that contextual modulation emerges to resolve the conflict caused by counter-mirror training – a laboratory treatment that participants are unlikely to encounter in their daily lives. However, neurophysiological studies indicate that ambiguity may be a typical feature of both the human and macaque MNS. The findings presented here suggest that contextual modulation may play a key role in ensuring that that the MNS contributes to coherent patterns of behaviour in spite of the conflicting responses of individual MNs.

REFERENCES:

Arnstein, D., Cui, F., Keysers, C., Maurits, N. M., & Gazzola, V. (2011). mu-Suppression during action observation and execution correlates with BOLD in dorsal premotor, inferior parietal, and SI cortices. *Journal of Neuroscience, 31*(40), 14243-14249.

Bien, N., Roebroeck, A., Goebel, R., & Sack, A. T. (2009). The brain's intention to imitate: the neurobiology of intentional versus automatic imitation. *Cerebral Cortex, 19*(10), 2338-2351.

Blakemore, S. J., & Frith, C. (2005). The role of motor contagion in the prediction of action. *Neuropsychologia, 43*(2), 260-267.

Bonini, L., Rozzi, S., Serventi, F. U., Simone, L., Ferrari, P. F., & Fogassi, L. (2010). Ventral premotor and inferior parietal cortices make distinct contribution to action organization and intention understanding. *Cerebral Cortex, 20*(6), 1372-1385.

Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin, 114*(1), 80-99.

Bouton, M. E. (1994). Context, ambiguity, and classical-conditioning. *Current Directions in Psychological Science, 3*(2), 49-53.

Bouton, M. E., Kenney, F. A., & Rosengard, C. (1990). State-dependent fear extinction with two benzodiazepine tranquilizers. *Behavioral Neuroscience, 104*(1), 44-55.

Bouton, M. E., & King, D. A. (1983). Contextual control of the extinction of conditioned fear: tests for the associative value of the context. *Journal of Experimental Psychology-Animal Behavior Processes, 9*(3), 248-265.

Bouton, M. E., & Peck, C. A. (1992). Spontaneous-recovery in cross-motivational transfer (Counterconditioning). *Animal Learning & Behavior, 20*(4), 313-321.

Brass, M., Bekkering, H., & Prinz, W. (2001). Movement observation affects movement execution in a simple response task. *Acta Psychologica, 106*(1-2), 3-22.

Brass, M., & Heyes, C. (2005). Imitation: is cognitive neuroscience solving the correspondence problem? *Trends in Cognitive Sciences, 9*(10), 489-495.

Brooks, D. C., Hale, B., Nelson, J. B., & Bouton, M. E. (1995). Reinstatement after counterconditioning. *Animal Learning & Behavior, 23*(4), 383-390.

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*(2), 400-404.

Calvo-Merino, B., Glaser, D. E., Grezes, J., Passingham, R. E., & Haggard, P. (2005). Action observation and acquired motor skills: An fMRI study with expert dancers. *Cerebral Cortex, 15*(8), 1243-1249.

Calvo-Merino, B., Grezes, J., Glaser, D. E., Passingham, R. E., & Haggard, P. (2006). Seeing or doing? Influence of visual and motor familiarity in action observation. *Current Biology, 16*(19), 1905-1910.

Catmur, C., Gillmeister, H., Bird, G., Liepelt, R., Brass, M., & Heyes, C. (2008). Through the looking glass: counter-mirror activation following incompatible sensorimotor learning. *European Journal of Neuroscience, 28*(6), 1208-1215.

Catmur, C., & Heyes, C. (2010). Time course analyses confirm independence of imitative and spatial compatibility. *Journal of Experimental Psychology: Human Perception and Performance*.

Catmur, C., Mars, R. B., Rushworth, M. F., & Heyes, C. (2010). Making mirrors: Premotor cortex stimulation enhances mirror and counter-mirror motor facilitation. *Journal of Cognitive Neuroscience*.

Catmur, C., Walsh, V., & Heyes, C. (2007). Sensorimotor learning configures the human mirror system. *Current Biology, 17*(17), 1527-1531.

Catmur, C., Walsh, V., & Heyes, C. (2009). Associative sequence learning: the role of experience in the development of imitation and the mirror system. *Philosophical*

Transactions of the Royal Society of London Series B: Biological Sciences 364(1528), 2369-2380.

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: the perception-behavior link and social interaction. *Journal of Personality and Social Psychology, 76*(6), 893-910.

Chong, T. T. J., Cunnington, R., Williams, M. A., Kanwisher, N., & Mattingley, J. B. (2008). fMRI adaptation reveals mirror neurons in human inferior parietal cortex. *Current Biology, 18*(20), 1576-1580.

Cook, R., Bird, G., Lunser, G., Huck, S., & Heyes, C. (2011). Automatic imitation in a strategic context: players of rock-paper-scissors imitate opponents' gestures. *Proceedings of the Royal Society B-Biological Sciences*.

Cook, R., Press, C., Dickinson, A., & Heyes, C. (2010). Acquisition of automatic imitation is sensitive to sensorimotor contingency. *Journal of Experimental Psychology: Human Perception and Performance, 36*(4), 840-852.

Cooper, R. P., Catmur, C., & Heyes, C. (under review). Are automatic imitation and spatial compatability mediated by different processes?

Craighero, L., Bello, A., Fadiga, L., & Rizzolatti, G. (2002). Hand action preparation influences the responses to hand pictures. *Neuropsychologia, 40*(5), 492-502.

Cunningham, C. L. (1979). Alcohol as a cue for extinction - state dependency produced by conditioned inhibition. *Animal Learning & Behavior, 7*(1), 45-52.

di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Experimental Brain Research, 91*(1), 176-180.

Dimberg, U., Thunberg, M., & Elmehed, K. (2000). Unconscious facial reactions to emotional facial expressions. *Psychological Science, 11*(1), 86-89.

Dinstein, I., Gardner, J. L., Jazayeri, M., & Heeger, D. J. (2008). Executed and observed movements have different distributed representations in human aIPS. *Journal of Neuroscience, 28*(44), 11231-11239.

Dinstein, I., Hasson, U., Rubin, N., & Heeger, D. J. (2007). Brain areas selective for both observed and executed movements. *Journal of Neurophysiology, 98*(3), 1415-1427.

Dinstein, I., Thomas, C., Behrmann, M., & Heeger, D. J. (2008). A mirror up to nature. *Current Biology, 18*(1), R13-18.

Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: a magnetic stimulation study. *Journal of Neurophysiology, 73*(6), 2608-2611.

Ferrari, P. F., Bonini, L., & Fogassi, L. (2009). From monkey mirror neurons to primate behaviours: possible 'direct' and 'indirect' pathways. *Philosophical Transactions of the Royal Society B-Biological Sciences, 364*(1528), 2311-2323.

Ferrari, P. F., Gallese, V., Rizzolatti, G., & Fogassi, L. (2003). Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *European Journal of Neuroscience, 17*(8), 1703-1714.

Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., & Rizzolatti, G. (2005). Parietal lobe: from action organization to intention understanding. *Science, 308*(5722), 662-667.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain, 119 (Pt 2),* 593-609.

Gazzola, V., Rizzolatti, G., Wicker, B., & Keysers, C. (2007). The anthropomorphic brain: the mirror neuron system responds to human and robotic actions. *Neuroimage, 35*(4), 1674-1684.

Gillmeister, H., Catmur, C., Liepelt, R., Brass, M., & Heyes, C. (2008). Experience-based priming of body parts: A study of action imitation. *Brain Research, 1217*, 157-170.

Grezes, J., Armony, J. L., Rowe, J., & Passingham, R. E. (2003). Activations related to "mirror" and "canonical" neurones in the human brain: an fMRI study. *Neuroimage, 18*(4), 928-937.

Grezes, J., & Morin, O. (2008). What is "mirror" in the premotor cortex? A review *Clinical Neurophysiology, 38*, 189-195.

Haslinger, B., Erhard, P., Altenmuller, E., Schroeder, U., Boecker, H., & Ceballos-Baumann, A. O. (2005). Transmodal sensorimotor networks during action observation in professional pianists. *Journal of Cognitive Neuroscience, 17*(2), 282-293.

Heiser, M., Iacoboni, M., Maeda, F., Marcus, J., & Mazziotta, J. C. (2003). The essential role of Broca's area in imitation. *European Journal of Neuroscience, 17*(5), 1123-1128.

Heyes, C. (2001). Causes and consequences of imitation. *Trends in Cognitive Sciences, 5*(6), 253-261.

Heyes, C. (2010a). Mesmerising mirror neurons. *Neuroimage, 51*(2), 789-791.

Heyes, C. (2010b). Where do mirror neurons come from? *Neuroscience and Biobehavioral Reviews, 34*(4), 575-583.

Heyes, C. (2011). Automatic imitation. *Psychological Bulletin, 137*(3), 463-483.

Heyes, C., Bird, G., Johnson, H., & Haggard, P. (2005). Experience modulates automatic imitation. *Cognitive Brain Research, 22*(2), 233-240.

Hickok, G., & Hauser, M. (2010). (Mis)understanding mirror neurons. *Current Biology, 20*(14), R593-594.

Iacoboni, M. (2009). Imitation, empathy, and mirror neurons. *Annual Review of Psychology, 60*, 653-670.

Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science, 286*(5449), 2526-2528.

Keysers, C., & Gazzola, V. (2010). Social neuroscience: Mirror neurons recorded in humans. *Current Biology, 20*(8), R353-R354.

Kilner, J. M., Neal, A., Weiskopf, N., Friston, K. J., & Frith, C. D. (2009). Evidence of mirror neurons in human inferior frontal gyrus. *Journal of Neuroscience, 29*(32), 10153-10159.

Leighton, J., & Heyes, C. (2010). Hand to mouth: automatic imitation across effector systems. *Journal of Experimental Psychology-Human Perception and Performance, 36*(5), 1174-1183.

Lingnau, A., Gesierich, B., & Caramazza, A. (2009). Asymmetric fMRI adaptation reveals no evidence for mirror neurons in humans. *Proceedings of the National Academy of Sciences of the United States of America, 106*(24), 9925-9930.

Longo, M. R., Kosobud, A., & Bertenthal, B. I. (2008). Automatic imitation of biomechanically possible and impossible actions: effects of priming movements versus goals. *Journal of Experimental Psychology: Human Perception and Performance, 34*(2), 489-501.

Malatesta, C. Z., & Haviland, J. M. (1982). Learning display rules: The socialization of emotion expression in infancy. *Child Development, 53*(4), 991-1003.

Molenberghs, P., Cunnington, R., & Mattingley, J. B. (2009). Is the mirror neuron system involved in imitation? A short review and meta-analysis. *Neuroscience and Biobehavioral Reviews, 33*, 975-980.

Molenberghs, P., Cunnington, R., & Mattingley, J. B. (2011). Brain regions with mirror properties: A meta-analysis of 125 human fMRI studies. *Neuroscience and Biobehavioral Reviews, 36*(1), 341-349.

Mukamel, R., Ekstrom, A. D., Kaplan, J., Iacoboni, M., & Fried, I. (2010). Single-neuron responses in humans during execution and observation of actions. *Current Biology, 20*(8), 750-756.

Nelson, J. B. (2002). Context specificity of excitation and inhibition in ambiguous stimuli. *Learning and Motivation, 33*(2), 284-310.

Nelson, J. B., Sanjuan Mdel, C., Vadillo-Ruiz, S., Perez, J., & Leon, S. P. (2011). Experimental renewal in human participants. *Journal of Experimental Psychology-Animal Behavior Processes, 37*(1), 58-70.

Newman-Norlund, R. D., Ondobaka, S., van Schie, H. T., van Elswijk, G., & Bekkering, H. (2010). Virtual lesions of the IFG abolish response facilitation for biological and non-biological cues. *Frontiers in Behavioral Neuroscience 4*, 5.

Newman-Norlund, R. D., van Schie, H. T., van Zuijlen, A. M. J., & Bekkering, H. (2007). The mirror neuron system is more active during complementary compared with imitative action. *Nature Neuroscience, 10*(7), 817-818.

Passingham, R. E. (1993). *The frontal lobes and voluntary action*. Oxford: Oxford University Press.

Pawlby, S. J. (1977). Imitative interaction. In H. Schaffer (Ed.), *Studies in mother-infant interaction* (pp. 203-224). New York: Academic Press.

Peck, C. A., & Bouton, M. E. (1990). Context and performance in aversive-to-appetitive and appetitive-to-Aversive Transfer. *Learning and Motivation, 21*(1), 1-31.

Press, C., Bird, G., Walsh, E., & Heyes, C. (2008). Automatic imitation of intransitive actions. *Brain and Cognition, 67*(1), 44-50.

Press, C., Gillmeister, H., & Heyes, C. (2007). Sensorimotor experience enhances automatic imitation of robotic action. *Proceedings of the Royal Society B-Biological Sciences, 274*(1625), 2509-2514.

Ray, E., & Heyes, C. (2011). Imitation in infancy: the wealth of the stimulus. *Developmental Science, 14*(1), 92-105.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience, 27*, 169-192.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research, 3*(2), 131-141.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience, 2*(9), 661-670.

Rozzi, S., Ferrari, P. F., Bonini, L., Rizzolatti, G., & Fogassi, L. (2008). Functional organization of inferior parietal lobule convexity in the macaque monkey: electrophysiological characterization of motor, sensory and mirror responses and their correlation with cytoarchitectonic areas. *European Journal of Neuroscience, 28*(8), 1569-1588.

Strafella, A. P., & Paus, T. (2000). Modulation of cortical excitability during action observation: a transcranial magnetic stimulation study. *Neuroreport, 11*(10), 2289-2292.

Sturmer, B., Aschersleben, G., & Prinz, W. (2000). Correspondence effects with manual gestures and postures: a study of imitation. *Journal of Experimental Psychology: Human Perception and Performance, 26*(6), 1746-1759.

Tiedens, L. Z., & Fragale, A. R. (2003). Power moves: Complementarity in dominant and submissive nonverbal behavior. *Journal of Personality and Social Psychology, 84*(3), 558-568.

Umilta, M. A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., et al. (2001). I know what you are doing: A neurophysiological study. *Neuron, 31*(1), 155-165.

Van Gucht, D., Vansteenwegen, D., Beckers, T., & Van den Bergh, O. (2008). Return of experimentally induced chocolate craving after extinction in a different context: divergence between craving for and expecting to eat chocolate. *Behaviour Research and Therapy, 46*(3), 375-391.

van Schie, H. T., van Waterschoot, B. M., & Bekkering, H. (2008). Understanding action beyond imitation: reversed compatibility effects of action observation in imitation

and joint action. *Journal of Experimental Psychology: Human Perception and Performance, 34*(6), 1493-1500.

Vansteenwegen, D., Hermans, D., Vervliet, B., Francken, G., Beckers, T., Baeyens, F., et al. (2005). Return of fear in a human differential conditioning paradigm caused by a return to the original acquistion context. *Behaviour Research and Therapy, 43*(3), 323-336.

Vu, K. P. (2007). Influences on the Simon effect of prior practice with spatially incompatible mappings: transfer within and between horizontal and vertical dimensions. *Memory & Cognition, 35*(6), 1463-1471.

Wang, Y., Newport, R., & Hamilton, A. F. (2011). Eye contact enhances mimicry of intransitive hand movements. *Biology Letters, 7*(1), 7-10.

Wise, S. P., & Murray, E. A. (2000). Arbitrary associations between antecedents and actions. *Trends in Neurosciences, 23*(6), 271-276.

FIGURES:

Figure 1:

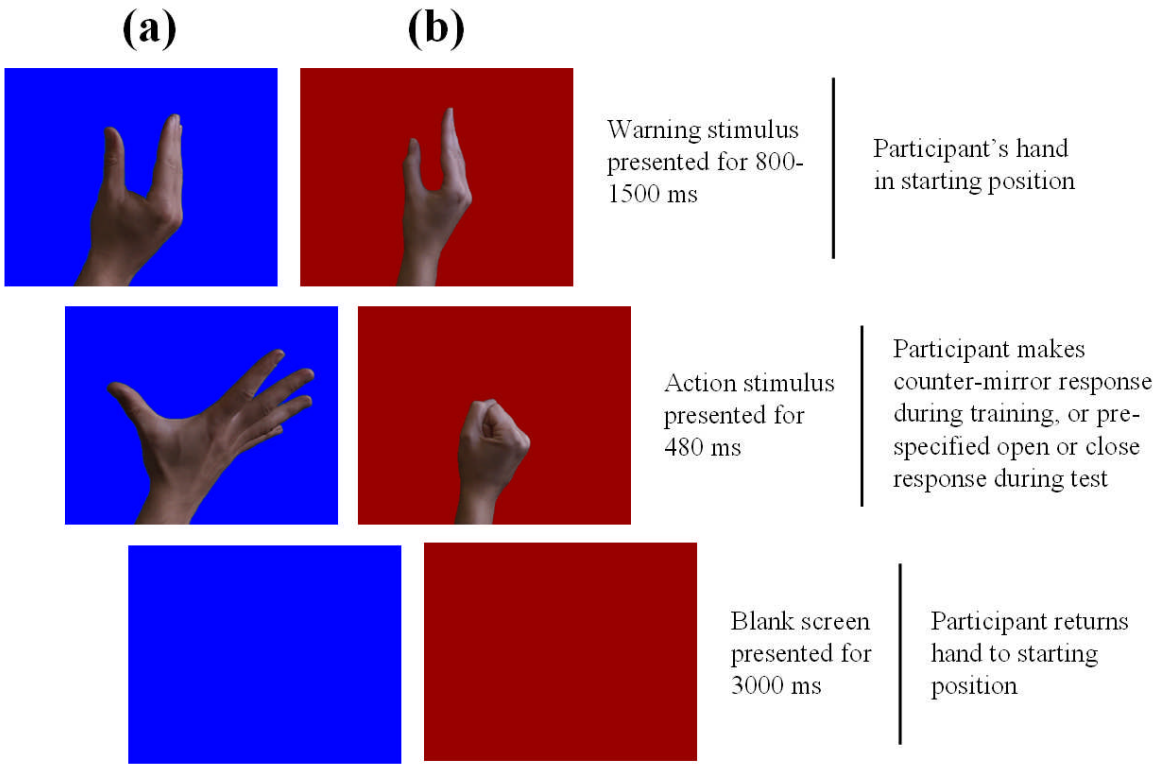

Figure 1: Display sequences for (a) male-hand open-stimulus trials in the blue context and (b) female-hand close-stimulus trials in the red context. During training participants made the counter-mirror response to the onset of the action stimulus (open responses to close stimuli; close responses to open stimuli). On test trials participants made pre-specified responses to the onset of the action stimulus; open-hand responses in one block, close-hand responses in the other.

Figure 2:

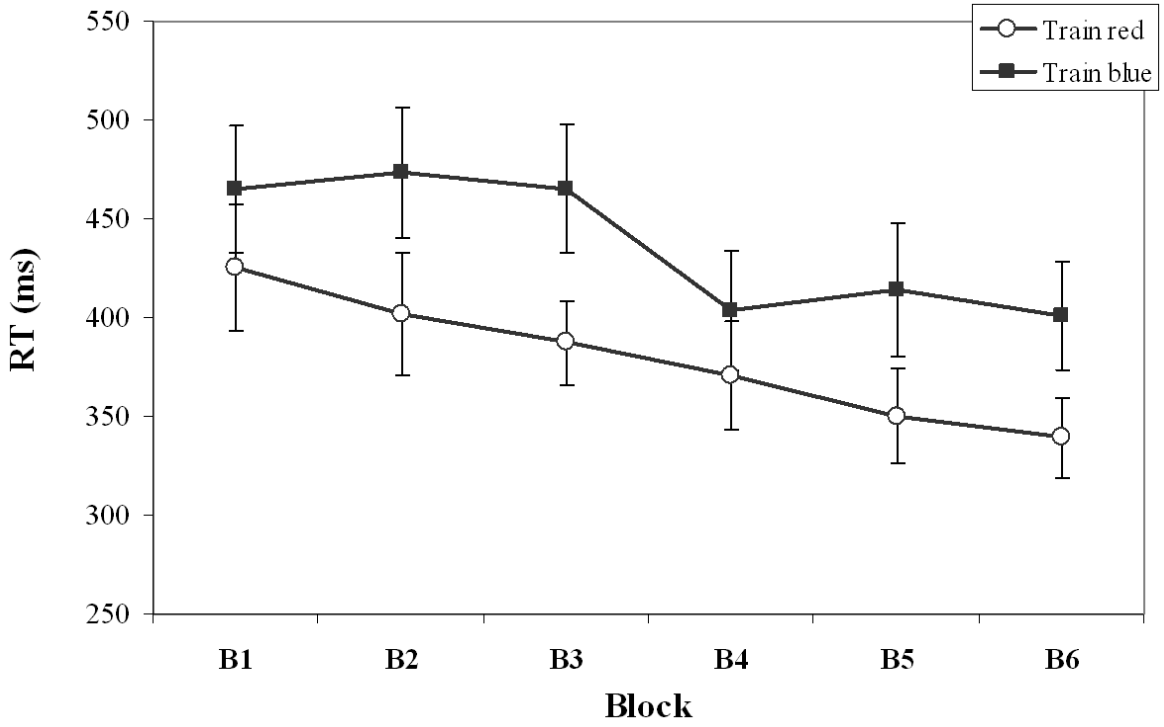

Figure 2: Mean RTs observed during training in Experiment 1, plotted across the six training blocks. Open circles represent the performance of the group trained in the red context, and filled squares represent the group trained in the blue context. Error bars represent standard error of the mean. RT = reaction time; B = block.

Figure 3:

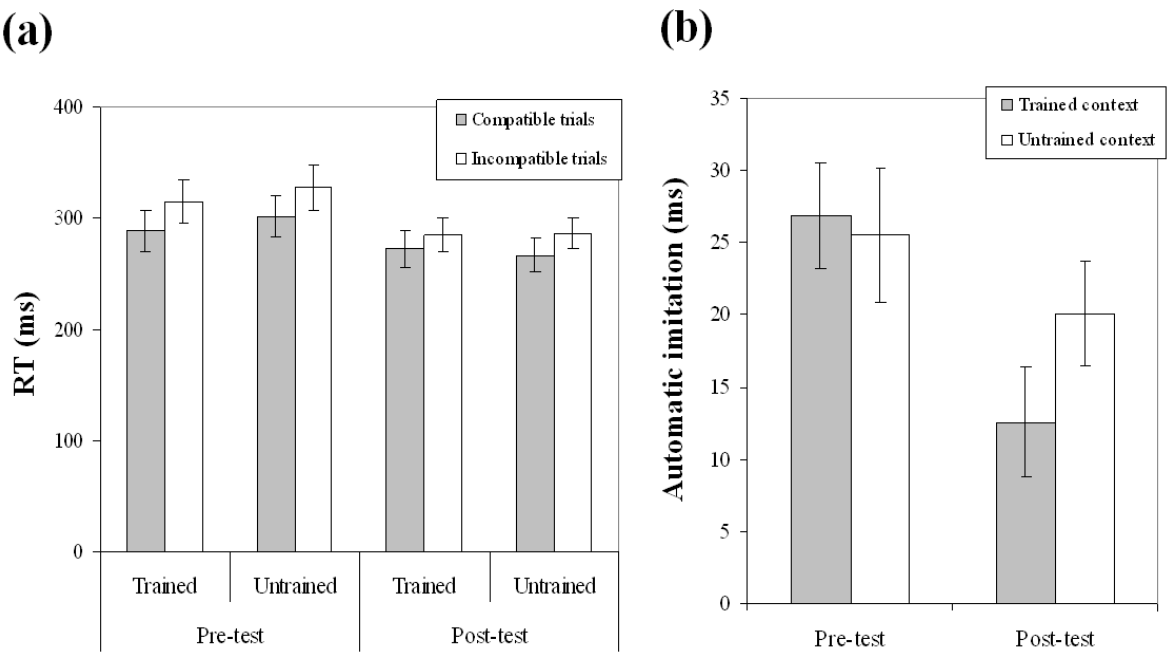

Figure 3: (a) Mean RTs observed at pre-test and post-test, in the trained and untrained contexts, on compatible and incompatible trials and (b) automatic imitation effects observed at pre-test and post-test in the trained and untrained contexts, calculated by subtracting mean RTs on compatible trials from mean RTs on incompatible trials. Error bars represent the standard error of the mean. RT = reaction time.

Figure 4:

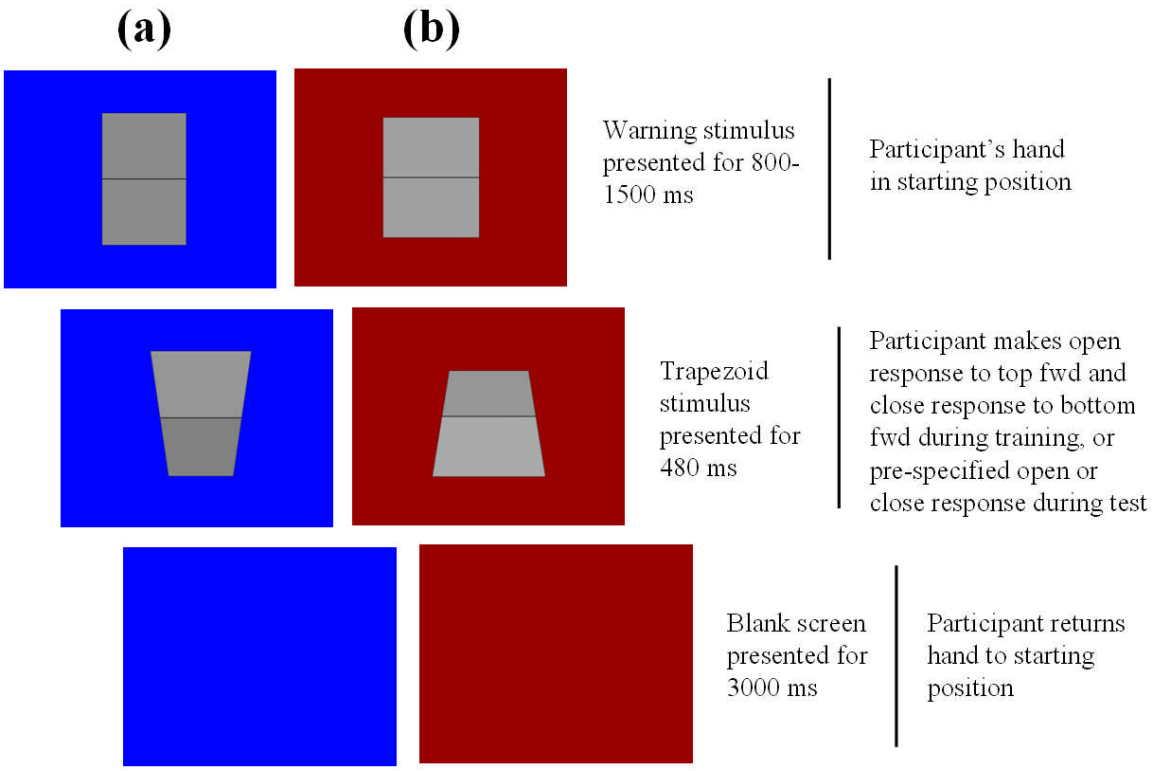

Figure 4: Display sequences for (a) larger aspect ratio top-heavy sequence in the blue context and (b) smaller aspect ratio bottom-heavy sequence in the red context. In each case the presentation of a top- or bottom-heavy trapezoid gave the appearance of the rectangle either falling towards or away from the participant. During training participants made open-hand responses when the top appeared to move forwards and close-hand responses when the bottom appeared to move forwards. On test trials participants made pre-specified responses to the onset of the trapezoid stimulus; open-hand responses in one block, close-hand in the other.

Figure 5:

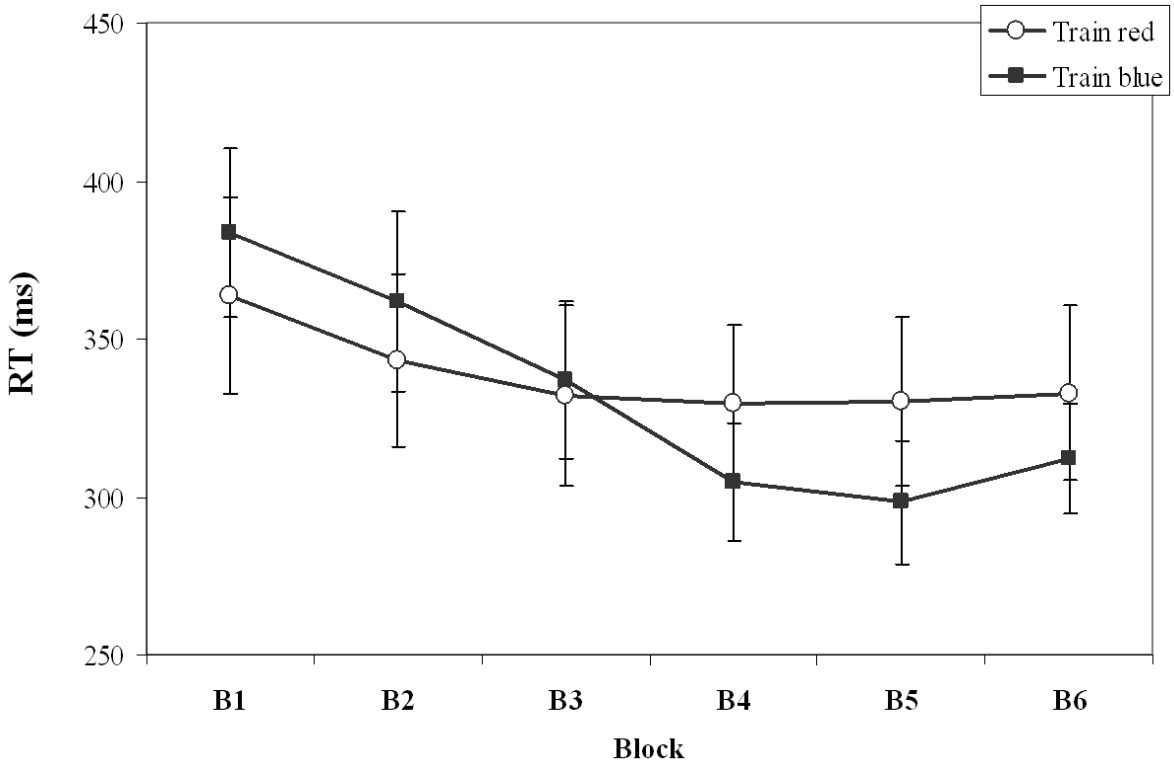

Figure 5: Mean RTs observed during training in Experiment 2, plotted across the six training blocks. Open circles represent the performance of the group trained in the red context, and filled squares represent the group trained in the blue context. Error bars represent standard error of the mean. RT = reaction time; B = block.

Figure 6:

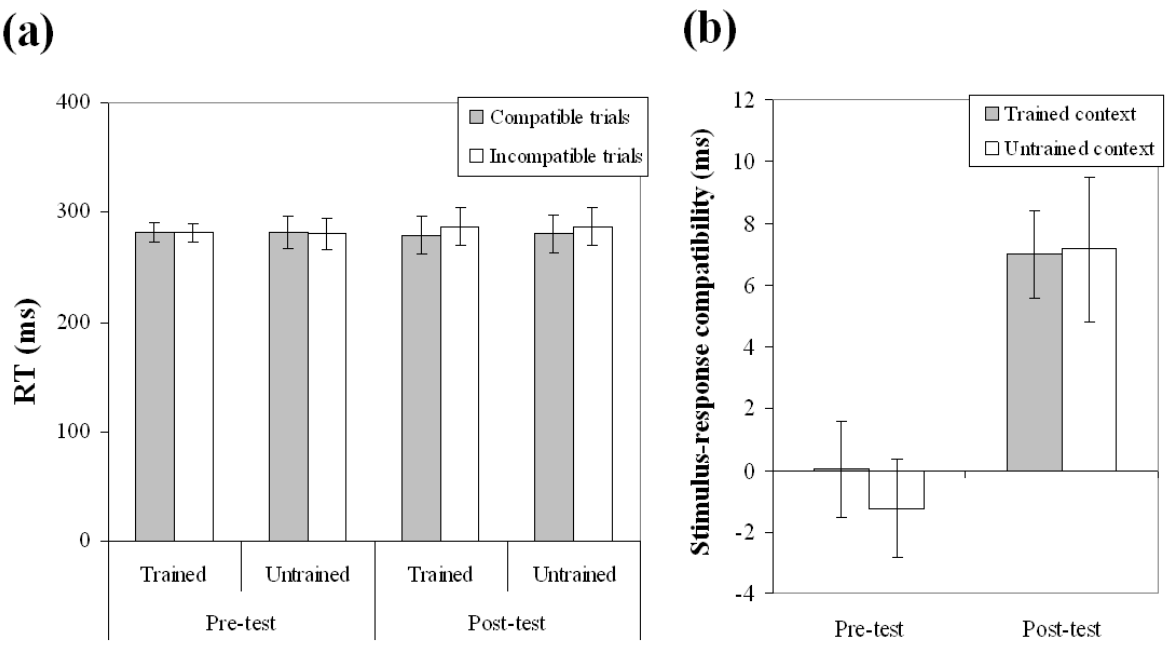

Figure 6: (a) Mean RTs observed at pre-test and post-test, in the trained and untrained contexts, on compatible and incompatible trials and (b) stimulus-response compatibility effects observed at pre-test and post-test in the trained and untrained contexts, calculated by subtracting mean RTs on compatible trials from mean RTs on incompatible trials. Error bars represent the standard error of the mean. RT = reaction time.

TABLES:

Table 1:

Table 1: Mean reaction times observed during Experiment 1 in the four tests shown for participants trained in the red and blue contexts. Standard deviations are presented in parenthesis.

	Red group		**Blue group**	
	Compatible	*Incompatible*	*Compatible*	*Incompatible*
Pre-test red	278.5 (*96.7*)	312.7 (*106.5*)	315.6 (*41.6*)	334.1 (*55.3*)
Pre-test blue	288.2 (*100.6*)	320.7 (*106.7*)	298.0 (*48.2*)	317.5 (*51.7*)
Post-test red	256.3 (*66.5*)	276.6 (*66.5*)	277.9 (*51.5*)	292.6 (*54.0*)
Post-test blue	254.9 (*62.6*)	280.3 (*62.6*)	288.2 (*62.5*)	293.0 (*57.8*)

Table 2:

Table 2: Mean reaction times observed during Experiment 2 in the four tests shown for participants trained in the red and blue contexts. Standard deviations are presented in parenthesis.

	Red group		**Blue group**	
	Compatible	*Incompatible*	*Compatible*	*Incompatible*
Pre-test red	286.2 (38.9)	284.9 (37.3)	265.4 (40.7)	264.9 (41.0)
Pre-test blue	297.3 (69.9)	295.3 (69.8)	277.4 (36.2)	278.8 (32.4)
Post-test red	297.1 (56.2)	302.9 (58.4)	258.2 (59.3)	264.9 (60.6)
Post-test blue	302.3 (72.6)	310.0 (74.1)	262.8 (78.5)	271.0 (78.3)