Comprehension of Action Sequences: The Case of Paper, Scissors, Rock

Patric Bach* (bach@mpipf-muenchen.mpg.de)
Günther Knoblich* (knoblich@mpipf-muenchen.mpg.de)
Angela D. Friederici° (orendi@cns.mpg.de)
Wolfgang Prinz* (prinz@mpipf-muenchen.mpg.de)

*Max Planck Insitute for Psychological Research; Amalienstr. 33, 80799 Munich, Germany
° Max Planck Institute for Cognitive Neuroscience, Stephanstr. 1a, 04303 Leipzig, Germany

Abstract

The basic idea of the present study is that it is useful to conceptualize the processes involved in action comprehension in a similar manner as the processes involved in sentence comprehension. One important question then is whether order and meaning of action sequences are processed sequentially or in parallel (analogous to syntactic and semantic processing in sentence comprehension). We conducted three experiments to address this issue. Participants were asked to detect violations of order, meaning, or both, in action sequences of the game Paper, Scissors, and Rock. The main results were that it took longer to detect single violations than double violations and that it was impossible to ignore any of the violations. This pattern of results suggests that the processes involved are highly automatic and that they are running in parallel.

Have you ever wondered why people in sports bars stare at the numerous TV-screens that are silently picturing sequences of actions like people in fancy dresses bumping forcefully into one another, or two people hitting a ball with their rackets? Obviously, in some situations observing other people's actions is more interesting than chatting. This is not only true for sports bars but also for French street cafés and many other places. But how do we understand what happens while we watch other people acting?

Whereas cognitive science has been making much progress regarding the processes involved in language comprehension (Friederici, Steinhauer, & Frisch, 1999; Garrod & Pickering, 1999; MacDonald et al., 1994; Rayner & Pollatsek, 1989), action comprehension has been hardly addressed. Although Schank and Abelson’s (1977) script theory seems relevant, it mainly deals with the comprehension of extended texts and not action comprehension as in the examples above. Another well-known approach, the grammar of action (Goodnow & Levine, 1973), addresses only the construction of complex movements from simpler units in writing. Our research is guided by the idea that it might be useful to conceptualize the processes involved in action comprehension in a similar fashion as the processes involved in sentence comprehension.

Sentence and Action Comprehension

In order to understand a sentence it is necessary to parse it according to the grammatical rules that governed its formation and to understand the meaning of the words given the context of the sentence. Likewise, to understand an action sequence it is often necessary to parse the sequence according to rules that constrain the order in which one action follows another and to understand the meaning of single actions given the context of the sequence. Although there might be action sequences the order of which is hardly constrained, most of them follow rules that can be spelled out clearly.

The main question for action comprehension (as for sentence comprehension) then becomes how processing of the order of action sequences (in analogy to syntactic processing) and the processing of its meaning are related. There are essentially four possibilities: (1) The analysis of order generally precedes and influences the analysis of meaning. (2) The analysis of meaning generally precedes and influences the analysis of order. (3) The two processes run in parallel and do not affect each other. (4) They run in parallel and they do affect each other. Controversies in sentence comprehension have mainly focused on the possibilities that are analogous to 1 and 4. The garden path model (Frazier & Rainer, 1982) suggests that meaning does not influence the selection of the initial syntactic structure (syntax first). The connectionist theory of MacDonald et al. (1994) states that syntactical and semantic constraints are narrowing the possible interpretations of a sentence in parallel and interactively. However, (2) and (3) are also viable possibilities for action comprehension. For instance, in favor of (2) one could say that the meaning of a single action of a peer might be most relevant for organisms (for instance because they can be dangerous or not) and that therefore the action comprehension system analyses meaning before order. The experiments that follow focused mainly on the question of whether the order and meaning of an action sequence are analyzed in parallel or whether there is any sequential order of the two processes.

Experimental Paradigm

In our experiments, we used a simple action domain, the game Paper, Scissors, and Rock. In this game, the order of consecutive actions is well defined and only a small set of gestures is meaningful in its context. Moreover, the game is well known (at least in Germany), so that people are quite familiar with the rules. The game has a fixed sequence (see Figure 1): In the beginning, the two players hold their hands close to the chest (upper position) and form a fist. In the next step, they drop their fists to a position near their waists (lower position) synchronously, and lift them again. This pattern is repeated once. During the third downward movement both players form one of three gestures: paper, scissors, or rock. The winner is determined by the following rules: Rock dulls scissors (rock wins), scissors cut paper (scissors win), paper wraps rock (paper wins). There are three meaningful gestures, that is, paper, scissors, and rock. Other gestures produce violations of context, as for instance the thumbs up gesture. It is also easy to see what a violation of order might look like. There is only one valid structure that is defined by alternating hand positions:

Upper, lower, upper, lower, upper, lower.

Therefore violations of order can be introduced by changing the sequence in the following way:

Upper, lower, upper, lower, upper, *upper.*

Obviously, violations of context and order are independent of one another. A violation of order might occur, although a meaningful gesture is formed (e.g. paper in the upper position, see leftmost picture in Figure 2). Alternatively, a meaningless gesture can be formed without violating order (e.g. thumbs up at the lower position (see picture in the middle of Figure 2). In addition, both violations can occur at the same time as when displaying a meaningless gesture at the wrong position (e.g. thumps up at the upper position, see rightmost picture in Figure 2). By comparing reaction times for different types of violations one can determine how order and meaning are analyzed in this type of action sequence.

Experiment 1

The aim of Experiment 1 was twofold. The first aim was to determine whether it is easier to detect violations of order than violations of context. If processing the order of an action sequence is similar to syntactic processing, it might be completed before the meaning of an action is fully retrieved or it might even be a necessary condition for the analysis of meaning. Hence people might be faster in detecting violations of order than violations of context. If, on the other hand, the processing of meaning is prevalent in action comprehension, violations of context should be faster detected.

The second aim was to determine whether order and meaning are processed in parallel. Faster detection of either violation does not necessarily imply that sequence and meaning are analyzed in a serial fashion. Rather, the differences could be due to different processing times taken by two parallel processes. If order and meaning are processed in parallel it should take less time to detect double violations than either single violation. The rationale behind this assumption is that the process that is completed first will trigger the reaction indicating that the violation was detected. Alternatively, evidence accumulated by each process might converge to trigger the wrong reaction. If order and meaning of action sequences are processed sequentially, detecting a double violation should take the same time as detecting the single violation (of either order or context) that is detected fastest.

One problem with measuring RTs for the detection of order and context violations in action sequences is that the onset of the violation is not well defined. When watching a video film displaying a person playing paper, scissors, and stone, the syntactic violation may be detected as soon as the expected downward movement does not occur. The semantic error can only be detected as soon as a gesture is formed. To avoid these problems, we used still frames that displayed the upper and lower position of the hand three times, respectively. The changing frames create the impression of a movement and they are readily interpreted as action sequences (Stürmer, Aschersleben, & Prinz, 2000). With still frames, syntactic, semantic, and double violations have the same onset, that is, the onset of the frame that completes the sequence at the lower position.

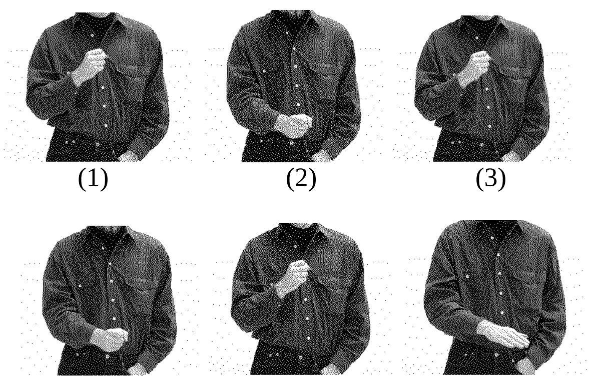
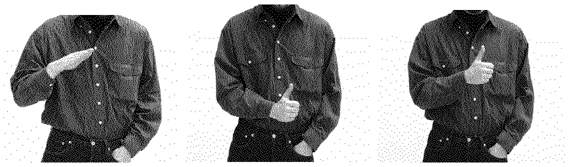


Figure 1: Sample sequence of pictures (correct).

Method

Participants. Sixteen participants, all students at the University of Munich, took part in the experiment, 9 of them male. They ranged in age from 24 to 35 years. The key assignment was counterbalanced across participants.



(a) v. of order (b) v. of context (c) double violation

Figure 2: Sample for last frame in action sequences with violation of order, meaning, or both.

Material. The material consisted of action sequences showing still frames of a person playing paper, scissors, and rock. There were six pictures displaying the actor forming legal gestures (Rock, Scissors, Paper), and six displaying him forming illegal gestures. The first five pictures of a sequence always showed a closed fist alternating between the upper and lower position. The sixth picture displayed one of the six possible gestures in the upper or lower position (see Figure 1 for a complete legal sequence). Hence, there were sequences ending with (1) a legal gesture at the lower position (correct), (2) an illegal gesture at the lower position (violation of context), (3) a legal gesture at the upper position (violation of order), and (4) an illegal gesture at the upper position. Figure 2 shows examples for the last frame in the violation conditions.

Procedure and Design. The experiment consisted of two blocks of 144 trials each. The order of action sequences was randomized in each block. To avoid response bias, correct action sequences were displayed as often as incorrect sequences. Hence there were 72 correct action sequences in each block, and 24 containing a violation of order, context, or both, respectively. The course of each trial was as follows: The first five frames were displayed for 500 ms, respectively. The sixth, critical frame was displayed for 1000 ms. The reaction time interval started with the onset of the last frame and lasted as long as the gesture was displayed (1000 ms). The participants judged whether there was a violation or not and pressed a left or a right key, accordingly. If they committed an error or did not react within the given time interval of 1000 ms, they were given error feedback.

Results and Discussion

Figure 3 displays the RTs for the detection of different types of violations during the first and second block of trials. Double violations (M = 540, S = 56) were faster detected than either violations of order (M = 569, S = 71) or violations of context (M = 602, S = 57). Moreover, violations of order were faster detected than violation of context. These differences were present in both blocks, although RTs were generally faster in Block 2 (M = 559 ms, S = 66) than in Block 1 (M = 580 ms, S = 55). The mean RTs for responses to correct action sequences were 566 ms (S = 58). They were not included in the statistical analysis because they are hard to compare to RTs for violations because of the different base rate and the different type of response.

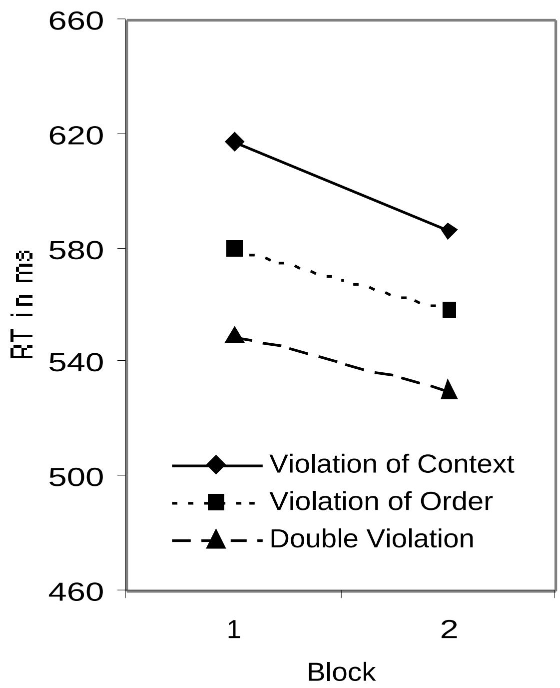


Figure 3: RTs for detecting different types of violations.

The RTs for the three error conditions were entered into a 3 x 2 repeated measurement ANOVA with the factors Type of Violation (Order, Context, vs. Double) and Block (1st vs 2nd). The analysis revealed significant main effects for Block, F (1, 15) = 9.9, p < .01, and Type of Violation, F(2, 30) = 28.4, p < .001, but no significant interaction. The RTs differed significantly for different types of violations. Duncan tests showed that double violations were detected significantly faster than either violations of order or context (both p < .001). In addition, there was also a significant difference between the two types of single violations (p < .001).

The pattern of RTs is in favor of a parallel processing explanation. Double violations were detected faster and more reliably than either violations of sequence or context. This is in favor of the claim that either of two parallel processes being completed first can trigger the reaction indicating that a violation was detected. Order violations were faster detected than context violations right from start. Hence, it seems that analyzing the order of a sequence is completed before the meaning of an action is fully processed. An alternative explanation for the different speed with which violations of order and context can be detected is perceptual saliency. Violations of order might have been easier to detect because a wrong hand position is perceptually more salient than a hand forming a different gesture.

Experiment 2

The second experiment was conducted to rule out the perceptual saliency explanation. To do so, we replicated Experiment 1 using a set of stimuli in which the perceptual saliency of order and context violations was more comparable. The deep structure of the action sequence, five alternations of the same gesture and formation of the target gesture with the last alternation, did not change. However, the alternations were now expressed as hand turnings instead of arm movements (see Figure 4). Hence both, violations of order and context now depended on the configuration of hand and fingers. Violation of order consisted in a missing turn of the hand, violations of context consisted in invalid gestures, as in Experiment 1 (see Figure 5).

If higher perceptual saliency caused the faster detection times for violations of order in Experiment 1 the detection times should be slower or equally fast as for violations of context under the new conditions. If faster processing of sequence information caused the difference the results should be the same as in Experiment 1. It is quite likely that such a difference would not occur right from start because people have to link the new type of alternation (hand turning) to the familiar deep structure of the sequence before optimal detection of order violations becomes possible. Hence, in order to allow this link to be established we added a further block of trials. For double violations we expected the same pattern as in Experiment 1, that is, faster detection than for either single violation.

Method

Participants. Fifteen participants, all students at the University of Munich, took part in the experiment, five of them male. They ranged in age from 24 to 35 years.

Material. A different set of stimuli was used. The fixed pattern of alternations consisted of hand turnings while the hand position remained fixed. The critical manipulation occurred with the last frame. The gesture displayed could either be correct, a violation of order, context, or both (see Figure 5 for examples). There were three different valid gestures and three that violated the context. There were two versions of each gesture, one displaying the face of the hand (correct) and one displaying the backside (violation of order).

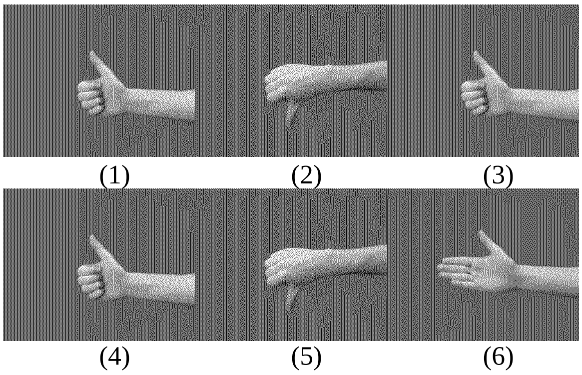


Figure 4: Sample sequence of pictures (correct).

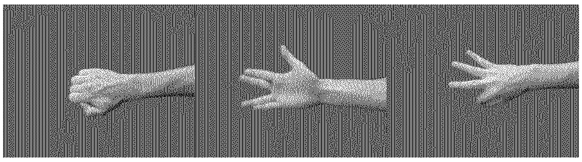


Figure 5: Examples for violation of order (left: wrong orientation), context (middle), and double violation (right) in Experiment 2.

Procedure and Design. The procedure was the same as in Experiment 1 with one exception. A third block of 144 trials was added.

Results and Discussion

Participants pressed the wrong button in about 3% of the cases and reacted too late in about 2% of the cases. Figure 3 displays the RTs for the detection of different types of violations during the first, second, and third block of trials. Overall, it took participants longer to detect violations of context (M = 615, S = 64) than violations of order (M = 596, S = 76). However, this difference occurred only after the first block. As expected, double violations were faster detected than either single violation right from start (M = 561, S = 70). Participants became faster during later blocks. Mean RTs were 617 (S = 56), 584 (S = 63), and 572 (S = 70) during Block 1, 2, and 3.

The RTs of the three error conditions were entered into a 3 x 3 repeated measurement ANOVA with the factors Type of Violation (Order, Context, vs. Double) and Block (1, 2, vs. 3). There were significant main effects for Type of Violation, $F(2, 28) = 14.7$, $p < .001$, and Block, $F(2, 28) = 20.9$, $p < .001$, the interaction between both factors was only marginally significant, $F(4, 56) = 2.0$, $p = 0.1$. Duncan-tests showed that the RTs for violation of order were significantly faster than for violations of context in Block 2 and 3, but not in Block 1. The difference between double violations and either single violation was significant in each Block.

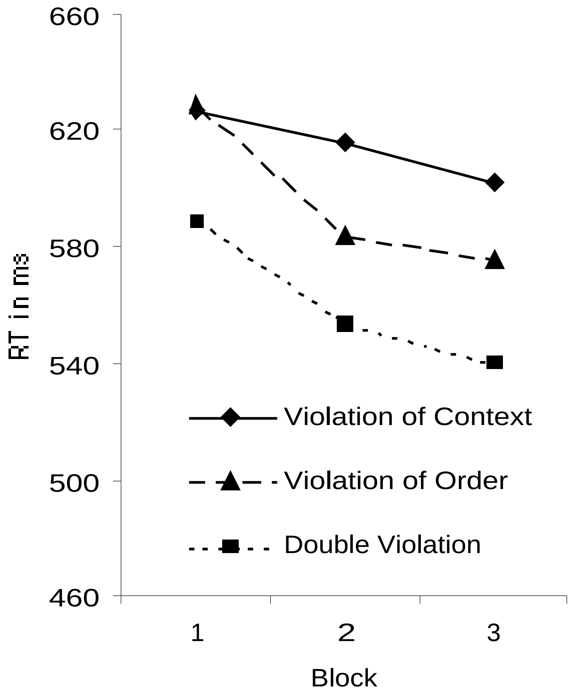


Figure 6: RTs for detecting different types of violations.

The results basically replicate Experiment 1. After Block 1, the RTs were almost numerically identical to those observed in the first experiment. This suggests that it was not perceptual saliency that created the RT differences between violations of order and context. Rather it seems that sequence information is faster processed, at least for simple sequences as in paper, scissors, and rock. The lack of difference during the first block is probably due to the necessity of establishing a link between the new perceptual cues and the familiar deep structure. The finding that double violations were faster detected right from start further supports the claim that order and meaning of action sequences are processed in parallel.

Experiment 3

It has often been claimed that parallel processes do not require attention and that they are highly automatic. Experiment 3 was conducted to determine whether this is also true for the processes involved in analyzing the order and meaning of action sequences. We used an interference paradigm to provide evidence for the automatic nature of the processes involved. The first block of Experiment 3 was identical to that of Experiment 2. In the following blocks we asked participants to ignore violations of order or context while detecting the other type of violation, respectively. In a first block, they were asked to detect violations of order and to ignore violations of context, that is, sequences containing the latter violation should be judged as correct. In the other block they were asked to detect violations of context and to ignore violations of order, that is, sequences containing a violation of order should be judged as correct.

If order and meaning of action sequences are analyzed automatically the violations to be ignored should interfere with judging a sequence as correct. If the processes are not automatic there should be no differences in RTs between action sequences containing an interfering violation and truly correct action sequences. Moreover, if sequence processing is faster than the processing of meaning a violation of order should interfere more strongly with the correct response than a violation of context.

Method

Participants. 20 participants, all students at the University of Munich, took part in the experiment, 8 of them male. They ranged in age from 24 to 35 years. The order of blocks was counterbalanced.

Material. The material used was identical to that of the second experiment.

Procedure and Design. The experiment consisted of three blocks of 144 trials each. The first block was the same as in Experiment 2. At the beginning of the second and third block participants received an instruction to attend to one of the single violations only and to ignore the other (e.g. detect violation of context/ignore violation of order). During the third block participants attended to the other type of violation and ignored the other (e.g. detect violation of order/ignore violation of context). The order of blocks 2 and 3 was counterbalanced across participants.

Results and Discussion

Participants pressed the wrong button in about 1% of the cases and reacted too late in about 3% of the cases. The RTs for different types of violations during the first block were indistinguishable from Experiment 2. Because of the space constraints, we will report only RTs for trials in which action sequences should be judged as correct. Figure 7 shows the results.

RTs were generally faster for correct sequences (M = 522, S = 59) than for sequences that contained an interfering violation (M = 580, S = 82). The difference was larger when a violation of order interfered (82 ms) than when a violation of context interfered (24 ms).

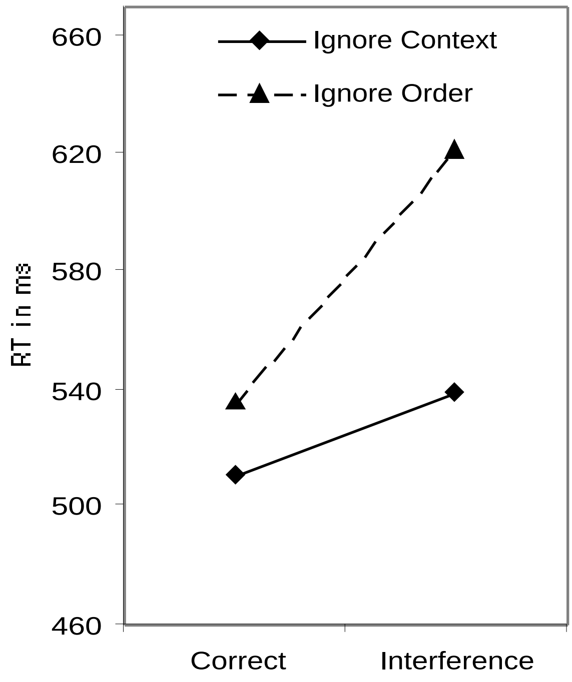


Figure 7: RTs with and without interfering violation.

The RTs were entered into a 2 x 2 repeated measurement ANOVA with the factors Type of Sequence (Correct vs. Interfering Violation) and Type of Interference (Order vs. Context). There were significant main effects for the factors Type of Sequence $F(1, 19) = 27.5$, $p < .001$ and Type of Interference $F(1, 19) = 86.6$, $p < .001$, as well as an interaction between the two factors $F(1, 19) = 23.1$, $p < .001$.

The pattern of results is consistent with the claim that the processes involved in analyzing order and meaning of action sequences are highly automatic. Otherwise one would not expect the huge interference effects that were observed in Experiment 3. It seems impossible to focus on one or the other aspect of an action sequence and to ignore the other. The stronger effect in the condition with interfering violations of order further supports the claim that analyzing the order of an action sequence takes less time than analyzing it's meaning.

General Discussion

The general pattern of results suggests that in action comprehension the analysis of order and meaning proceeds in parallel. Moreover, both processes seem to be highly automatic, that is, none of the two aspects of action sequences can be easily ignored. A further result is that analyzing the order of a sequence might be completed before its meaning is fully analyzed. However, further experiments are needed to determine whether the faster processing of order is a general phenomenon or whether it depends on the complexity of the task at hand.

The possibility to conceptualize action comprehension in an analogous manner to language comprehension does not necessarily imply that the cognitive processes involved in both domains are the same or that they are governed by the same principles. However, this strategy provides an opportunity to relate findings from both domains and to detect interesting parallels or differences. Currently, we are conducting ERP studies to determine whether violations of order and context in action sequences evoke the same ERP components as syntactic and semantic violations in spoken sentences (Friederici et al., 1999). The results of such studies might provide more conclusive answers to the question of whether there are cognitive processes which are actually contributing to both, action and language comprehension (Rizzolatti & Arbib, 1998).

References

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology, 14*, 178-210.

Friederici, A. D., Steinhauer, K., & Frisch, S. (1999). Lexical integration: Sequential effects of syntactic and semantic information. *Memory & Cognition, 27,* 438-453.

Garrod, S., & Pickering M. J. (1999). *Language Processing.* Hove: Psychology Press.

Goodnow, J. J., Levine, R. A. (1973) “The grammar of action”: Sequence and syntax in children's copying. *Cognitive Psychology. 4*, 83-98.

MacDonald, M. C., Pearlmutter, N. J., Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review. 101*, 676-703

Rayner, K., Pollatsek, A. (1989). *The psychology of reading.* Englewood Cliffs,: Prentice-Hall.

Rizzolatti, G., & Arbib, M. (1998). Language within our grasp. *Trends in Neurosciences, 21,* 188-194.

Schank, R. C., Abelson, R. P. (1977). *Scripts, plans, goals and understanding: An inquiry into human knowledge structures.* Hillsdale: Lawrence Erlbaum.

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