Title: Actions speak louder than words: comparing automatic imitation and verbal command

Authors: Helge Gillmeister, Arnaud Badets and Cecilia Heyes

University College London, London, UK

Corresponding author: Helge Gillmeister

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

Tel.: +44 207 679 5379

E-mail: h.gillmeister@ucl.ac.uk

Running head: Actions speak louder than words

Word count: 3904

Abstract

Automatic imitation – copying observed actions without intention – is known to occur, not only in neurological patients and those with developmental disorders, but also in healthy, typically-developing adults and children. Previous research has shown that a variety of actions are automatically imitated, and that automatic imitation promotes social affiliation and rapport. We assessed the power of automatic imitation by comparing it with the strength of the tendency to obey verbal commands. In a Stroop interference paradigm, the stimuli were compatible, incompatible and neutral compounds of hand postures and verbal commands. When imitative responses were required, the impact of irrelevant action images on responding to words was greater than the effect of irrelevant words on responding to actions. Control group performance showed that this asymmetry was not due to modality effects or differential salience of action and word stimuli. These results indicate that automatic imitation was more powerful than verbal command.

Introduction

Even when we do not intend to imitate others, we are inclined to copy their body movements. This tendency, known as ‘mimicry’ or ‘automatic imitation’, was once thought to be confined to patients with frontal brain damage (Lhermitte, Pillon, & Serdaru, 1986), atypically-developing individuals (e.g. Charman & Baron-Cohen, 1994), ‘savages’ (Darwin, 1989) and nonhuman animals (Thorndike, 1898). More recent research has shown that automatic imitation is also common in healthy, typically-developing adults (e.g. Wallbott, 1991; Lakin & Chartrand, 2003; Brass, Bekkering, Wohlschläger, & Prinz, 2000) and children (Simpson & Riggs, 2007). The purpose of the present study was to estimate the strength of our tendency automatically to imitate the behavior of others by comparing it with the strength of our tendency to do what we are told; to perform actions on verbal command.

Most previous research on automatic imitation has been concerned, not with the strength of this tendency, but with its pervasiveness and effects on social attitudes. Carefully controlled laboratory studies have found automatic imitation of facial expressions (e.g. Wallbott, 1991), as well as finger (e.g. Brass et al., 2000), hand (Heyes, Bird, Johnson, & Haggard, 2005) and arm movements (e.g. Kilner, Paulignan, & Blakemore, 2003). Studies investigating the ‘chameleon effect' in semi-naturalistic social situations have shown that gestures such as ear-touching and foot-wagging are automatically imitated, that this kind of mimicry can occur without the imitator's conscious awareness, and that it promotes affiliation and rapport between social partners (e.g. Lakin & Chartrand, 2003).

Indirect evidence of the pervasiveness of automatic imitation has been provided by functional imaging and transcranial magnetic stimulation (TMS). For example, imaging has shown that the observation of hand, foot and mouth movements activates the same areas of premotor cortex active during their execution (Buccino et al., 2001). Revealing yet further specificity, the observation of hand and arm movements selectively increases TMS-induced motor evoked potentials from the particular muscles involved in executing these movement (e.g. Strafella & Paus, 2000).

In behavioral studies, stimulus-response compatibility (SRC) procedures are often used to detect automatic imitation. These procedures provide some indication of the strength of the automatic imitation tendency by showing that it can interfere with performance based on task instructions. For example, Kilner et al. (2003) instructed participants to make sinusoidal arm movements in a vertical plane while observing a model perform the same vertical movements (compatible condition) or sinusoidal arm movements in a horizontal plane (incompatible condition). Although participants were, presumably, equally motivated to obey instructions in the two conditions, their movements showed more, counter-instructional deviation from the vertical plane in the incompatible than in the compatible condition. Other SRC studies have shown that automatic imitation interferes, not only with the spatial properties of movement, but also with its timing. Participants instructed in a simple reaction time (RT) task to open their hand as soon as an observed hand began to move, initiated the opening movement faster when the stimulus hand opened than when it closed (Heyes et al., 2005). Similar studies have shown that automatic imitation can influence the timing of hand and finger movements even when the observed movements are task-irrelevant, i.e. when participants are instructed to respond, not to the observed movements, but to arbitrary stimuli such as digits (Brass et al., 2000), crosses (Bertenthal et al., 2006) or colors (Stürmer, Aschersleben, & Prinz, 2000).

As far as we are aware, only one study has explicitly compared the strength of automatic imitation with that of other response tendencies (Brass et al., 2000). This study found that the impulse to imitate finger movements was stronger than the tendency to respond with finger movements to arbitrary symbols and to static spatial markers. The results were important in providing evidence that automatic imitation is genuinely automatic (i.e. that it occurs contrary to task instructions), and that it is not reducible to spatial compatibility (see also Heyes et al., 2005;

Bertenthal et al., 2006), but Brass et al. (2000) provided only a very conservative estimate of the strength of automatic imitation. Theories of imitation assume that it is based on stimulus-response connections that are either innate (e.g. Meltzoff & Moore, 1997) or the product of long-term learning (e.g. Heyes & Ray, 2000). If this is the case, it is not surprising that the tendency to imitate is stronger than the tendency, based solely on task instructions, to respond differentially to symbolic cues. Like imitation, spatial compatibility effects depend on innate or learned response tendencies (Tagliabue, Zorzi, Umilta, & Bassignani, 2000). However, Brass et al.'s study did not show that automatic imitation is generally stronger than the tendency to respond to the site of stimulation; only that automatic imitation is stronger than spatial compatibility when the spatial cue is smaller and less dynamic than the body movement cue.

The present study provided a more stringent test of the strength of automatic imitation by comparing it with that of the tendency to obey verbal commands. Like imitation, verbal command is a common method of instruction in everyday life, and the power of words to evoke actions is a product of deeply engrained mechanisms. Indeed, one theory of imitation, the associative sequence learning (ASL) model (e.g. Heyes & Ray, 2000), suggests that the two response tendencies become engrained in the same way; that we learn to imitate through correlated experience of observing and executing action units, just as we learn the meanings of words through correlated experience of the words and their referents.

We used a Stroop procedure to compare the strengths of automatic imitation and verbal command. There were four groups of participants. In the focal group (Manual-Auditory), participants were required in each trial to open or to close their hand in response to a compound stimulus. The compound consisted of an image of a hand in an open, closed or neutral posture, and the sound of a word: 'open', 'close' or a neutral nonword. In one condition, participants were instructed to imitate the action and to ignore the word (action-relevant task), and in the other condition they were told to obey the verbal command and to ignore the action (word-relevant task). In any given trial, the stimulus on the task-irrelevant dimension (the word in the action task, and the action in the word task) was compatible, incompatible or neutral with respect to the stimulus on the task-relevant dimension. For example, in the action task, an image of an open hand was accompanied equally often by the word 'open' (compatible), the word 'close' (incompatible) and by a nonword (neutral).

If the tendency to imitate is stronger than the tendency to obey verbal commands, then, in this focal group, one would expect the impact on performance of action stimuli in the word task to be greater than the impact of word stimuli in the action task. More specifically, one would expect the compatible task-irrelevant stimulus to speed responding, and /or the incompatible task-irrelevant stimulus to slow responding, more in the word task than in the action task. However, an effect of this kind would not be sufficient to show that automatic imitation is stronger than the tendency to obey verbal commands, for two reasons. First, it could be that the action images used in this experiment were more salient or easier to discriminate than the word stimuli. In this case, one would expect action images to be more potent stimuli, not only for automatic imitation, but also for nonimitative responding. To address this issue, we included a second group of participants (Vocal-Auditory) who were presented with exactly the same stimuli as the focal group, action images in compound with word sounds, but they were required to make vocal rather than imitative responses. For example, in the action task, this group said ‘open’ when they saw an opened hand, and ‘close’ when they saw a closed hand. Langton, O’Malley, & Bruce (1996, Experiment 5) found that irrelevant gestures affected vocal responses to words to the same extent as irrelevant words affected vocal responses to gestures. Therefore, we expected that, in contrast with the focal group, the performance of the Vocal-Auditory group would be affected equally by irrelevant actions in the word task, and by irrelevant words in the action task.

The second issue concerns modality of stimulus presentation. In the focal group, actions were presented visually and words were presented in the auditory modality because those conditions are typical of everyday life. In the course of development, it is likely that simple verbal instructions, consisting of a single word, are more often heard than seen. However, because spoken words unfold over time, whereas images are instantaneously available for processing, auditory presentation of verbal commands could put them at a disadvantage. In other words, if irrelevant actions have a greater impact than irrelevant words in the focal group, this could reflect, not the relative strengths of automatic imitation and verbal command, but faster processing of visual than auditory stimuli. To address this issue we included two further groups in which the word stimuli were written rather than spoken. One of these groups (Manual-Visual) made hand movement responses, and the other (Vocal-Visual) made vocal responses.

Thus, there were four groups: Manual-Auditory, Vocal-Auditory, Manual-Visual and Vocal-Visual. We predicted that in the focal Manual-Auditory group the effect of irrelevant actions on speed of responding to words would be greater than the effect of irrelevant words on responding to actions. If this asymmetric effect indicates that the automatic tendency to imitate is stronger than the tendency to obey verbal commands, rather than an effect of nonspecific features of the stimuli or stimulus-response mapping, then it should also be present in the Manual-Visual group, but not in the Vocal-Auditory or Vocal-Visual groups.

Method

Participants

Forty-eight right-handed volunteers (15 men, mean age: 22.8±7.5 years) were randomly assigned to one of four groups: Manual-Auditory, Vocal-Auditory, Manual-Visual and Vocal-Visual. All had normal or corrected-to-normal vision and normal hearing. The experiment was carried out with local ethical approval and written consent.

Stimuli and Apparatus

Warning and imperative stimuli were compounds of hand actions and words with coincidental onsets. Hand actions were life-sized images of postures made by a male right hand, taken from the angle at which one normally views one's own hand, and presented on a laptop computer screen (60Hz, 400mm, 96DPI) in color on a black background. For the warning stimulus, the hand was in a neutral posture, with the fingers closed and pointing upwards in parallel with the thumb (visual angle: 6.96° x 13.33°), and was shown for a variable duration between 800ms and 1520ms. For the imperative stimuli, the hand was in an opened (15.5° x 13.5°), closed (7.0° x 11.2°) or inverted neutral posture (see Figure 1D for examples), and was shown for 640ms. Word stimuli were either sound files presented via the laptop's internal speaker (auditory) or superimposed in white ink on the hand stimuli in the centre of the screen (visual; 6.5° to 7.1° x 2.6° to 3.1°). For the warning stimulus, the nonword *clepo* was presented for 650ms (auditory) or between 800 and 1520ms (visual). For the imperative stimuli, the word 'open', 'close' or the nonword *pocle* (see Figure 1C for examples) were presented for 640ms (visual) or between 600ms and 640ms (auditory).

The nonwords *clepo* (warning stimulus) and *pocle* (neutral stimulus) were phonotactic amalgams of phonemes contained in the two words 'open' and 'close'. *Pocle* contained the same syllables as *clepo*, presented in reverse order.

For the manual response groups, response onset of opening and closing hand movements was measured by recording the electromyogram (EMG) from the first dorsal interosseus muscle of the right hand (see Heyes et al., 2005). For the vocal response groups, onset of voice responses was measured via a free-standing electret microphone (Vivanco EM 32, Vivanco-direct.com). The RT interval began with the onset of the imperative stimulus, and ended with EMG onset (manual responses) or the activation of the microphone (vocal responses).

Design and Procedure

Participants sat at a viewing distance of approximately 700mm from the stimulus presentation screen. For the manual response groups, the participant's right forearm lay in a horizontal position across his/her body, supported from elbow to wrist by an armrest. The wrist was rotated so that the fingers moved upwards during opening responses, and downwards when closing. Thus, the plane of response movement (up-down) was orthogonal to the plane of action stimulus movement (left-right), controlling for any effects of left-right spatial compatibility. After making each response, participants returned their hand to the neutral starting position; their fingers closed and parallel to the thumb.

Each trial began with the presentation of the warning stimulus. After a variable duration it was replaced by the imperative stimulus. Participants were instructed to respond to the imperative stimulus as quickly as possible, without making errors, by opening or closing their hand (manual response groups) or by saying ‘open’ or ‘close’ (vocal response groups) as soon as they saw an open or closed hand posture (action-relevant task), or heard or saw the word ‘open’ or ‘close’ (word-relevant task). They were instructed to ignore the irrelevant dimension. After the presentation of the imperative stimulus, the screen went black for 3000ms before the next trial.

Four action-relevant and four word-relevant task blocks of 60 trials were presented in alternating order, counterbalanced between participants. Relevant and irrelevant stimulus compounds were *compatible* (e.g. an open hand accompanied by the word ‘open’), *incompatible* (e.g. an open hand accompanied by the word ‘close’) or *neutral* (e.g. an open hand accompanied by the nonword *pocle*). The six trial types, defined by compatibility (compatible, neutral or incompatible) and relevant stimulus (open or close), were equiprobable and randomly intermixed within each block.

Results

Mean RTs are plotted as a function of task and compatibility in Figures 1A-D. Incorrect responses and RTs less than 100ms or greater than 1500ms were removed (3.1%).

Figure 1 about here

As predicted, in the focal Manual-Auditory group (A) the impact of irrelevant actions on responding to words was greater than the impact of irrelevant words on responding to actions; there was an asymmetry favoring actions over words. This asymmetry was not observed in the Vocal-Auditory group (B), who responded to exactly the same stimuli using vocal responses rather than hand actions, suggesting that the asymmetry was not due to greater salience of the action than of the word stimuli. The asymmetry favoring actions over words was present in the Manual-Visual group (C), who saw rather than heard the word stimuli, indicating that it did not depend on faster processing of visual than auditory stimuli. Providing further confirmation that this asymmetry was not due to nonspecific factors, the Vocal–Visual group (D) showed the reverse asymmetry; irrelevant actions had a lesser effect on responding to words than did irrelevant words on responding to actions.

These impressions were confirmed by an initial ANOVA, in which task (action-relevant, word-relevant) and compatibility (compatible, neutral, incompatible) were within-subject factors, and response mode (manual, vocal) and word modality (auditory, visual) were between-subject factors, and by subsequent analyses in which a 2x3 ANOVA (task x compatibility) was applied to the RT data from each group separately. The initial analysis indicated a significant three-way interaction (task x compatibility x response mode: $F(2, 94) = 35.6$, $p < .001$), and a nonsignificant four-way interaction (task x compatibility x response mode x word modality: $F(2, 94) = 1.1$, $p = .341$). The separate analysis of the data from the focal Manual-Auditory group yielded a significant interaction between task and compatibility ($F(2, 22) = 20.8$, $p < .001$), confirming that there was an asymmetry favoring actions over words. This interaction was also significant in the Manual-Visual group ($F(2, 22) = 25.5$, $p < .001$), but it was absent in the Vocal-Auditory group ($F(2, 22) = 1.5$, $p = .252$), and reversed in the Vocal-Visual group ($F(2, 22) = 5.5$, $p = .017$).

In the two groups where there was an asymmetry favoring actions over words, mean RT in the action-relevant task was shorter than in the word-relevant task (Manual-Auditory: $F(1, 11) = 48.7$, $p < .001$; Manual-Visual: $F(1, 11) = 172.3$, $p < .001$). To check whether the action-dominant asymmetry was dependent on this main effect of task on RT, the data from these groups were subjected to bin analyses. For each group, RTs of each participant in each task were divided into five bins of equal size (Ratcliff, 1979). Three quintiles were selected in which, within group, mean RT on neutral trials was approximately equal in action-relevant and word-relevant tasks. The data from these quintiles were subjected to 2x3x3 ANOVAs (task x compatibility x bin). These analyses showed that, in each group, although there was no main effect of task on RT (Manual-Auditory: $F < 1$; Manual-Visual: $F(1, 11) = 1.1$, $p = .316$), there was a significant task x compatibility interaction (Manual-Auditory: $F(2, 22) = 11.8$, $p < .001$; Manual-Visual: $F(2, 22) = 11.9$, $p = .001$). Thus, the action-dominant asymmetry observed in the Manual-Auditory and Manual-Visual groups did not depend on faster responding in the action task than in the word task.

Discussion

Previous research has shown that healthy adult humans have a pervasive and automatic tendency to imitate the actions of others, but this is the first study to provide a stringent test of the strength of this tendency. Using hand actions in a Stroop procedure, the power of actions to elicit imitative responses was compared with the strength of our tendency to obey verbal commands. The results from the focal group, who made manual responses to simultaneously presented actions and spoken words, showed that the impact of irrelevant actions on responding to words was greater than the impact of irrelevant words on imitative responding to actions. The same asymmetry was observed when written, rather than spoken, words were presented, indicating that it was not due to faster processing in the visual modality. The same asymmetry was *not* observed when participants made vocal, rather than imitative, responses, indicating that the action-dominant asymmetry was not due to greater salience or discriminability of the action images than of the verbal stimuli. Therefore, these findings suggest that the human tendency to imitate is stronger than the tendency to obey verbal commands.

Previous studies have indicated that irrelevant actions influence the control of movements made in response to color, spatial and symbolic cues (Stürmer et al., 2000; Bertenthal et al., 2006; Brass et al., 2000). The present findings show for the first time that automatic imitation effects occur, not only when the imperative stimuli bear an arbitrary or purely spatial relationship with responses, but also when they are verbal commands; that is, when the relationship between the imperative stimulus and the response is both specific and overlearned.

Langton, O'Malley, & Bruce (1996, Experiment 5) used a Stroop procedure to compare the power of actions and words, but they did not examine imitative responding. Instead, they required participants to make vocal responses to directional gestures (a person pointing up, down, left and right) and to their verbal equivalents, and found symmetrical compatibility effects; irrelevant gestures affected vocal responses to words to the same extent as irrelevant words affected vocal responses to gestures. We found the same symmetrical pattern in our Vocal-Auditory group, when participants were making nonimitative responses, but a contrasting pattern, indicating action dominance, when participants were making imitative responses. Thus, comparison of the two studies i) confirms that action dominance is specific to imitation, and ii) indicates that, in the case of nonimitative vocal responding, actions and words have comparable impact both when the action stimuli are pointing gestures and when they are opening and closing hand movements.

In a variant of the game ‘Simon says’, played at teatime in Victorian England, children were required to grip the tablecloth when an adult, gripping or releasing the cloth, said ‘Hold tight!’, and to release the cloth, regardless of the adult’s action, when he said ‘Let go!’. Presumably, amusement derived from the fact that, like the participants in the present experiment, children could not resist the influence of automatic imitation, and were therefore compelled flagrantly to disobey the authority of verbal command. However, the results of the present study do not merely vindicate the disobedient behavior of Victorian children. They show that automatic imitation is much more than a parlour game, or a device that experimental psychologists can use to investigate the processes involved in stimulus-response translation. These findings show that automatic imitation is not only pervasive but also powerful. Even among healthy, typically-developing adults, it is more powerful than the tendency to obey verbal commands. In this context, actions do indeed speak louder than words.

References

Bertenthal, B.I., Longo, M.R., & Kosobud, A. (2006). Imitative response tendencies following observation of intransitive actions. *Journal of Experimental Psychology: Human Perception and Performance, 32*, 210–225.

Brass, M., Bekkering, H., Wohlschläger, A., & Prinz, W. (2000). Compatibility between observed and executed finger movements: comparing symbolic, spatial, and imitative cues. *Brain and Cognition, 44*, 124-43.

Buccino, G., Binkofski, F., Fink, G.R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R.J., Zilles, K., Rizzolatti, G., & Freund, H.J. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *European Journal of Neuroscience, 13,* 400-404.

Charman, T., & Baron-Cohen, S. (1994). Another look at imitation in autism. *Development and Psychopathology, 6*, 403-413.

Darwin, C. (1989). *Voyage of the Beagle*. London: Penguin Books.

Heyes, C.M., & Ray, E.D. (2000). What is the significance of imitation in animals? *Advances in the Study of Behavior*, 29, 215–245.

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

Kilner, J.M., Paulignan, Y., & Blakemore, S.J. (2003). An interference effect of observed biological movement on action. *Current Biology, 13*, 522–525.

Lakin, J.L., & Chartrand, T.L. (2003). Using nonconscious behavioral mimicry to create affiliation and rapport. *Psychological Science, 14*, 334-339.

Langton, S.R.H., O'Malley, C., & Bruce, V. (1996). Actions speak louder than words: Symmetrical cross-modal interference effects in the processing of verbal and gestural information. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 1357–1375.

Lhermitte, F., Pillon, B., & Serdaru, N. (1986). Human autonomy and the frontal lobes. Part I: Imitation and utilization behavior: a neuropsychological study of 75 patients. *Annals of Neurology, 19*, 326-334.

Meltzoff, A. N. & Moore, M. K. (1997). Explaining facial imitation: A theoretical model. *Early Development and Parenting, 6,* 179-192.

Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin, 86*, 446–461.

Simpson, A. & Riggs, K.J. (2007). Under what conditions do young children have difficulty inhibiting manual actions? *Developmental Psychology, 43,* 417-428.

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

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

Tagliabue, M., Zorzi, M., Umilta, C., & Bassignani, F. (2000). The role of long-term-memory and short-term-memory links in the Simon effect. *Journal of Experimental Psychology: Human Perception & Performance, 26,* 648-670.

Thorndike, E.L. (1898). *Animal Intelligence: An Experimental Study of the Associative Processes in Animals* (Psychological Review, Monograph Supplements, No. 8). New York: Macmillan.

Wallbott, H.G. (1991). Recognition of emotion from facial expression via imitation? Some indirect evidence for an old theory. *British Journal of Social Psychology, 30*, 207-219.

Author note

AB is now at Faculté de Psychologie et des Sciences de l'Education, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.

Figure caption

Figure 1. RTs in compatible, neutral and incompatible trials for word-relevant (solid line) and action-relevant (broken line) task conditions. Results are presented separately for the four different participant groups: (A) Manual-Auditory, (B) Vocal-Auditory, (C) Manual-Visual and (D) Vocal-Visual. Vertical bars indicate standard error of the mean. Images show compatible, neutral and incompatible stimulus compounds in action-relevant (Panel C) and word-relevant (Panel D) task conditions for the visual word modality groups (C and D). For the auditory word modality groups (A and B), words were spoken.