Running head: METHODOLOGY OF AUTOMATIC IMITATION RESEARCH

An empirical analysis of the methodology of automatic imitation research in a strategic
context

Balazs Aczel^{1†}

Zoltan Kekecs^{2‡}

Bence Bago^{3‡}

Aba Szollosi^{1‡}

Andrei Foldes^{1‡}

¹Eotvos Lorand University, Budapest (Hungary)

²Baylor University, Waco (US)

³University of Edinburgh, Edinburgh (UK)

Author notes

Authors. [†] Balazs Aczel designed the experiments and wrote the manuscript. [‡] Zoltan Kekecs, Bence Bago, Aba Szollosi and Andrei Foldes performed the experiments and analyzed the data. Correspondence: Balazs Aczel, ELTE Pszichológiai Intézet, Izabella u. 46. 1064. Budapest, Hungary. Telephone: +36-20-4143124. E-mail: aczel.balazs@ppk.elte.hu

Abstract

Since the discovery of the Mirror Neuron System it has been proposed that the automatic tendency to copy observed actions exists in humans and that this mechanism might be responsible for a range of social behavior. A strong argument for automatic behavior can be made when actions are executed against motivation to do otherwise. Strategic games where imitation is disadvantageous serve as ideal designs for studying the automatic nature of the participants' behavior. Most recently, Belot, Crawford, and Heyes (2013) conducted an explorative study using a modified version of the Rock-Paper-Scissors game and suggested that in the case of asynchrony in the execution of the gestures automatic imitation can be observed early on after the opponent's presentation. In our study, we video-recorded the games, which allowed us to examine the effect of delay on imitative behavior as well as the sensitivity of the previously employed analyses. The examination of the recorded images revealed that more than 80 % of the data was irrelevant to the study of automatic behavior. Additional bias in the paradigm became apparent, as previously presented gestures were found to affect the behavior of the players. After noise filtering, we found no evidence of automatic imitation neither in the whole filtered data-set, nor in selected time-windows based on delay length. Besides questioning the strength of the results of previous analyses, we propose several experimental and statistical modifications for further research on automatic imitation.

Keywords: automatic imitation, action control, strategic context, mirror neurons, Matching Pennies game

Automatic imitation refers to the phenomenon that the observation of the topographical features of an action facilitates the execution of a similar action in the observer. The study of automatic imitation received considerable interest over the last two decades as, if the effect truly exists in humans, it might offer a solution to the “correspondence problem” (Brass & Heyes, 2005) and thereby provide an explanation of how the imitator is able to map the action of the observed model into the same pattern of motor activation. One suggested solution to this vexed question assumes that people learn these contingencies through repeated experience (Prinz, 2002); others argue that dedicated neurological mechanisms can also enhance imitation (Iacoboni et al., 1999). The so-called mirror neurons in the ventral premotor area F5 of monkeys have been found to discharge both when the animals observed and when they executed the same motor action (Fadiga, Fogassi, Gallese, & Rizzolatti, 2000). This Mirror Neuron System was proposed to mediate between observed and performed actions and to provide functional mechanism to a rich set of phenomena in social cognition, such as social learning (Meltzoff & Decety, 2003) or empathy (Iacoboni, 2005).

Whereas the topic has become of the interest for interrelated theoretical fields, two sets of questions have been raised to challenge the concept of automatic imitation. The first set of questions centers on whether the effect truly reflects processes of imitation or could be explained as a product of more general long-term S-R associations. The second set of questions asks whether the automatic nature of the effect has been convincingly demonstrated.

In the first report of automatic imitation of body movements, Stürmer, Aschersleben, and Prinz (2000) instructed the participants to open or close their own hand for the signaling color of a presented opening or closing hand stimulus. Only the changing color and not the movement of the hand stimulus was relevant for the task of the participant. Although the

opening or closing feature of the stimulus were task irrelevant, people were found to be faster when viewing a compatible and slower when viewing an incompatible movement. Further experiments indicated that people are not just faster, but more accurate when viewing body movements compatible to their task (Brass, Bekkering, & Prinz, 2001; Edwards, Humphreys, & Castiello, 2003; Kilner, Hamilton, & Blakemore, 2007), suggesting that the perception of body movements inherently facilitates the execution of the same movements. Whether the vision of familiar biological body movements provokes more automatic imitation than non-biological movements was also investigated to assess the degree to which the mirror system is a product of phylogenetic or ontogenetic learning. Earlier studies suggested that automatic imitation is reduced, but not eliminated when robotic stimuli are observed instead of human actions (Press, Bird, Flach, & Heyes, 2005); however, this “human bias” can be eliminated by sensorimotor experience of robotic stimuli (Press, Gillmeister, & Heyes, 2007). These results support the Associative Sequence Learning hypothesis which could explain the imitating effect as primarily the result of previous learning, correlated experience of observation and execution of the same actions (Heyes, Bird, Johnson, & Haggard, 2005).

Nevertheless, objections were raised about the interpretation of these earlier results of the automatic imitation studies, as they were confounded by spatial correspondence between the task-irrelevant stimulus and the response (Jansson, Wilson, Williams, & Mon-Williams, 2007). For example, a presented finger movement can be imitated faster either because of its left-right spatial correspondence to the participant’s finger or because of its anatomical identity. Indeed, spatial compatibility was found to have a strong effect on reaction times (Bertenthal, Longo, & Kosobud, 2006; Jiménez et al., 2012) supporting an alternative explanation to the previous observations. In later experiments, the possible effect of automatic imitation and spatial correspondence were controlled by different techniques, such

as orthogonally aligned stimulus and response hands (Heyes et al., 2005). Automatic imitation was found to be present regardless of spatial compatibility in behavioral (Bertenthal et al., 2006; Catmur & Heyes, 2011; Jiménez et al., 2012) and neuropsychological (Mengotti, Ticini, Waszak, Schütz-Bosbach, & Rumiati, 2013) studies.

A related challenge for automatic imitation is that S-R compatibility provides very similar predictions. It is a robust finding in the associative learning literature that perceptual or structural similarity between the stimulus and the required response can increase the speed and decrease the error in the response (Kornblum, Hasbroucq, & Osman, 1990). Indeed, a tapping or lifting pen was found to effect the participants' finger movement to the same degree as tapping or lifting finger (Jansson et al., 2007). Still, when the two effects were pitted against each other in a task with opposite stimulus-response instructions, participants tested for automatic imitation did not show reverse compatibility effect as a single route S-R compatibility model would have predicted (Boyer, Longo, & Bertenthal, 2012).

These converging findings, therefore, provide support to the claim that the sight of an action facilitates the motor replication of that action; however, the second main question of this field, the automatic nature of the effect, is not self-evidential. In recent empirical studies of motor imitation (or in naturalistic interpersonal settings, mimicry) has been labeled as automatic (Hogeveen & Obhi, 2013), unconscious (Belot et al., 2013), nonconscious (Van Baaren, Maddux, Chartrand, de Bouter, & van Knippenberg, 2003), or unintentional (Richardson, Marsh, Isenhower, Goodman, & Schmidt, 2007), generally implying that the copying action is independent of the actor's intention (Heyes, 2011). It is commonly assumed that imitation is automatic when the participants copy an action without (Heyes, Bird, Johnson, & Haggard, 2005), or against (Leighton & Heyes, 2010) the instructions of the experiment. In these cases automaticity is often implied when the imitating participants are

unable to report or guess the purpose of the experiment in a post-hoc questionnaire (Leighton & Heyes, 2010).

Evidence for automaticity in human behavior, however, has been repeatedly questioned in various domains of psychology. In particular, it has been convincingly argued that post-experimental verbal reports can only indicate that someone *is* aware of something; they cannot verify the lack of awareness (Stadler & Roediger III, 1998). Any retrospective test of awareness can be challenged along issues of sensitivity (e.g., participants may not completely understand the question; participants may be unable or unmotivated to retrieve or verbalize their memories); or they can fall short of the information criterion (i.e., the question does not tap into the same knowledge as used during the action) (Shanks & St John, 1994).

Traditionally it is assumed that automatic actions are of bottom-up nature, hence their execution does not require attention and does not rely on higher order cognitive resources. Some studies asked the participants to direct their attention away from the task-irrelevant movements (e.g., Bertenthal et al., 2006; Brass, Bekkering, Wohlschläger, & Prinz, 2000), which could suggest that the imitated features were not in the focal attention. However, it can be argued that the moving stimuli are attentionally salient and can involuntarily capture the participant's attention without imposing a high processing load on cognitive resources (Chong, Cunnington, Williams, & Mattingley, 2009). While functional neural dissociation between automatic and intentional imitation has been suggested by fMRI results (Bien, Roebroek, Goebel, & Sack, 2009), behavioral results show strong top-down control in the imitation of observed actions. For example, Chong et al. (2009) observed that the imitation of grasp could be eliminated by directing the participant's attention to a task-irrelevant feature. Others argue that imitation is always goal directed (Csibra & Gergely, 2007; Wohlschläger & Bekkering, 2002). According to the Theory of Goal-Directed Imitation (Bekkering,

Wohlschläger, & Gattis, 2000; Wohlschläger, Gattis, & Bekkering, 2003), the perceived act is decomposed into hierarchically ordered goal aspects based on the functionality of the action. These goals can elicit a relevant motor program, which can or cannot be matching movements. According to Csibra (2007), imitation could be better understood by action reconstruction rather than direct matching. As the observed action is always decomposed into units and some aspects are judged to be relevant and others remain disregarded, this analysis of actions can also extend to the goal of the observed action. If the observer and the model have similar biological and environmental constraints then the observer, to achieve her goals, might reconstruct the observed action in her motor system. Under different circumstances, the observer might choose different means to reach her goals. In this sense, all imitations are emulations, reproductions of an outcome rather than the observed action. This flexible model of action-mirroring is supported by certain empirical observations in cooperative setting. For instance, in the study of van Schie, Waterschoot, and Bekkering (2008) the participants were presented with an image where an actor grasped an object either with full grip or with precision grip. The direct matching hypothesis would predict that the perception of the action would result in shorter reaction times when the task is to imitate that action compared to when the task is to execute the complimentary action. Contrary to this hypothesis, the results showed a reversed compatibility effect, the participants appeared to be more primed to execute the complimentary action. A possible explanation of this result is that the joint action context might have evoked a goal to cooperate, which in turn facilitated the complimentary motor reaction. Others also found that interpretation is part of action mirroring. For example, in the experiment of Longo and Berenthal (2009) imitation was reduced when the participants were reminded about the virtualness of the hand stimulus. Campione and Gentilucci (2011) found that the participants imitated not the specific values of kinematic landmarks of reaching-grasping movements, but the kinematics typical of a given goal, suggesting that

imitation is guided rather by the specific goals than the kinematic values of the observed model.

Unfortunately, most of the studies of automatic imitation do not allow us to evaluate what role the participants' goals played in the behavior. The automatic nature of the observed imitation has been typically inferred from the observation that participants execute the instructed movement faster when a presented task-irrelevant motion is similar as compared to when it is dissimilar to the instructed movement (Brass et al., 2001; Heyes et al., 2005; Stürmer et al., 2000). Nevertheless, participants had no naturalistic goals with the executed actions and had very little incentive to obey the task instructions. It is, therefore, not clear to what degree was the observed imitative behavior automatic or involuntary in these studies.

One experimental technique that is often regarded as a reliable test of automatic (or implicit) human processing is to provide a context where the participants are motivated to avoid a certain behavior. Actions executed against strong incentives to avoid them serve as strong arguments for automatic or non-conscious behavior (Jacoby, 1991). Strategic games provide an ideal context for motivating the execution or control of a selected behavior. The first attempt to employ a task with a strategic context to distinguish automatic and intentional imitation has been supplied by Cook et al. (Cook, Bird, Lünser, Huck, & Heyes, 2012). Their participants played Rock-Paper-Scissors (RPS) games where imitating the opponent (draw) is not a rewarding strategy. During the game, either one or the other player was blindfolded (blind) and the winners of each 60-game match received monetary reward. Due to naturally occurring asynchrony between the onset of the gesture movements, on some of the trials sighted participants could have seen the gesture of the blind players shortly before initiating their own. The authors argued that on these trials the sighted players automatically imitated the blind players, because the proportion of draws in the sighted-blind condition was higher

than chance, and than the proportion of draws in the blind-blind condition. In a subsequent study, Aczel, Bago, & Foldes (2012) reanalyzed these results and replicated the original experiment and found that automatic imitation is only one of the several reasons why the proportion of draws in the RPS games could be higher than pure chance. For example, repeated choices can be sensitive to sequence or priming effects. A recent analysis of over 100,000 games of the RPS task revealed that winners tend to repeat their winning gesture while losers tend to change their gesture (Wang, Xu, & Zhou, 2014) which increases the chance that in the proceeding game they would play the same gesture. The players of the same study also preferred to use one of the gestures over the other ones (rock: 36%). Aczel et al. (2012) also found in their replication, as well as in the original study, that the three gestures were not equally shown in the games, players presenting higher frequency of scissors. These frequency differences can be problematic since even without being significant they can result in a rate of draws significantly higher than chance. A further bias towards draws can be speculated in the experimental setting of Cook et al.'s design: those sighted participants who saw the gesture of the blind player had no strong incentive to avoid imitation, as draw outcome was better than losing the game. One could argue that draws are safer options in RPS games than avoiding imitation where an additional decision is needed to select the winning gesture and avoid the losing gesture. These effects can all result in higher than $33\frac{1}{3}\%$ of draws without automatic imitation. Reanalyzing the comparison of the proportion of draws in the sighted-blind and blind-blind conditions in the results of Cook et al, Aczel et al. found no support for the original conclusions, there was no significant difference between the two conditions. Also, their replication resulted in the same conclusion of no evidence for automatic imitation in the RPS game.

A main reason why RPS games might not be ideal for testing automatic imitation is that participants are not strongly motivated to avoid imitating the gestures of the opponent. In a more recent work, Belot et al. (2013) employed a modified design of the game which is more suited to study automatic imitation in a strategic context. The Matching Pennies game is similar to RPS with a few alterations. The players start with a neutral gesture and then show either open-hand or closed-hand following a count of three. During the game, either one or the other player was blindfolded and one of the players won when the gestures were the same, the other player won when the gestures were different. There are no draws in this game and the players receive reward after each game, so one of the players always has strong incentive to avoid imitation. When analyzing the overall frequency of matches for sighted players with an incentive to mismatch, Belot et al. (2013) did not find evidence that these players would imitate. This result might have occurred due to the fact that automatic and intentional responses work in opposite direction, offsetting each other.

Belot et al. (2013) suggests that there is a range of time delay after the gesture of the blind player in which the sighted player is already able to respond, but only automatically. When the delay is sufficiently longer the intentional control can be executed as a response to the observed gesture. We plotted their hypothesis on Figure 1, describing how automatic, then intentional control can result in matching or mismatching behavior with the increase of time delay.

[Insert Figure 1. about HERE]

Following this hypothesis, Belot et al. estimated the intentional and automatic effects on the responses performing probit regressions on the data at the individual game level. Neither showed support for automatic imitation when regarding the whole sample, but the authors argued that when taking into account timing and experience the results are more

revealing. Therefore, they split the sample into games played early and late in a round, and then split the sample further into games played in early and late rounds played against a particular opponent. After conducting six one-tailed t-tests on the split data, they found two significant results: the main effect of the opponent's gesture (interpreted as the automatic imitation effect) was significant in the early games played in the rounds, and especially in early rounds played against a particular player, when the incentives were still new to the players. From these findings the authors inferred that "the results strongly confirm the occurrence of automatic imitation" (p. 2763).

While it is reasonable to expect effects of timing and experience, we find it confounding to split the data into early 10 games and late 10 games within the 12 rounds. As a result of this separation, games played later in the experiment (e.g., games 41-50) would count as early, and games played in earlier rounds (e.g., games 11-20) would count as late games. This post-hoc separation of the data and the conduct of multiple t-test analyses undermine the strength of evidence for automatic imitation in this study, since testing the same hypothesis multiple times increases alpha error probability. Furthermore, the statistical approach of Belot et al. relies on the assumption that including the interaction of the blind player's gesture and the task in the model would somehow control for the effect of intentional imitation, and thus the coefficient corresponding to the opponent's gesture would capture the pure automatic imitation effect. However, in this type of regression, the variance explained by the interaction of two categorical variables is not controlled for in the coefficients corresponding to the predictors. Instead, these coefficients only display the effect of the predictor in the subset of data in which the other categorical predictor is at its reference level.

In summary, the Matching Pennies game is a potentially suitable paradigm in which to test the presence of automatic imitation in a strategic context. However, we argue for several

experimental and statistical modifications for a more convincing analysis of the question.

Firstly, to obtain more informative output data, it is important to reduce the level of noise in the measurements. We suspect that presently at least three major sources of noise may hinder the analysis of the Matching Pennies experiment. The first source of noise is that the outcome of the games is analyzed irrespective whether the blind players executed the gestures before or after the sighted players, however, the latter is extraneous for our study of automatic imitation since an event cannot have effect on a past event. Another source of noise is that visual information is unlikely to have any effect on behavior when the delay between the two gestures is too short. Based on previous studies of human motor control, we argue that imitation is not physically possible until at least 150 ms if visual categorization is needed (Vanrullen & Thorpe, 2001). Even in a simple reaction time task measured through finger tapping¹, the first quintile of response times were always above 150 ms in a study by Brass, Bekkering and Prinz (2001). Although faster than 150 ms responses (the lowest being 120 ms) have been demonstrated by Kirchner and Thorpe (2006) in a discrimination reaction time task, these results originate from saccadic eye movement responses. Even though saccadic movements are considerably faster than hand movements, and average minimum response time across participants was 150 ms here as well. Thus, it is reasonable to argue that games with shorter than 150 ms delay of the sighted player is irrelevant for imitation research where hand movement is required. A third source of noise is possible if the sighted player accidentally or voluntarily would not keep his or her eyes on the gesture of the blind player. This source of noise is difficult to control in group sessions when solely the participant umpire, who is uninformed about the aim of the experiment, is in charge to enforce the rules of the game. To increase the sensitivity of our analysis in this study, we doubled the sample

¹ In this study, the initiating stimuli were also finger tapping motions. Reaction times were above 150 ms even in response to ideomotor-compatible finger movement stimuli.

size and to study the behavior of the participants with greater scrutiny the experiments were recorded by a hidden video camera. Measuring the delay between the initiations of the two gestures in each game allowed us to select only the informative trials for the analyses.

Analyzing the data in triads in the previous studies provided only limited insight into the behavior of the individuals. Especially for the exploration of the effect of experience and timing, mixed effect analysis provides more sensitive model to analyze the recorded data.

Another deviation in our analysis from previous practice concerns the operational definition of automatic imitation. Belot et al. proposed a technique to calibrate the estimates of the effect of intentional and automatic imitation from aggregated matching frequencies of both players with incentive to match and mismatch. However, it cannot be excluded that people may be better at intentional matching than at intentional mismatching. In which case, we could expect an overall deviation towards matching, even without automatic processes.

Therefore, we focused our analysis mostly on the incentive-to-mismatch condition, as this condition is a clear indicator of automatic imitation since matching behavior in this condition should not be in the participants' interest.

Methods

102 participants (67 females) were recruited among university students ($M = 21.91$ years), all native speakers of Hungarian and naïve to the purpose of the experiment. Each participant could win 10 HUF (approx. 0.05 USD) per game, so that one player could win a maximum of 1600 HUF (approx. 7.25 USD) in the whole experiment². Informed ethical consent was obtained from all participants.

² Although we provided lower nominal monetary incentives than Belot et al. (2013), based on the average household net-adjusted disposable income per capita (which is already adjusted for Purchasing Power Parities), the two honorariums had approximately equivalent value.

The experiments followed closely the procedure described in Belot et al. (Belot et al., 2013). The participants were tested in triads in succession (34 triads sessions in total), so that only the members of one triad and the experimenter were present at any particular session, each session lasting for approximately 40 min. After providing informed ethical consent, labels X, Y and Z were randomly assigned to the members of the triad. Detailed instructions were presented as well as read aloud to the participants, who then had the opportunity to ask questions.

During the experiment, the triads were required to play 12 rounds of Matching Pennies games in 3 series. Each series consisted of 4 rounds being 20 games in each round. Within each round one player was always blindfolded, one was sighted and the third one took the role of umpire. The roles alternated as indicated in Table 1.

[Insert Table 1. about HERE]

In each game, the two players sat at opposing sides of a table, positioning their elbows on an assigned area 55 cm apart. The two acting players started each game with a neutral hand position, as in Figure 2/a. The players were asked to show either Open Hand or Closed Hand gesture (as demonstrated in Figure 2/b and Figure 2/c) immediately after the umpire gave a count of three. In the case of apparent asynchrony in the execution of the gestures, the umpire required the players to replay that game. Each series was preceded by three unpaid practice games where both players were sighted. Throughout a given series, one of the players won if they showed an identical gesture to the other player (Match), whereas the other player won the game when the two gestures were different (Mismatch). These roles changed at the end of a series (see Table 1). After each game, the umpire informed the players about the gestures and announced the outcome of the games, so that both the sighted and blindfolded players had the same information. The experimenter ascertained that the

participants followed the instructions and he also recorded the gestures and outcomes in each game. After the experiment, the participants were interviewed about the strategies that they could recall using.

[INSERT Figure 2. HERE]

Unbeknownst to the participants, the sessions were recorded by a hidden video camera. Following the completion of all the experiments, the participants were provided with the opportunity to learn about the purpose of the study, at which time they had the right to withdraw their data from the records.

Thus, there were only two deviations from the experimental design of Belot et al. (2013): 1) the sessions were recorded by a hidden video camera, and 2) the triads were tested in succession and not in group testing (covert high definition video recording of several triads at once was not achievable).

The video recordings were assessed frame-by-frame by one rater without time-pressure to determine the starting time of each gesture. The frames in which the players started to express their gestures were recorded. During the assessment of the recordings, the raters remained blind to each player's incentives (match vs. mismatch). A random half of the recordings were selected to be re-assessed by several raters for any significant inter-rater deviations. The raters also assessed whether subjects complied with the instructions.

A crucial element of the experiment is that the sighted player has to be able to see the gesture of the blindfolded player so that automatic imitation can take place. During the assessment of the recordings, the raters reported that some of the sighted players averted their gaze from the blind player's hand (for example facing away from the opponent or covering their own eyes with their hand). Trials in which no visual information is received by the

sighted player about the blind player's gesture cannot result in automatic imitation and so it was deemed necessary to exclude these trials from the analysis. Each video was rated by four raters who were blind to the study hypotheses and individually rated each round according to whether or not the sighted player's gaze was averted from the opponent's hand when the opponent showed his or her gesture. The ratings were aggregated and games which were deemed to have averted gaze were excluded from further analysis. (For details of the rating procedure, see the Supplemental Material.)

In previous studies observations had to be aggregated before analysis, because classical statistical tests assume independence of the observations. Aggregation of data into counts or proportions and analyzing them via t-tests or ANOVAs (as done in Cook, et al., 2012 or Belot et al., 2013) is an inappropriate statistical method to deal with clustered binomial data as it violates the assumption of normality, leads to incorrect estimation of event probabilities and reduction in error variances, thus inflating the risk of Type I error (Jaeger, 2008; Quené & Van den Bergh, 2008). More recent statistical methods, the mixed effect models, can overcome this problem by incorporating random effects into the statistical models. A further advantage of mixed effect logistic models over ANOVAs on aggregated proportions are their superiority in detecting real effects and thus in retaining statistical power (Dixon, 2008). We used such mixed effect models in our current study. For this purpose, we used the lme4 package in R (Bates, Maechler, & Dai, 2008). This package is specifically designed to fit random effect models in which data are clustered (such as in our case, where our observations are not independent from each other, as trials are performed by the same individuals and the pair of players several times). Unlike previous packages dedicated to random effect models, lme4 handles binary outcomes – such as matching or mismatching of gestures – and crossed random effects – such as players appearing in more than one pair – gracefully. As lme4 can

handle crossed random effects, there is no need to control for the effect of the identity of the sighted player as a fixed effect (such as in Belot et al. (2013)), rather this can be entered into the model as a random effect as well as the random effect of pairs.³

To confirm the validity of our results, we obtained the data from the study of Belot et al. (2013) and we ran every analysis on it that was compatible with the dataset⁴. To distinguish the results of the two studies, we refer to our own data as the Aczel dataset, while data from the study of Belot et al. (2013) is referred to as the Belot dataset.

Results

From the 34 sessions, 2 were excluded because of a recording error (the battery in the camera went flat during the session). In total, 14 hours of video was assessed (over 1 500 000 frames). The intra-class correlation of the starting frame of a gesture was very high ($ICC > 0.99$). In 87.4% of the games the difference between the starting frame of the gesture coded by the two raters was 0 or 1 frame. From the 7680 trials of the 32 remaining sessions, 554 were excluded from further analysis because of averted gaze, and 438 further trials were omitted because of undiscernible or missing gestures.

Effect of the opponent's current gesture

Overall, we observed matches in 50.09% and mismatches in 49.91% of the Matching Pennies games. According to the logistic regression, there was no statistical difference in the frequency of these two outcomes, $OR = 1.00$, $Z(6686) = -0.15$, $p = .878$. Performing the same

³ The R script for the analyses and further details on model selection and the use of one- and two-tailed tests are accessible in the Supplemental Material.

⁴ Because Belot et al. (2013) did not record data about the length of delay between the players' presentation of gestures, analyses involving segmentation by delay length or the effect of delay were not repeated on the Belot dataset. Specifically, the following upcoming sections involve analysis only on the Aczel dataset: Effect of delay, Effect of experience, Sensitivity analyses.

analysis on the Belot dataset revealed a slightly higher percentage of matches compared to mismatches (51.41% vs. 48.59%); however, this difference was still not statistically significant, $OR = 1.06$, $Z(4318) = 1.70$, $p = .089$, and there was no significant difference between matching probabilities in the two datasets $OR = 1.06$, $Z(11005) = 1.44$, $p = .149$.

Based on the above mentioned reaction-time results we suspected that a significant portion of the full dataset (games with less than ~150 ms delay) is actually unlikely to be relevant for automatic imitation as it is physically not possible for the sighted player to react to the gesture of the blind opponent within such short time period. Thus we decided to use a cutoff of around 150 ms delay to test our hypotheses on a less noisy subset of the data. Because of the frame-rate of our camera, the closest to this cutoff were the 133ms and 166ms frames. The examination of our data displayed in Figure 3 indicated that in sighted players with incentive to match imitation effects can be observed as soon as 133 ms after the blindfolded player's gesture so we chose this delay length as a cutoff point for our analyses. This noise filtering revealed that in 81.52 % of the data either the blind players initiated the gesture later than the sighted players or the delay after their gesture was too short for the sighted players to react. We have noticed during this analysis that the mean of delay in the unfiltered data is significantly lower than zero ($t(85.29) = -7.09$, $p < .001$, mean = -29 ms; $SD = 172$ ms, pair ID as a random effect), which indicates that the blind players tended to present their gestures later than the sighted players.

In the games in which the sighted player's gesture was delayed by at least 133 ms, matches occurred significantly more often than mismatches, $OR = 1.15$, $Z(1234) = 2.50$, $p = .013$. Also, it appears that this effect had been masked by the noise in the full dataset. When we introduced task as a predictor in this model, we found that players with an incentive to match had a higher match frequency than players with an incentive to mismatch. However,

this difference did not reach statistical significance, $OR = 0.83$, $Z(1233) = -1.57$, $p = .059$ (one tailed). Detailed analysis showed that the effect of bias toward matches was detectable only for those cases where the aim of the sighted participant was to match the opponent's gesture, $OR = 1.28$, $Z(542) = 2.82$, $p = .002$ (one tailed), while this was not significant for sighted players with an incentive to mismatch, $OR = 1.07$, $Z(690) = 0.84$, $p = .403$, which does not support automatic imitation.⁵

Effect of delay

Belot et al. hypothesize that early on after the presentation of the blindfolded player the sighted player is able to respond only automatically, while intentional action could occur only after a longer delay. For participants with the incentive to match both automatic and intentional effects favor matching over mismatching behavior thus we would expect an overall bias for matching that increases with time, as with the increasing delay the sighted player has a higher chance of consciously realizing the other player's gesture and thus act according to his or her task. (We approximate this with a positive linear effect of the length of delay on the probability of matching.) On the other hand, automatic imitation would predict that participants with the incentive to mismatch would be inclined to execute mismatch only when the delay is sufficiently long for voluntary control, while they would tend to match the opponent's gesture in the early stages of the delay. (We approximate this with a quadratic effect of the length of delay on the probability of matching). In the case of matching incentive of the sighted participants the results did not support the prediction of this positive linear effect, $OR = 1.00$, $Z(541) = 0.02$, $p = .494$ (one tailed). Similarly, for sighted players with an incentive to mismatch, we found neither a quadratic, $OR = 0.95$, $Z(688) = 0.68$, $p = .496$, nor

⁵ As argued in the Introduction, we only regard evidence of higher than chance matching behavior of players with an incentive to mismatch as supporting automatic imitation.

a linear effect of the delay, $OR = 1.00$, $Z(688) = -0.03$, $p = .706$, on matching behavior (Figure 3). As a consequence, this fine-grained analysis could not find evidence for automatic imitation in this strategic game situation.

[INSERT Figure 3. HERE]

Effect of experience

A possible reason for not finding evidence for automatic imitation in this game might be that, as Belot et al. conjecture, learning occurred in the course of the whole game and the participants followed a different strategy in the earlier and later parts. For example, players could have learned to focus on the gesture of the opponents only when they had an incentive to match and could avoid looking at the gesture of the opponent when they had an incentive to mismatch, thereby optimizing the use of automatic imitation. According to this, we should find that the quantity of experience of the player has an effect on the frequency of mismatching behavior. Only trials where the player was sighted and had incentive to mismatch were counted as relevant experience. As a matter of fact, this meant a pair-specific learning, as a particular player had incentive to mismatch only in one of the two dyads he/she participated in. We found no effect of the gesture of the blind player, $OR = 1.15$, $Z(688) = 0.94$, $p = .348$, and the quantity of experience on the gesture of the player with an incentive of mismatch, $OR = 1.00$, $Z(688) = 0.16$, $p = .877$, indicating no evidence for learning in the course of the game (as the effect of experience on learning is often not linear, we also tested for a logarithmic and a quadratic trend to model learning, none of which indicated a significant learning effect, $ps > .100$).

Sensitivity analyses

Because we have only indirect empirical indications (see Introduction) as to how much delay is sufficient for automatic imitation effects to emerge, 100, 166 and 200 ms were also considered as delay cutoffs in this study and all analyses were repeated using each cutoff point separately (results are displayed in Supplemental Material, Table 1). These results are consistent with the ones using 133 ms as a cutoff. There are a few minor differences in the results, namely the probability of matching is not significantly higher than zero in the dataset with the 100ms cutoff, and there is a significant linear and quadratic effect of delay length on the matching probability in the incentive to mismatch subgroup. However, the linear effect is negative, showing a decreasing probability of matching with delay length, and the quadratic effect is also not in the direction what would indicate automatic imitation. –Thus, our interpretations remain unchanged: we found no evidence for automatic imitation on these trials.

We also conducted an analysis on the matching frequencies after cutting the dataset into four segments according to the time delay of the gesture. The cutting point was at the upper end of each quartile of the time delay (the four quartiles included games with delay of 133 ms, 166 ms, 200 ms and 233-1000 ms respectively). Although this segmentation of the data results in a critically low number of trials per quartile (236, 176, 96 and 184 respectively), the analysis of each of the quartiles of delay for sighted players with an incentive to mismatch indicated no evidence that the matching frequency would differ significantly from the level of chance, all uncorrected $ps > .05$ (see Table 2).

Additionally, we performed a moderator analysis, to see whether there was an effect of gender on the results. We found no effect of the gender of the umpire on the matching probability in the trials where the sighted player responds with at least 133 ms delay, $OR = 1.06$, $Z(1233) = 0.51$, $p = .611$, nor was there any interaction effect between the gender of the

umpire and the gender of the blind, $OR = 1.07$, $Z(1231) = 0.28$, $p = .779$, or the sighted player, $OR = 0.93$, $Z(1231) = -0.30$, $p = .764$.

[INSERT Table 2. HERE]

Effect of previous gestures

In a search for possible strategies and sequential patterns in the behavioral data, we implemented a model involving the effect of outcome of the preceding game (win or lose), and role (sighted or blind) and task (to match or to mismatch) in the current game to predict whether the player repeated or changed his or her own preceding gesture. These analyses were performed on the whole dataset instead of the dataset filtered by delay length⁶. Task was dropped from the final model during model comparisons based on Akaike information criterion (AIC) and Bayesian information criterion (BIC) values (model fit statistics). As apparent in Figure 4 (Aczel dataset panel), we found a significant interaction of role and outcome in the final model, $OR = 1.21$, $Z(12653) = 2.70$, $p = .007$. Post-hoc analysis revealed a significant effect of role, $OR = 0.86$, $Z(12655) = -4.14$, $p < .001$, resulting from the fact that blind players had a significant tendency to repeat their preceding gestures. The effect of outcome was also significant, in line with the predictions of the Conditional Response model (CR) of game theory (see e.g., Wang et al., 2014), in that winners of the previous game tended to stay with their previous gesture more readily than losers, $OR = 1.23$, $Z(12655) = 5.81$, $p < .001$. The significant interaction of role and outcome meant that sighted players behaved as predicted by CR: they switched gestures when they lost, while they stuck to their previous gesture after winning. On the other hand, while blind players also persisted with their previous gesture when they won, they did not seem to have a preference for switching or

⁶ First games of the rounds were omitted because the games preceding them were further back in time and in those games the player had a different role (sighted or blind).

remaining with the previous gesture when they lost. This effect resulted in an overall significant bias toward matching the player's own previous gesture for blind players, $OR = 1.08$, $Z(6327) = 2.04$, $p < .041$, while no such bias can be observed for sighted players, $OR = 0.93$, $Z(6327) = -1.74$, $p = .081$.

[INSERT Figure 4. HERE]

To confirm the existence of these effects, we conducted the same analysis on the Belot dataset. We found the same patterns (compare the two panels of Figure 4). The analysis revealed a significant main effect of the outcome of the previous game, $OR = 1.16$, $Z(8205) = 3.46$, $p < .001$, again, consistent with CR. The main effect of role, and the interaction of role and previous game outcome also showed the same pattern as in our dataset, although, they fell short of statistical significance, $OR = 0.92$, $Z(8205) = -1.87$, $p = .061$ and $OR = 1.16$, $Z(8205) = 1.61$, $p = .107$ respectively. Nonetheless, there was no significant difference between the Aczel and the Belot dataset neither in the effect of role, $OR = 1.07$, $Z(20861) = 1.15$, $p = .248$, nor in the interaction of role and previous outcome, $OR = 0.95$, $Z(20857) = -0.44$, $p = .661$. Thus, we can conclude that the Aczel and the Belot datasets display the same sequential patterns: 1) sighted players tend to stick with their previous gesture if they have won in the previous game, and they change their gesture if they have lost, while 2) blind players tend to repeat their previous gesture, especially when they have won in the previous game.

To further investigate this issue we built another model to see whether players matched their opponent's previous gesture. This model contained the effect of the outcome of the previous game (win or lose), and the effect of role (sighted or blind) and task (to match or to mismatch) in the current game, just as in the case of the previous model. During the model selection based on AIC and BIC values all three predictors, and their interactions were

retained in the final model. The final model revealed a significant three-way interaction between the task and role of the player and the outcome of the preceding game, $OR = 1.83$, $Z(12646) = 4.19$, $p < .001$. As also apparent in the Aczel dataset panel of Figure 5, post-hoc analysis showed that sighted players are inclined to match their opponent's previous gesture if they have a task to match, while they tend to mismatch the opponent's previous gesture when they have an incentive to mismatch (similar effect of task is present on the subsample of sighted players who won the previous game, $OR = 0.84$, $Z(3213) = -2.43$, $p = .015$, and who lost previously, $OR = 0.62$, $Z(3108) = -6.39$, $p < .001$). The same significant effect of task can be seen for blind players who won the previous game, $OR = 0.76$, $Z(3108) = -3.76$, $p < .001$, but there was no such task-specific imitation of the opponent's previous gesture for blind players who lost previously, $OR = 1.05$, $Z(3213) = 0.68$, $p = .500$.

[INSERT Figure 5. HERE]

Again, we tested the same model on the Belot dataset, and found the same pattern as in our dataset (compare the two panels of Figure 5). That is, we found a significant three way interaction of role and task in the current game and outcome of the previous game on the matching probability of the opponent's previous gesture, $OR = 1.64$, $Z(8196) = 4.44$, $p < .001$. Just as in our dataset, sighted players who won ($OR = 0.83$, $Z(2082) = -2.00$, $p < .046$), and who lost in the previous game ($OR = 0.75$, $Z(2014) = -3.15$, $p < .002$), and blind players who won in the previous game ($OR = 0.82$, $Z(2014) = -2.18$, $p < .029$) all had a tendency to repeat their opponent's previous gesture when they had an incentive to match, and tended to avoid their opponent's previous gesture when they had an incentive to mismatch. Also in accord with finding in our dataset, there was no effect of task on the matching of the opponent's previous gesture, on blind players who lost in the previous game, $OR = 1.04$,

$Z(2082) = 0.46, p < .646$. Thus once again we draw the same conclusions from the two datasets.

Note that the two phenomena found here, the general tendency of blind players to imitate their own previous gestures and the tendency of sighted players to imitate their opponent's previous gesture when they have a task to match and to mismatch their opponent's previous gesture when they have an incentive to mismatch can result in systematic bias toward matching or mismatching depending on the task of the sighted player: The tendency of a sighted player with an incentive to match to repeat the opponent's previous gesture, together with the tendency of blind players to stay with their previous gesture would result in a higher probability of matching in the current game. Conversely, the tendency of sighted players with a task to mismatch to avoid the opponent's previous gesture together with blind players' tendency to imitate their own preceding gestures can result in a deflation of matching probability, which would make it more difficult to detect automatic online imitation in sighted players with an incentive to mismatch.

Controlling for effects of previous gestures

Based on our analysis it seems that tendency of blind players to stick with their previous gesture is mostly present when they have won in the preceding game, $OR = 1.14$, $Z(3110) = 2.42, p < .015$, while there is no significant bias toward matching or mismatching their own previous gesture in blind players who lost previously, $OR = 1.03$, $Z(3215) = 0.60, p < .546$. Similarly, it seems that blind players who lost in the previous game have no tendency to match the opponent's gesture either, $OR = 0.99$, $Z(3215) = -0.27, p < .786$. These results suggest that blind players who have lost in the previous game are not affected by their own or their opponent's previous gesture. Thus, we used a subsample of our data in which the blind player lost in the preceding game, to control for the above mentioned task-relevant bias in

matching. In this subsample of our data, after filtering for irrelevant trials for online imitation (trials where the sighted player responded with less than 133 ms delay) we found that the significant bias toward matching of the current gestures was no longer present, $OR = 1.03$, $Z(606) = 0.36$, $p < .717$. Furthermore, a significant effect of task was revealed on the matching probability, $OR = 0.74$, $Z(605) = -1.79$, $p < .037$ (one tailed), indicating that games in which sighted player had an incentive to match displayed a higher matching probability compared to games in which the sighted players had to mismatch. Just like in the previous analysis, in this subsample of trials there was again no indication of online imitation in games where the sighted player had the objective to mismatch, $OR = 0.90$, $Z(338) = -0.98$, $p = .329^7$. In games where the sighted player had an incentive to match, the matching frequency decreased in this subsample of data compared to the analysis where the biasing effect of previous gestures were not controlled for, OR from 1.28 to 1.22, and the matching probability was no longer statistically significantly higher than 50% ($OR = 1.22$, $Z(226) = 1.56$, $p = .060$, one tailed).

Additional Bayesian analysis was conducted on these data, since Bayes Factor can indicate the level of confidence that one can place in the null-hypothesis based on the data, thus helping the interpretation of null-significant results (Dienes, 2011). According to this analysis⁸, the results strongly argued against the presence of automatic imitation, $B_{(mismatching < matching)} = 0.186$, where a value below 1/3 indicates that the data provide substantial evidence for the null (Jeffreys, 1961).

Discussion

⁷ Nor was there any indication for automatic online imitation in the sub-sample analyses by delay-quartiles in the subsample of trials in which the blind player lost in the preceding round (all uncorrected $ps > .05$).

⁸ The prior of our model was calculated based on the method of Rouder, Morey, Speckman, & Province (2012), using their 'BayesFactor' R package (Morey & Rouder, 2014).

In this study we analyzed the current methodology used in automatic imitation research of strategic context situations. Strategic context games are ideal tests of automatic imitation as automatic and intentional control can be pitted against each other. We introduced video recording in our experiment, which followed the methodology of Benet et al.'s (2013) Matching Pennies experiment closely, where participants in certain conditions had explicit incentives against imitating the other player. As argued above, only imitative behavior produced against incentives to not imitate can be regarded as convincing evidence for automaticity.

In the analysis of the video recordings we identified several previously unrecognized factors that can result in a severe decrease of sensitivity in the measures and loss of power in the analysis. Firstly, it became apparent from the recordings that some of the sighted players consistently averted their gaze from the blindfolded player for a prolonged period of time, leaving the records unusable for the analysis of imitative behavior. Whether not keeping one's eyes on the blind opponent is a cultural or strategic behavior is unclear, but further studies should try to decrease this source of noise in the experiments.

Secondly, the frame-by-frame analysis of the recordings revealed that 81.52 % of the data was irrelevant for the study of automatic behavior, as either the blind players initiated the gesture later than the sighted players or the delay after their sign was too short for the sighted players to react. Based on previous explorative studies, we argue that it is not physically possible to detect the gesture of the opponent and execute motor reaction within ~150 ms. Trials with delays longer than this limit from when the blind players initiated their gestures were found in only 18.48 % of all the trials (still hundreds of games). This result highlights the importance of filtering out the irrelevant records from the dataset, as it is an

effective way of decreasing the probability for beta-error and thus to increase power. In our case, we decreased the probability of beta-error this way by a magnitude of 5.7.

Thirdly, we identified several biasing effects independent of automatic imitation potentially affecting the matching behavior of the players in the Matching Pennies game. One of these effects is the win-stay lose-shift (also called Pavlov) conditional response strategy (Nowak & Sigmund, 1993), which has been already found to give a good prediction of player behavior in the RPS game (Wang et al., 2014). However, the blind players tended to follow a different strategy and did not show a tendency to shift their previous gesture when they lost, while they still tended to stay with their previous gesture when they won, which produced an overall tendency of self-repetition in the blind players' gestures. Another effect influencing the players' current gesture was that blind players who won the previous game and sighted players (regardless of the outcome of the preceding game) seemed to repeat the previous gesture of the opponent when they had an incentive to match, while in the incentive to mismatch condition they mismatched the previous gesture of the opponent. This behavior is more likely the result of a priming effect rather than an overall tendency for imitation proposed by Cook, Bird, Lünser, Huck, & Heyes (2012) or Belot et al. (2013), as it works both in the direction of matching and mismatching. These two biasing effects can possibly result in an increase of frequency of matches in the 'to match' condition and an increase in the frequency of mismatches in the 'to mismatch' condition, both independently of automatic online imitation. These same effects of previous gestures on current gestures were verified on the dataset from the study of Belot et al. (2013), which makes it unlikely that these results would be chance findings. Altogether, we found that the previous gesture presented by either the player or the opponent affects the gesture of the player and these effects can either mask

the effect of automatic online imitation or produce matching patterns in the next games artificially.

Finally, traditional statistical methods employed on aggregated data of participant triads can lead to multiple statistical problems, among others, a massive loss of degrees of freedom and, thus, statistical power. More recent statistical methods (e.g., mixed effect models) can overcome these problems by incorporating random effects into the statistical models so that performance can be assessed at the level of the individual. Had previous studies detected a true effect of automatic imitation on aggregated and unfiltered data, we would have been more likely to find evidence for it with increased power.

The analysis of the filtered data firstly revealed that sighted players were able to react to gestures of the blind players, albeit only when they were instructed to match and only when the delay following the gesture of the blind player was at least ~133 ms, and even this effect decreased and lost statistical significance in the analysis where we controlled for effects of gestures from the preceding game. We tested Belot et al.'s hypothesis that automatic responses would occur early on after the presentation of the gesture of the opponent, while intentional reaction requires longer delay. However, no such effect of the length of the delay was found in our study. We also analyzed whether the quantity of experience in the game has an effect on the behavior of the participant. One could expect that explicit strategies can be learned during the game of how to avoid imitating the opponent in the mismatching condition. The results did not provide evidence for learning during the course of the game, thus the effect of experience could not explain why we found no evidence for automatic imitation. Bayesian analysis suggests that the data should strengthen our confidence in the null-hypothesis of no effect.

These findings are in accord with the predictions of the Theory of Goal-Directed Imitation (Wohlschläger et al., 2003) which assumes that when higher-level action-interpretation is available then imitation is goal-directed. In this sense, the action of the model is never an uninterpreted signal; instead it is mirrored only if the actor does not find a better means to achieve her goal. Evidence that the participants with incentive to mismatch show imitative reactions would count against the theory since action mirroring would not correspond with the actor's goal. Although this theory of imitation does not declare whether quick reactions could escape from this top-down control allowing for direct matching, the present results do not indicate this interpretation.

A message of the present study is that strategic games provide a potentially promising method of exploration for this topic; however certain methodological considerations are to be highlighted. Our analysis drew attention to the phenomenon that the presently shown gesture can be affected by the previously shown gestures. This sequential effect might not be fully intentional; yet our post-experimental interviews suggest that the participants were often trying to outwit their opponents by predicting their choice. Following an explicit strategy of choosing the gesture before the signal of the umpire could have worked against the effect of imitation. We believe that the biasing effect of the previously presented gestures could be decreased if both players were sighted in the experiment. Blindfolding one of the players in the previous experimental designs was to control which player could take advantage of the gesture asynchrony. Video recording the behavior of the players can prevent this problem and could also provide several benefits in the data collection of future studies. Firstly, without blindfolding one of the players sequential effects should decrease or equally affect both players. Secondly, this change in the design could more than double the number of trials in the dataset, since blindfolded players usually presented their gestures earlier than sighted

players, thus these data had to be excluded from the analysis. Thirdly, this change would also increase the ecological validity of the game, so we could test the appearance of automatic imitation in more naturalistic settings. Nevertheless, one disadvantage of the video recording method is that the evaluation of the recordings is very cumbersome. The application of motion capture systems could probably decrease the required workload.

For further studies, we also suggest that the time delay between the action of the observed and the observer should be measured and controlled. Video recording appeared to provide sufficient information to narrow our focus to the most relevant data. Even though we found no evidence for automatic imitation in this experiment, Figure 3 would suggest that further investigations could attempt to locate the effect at around 200ms. Besides recording the game of two sighted players, the increase of the number of relevant trials could be achieved by delaying the signal to initiate action for one of the participants⁹. This way, through a significant increase in trials with delay it would also be possible to perform even more fine-grained analyses on the effect of delay length on imitation.

In conclusion, despite our attempts to execute a meticulous search for automatic imitation, our present results failed to support evidence that the observed imitative behavior would be of automatic nature. Results supporting the assumption that, as other animals, humans also show an unintentional tendency to imitate others could truly become an important step in understanding the underlying mechanisms behind a range of cognitive functions. Nevertheless, critical examination of previous studies and thorough analysis of the empirical results suggests that it would be premature to conclude the presence of automatic imitation in adult humans. While subsequent studies with advanced sensitivity could seek further for evidence of automatic imitation in human behavior, the question remains whether

⁹ For example by pre-recorded instructions displayed via headphones.

robust human functions such as social learning or language acquisition could rely on such an evanescent phenomenon.

References

- Aczel, B., Bago, B., & Foldes, A. (2012). Is there evidence for automatic imitation in a strategic context? *Proceedings of the Royal Society B: Biological Sciences*, 279(1741), 3231–3233.
- Bates, D., Maechler, M., & Dai, B. (2008). The lme4 package. *Computer Software Manual*. Retrieved from [Http://cran.R-Project.org/web/packages/lme4/lme4.Pdf](http://cran.R-Project.org/web/packages/lme4/lme4.Pdf) on 2014-06-01.
- Bekkering, H., Wohlschläger, A., & Gattis, M. (2000). Imitation of gestures in children is goal-directed. *The Quarterly Journal of Experimental Psychology: Section A*, 53(1), 153-164.
- Belot, M., Crawford, V. P., & Heyes, C. (2013). Players of Matching Pennies automatically imitate opponents' gestures against strong incentives. *Proceedings of the National Academy of Sciences*, 110(8), 2763–2768.
- 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(2), 210.
- Bien, N., Roebroek, 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.
- Boyer, T. W., Longo, M. R., & Bertenthal, B. I. (2012). Is automatic imitation a specialized form of stimulus–response compatibility? Dissociating imitative and spatial compatibilities. *Acta Psychologica*, 139(3), 440–448.
- Brass, M., Bekkering, H., & Prinz, W. (2001). Movement observation affects movement execution in a simple response task. *Acta Psychologica*, 106(1), 3–22.

- 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(2), 124-143.
- Brass, M., & Heyes, C. (2005). Imitation: is cognitive neuroscience solving the correspondence problem? *Trends in Cognitive Sciences*, 9(10), 489–495.
- Campione, G. C., & Gentilucci, M. (2011). Is the observation of the human kinematics sufficient to activate automatic imitation of transitive actions? *Behavioural Brain Research*, 225(1), 201–208.
- Catmur, C., & Heyes, C. (2011). Time course analyses confirm independence of imitative and spatial compatibility. *Journal of Experimental Psychology: Human Perception and Performance*, 37(2), 409.
- Chong, T. T.-J., Cunnington, R., Williams, M. A., & Mattingley, J. B. (2009). The role of selective attention in matching observed and executed actions. *Neuropsychologia*, 47(3), 786–795.
- Cook, R., Bird, G., Lünser, G., Huck, S., & Heyes, C. (2012). Automatic imitation in a strategic context: players of rock–paper–scissors imitate opponents’ gestures. *Proceedings of the Royal Society B: Biological Sciences*, 279(1729), 780–786.
- Csibra, G. (2007). Action mirroring and action understanding: An alternative account. *Sensorimotor Foundations of Higher Cognition: Attention and Performance*, 22, 435–480.
- Csibra, G., & Gergely, G. (2007). “Obsessed with goals”: Functions and mechanisms of teleological interpretation of actions in humans. *Acta Psychologica*, 124(1), 60–78.
- Dienes, Z. (2011). Bayesian versus orthodox statistics: Which side are you on? *Perspectives on Psychological Science*, 6(3), 274–290.

- Dixon, P. (2008). Models of accuracy in repeated-measures designs. *Journal of Memory and Language*, 59(4), 447–456.
- Edwards, M. G., Humphreys, G. W., & Castiello, U. (2003). Motor facilitation following action observation: A behavioural study in prehensile action. *Brain and Cognition*, 53(3), 495–502.
- Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (2000). Visuomotor neurons: Ambiguity of the discharge or “motor” perception? *International Journal of Psychophysiology*, 35(2), 165–177.
- Heyes, C., Bird, G., Johnson, H., & Haggard, P. (2005). Experience modulates automatic imitation. *Cognitive Brain Research*, 22(2), 233–240.
- Iacoboni, M. (2005). Understanding others: Imitation, language, empathy. *Perspectives on Imitation: From Cognitive Neuroscience to Social Science*, 1, 77–99.
- 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.
- Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, 30(5), 513–541.
- Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, 59(4), 434–446.
- Jansson, E., Wilson, A. D., Williams, J. H., & Mon-Williams, M. (2007). Methodological problems undermine tests of the ideo-motor conjecture. *Experimental Brain Research*, 182(4), 549–558.
- Jeffreys, H. (1961). *Theory of probability*. Oxford, UK: Oxford University Press.
- Jiménez, L., Recio, S., Méndez, A., Lorda, M. J., Permuy, B., & Méndez, C. (2012). Automatic imitation and spatial compatibility in a key-pressing task. *Acta Psychologica*, 141(1), 96–103.

- Kilner, J., Hamilton, A. F. de C., & Blakemore, S.-J. (2007). Interference effect of observed human movement on action is due to velocity profile of biological motion. *Social Neuroscience*, 2(3-4), 158–166.
- Kirchner, H., & Thorpe, S. J. (2006). Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. *Vision Research*, 46(11), 1762–1776.
- Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: cognitive basis for stimulus-response compatibility--a model and taxonomy. *Psychological Review*, 97(2), 253.
- Longo, M. R., & Bertenthal, B. I. (2009). Attention modulates the specificity of automatic imitation to human actors. *Experimental Brain Research*, 192(4), 739–744.
- Meltzoff, A. N., & Decety, J. (2003). What imitation tells us about social cognition: a rapprochement between developmental psychology and cognitive neuroscience. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 358(1431), 491–500.
- Mengotti, P., Ticini, L. F., Waszak, F., Schütz-Bosbach, S., & Rumiati, R. I. (2013). Imitating others' actions: transcranial magnetic stimulation of the parietal opercula reveals the processes underlying automatic imitation. *European Journal of Neuroscience*, 37(2), 316–322.
- Nowak, M., & Sigmund, K. (1993). A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. *Nature*, 364(6432), 56–58.
- Press, C., Bird, G., Flach, R., & Heyes, C. (2005). Robotic movement elicits automatic imitation. *Cognitive Brain Research*, 25(3), 632–640.
- 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.

- Prinz, W. (2002). Experimental approaches to imitation. In W. Prinz & A. N. Meltzoff, *The imitative mind: Development, evolution and brain bases* (Vol. 6, pp. 143–162). Cambridge University Press.
- Quené, H., & Van den Bergh, H. (2008). Examples of mixed-effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language*, 59(4), 413–425.
- Shanks, D. R., & St John, M. F. (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences*, 17(03), 367–395.
- Stadler, M. A., & Roediger III, H. L. (1998). The question of awareness in research on implicit learning. In M. A. Stadler & P. A. Frensch (Eds.), *Handbook of implicit learning* (pp. 105–132). Thousand Oaks, CA, US: Sage Publications, Inc.
- Stürmer, 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.
- Vanrullen, R., & Thorpe, S. J. (2001). The time course of visual processing: from early perception to decision-making. *Journal of Cognitive Neuroscience*, 13(4), 454–461.
- 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.
- Wang, Z., Xu, B., & Zhou, H.-J. (2014). Social cycling and conditional responses in the Rock-Paper-Scissors game. *arXiv Preprint arXiv:1404.5199*.
- Wohlschläger, A., & Bekkering, H. (2002). The role of objects in imitation. *Advances in Consciousness Research*, 42, 101–114.

Wohlschläger, A., Gattis, M., & Bekkering, H. (2003). Action generation and action perception in imitation: an instance of the ideomotor principle. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 358(1431), 501–515.

Figure Captions

Figure 1. Hypothetical effects of automatic and intentional control over matching and mismatching behavior in the Matching Pennies game, based on Belot et al. (2013).

Figure 2. The neutral position (a), the Open Hand (b) and the Closed Hand (c) gestures.

Figure 3. The estimated imitation effects in the 'to match' and 'to mismatch' conditions by delay length using logit regression coefficients with corresponding ± 1.96 SE. The figure shows a matching tendency emerging in gestures of the sighted player presented as early as 133ms after the blind player's gesture, when the sighted player has an incentive to match the opponent's gesture.

Figure 4. The effect of the players' own previous gesture on their current gesture with the effect of role (sighted or blind) in the current game and outcome of the previous game (win and loss). Positive estimate indicate the players' tendency to repeat their previous gesture. Error bars depict ± 1.96 SE. The two panels represent results from the current study (Aczel dataset), and the study performed by Belot et al. (2013) (Belot dataset). The figure shows that both sighted and blind players tend to stick with their previous gesture if they have won in the previous game, and that sighted players tend to change their gesture if they have lost, while blind players have no such preference for avoiding their previous gesture when they have lost in the preceding game.

Figure 5. The effect of the opponent's previous gesture on the player's current gesture with the effect of role (sighted or blind) and task (to match or to mismatch) in the current game and outcome of the previous game (win and loss). Positive estimate indicate the players' tendency to repeat the opponent's previous gesture. Error bars depict ± 1.96 SE. The two panels represent results from the current study (Aczel dataset), and the study performed by

Belot et al. (2013) (Belot dataset). Blind players who won the previous game and sighted players (regardless of the outcome of the preceding game) tend to repeat the previous gesture of the opponent when they have an incentive to match, while they avoid the opponent's previous gesture in the incentive to mismatch condition. This effect is not present in blind players who lost in their preceding game.