

# Spontaneous gestures during mental rotation tasks: Insights into the microdevelopment of the motor strategy

Chu, Mingyuan; Kita, Sotaro

DOI:  
[10.1037/a0013157](https://doi.org/10.1037/a0013157)

*Document Version*  
Peer reviewed version

*Citation for published version (Harvard):*  
Chu, M & Kita, S 2008, 'Spontaneous gestures during mental rotation tasks: Insights into the microdevelopment of the motor strategy', *Journal of Experimental Psychology: General*, vol. 137, no. 4, pp. 706-723.  
<https://doi.org/10.1037/a0013157>

[Link to publication on Research at Birmingham portal](#)

## **Publisher Rights Statement:**

This article may not exactly replicate the final version published in the APA journal. It is not the copy of record. DOI:  
<http://dx.doi.org/10.1037/a0013157>.  
Copyright APA Journals®

## **General rights**

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

## **Take down policy**

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact [UBIRA@lists.bham.ac.uk](mailto:UBIRA@lists.bham.ac.uk) providing details and we will remove access to the work immediately and investigate.

Running head: MENTAL ROTATION AND GESTURE

Appeared in *Journal of Experimental Psychology: General*

< <http://www.apa.org/pubs/journals/xge/index.aspx> >

Chu, M., & Kita, S. (2008). Spontaneous gestures during mental rotation tasks: Insights into the microdevelopment of the motor strategy. *Journal of Experimental Psychology: General*, 137, 706-723.

(c) American Psychological Association. This article may not exactly replicate the final version published in the APA journal. It is not the copy of record.

Spontaneous gestures during mental rotation tasks: Insights into the microdevelopment of the  
motor strategy

Mingyuan Chu and Sotaro Kita

University of Birmingham

## Abstract

This study investigated the motor strategy in mental rotation tasks by examining two types of spontaneous gestures (hand-object-interaction gestures, representing the agentive hand action on an object, vs. object-movement gestures, representing the movement of an object by itself), and different types of verbal descriptions of rotation. Hand-object-interaction gestures were produced earlier than object-movement gestures, the rate of both types of gestures decreased, and gestures became further away from the stimulus object over trials (Experiment 1 & 3). Furthermore, in the first few trials, object-movement gestures increased while hand-object-interaction gestures decreased, and this change of motor strategies was also reflected in the type of verbal description of rotation in the concurrent speech (Experiment 2). This change of motor strategies was hampered when gestures were prohibited (Experiment 4). We concluded that the motor strategy becomes less dependent on agentive action on the object, and also becomes internalized over the course of the experiment, and gesture facilitates the former process. When solving a problem regarding the physical world, adults go through developmental processes similar to internalization and symbolic distancing in young children, albeit within a much shorter time span.

*Keywords:* gesture, mental rotation, cognitive development, problem solving

## Spontaneous gestures during mental rotation tasks: Insights into the microdevelopment of the motor strategy

Gestures that spontaneously accompany speech can be a window into speakers' mind, especially their analogue imagistic thinking (McNeill, 1992). It has been argued that speech production processes are linked to gesture production processes at the level of conceptual planning (Kita, 2000; but see e.g., Krauss, Chen, & Gottesman, 2000 for an alternative) as conceptually more complex speaking tasks trigger more gestures (Alibali, Kita, & Young, 2000; Hostetter, Alibali, Kita, 2007; Melinger & Kita, 2007). Consistent with this view that gestures are involved in conceptualization processes, various studies have shown that gestures can reveal important aspects of problem solving and learning processes. For example, discrepancy between the contents of gesture and concurrent speech indicates that children are in a transitional phase in the understanding of Piagetian conservation tasks (Church & Goldin-Meadow, 1986) or arithmetic equations (Perry, Church, & Goldin-Meadow, 1988). Similar discrepancy in adults indicates that they are considering alternative strategies in a Tower of Hanoi problem (Garber & Goldin-Meadow, 2002).

Gestures can provide insights into the choice of problem solving strategies used by adults. Alibali, Bassok, Solomon, Syc, and Goldin-Meadow (1999) found that when people were asked to describe and then solve a mathematical problem, their gestures could predict the strategy they used in the solution. Schwartz and Black (1996) showed that gestures revealed how the type of problem solving strategies chosen by the problem solver changed over the course of an experiment. They presented people with a problem concerning a physical system (interlocking gears), which could be solved either by mental simulation of gear movement or by an abstract rule based on whether the number of gears was odd or even. When people were using the mental simulation strategy (as revealed by the verbal protocol

and solution latency), they produced more spontaneous gestures representing gear movement than when they were using an abstract strategy. They also found that participants' strategy typically changed from mental simulation to the abstract rule over the course of trials, and this change was reflected in the decrease of gestural depictions of gear movement.

Gestures can not only reflect the strategy change, but also play a causal role in solving problems regarding the physical world. Alibali and Kita (under review) showed that strategies for solving a physical problem differed depending on whether participants were allowed to gesturally depict physical features of the problem. In their study, children were asked to explain Piagetian conservation tasks, and they were more likely to use information that was not perceptually present when gesture was prohibited than when it was allowed. Similarly, Schwartz and Black (1999) claimed that acting on objects could help adult participants solve a novel problem regarding a physical event. In this study, the participants were shown two glasses that had different widths but equal heights and were asked to imagine that the glasses were filled to the same level with water. The participants had to judge whether the two glasses would spill at the same or different angles. They found that people rarely answered the question correctly verbally using their explicit knowledge. However, when closing their eyes and rotating the empty glasses by hand, participants could indicate the answer correctly more frequently.

Because gestures are particularly frequent when people solve problems regarding spatial transformations (Trafton et al., 2006), a mental rotation task, as a typical type of spatial transformation, provides us a good opportunity to investigate the role of gestures in problem solving. In the present study, we examined the spontaneous gestures in two types of mental rotation tasks to see how the motor strategy changes over trials and whether gestures play a causal role in this strategy change.

Since the seminal studies by Shepard and colleagues (Shepard & Metzler, 1971; Cooper & Shepard, 1973), the exact underlying mechanism for mental rotation tasks has been a heavily debated issue. One of the important proposals is that motor processes are crucially involved in mental rotation. Sekiyama (1982) provided some of the first evidence for the link between motor processes and mental rotation. In her study, the participants were asked to judge whether a line drawing of a hand presented in different orientations was a left or right hand. She found that reaction time as a function of rotation angles differed for the left and right hand stimuli, which reflected the extent to which clockwise or counter-clockwise rotation was anatomically constrained for a given hand. Similarly, Parson (1987) also found that when using body parts as the stimulus in a mental rotation judgment task, reaction time to perform left-right judgments was strongly affected by anatomical constraints on motion to the orientation of the stimulus. Wohlschläger and Wohlschläger (1998) showed that motor processes were involved in mental rotation even when an abstract geometric object was rotated. In the first experiment involving Shepard-Metzler type problem, one group of participants solved the problem by mentally rotating the object, and another group of participants solved the problem by turning a knob that rotated the object on the computer screen in the same direction. They found that the response time (RT) across different rotation angles was not significantly different between the two groups. Thus, they concluded that rotary object manipulation was commensurate with mental rotation. In the second experiment, they further investigated whether the rotational hand movements could influence the performance of mental rotation. The participants were asked to turn the knob either in the same direction as the direction of the shortest angle or in the opposite direction. Unlike the first experiment, turning the knob did not rotate the object on the screen. Nevertheless, the RT was considerably shorter when the rotational hand movements were in the direction

congruent with mental rotation than in the opposite direction. Thus, the execution of rotational hand movements facilitated the simultaneously performed mental rotation when the directions of rotation matched.

Wexler, Kosslyn, and Berthoz (1998) provided corroborating evidence. In their study, the participants were asked to mentally rotate two-dimensional geometric figures (used in Cooper and Shepard, 1973) while the hand holding a joy-stick made rotary movement. They found that the reaction time was shorter and the error rate was lower when the direction of manual rotation was congruent with that of mental rotation. Furthermore, the degrees of rotation of the joy-stick from the beginning of the trial to the response correlated with the degree of mental rotation required to respond. It should be, however, noted that they found these effects only in the first half of the experiment, but not in the second half.

Schwartz and Holton (2000) showed that motor facilitation of mental rotation is not simply due to shared representation of rotation. In their experiments, the stimulus was actually a three-dimensional object (analogous to the ones in Shepard and Metzler, 1971) on a spool, which could be rotated by pulling a string. During the mental rotation task, participant pulled the string to rotate the visually occluded stimulus object. Even though the manual action was not rotary (the string was pulled straight), when the object rotated in the direction congruent with mental rotation, the reaction time was shorter than when the object rotated in the incongruent direction. They concluded that the motor facilitation of mental rotation is due to mental simulation based on a mental model that incorporates not only the spatial information about the rotating object but also other non-spatial information (e.g., the mechanical interaction between the spool and the string).

It has also been noted that, in mental rotation tasks, change in participants' behaviour over the course of trials is substantial. Kail (1986) found that mental rotation became faster

over trials. Furthermore, the aforementioned study by Wexler et al. (1998) found the influence of manual rotation on mental rotation only in the first half of the experiment, but not in the second half. They gave two possible explanations for this change. First, in the second half, the participants might have taken a strategy not involving rotation of the stimulus figures at all. Second, the mental rotation task might become more automatic and does not involve motor planning processes as strongly. The latter explanation can be further extended and related to theories of the cognitive development in children (Piaget, 1968; Werner & Kaplan, 1963).

It has been proposed that children's representation of the physical world becomes increasingly detached from the physical world itself in the course of development. For example, Piaget (1968) proposed that, young children form conceptual understanding of the physical world through bodily interaction with the physical world<sup>1</sup>. For example, only after acting on objects repeatedly, the child becomes able to represent these objects internally. That is, repeated sensorimotor experiences lead to an internalized schema of how physical action and objects interact. Werner and Kaplan (1963) suggested a symbolic distancing process in children's cognitive development. That is, children start out with representations in which the "symbols" (depicting element) are closely linked to the "referents" (depicted content) both physically and representationally. In the course of development, children increasingly physically separate symbols from referents, and start to use symbols independently from their referents. Children also increasingly separate properties of symbols from properties of referents, and start to use arbitrary symbols to represent referents. Thus, in both physical and representational senses, the symbolic distance between the "symbols" and the "referents" becomes larger and larger.



We propose that an analogous process exists also in adults, albeit within a much shorter time span. That is, when solving novel problems concerning the physical world, adults may start with bodily exploration of the physical world. The knowledge gained through the bodily interaction with the physical world is gradually transformed into the format that is more detached from the physical world, and eventually into entirely internal representation. In this process, the problem-solving strategy becomes less and less constrained by the external physical world so that people can solve the problem in a more efficient way.

More specifically, in relation to mental rotation type tasks, we hypothesize three different stages in this process. In the first stage, people try to solve the problem by bodily manipulating the physical object or by gesturally simulating such action. Like in children, this strategy can provide people with first hand experience about how the physical object can interact with action. This strategy, however, is restricted by both the physical feature of the object, such as the size, location, orientation and etc., and the anatomical restriction of body parts. In the second stage, the strategy still depends on body movement (such as gesture), but the representation in the body movement is ‘deagentivized’. That is, the agent of the action disappears. Now people do not need to actually bodily manipulate the physical world (or gesturally simulate it), but their body part, especially the hand, represents the relevant object, and the body movement (i.e. gesture) represents the movement of the object. Thus, the body movement becomes more self-contained as a representation, and detached from the object in the physical world. In this stage, the restriction from the feature of the object in the physical world goes away, and the strategy is now only limited by the anatomical restriction of body parts. In the third stage, the knowledge gained from the first two stages becomes ‘internalized’, and people no longer depend on overt bodily manipulation or representation to

solve the problem. Now people are finally liberated from the restriction of the physical world, so that they can solve the problem with great efficiency. Such a process might have been responsible for the differences in the participants' behaviours between the first and second halves of the experiment in Wexler et al. (1998).

It is also possible to hypothesize that gesture facilitates the deagentivization process in adults' problem solving. Gesture, as a simulation of the actual action on the physical world, may greatly enrich the sensorimotor experiences. This rich information may facilitate people to transform their strategies from bodily manipulation of the physical world into more self-contained and detached strategies that focus on movement of the object. In addition, the unstable nature of gesture execution may help people discover new strategies. For example, at the beginning participants may use a grasp handshape in the gesture to simulate the manipulation of an object in the physical world. However, the grasp handshape may become looser and looser over time, and sometimes change to a flat handshape. This new handshape may lead to a new strategy in which there is no need for an agent to manipulate the object, but the hand itself can represent the object in the physical world. For example, the flat hand can be rotated (away from the object) to represent rotation of the object. However, when participants are not allowed to gesture, this process might be hampered, and people may be stuck at the initial strategy, involving an agent acting upon the physical world.

To confirm these hypotheses, we examined spontaneous gestures and speech during two types of mental rotation tasks: a description task in which the participants were required to verbally describe rotation of an Shepard-Metzler (1971) style three-dimensional object and a judgment task (similar to those used in Shepard and Metzler, 1971 and in Wohlschläger and Wohlschläger, 1998) in which the participants were asked to choose one of the two mirror three-dimensional objects to match the stimulus object. In the judgment task, the participants

responded with foot pedals, leaving the hands free for possible spontaneous gestures.

Participants spontaneously produced gestures that simulated the manipulation and rotation of the object in both mental rotation tasks. As shown in previous studies (Alibali et al., 1999; Church & Goldin-Meadow, 1986; Garber & Goldin-Meadow, 2002; Perry, Church, & Goldin-Meadow, 1988; Schwartz & Black, 1996), gestures can serve as a window into learning and problem solving processes. In the current study, we observed how the type and rate of gestures changed over the course of trials (Experiments 1, 2 and 3), as well as the change of the verbal description mode of rotation (Experiment 2), in order to gain insights into how the nature of motor strategies changed over the course of trials. In Experiment 4, we examined whether gesture played a causal role in this strategy change in mental rotation by comparing motor strategies expressed in the verbal description of rotation between the gesture-allowed and gesture-prohibited conditions.

### Experiment 1

The main goal of Experiment 1 is to examine the hypothesis that the external motor strategy, in the form of spontaneous gestures, becomes ‘deagentivized’ and ‘internalized’ over the course of the experiment. If some of the spontaneous gestures can represent the external motor strategy used in solving mental rotation problems, and if such strategy becomes ‘deagentivized’, gestures that represent an agent manipulating the stimulus object (e.g., gestures with a grasping handshape as if to grasp the object on the computer screen) should occur earlier than that merely represent the movement of the stimulus object (e.g., a flat hand, which stands for the object, is rotated). In addition, if the external motor strategy gradually becomes ‘internalized’ over the course of the experiment, gesture frequency should decrease over trials as more efficient and fully internal strategy takes over. Finally, if some gestures were indeed produced to simulate an agent manipulating the stimulus object, they

should be physically more anchored to the object on the computer screen than those that only represent the movement of the stimulus object. In addition, if the deagentivization and internalization processes can be seen as symbolic distancing (Werner & Kaplan, 1963) of gestural simulation from the stimulus object, it should manifest itself as an increase in the physical distance between the gesture hand and the stimulus object over the course of the experiment. Thus, we examined 1) the order in which the two types of gestures appeared within a trial and over the course of the experiment, 2) how gesture rate changed over the course of the experiment, 3) how close to the stimulus object the two types of gestures were produced, and how this changed over the course of the experiment.

## Method

### *Participants*

Forty two right-handed native English speakers, 27 women and 15 men, took part in the study. All participants had normal or corrected-to-normal vision. They were paid either course credit or £4 for participation. The participants' age ranged from 18 to 56 years ( $M = 24.76$ ,  $SD = 9.65$ ). We excluded data from seven of the forty two participants who did not produce any gesture throughout the experiment. Thus, the final sample consisted of 35 individuals, 22 women and 13 men.

### *Stimuli*

The three-dimensional object used in the current experiment was based on the stimulus used by Shepard and Metzler (1971) (see Figure 1). The stimuli were created by software, "Blender". The surfaces of the object were shaded grey, and lamp light-sources were put 250 cm above, 10 cm in front of, and 30 cm to the left of the object centre.

\*\*\* Figure 1 about here \*\*\*

Each stimulus consisted of two line drawings of the same three-dimensional object at different orientations. The right object was always in the canonical position in the sense that its sides were parallel to either the horizontal axis, or the vertical axis, or the axis pointing to depth. Thirty stimuli were created by rotating the left object in 60° steps around an axis that went through the object's centre (60°, 120°, 180°, 240°, and 300°) to create the left objects. The Cartesian rotational axes (horizontal, vertical, and depth) and the figural axes of the object were parallel to each other at 0° orientation. At each angle for each axis, we had two stimuli, varying in size, either small or big (the smaller objects' size was 1/3 of the bigger ones). The edge length of each cube on the computer screen was 1.5 cm for the bigger size and 0.5 cm for the smaller size. The distance between the centres of the two objects was 14.5cm for the bigger stimuli, and 7cm for the smaller stimuli. In the present study, this size variable was not investigated.

Three more stimuli were generated for the three practice trials. The rotation angles in the practice trials were different from any of the stimuli used in the experimental trials. In the first practice trial, the object was rotated on the horizontal axis by 45 degrees. In the second practice trial, it was a 135 degrees rotation on the vertical axis. In the third practice trial, the object was rotated on the depth axis by 30 degrees.

### *Apparatus*

Stimuli were presented centrally on a 15-inch CRT monitor. The participants' gestures and verbal descriptions were captured by two cameras (one from the left side and the other from the back over the participants' right shoulder). Video was recorded on PAL DV VCRs (at 25 frames per second).

### *Design*

All analyses had a within-participant design. The total experiment consisted of 3 practice trials and 30 experimental trials. The experimental trials used a pseudo-randomised order with no repetition of the same axis within the two consecutive trials. The size counterparts were separated by at least 5 intervening trials, and the same size did not repeat more than four times in a row. The order of the practice trials was the same for all participants, but the order of experimental trials was reversed for half of the participants.

### *Procedure*

All participants were tested individually. They were seated approximately 70 cm in front of the monitor. The experimenter was seated next to the participants. They were instructed to describe how the left three-dimensional object could be rotated to the position of the right one. They were also told that their response time would not be recorded so that they did not need to solve the problem under time pressure. In principle, they were allowed to produce any kind of description of rotation. However, in the practice trials, they were asked to describe the axis, the direction and angles of rotation if their descriptions did not clearly include these pieces of information. As they did not know the exact rotation angles of the stimuli, they were told to estimate the rotation angles. For each trial, the experimenter pressed the “space” bar on the keyboard to display the stimulus. No feedback was given concerning the accuracy of their responses.

### *Gesture Coding*

Gestures coding was carried out, using video annotation software, ELAN, developed by Max Planck Institute for Psycholinguistics. Gestures were segmented into series of gesture strokes (Kendon, 1980; McNeill, 1992) and "independent holds" (Kita, van Gijn, & van der Hulst, 1998), which were “holds” not following or preceding any strokes and expressed meaning by themselves. The segmentation was carried out, following the procedure in Kita,

van Gijn, van der Hulst, and Harry (1998). Gesture strokes are performed more forcefully than other phases of gestures (e.g., preparation), and they express meaning of gestures.

Each gesture was coded according to the following classification system (developed on the basis of the classification system in McNeill, 1992). Hand-object-interaction gestures were the gestures that could be interpreted, in the context of concurrent speech, as depicting physical manipulation of the stimulus object by hands (e.g., the index finger and the thumb are opposed as if to grasp the object). Object-movement gestures were the gestures that could be interpreted, in the context of concurrent speech, as depicting the axis, angle, and direction of rotation without any grasping handshape (e.g., a flat hand, representing the object, may rotate around the wrist, or a hand with the extended index finger may draw a circle in the air). Tracing gestures depicted the outlines of the stimulus object (e.g., the index finger traces the edge of the object). Rotation direction gestures depicted a straight vector indicating the direction of rotation. Relative-location gestures depicted the relative locations of the two objects on the computer screen. Object angle gestures represented the angle between the rotated and the canonical objects. Viewpoint gestures indicated the viewpoint from which rotation was described. Deictic gestures pointed at a location of an object or pointed toward the direction to which the object was facing. Beat gestures were two-phase movement with rapid flicks of the fingers or hand, but they did not present any discernible meaning. Emblem gestures were conventionalized gestures, which conveyed some known meaning, such as “maybe” (e.g., a flat hand with the palm down, wavering), “you know” (e.g., a flat hand with the palm up, possibly with a shoulder shrug), and etc.

The locations of the gestures were also coded in terms of the distance between the hand and the monitor. Near-Screen gestures were those gestures in which the distance

between hand and computer screen was less than 20 cm. Far-from-screen gestures were those gestures in which the distance between the hand and the monitor was more than 20 cm.

In order to establish inter-coder reliability of gesture coding, three trials per participant were randomly selected, and a second independent coder classified all gestures that occurred in these trials ( $N = 205$ ). The two coders' decisions matched 92.20% for the gesture type coding (Cohen's  $k = .86$ ,  $p < .001$ ), 93.17% (Cohen's  $k = .80$ ,  $p < .001$ ) for the location coding.

## Results and Discussions

The participants produced overall 341 gestures in the practice trials and 2084 gestures in the experiment trials. The following analyses focused only on the hand-object-interaction gestures and object-movement gestures because these two types of gestures both encoded all three parameters of rotation: the axis, the angle, and the direction, and these two types of gestures were the two most frequent gestures, comprising 62.06% of all gestures.

### *Appearance Order of Different Types of Gestures*

According to our hypothesis, participants should produce hand-object-interaction gestures earlier than object-movement gestures as the external motor strategy becomes deagentivized. We examined the appearance order of these two types of gestures both across trials and within a single trial.

*Gesture type change over the course of the experiment.* In this analysis, we focused on two types of trials, that is, hand-object-interaction trials and object-movement trials. Hand-object-interaction trials had at least one hand-object-interaction gesture but no object-movement gesture, whereas object-movement trials had at least one object-movement gesture but no hand-object-interaction gesture. Trial numbers were used to indicate where in the experiment these two types of trials appeared. The lower the trial number, the earlier the



trial occurred. We then compared the mean trial number of hand-object-interaction trials and object-movement trials. The mean trial number of hand-object-interaction gesture trials ( $M = 13.27$ ,  $SD = 4.82$ ) was significantly lower than that of object-movement gesture trials ( $M = 16.29$ ,  $SD = 3.80$ ),  $t(19) = 2.48$ ,  $d = 0.70$ ,  $p < .05$ . Thus, hand-object-interaction gestures were produced in significantly earlier trial in the experiment than object-movement gestures. This result supports our idea that the external motor strategy becomes deagentivized over the course of the experiment.

*Gesture type change within a single trial.* The goal of this analysis is to provide evidence that deagentivization can even occur within a single trial. If participants deagentivized their external motor strategy in a single trial, they should produce hand-object-interaction gestures earlier than object-movement gestures. In this analysis, we focused on the trials that have at least one hand-object-interaction gesture and one object-movement gesture. We then gave a score to each gesture according to its position in the trial. For example, if a participant produced 3 gestures in one trial, score 1 would be given to the first gesture and score 3 would be given to the last gesture. Thus, the lower the score, the earlier in the trial the gesture was produced. We compared the mean position score of hand-object-interaction gestures and object-movement gestures. The mean position score of hand-object-interaction gestures ( $M = 2.24$ ,  $SD = 0.85$ ) was significantly lower than that of object-movement gestures ( $M = 2.82$ ,  $SD = 1.06$ ),  $t(16) = 2.98$ ,  $d = 0.60$ ,  $p < .01$ . Namely, hand-object-interaction gestures were produced significantly earlier in a single trial than object-movement gestures did. This result again supports our deagentivization hypothesis.

*Discussion.* In the above analyses, we investigated the appearance order of the hand-object-interaction gestures and object-movement gestures. We found that the participants produced hand-object-interaction gestures significantly earlier than

object-movement gestures both across trials and within a single trial. This suggested that, when solving a mental rotation task, participants initially imagined holding the object on the computer screen with their hand, and the gestural simulation of rotation took a more concrete and object-anchored form. As participants became familiar with the object and the task, the gestural simulation of rotation became more self-contained in the sense that there was no longer overt depiction of hand-object interaction in gestures, but the gesture hand itself became the object, and gestures only represented the movement of the object. This change reflected the deagentivization process in which the agent of the hand-object interaction disappeared, and the gesture form became more self-contained and detached from the object.

*Change in Gesture Rates over Experimental Trial Halves and Practice Trials.*

According to our hypothesis, participants' external motor strategy, in the form of spontaneous gestures, should gradually become internalized as they became familiar with the experiment task. We examined how gesture rates (number of gestures per minute) changed over the two trial halves of the experiment. We also extended the gesture rate analysis to the practice trials as we found interesting trends in our exploratory data analysis.

*Change in Gesture Rates over Trial Halves (First Half vs. Second Half).* Gesture rates (number of gestures per minute) were submitted to a 2\*2 repeated measures ANOVA with gesture types (hand-object-interaction vs. object-movement) and trial halves (first half vs. second half) as independent variables (see Figure 2 for the means). There was a main effect of gesture types, that is, the rate of object-movement gestures was higher than that of hand-object-interaction gestures,  $F(1, 34) = 7.63$ ,  $MSE = 12.51$ ,  $p < .01$ ,  $\eta_p^2 = 0.18$ . There was a main effect of trial halves, that is, gesture rates were lower in the second half than in the first half,  $F(1, 34) = 8.04$ ,  $MSE = 0.76$ ,  $p < .01$ ,  $\eta_p^2 = 0.19$ . The interaction between gesture types and trial halves was not significant<sup>2</sup>,  $F(1, 34) = 0.61$ ,  $MSE = 0.95$ ,  $\eta_p^2 = 0.02$ .

\*\*\* Figure 2 about here \*\*\*

*Change in Gesture Rates over Three Practice Trials.* Gesture rates (number of gestures per minute) were submitted to a 2\*3 repeated measures ANOVA with gesture types (hand-object-interaction vs. object-movement) and trials (first vs. second vs. third practice trials) as independent variables (see Figure 3 for the means). There was a main effect of gesture types, that is, the rate of object-movement gestures was higher than that of hand-object-interaction gestures,  $F(1, 29) = 13.25$ ,  $MSE = 11.28$ ,  $p < .01$ ,  $\eta_p^2 = 0.31$ . There was a main effect of trials,  $F(2, 58) = 9.92$ ,  $MSE = 6.90$ ,  $p < .01$ ,  $\eta_p^2 = 0.26$ . The interaction between gesture types and trials was significant,  $F(2, 58) = 18.93$ ,  $MSE = 10.49$ ,  $p < .01$ ,  $\eta_p^2 = 0.40$ .

Tukey posthoc tests showed that the rate for object-movement gestures was higher for the third practice trial than the first and second practice trials (both  $p < .01$ ). For hand-object-interaction gestures, there was no significant difference between any of the practice trials, though there was a trend that the rate decreased over the three practice trials. Furthermore, the rate for object-movement gestures was higher than that for hand-object-interaction gestures in the third practice trial ( $p < .01$ ), but not in the first two practice trials. Thus, the interaction arose from the fact that the rate increased for object-movement gestures, but not for hand-object-interaction gestures.

\*\*\* Figure 3 about here \*\*\*

*Discussion.* The purpose of these analyses was to investigate how the rates of hand-object-interaction gestures and object-movement gestures changed with the progress of the experiment. During the 30 experimental trials, the rates of both hand-object-interaction gestures and object-movement gestures decreased over trials. This suggested that as participants became more experienced in the task, the external motor strategy became

internalized, and no longer required overt hand movements. However, it was also interesting to see that in the first few practice trials, the rate of hand-object-interaction gestures and that of object-movement gestures showed different patterns of change. The rate of object-movement gestures, whose representation was self-contained and not anchored to the stimulus object, significantly increased over the three practice trials, while the rate of hand-object-interaction gestures decreased, though not significantly. The decrease of hand-object-interaction gestures and the increase of object-movement gestures in the first three practice trials also support our deagentivization hypothesis.

It should be noted that all participants performed the three practice trials in the same order. Thus, there was a confounding of the problems they solved and the trial order. This problem will be addressed in Experiment 2.

#### *Gesture Location Analyses*

In these analyses, we investigated the locations at which gestures were performed. In the previous analyses, we treated hand-object-interaction gestures as being more object-anchored whereas object-movement gestures as being more self-contained and more detached from the object. It would be useful to test the validity of our gesture categorization by examining whether hand-object gestures were indeed performed closer to the object on the computer screen than object-movement gestures. In addition, according to the symbolic distancing theory, namely symbols become further away from referents, it would be interesting to see how the physical distance between gesture hand and the stimulus object changed over the course of the experiment.

First, we analysed whether hand-object-interaction gestures and object-movement gestures differed in terms of the proportion of near-screen gestures, in general (data for the first and second halves combined). Hand-object-interaction gestures were more likely to be

performed near the stimulus objects on the screen ( $M = .09$ ,  $SD = 0.22$ ) than object-movement gestures ( $M = .03$ ,  $SD = 0.09$ ),  $t(24) = 2.09$ ,  $d = 0.36$ ,  $p < .05$ . Next, the proportion of near-screen gestures was submitted to a 2\*2 repeated measures ANOVA with gesture types (hand-object-interaction vs. object-movement) and trial halves (first half vs. second half) as independent variables (see Figure 4 for the means). There was a main effect of trial halves,  $F(1, 13) = 5.08$ ,  $MSE = 0.01$ ,  $p < .05$ ,  $\eta_p^2 = 0.28$ . There was no main effect of gesture types,  $F(1, 13) = 2.70$ ,  $MSE = 0.01$ ,  $ns$ . The interaction between gesture types and trial halves was significant,  $F(1, 13) = 4.69$ ,  $MSE = 0.01$ ,  $p < .05$ ,  $\eta_p^2 = 0.27$ .

Tukey posthoc tests showed that for hand-object-interaction gestures the proportion of near-screen gestures was significantly higher in the first half of the experiment than in the second half of the experiment ( $p < .05$ ). For object-movement gestures, there was no significant difference between the first half and the second half of the experiment. Thus, the interaction arose from the fact that the proportion of near-screen gestures decreased for hand-object-interaction gestures, but not for object-movement gestures.

\*\*\* Figure 4 about here \*\*\*

*Discussion.* The above results indicated that hand-object-interaction gestures were anchored to the stimulus object, but object-movement gestures were not, thus the former was more readily performed near the stimulus object than the latter, in general. Furthermore, as participants repeated the same task, hand-object-interaction gestures became less anchored to the stimulus objects and moved towards internalization. The increase of the physical distance between the stimulus object and hand-object-interaction gestures suggested that symbolic distancing can also be seen in adults' learning process as well. For object-movement gestures, the proportion of near-screen gestures did not significantly decrease in the second half. This is probably due to the floor effect as object-movement gestures were less anchored to the

stimulus object, thus it was relatively hard to find object-movement gestures near the computer screen even in the first half of the experiment.

## Experiment 2

The first goal of Experiment 2 is to replicate the findings on the practice trials in Experiment 1 (Figure 3) with fully counterbalanced item orders, thereby eliminating the confounding between trials and items. The second and main goal of Experiment 2 is to examine if different motor strategies identified in gestures are also reflected in different types of verbal descriptions of rotation.

In Experiment 1, we inferred deagentivization of the motor strategy from the earlier appearance of hand-object-interaction gestures as well as the decrease of hand-object-interaction gestures and the increase of object-movement gestures in the first three trials. In the current experiment, we investigated the participants' verbal descriptions of rotation in order to see whether we could obtain converging evidence for the deagentivization process as we found in gestures. One important difference between the hand-object-interaction gesture and the object-movement gesture is that the former represents an agent manipulating an object and the latter represents just the movement of an object. Similarly, a distinction as to the degree of agent-salience can also be observed in the verbal descriptions of rotation. A description with a transitive verb in active voice such as "I would rotate it clockwise for 60 degrees" highlights the agent more than a description with a transitive verb in passive voice such as "it is rotated clockwise for 60 degrees", in which the agent is merely implied. The agent disappears in a description without any transitive verb such as "it rotates clockwise for 60 degrees" or "clockwise 60 degrees". Thus, we have the following deagentivization cline in verbal descriptions of rotation from the most agent-salient

to the least agent-salient: an active transitive verb, a passive transitive verb, no transitive verb.

In the following speech analyses, we first compared the speech mode between the participants who gestured and those who did not produce any gesture. Furthermore, among gesturers, we investigated whether we could find converging evidence for the deagentivization process from the participants' gestures and speech.

## Method

### *Participants*

Forty one right-handed native English speakers, 26 women and 15 men, took part in the study. All participants had normal or corrected-to-normal vision. They were paid either course credit or £4 for participation. The participants ranged in age from 18 to 51 years ( $M = 22.70$ ,  $SD = 6.57$ ). There were 29 gesturers, 18 women and 11 men, who produced at least one gesture in the experiment.

### *Stimuli & Apparatus*

We used the same three items in the practice trials of Experiment 1 and the same apparatus as in Experiment 1.

### *Design*

The experiment consisted of three trials. The order of the three trials was counterbalanced across the participants in such a way that each item occurred equally often in each of the three trials.

### *Procedure*

The procedure was the same as Experiment 1 except that the three trials were not presented as practice trials.

### *Gesture Coding*

Gesture coding categories were the same as Experiment 1. In order to establish the intercoder reliability, one trial per participant was randomly chosen and a second independent coder classified all gestures that occurred in these trials ( $N = 63$ ). The same three categories, that is, hand-object-interaction, object-movement, and other, were used in the reliability check. The two coders matched 95.24% of the gestures (Cohen's  $k = .92$ ,  $p < .001$ ). A third independent coder classified the same gestures only based on the handshape and the physical movement of the hand without listening to the speech. The two coders matched 89.23% of the gestures (Cohen's  $k = .83$ ,  $p < .001$ ).

### *Speech Coding*

The verbal descriptions of rotation were categorized in an analogous way to the distinctions we made in the gesture behaviour that reflected the different degrees of deagentivization of the motor strategy: hand-object-interaction gestures (as if an agent manipulated the object) vs. object-movement gestures (self-contained depiction of the object's rotation), as in Experiment 1. The following categories for the verbal description modes are listed from that indicative of the weakest deagentivization to that indicative of the strongest deagentivization. *Agent-explicit* descriptions (e.g. Rotate it clockwise 60 degrees; I would rotate it clockwise 60 degrees) were those in which the participant used a transitive verb in the active voice. *Agent-implicit* descriptions (e.g. It needs to be rotated clockwise 60 degrees; It is rotated clockwise 60 degrees.) were those in which the participant used a passive form of a transitive verb. *Agent-less* descriptions (e.g. it rotates clockwise 60 degrees; Rotate clockwise 60 degrees; It is a clockwise rotation 60 degrees; Clockwise 60 degrees) were those in which the participant did not use any transitive verb. All descriptions can be categorized into one of these three speech modes (see more sample excerpts in Appendix A).

## Result and Discussions



*Change in Gesture Rates over Three Trials.*

The participants produced 211 gestures, among which 75.83% were hand-object-interaction and object-movement gestures. Gesture rates (number of gestures per minute) were submitted to a 2\*3 repeated measures ANOVA with gesture types (hand-object-interaction vs. object-movement) and trials (first vs. second vs. third trials) as the independent variables (See Figure 5 for the means). More object-movement gestures were produced than hand-object-interaction gestures,  $F(1, 28) = 8.66$ ,  $MSE = 32.17$ ,  $p < .01$ ,  $\eta_p^2 = 0.24$ . The main effect of trials was not significant,  $F(2, 56) = 0.56$ ,  $MSE = 10.39$ ,  $\eta_p^2 = 0.02$ . The interaction between gesture types and trials was significant,  $F(2, 56) = 11.93$ ,  $MSE = 9.94$ ,  $p < .01$ ,  $\eta_p^2 = 0.30$ .

Tukey posthoc tests showed that the rate for object-movement gestures was higher for the third trial than the first trial ( $p < .01$ ). The rate for hand-object-interaction gestures was lower for the third trial than the first trial ( $p < .05$ ). Furthermore, the rate for object-movement gestures was higher than the rate for hand-object-interaction gestures in the third trial ( $p < .01$ ), but not in the first and second trials. Thus, the interaction arose from the fact that the rate for object-movement gestures increased, while that for hand-object-interaction gestures decreased.

\*\*\* Figure 5 about here \*\*\*

*Discussion.* Thus, we obtained essentially the same pattern of results as reported in Figure 3 from Experiment 1 with full counterbalancing of items. The significant interaction between gesture type and trial and non-significant main effect of trials indicated that object-movement gestures took over hand-object-interaction gestures in the three trials. This is consistent with our claim that the motor strategy becomes deagentivized over the course of the experiment, as a step towards a larger symbolic distance.

*Speech Analyses.*

In the first analysis, we compared the verbal description modes between the gesturers and non-gesturers. In the second and third analyses, we focused on the participants who made at least one hand-object-interaction gesture or object-movement gesture during the experiment, and analyzed their verbal description modes.

*Comparison between gesturers and non-gesturers.* In order to give an account of the strategies used by those non-gesturers, we compared the verbal description modes between the gesturers ( $n = 30$ ) and non-gesturers ( $n = 11$ ). In the analysis, we focused on the gesturers who produced hand-object-interaction gestures and/or object-movement gestures ( $n = 29$ ). One gesturer who did not produce either of these two types was excluded from the analysis. Score 1 to 3 was given to each participant's description in each trial (Agent-explicit = 1; Agent-implicit = 2, Agent-less = 3). The higher the score, the more deagentivized the verbal description was. We treated the speech mode score as ordinal measurement for the following reasons. The agent-explicit description mode was more agent-salient than the agent-implicit description mode, and the agent-implicit mode was more agent-salient than the agent-less mode. However, it was not sensible to treat them as an interval measurement because we could not conceptually equate the interval between the agent-explicit mode and the agent implicit and the interval between the agent implicit and the agent-less mode, though they were both numerically one. Thus, the median score of each participant's description modes across three trials was calculated and the Mann-Whitney test was performed. The median score for the verbal description modes was significantly higher (indicating more deagentivization) for the non-gesturers (*Median* = 3, *Interquartile Range* = 1) than the gesturers (*Median* = 2, *Interquartile Range* = 0), Mann-Whitney,  $U = 90.50$ ,  $p < .05$ . Namely, the non-gesturers used a more deagentivized description mode than the gesturers did. There

are at least two possible explanations of this result, depending on the assumption as to why non-gesturers did not produce gestures. If we assume that the lack of gesturing in non-gesturers is related to the deagentivization and internalization processes, a possible explanation for the result is as follows. The non-gesturers' motor strategies had already gone through these two processes, thus they did not need the external motor strategy any more. In other words, they did not produce gestures because their strategy had already been deagentivized and internalized. Alternatively, if we assume that the lack of gesturing in non-gesturers was totally independent of the deagentivization and internalization processes, the result could be interpreted in other ways. For example, the non-gesturers might have had a different communication style from the gesturers, and perhaps the non-gestures were shier to use gestures than gesturers. In this case, one might conclude that the non-gesturers' suppression of gestures lead to more deagentivized description. In other words, because they did not produce gestures, their descriptions were in a more deagentivized mode. We prefer the former explanation. However, in the current experiment, we could not rule out the latter alternative explanation. In Experiment 4, we will use a more direct empirical test for the role of gestures by manipulating the availability of gestures.

*Gesturers whose verbal description mode did not change.* In this analysis, we focused on the gesturers who did not change their verbal description mode throughout the three trials. We divided them into two groups. One is the agent-explicit description group, that is, the participants who used active transitive description (i.e., the least deagentivized description) in all three trials ( $n = 5$ ). The other group is the non-agent-explicit description group, that is, the participants who used either agent-implicit or agent-less descriptions in all three trials ( $n = 8$ ). The mean proportion of hand-object-interaction gestures (out of hand-object-interaction gestures and object-movement gestures) was significantly higher in the agent-explicit

description group ( $M = .61$ ,  $SD = 0.28$ ) than that of the non-agent-explicit description group ( $M = .18$ ,  $SD = 0.21$ ),  $t(11) = 3.11$ ,  $d = 1.74$ ,  $p < .05$ . Thus, the participants who used agent-explicit description mode throughout produced hand-object-interaction gestures more often than did those who used agent-implicit or agent-less description modes throughout. This suggested that verbal description modes and gesture types did give a converging picture on the degree of deagentivization.

*Gesturers whose verbal description mode changed.* In this analysis, we focused on those gesturers who changed their verbal description modes over the three trials. We divided these participants into four groups (2\*2) based on how they changed their gesture types and verbal description modes. According to the pattern of change in gesture types, we divided the participants into two groups. The first group showed a change in the gesture types that was unequivocally compatible with deagentivization of the motor strategy (i.e., compatible under the most stringent and conservative criteria). The participants in this group produced hand-object-interaction gestures either in the first trial or in both first and second trials but not in the third trial, and they did not produce any object-movement gesture preceding hand-object-interaction gestures. The second group consisted of all other participants, who did not meet the criteria for the first group. According to the pattern of change in verbal description modes, we also divided the participants into two groups. The first group showed a change in verbal description modes that was unequivocally compatible with deagentivization of the motor strategy (i.e., compatible under the most stringent and conservative criteria). The participants' verbal description changed monotonically from the mode indicative of weaker deagentivization to the mode indicative of stronger deagentivization along the cline from an agent-explicit description mode to agent-less description mode. The second group consisted of all other participants, who did not meet the criteria for the first group. The combination of

gesture-based and speech-based divisions created four groups (see Table 1). There was a significant association between the indication of deagentivization of the motor strategy in gesture and that in speech, Fisher's exact test,  $p = .008$ . More specifically, the participants who showed a clear sign of deagentivization in gestures tended to do so also in speech, and those who did not show a clear sign in gesture tended not to do so in speech either.

\*\*\* Table 1 about here \*\*\*

*Speech-gesture timing and verbal description modes in the first trial.* In the above analyses, we have shown that the deagentivization process can be reflected in the change gesture types as well as in the change of verbal description modes. It is still unclear whether gestures merely reflected deagentivization of the motor strategy or they actually facilitated the deagentivization process. As the availability of gesture was not manipulated in this experiment, it was not possible to obtain direct evidence for gestural facilitation of deagentivization. However, indirect evidence could be obtained by investigating how speech-gesture timing predicts the verbal description mode used in the trial. More specifically, whether a preceding gesture could influence the following description mode, as compared to when gestures started after the verbal response. In this analysis, we focused on the verbal description modes in the first trial to eliminate any influence from gesture and speech in the preceding trials. We divided the participants who gestured in the first trial into two groups, based on whether they initiated a gesture (i.e., initiated the preparation phase of a gesture (McNeill, 1992; Kita, van Gijn, & van der Hulst, 1998)) before the onset of the verbal description of rotation ( $n = 14$ ) or they initiated a gesture after the onset of the verbal description ( $n = 13$ ). We compared the verbal description modes between these two groups of participants. Again, score 1 to 3 was given to each participant's description mode (Agent-explicit = 1; Agent-implicit = 2, Agent-less = 3). The higher the score, the more

deagentivized the verbal description was. The median score was calculated and Mann-Whitney test was performed. The score for the verbal description modes was significantly higher (indicating more deagentivization) for the participants who gestured before the onset of their verbal description (*Median* = 3, *Interquartile Range* = 1) than the participants who gestured after the onset of their verbal description (*Median* = 2, *Interquartile Range* = 1.5), Mann-Whitney,  $U = 51.00$ ,  $p < .05$ . Namely, the participants who initiated a gesture before their verbal description used a more deagentivized form of verbal description modes than those who initiated a gesture after their verbal description. In order to further examine whether gesture facilitates the deagentivization of the motor strategy, we can prohibit participants from gesturing to see whether deagentivization process becomes slower or even disappears. This will be addressed in Experiment 4.

*Discussion.* The main goal of the speech analyses was to analyze the verbal descriptions of rotation and provide converging evidence for deagentivization of the motor strategy as observed in gestures. We found the degree of deagentivization inferred from the verbal description of rotation was consistent with that inferred from the gesture behaviour. Among gesturers, those who consistently described rotation with an active transitive verb (i.e., the least deagentivized mode) in all three trials tended to use hand-object-interaction gestures more often than those who consistently used either a passive transitive verb or no transitive verb. For those who changed their verbal description modes over the three trials, speech and gesture provided a converging picture on whether or not deagentivization of the motor strategy happened to a given participant. Thus, both gesture types and verbal description modes provided a converging picture as to how explicitly the agent of an action was represented, and the gesture type and the verbal description mode both changed in the direction of deagentivization over the trials.

The comparison of gesturers and non-gesturers yielded an interesting result. It was found that the non-gesturers' description modes were more deagentivized than the gesturers' description modes. One possible interpretation is that non-gesturers had already gone through the deagentivization and internalization process before the first response.

In the last speech analysis, we provided some indirect evidence that gestures can facilitate the deagentivization of the motor strategy. We found that, in the first trial, those who initiated a gesture before the onset of their verbal description of the rotation used more deagentivized description modes than those who initiated a gesture after the onset of their verbal description did.

An alternative account for our deagentivization and internalization claims must be mentioned here because both Experiment 1 and Experiment 2 took place in a conversational situation. According to Grice (1975)'s Cooperative Principle and Maxims in effective communication, the conversation between a speaker and a listener should be brief and avoid unnecessary prolixity. Note that in both experiments the experimenter sat beside the participants and listened to their verbal description of rotation. Obviously, some kind of common knowledge of the stimulus object had been built between the participant and the experimenter over the course of the experiment. Thus, deagentivization of gesture and speech might have simply been due to the inappropriateness of referring the stimulus object in the same way repeatedly. Furthermore, the internalization could also be explained as the result of the increasing common ground between the participant and the experimenter. For example, it might have been unnecessary to refer the stimulus object by hand repeatedly after it had been introduced to the conversation. Thus, a mental rotation task without any communication is needed to rule out this alternative pragmatic account. In Experiment 3, a judgment task was used instead of a description task, and the participants were seated alone in an experimental

room, and responded with two foot-pedals in order to leave their hands free for possible gesturing. They did not talk during the experiment and their spontaneous gestures were recorded by a hidden camera.

### Experiment 3

The main goal of Experiment 3 is to replicate the main findings of Experiment 1 in a non-communicative mental rotation task in order to rule out the pragmatic account for the changing pattern observed in spontaneous gestures. If participants' external motor strategy, in the form of spontaneous gestures, deagentivized and internalized over trials, we should, in the current experiment, observe essentially the same changing pattern of the gesture type, frequency and location in Experiment 1, that is, (1) hand-object-interaction gestures should appear earlier than object-movement gestures; (2) the gesture frequency should in general decrease over the course of the experiment; (3) the gesture location should become further away from the object over trials.

### Method

#### *Participants*

One hundred and thirty two participants, 98 women and 34 men, took part in the study. All participants had normal or corrected-to-normal vision. They were paid course credit for participation. The participants ranged in age from 18 to 33 years ( $M = 20.12$ ,  $SD = 2.27$ ). Among these 132 participants, 65 participants, 54 women and 11 men, produced at least one gesture during the experiment.

#### *Stimuli*

The three-dimensional object used in the current experiment was very similar to those used in Experiment 1 and 2 (see Figure 6). In the current experiment, however, all stimuli had the same size, and the edge length of each cube on the computer screen was 1 cm.



\*\*\* Figure 6 about here \*\*\*

Each stimulus consisted of two three-dimensional objects on the upper screen and one on the lower screen. The upper left and upper right objects were mirror images of each other on the vertical axis, and they were always in the canonical position in the sense that their sides were parallel to either the horizontal axis, or the vertical axis, or the axis pointing to depth. The lower object was rotated from the upper left object in 50% of trials and from the upper right object in the other 50% of trials. The lower object was rotated in four angles (60°, 120°, 240°, and 300°) around the bisector that went through the object's centre between the horizontal and vertical axis, the horizontal and in-depth axis, and the vertical and in-depth axis.

### *Apparatus*

Stimuli were presented centrally on a 15-inch LCD monitor. The participants' performance was captured by a hidden camera located on the left side and about 2.5 metres away. The video was recorded on a SONY DCR-HC19E PAL camcorder (at 25 frames per second).

### *Design*

The total experiment consisted of 24 experimental trials (left vs. right \* 4 angles \* 3 axes) and no practice trials. Stimuli were randomly presented by the computer. The relative position of the two mirror images on the upper screen was balanced across the participants.

### *Procedure*

The participants were tested individually. In order to maximally reduce the communicative environment, the experimenter left the room before the stimulus presentation started, and thus the participants were left alone in the room. Their behaviour during the experiment was video-recorded by a hidden camera. They were debriefed of a hidden video

camera and its purpose after the experiment, and the participants were given the opportunity to request erasing the recording, which none of the participants requested. None of the participants reported that they were aware of the hidden camera.

The participants responded with two foot-pedals silently, leaving their hands free for spontaneous gestures. They were seated approximately 70 cm in front of the monitor. They were told that accuracy was the first priority, and it was not important to respond quickly. We de-emphasized quickness of responses so that spontaneous gestures were not suppressed due to the time pressure. Each trial began with a white fixation cross in the center of the screen for 1000ms, followed by the stimulus. The task was to make a judgment whether the lower three-dimensional object was the same as the upper left object or the upper right object by pressing the correspondent foot pedal (left or right). When the response was given, the next trial started automatically. No feedback was given concerning the accuracy of the response.

### *Gesture Coding*

Gesture categories and location coding were the same as Experiment 1 except that the linguistic information was not used in coding as the participants did not speak. In order to establish the intercoder reliability, 15% of all gestures were randomly chosen and a second independent coder classified these gestures ( $N = 117$ ). The same three gesture categories, that is, hand-object-interaction, object-movement, and other, were used in the reliability check. The two coders' decisions matched 89.74% for the gesture type coding (Cohen's  $k = .79$ ,  $p < .01$ ), 94.87% (Cohen's  $k = .84$ ,  $p < .01$ ) for the location coding.

### Result and discussions

Participants produced a total of 790 gestures. We focused only on hand-object-interaction gestures and object-movement gestures that comprised 41.52% of all gestures<sup>3</sup>.

### *Appearance Order of Different Types of Gestures*

According to our hypothesis, participants should produce hand-object-interaction gestures earlier than object-movement gestures as their external motor strategy, in the form of spontaneous gestures, became deagentivized.

*Gesture type change over the course of the experiment.* The mean trial number of hand-object-interaction gesture trials (i.e., trials with at least one hand-object-interaction gesture but no object-movement gesture) ( $M = 8.77$ ,  $SD = 4.03$ ) was significantly lower than that of object-movement gesture trials (i.e., trials with at least one object-movement gesture but no hand-object-interaction gesture) ( $M = 13.16$ ,  $SD = 5.13$ ),  $t(12) = 3.51$ ,  $d = 0.95$ ,  $p < .01$ . Namely, hand-object-interaction gestures were produced in significantly earlier trials in the experiment than object-movement gestures.

*Gesture type change within a single trial.* This analysis focused on the trials that included both hand-object-interaction gestures and object-movement gestures. The mean position score of hand-object-interaction gestures ( $M = 2.20$ ,  $SD = 0.99$ ) was significantly lower than that of object-movement gestures ( $M = 3.08$ ,  $SD = 1.17$ ),  $t(13) = 3.15$ ,  $d = 0.81$ ,  $p < .01$ . Namely, hand-object-interaction gestures occurred significantly earlier than object-movement gestures within a single trial.

*Discussion.* Thus, we replicated our findings in Experiment 1 about appearance order of hand-object-interaction gestures and object-movement gestures in a non-communicative mental rotation task. The participants produced hand-object-interaction gestures significantly earlier than object-movement gestures both across trials and within a single trial. This deagentivization process could not be due to establishment of common ground between the participant and the experimenter. We argue that the change in the gesture type rather reflected the change in the motor strategy for solving the mental rotation task. Though the explanation

based on common ground cannot be ruled out for the results in Experiments 1 and 2, the most parsimonious account is that the same deagentivization of the motor strategy is responsible for the equivalent findings in Experiments 1, 2, and 3.

*Change in Gesture Rates over Trial Halves (First Half vs. Second Half)*

Gesture rates (number of gestures per minute) were submitted to a 2\*2 repeated measures ANOVA with gesture types (hand-object-interaction vs. object-movement) and trial halves (first half vs. second half) as independent variables (see Figure 7 for the means). There was no main effect of gesture types,  $F(1,40) = 0.01$ ,  $MSE = 2.00$ ,  $ns$ . There was a main effect of trial halves, that is, gesture rates were lower in the second half than in the first half,  $F(1,40) = 7.63$ ,  $MSE = 0.42$ ,  $p < .01$ ,  $\eta_p^2 = .16$ . There was no interaction between gesture types and trial halves,  $F(1,40) = 0.35$ ,  $MSE = 0.30$ ,  $ns$ .

\*\*\* Figure 7 about here \*\*\*

*Discussion.* We replicated the findings about the gesture rate change across the two trial halves in Experiment 1. Over the course of the experiment, the rate of both hand-object-interaction gestures and object-movement gestures significantly decreased. This suggested that the external motor strategy, in the form of spontaneous gestures, became internalized and replaced by internal strategies.

We could not perform the same first three trials analysis as we did in Experiment 2 due to lack of data as, in the first three trials, the rate (number of gestures per minute) of hand-object-interaction gestures ( $M = 0.41$ ,  $SD = 1.68$ ) and object-movement gestures ( $M = 0.40$ ,  $SD = 1.06$ ) was much lower in the silent mental rotation task than the rate of hand-object-interaction gestures ( $M = 2.05$ ,  $SD = 2.78$ ) and object-movement gestures ( $M = 4.43$ ,  $SD = 3.26$ ) in the descriptive mental rotation task. The lower rate of representational gestures in the less communicative setting is compatible with previous literature (e.g., Alibali,

Heath, & Meyer, 2001; Cohen, 1977). Nevertheless, we already provided evidence for the deagentivization process in the analyses of the appearance order of the two gesture types in the preceding subsection.

### *Gesture Location Analyses*

First, we analyzed whether hand-object-interaction gestures and object-movement gestures differed in terms of the proportion of near-screen gestures, in general. Hand-object-interaction gestures were more likely to be performed near the stimulus objects on the screen ( $M = .19$ ,  $SD = 0.30$ ) than object-movement gestures ( $M = .15$ ,  $SD = 0.29$ ),  $t(17) = 2.21$ ,  $d = 0.14$ ,  $p < .05$ . Next, the proportion of the near-screen gestures was submitted to a 2\*2 repeated measures ANOVA with gesture types (hand-object-interaction vs. object-movement) and trial halves (first half vs. second half) as independent variables. There was no main effect of gesture types,  $F(1, 6) = 0.62$ ,  $MSE = 0.02$ , *ns*. There was no main effect of trial halves,  $F(1, 6) = 4.21$ ,  $MSE = 0.12$ , *ns*. The interaction between gesture types and trial halves was not significant,  $F(1, 6) = 3.06$ ,  $MSE = 0.02$ , *ns*. The lack of significant results was probably due to the small number of participants ( $N = 7$ ) included in the ANOVA because only the participants who produced both gesture types in both halves were included. Thus we performed two separate t-tests for each gesture type so that more participants can be included in the analyses. The proportion of the near-screen gestures was significantly higher in the first half ( $M = .27$ ,  $SD = 0.34$ ) than in the second half ( $M = .08$ ,  $SD = 0.20$ ), for hand-object-interaction gestures,  $t(10) = 2.42$ ,  $d = 0.68$ ,  $p < .05$ . The proportion of the near-screen gestures was not significantly different in the first half ( $M = .22$ ,  $SD = 0.40$ ) and the second half ( $M = .12$ ,  $SD = 0.31$ ), for object-movement gestures,  $t(20) = 1.33$ , *ns*.

*Discussion.* We essentially replicated the findings that hand-object-interaction gestures were anchored to the stimulus object, but object-movement gestures were not, and

that hand-object-interaction gestures became less anchored to the stimulus objects and moved towards internalization. For object-movement gestures, the proportion of near-screen gestures was not significantly higher in the first half than in the second half of the experiment.

### Experiment 4

The main goal of Experiment 4 is to directly manipulate the availability of gesture in order to provide direct evidence for our claim that gesture helps deagentivization. We randomly assigned the participants in to the gesture-allowed group and gesture-prohibited group, and compared their verbal description modes in the two conditions. If gesture helps deagentivization, the motor strategy expressed in the verbal response should be in a more deagentivized mode (i.e., less agent salient) when gestures are available. Thus, the overall verbal description modes should be more deagentivized in the gesture-allowed condition than in the gesture-prohibited condition.

### Method

#### *Participants*

Forty nine native English speakers, 43 women and 6 men, took part in the study. All participants had normal or corrected-to-normal vision. They were paid course credit for participation. The participants ranged in age from 18 to 35 years ( $M = 19.51$ ,  $SD = 2.98$ ).

#### *Stimuli & Apparatus*

We used the same three items and apparatus in Experiment 2.

#### *Design*

The order of the three trials was counterbalanced across the participants as Experiment 2. Each individual was assigned randomly to either gesture-allowed group or gesture-prohibited group.

#### *Procedure*

The procedure was exactly the same as Experiment 2, except that the participants in the gesture-prohibited group were asked to sit on their hands in order to prohibit them from gesturing.

### *Speech Coding*

Speech coding was the same as Experiment 2.

### Result and discussions

In the first analysis, we compared the overall level of deagentivization in the verbal description between the gesture-allowed and gesture-prohibited conditions. In the second and third analyses, we compared the two conditions in term of the likelihood of producing agent-explicit description of rotation in the first trial and the likelihood of further deagentivization in the second and third trials.

#### *Analysis of the Overall Level of Deagentivization Indicated by the Verbal Description Modes.*

In this analysis, we compared the overall level of deagentivization in the verbal description modes between the gesture-allowed condition ( $n = 25$ ) and the gesture-prohibited condition ( $n = 24$ ). According to our hypothesis that gesture helps deagentivization of the motor strategy, the description modes in the gesture-allowed condition should be more deagentivized than in the gesture-prohibited condition. Once again, a score of 1 to 3 was given to each participant's description in each trial (Agent-explicit = 1; Agent-implicit = 2, Agent-less = 3). The higher the score, the more deagentivized the verbal description was. For each participant, the median score over the three trials was calculated. The score for the verbal description modes was significantly higher (indicating more deagentivization) in the gesture-allowed condition (*Median* = 2, *Interquartile Range* = 2) than in the gesture-prohibited condition (*Median* = 1, *Interquartile Range* = 1.75), Mann-Whitney,  $U = 205.5$ ,  $p < .05$ .

*Analysis of the Description Modes in the First Trial.*

In this analysis, we focused on the participants' description modes in the first trial. In Experiment 1, we provided evidence that deagentivization occurred even within a single trial (within a trial, a hand-object-interaction gesture tended to precede an object-movement gesture). Further in Experiment 2, we provided indirect evidence that gesture could facilitate deagentivization of the motor strategy in the first trial. Thus, we examined whether deagentivized descriptions occurred in the first trial more often in the gesture-allowed condition than in the gesture-prohibited condition. We divided the participants into four groups (see Table 2) based on whether they used agent-explicit description (i.e., the least deagentivized description) in the first trial or not and whether their gestures were prohibited or not. There was a significant association between the use of agent-explicit description in the first trial and the availability of gesture, Fisher's exact test,  $p = .046$ . More specifically, people were less likely to use agent-explicit description in the first trial when gestures were allowed. In other words, people in the gesture allowed condition were more likely to use the deagentivized forms of verbal descriptions (agent-implicit or agent-less) in the first trial, as compared to those in the gesture-prohibited condition.

\*\*\* Table 2 about here \*\*\*

*Analysis of the Description Modes in the Second and Third Trials.*

In this analysis, we only focused on the participants who used agent-explicit description in the first trial. We examined whether more people showed deagentivization in their description modes in the following two trials in the gesture-allowed condition than in the gesture-prohibited condition. In this analysis, we divided these participants into four groups (see Table 3) based on whether they deagentivized their verbal description modes and whether their gestures were prohibited. For the grouping based on the verbal description



modes, the first group showed a change in verbal description modes that was unequivocally compatible with deagentivization of the motor strategy (i.e., compatible under the most stringent and conservative criterion). Namely, the participants' verbal description changed monotonically along the cline from agent-explicit description mode to agent-less description modes. The second group consisted of all other participants. There was a significant association between the deagentivization of the description mode and the availability of gesture, Fisher's exact test,  $p = .032$ . More specifically, people were more likely to deagentivize their verbal descriptions in the second and third trials when gestures were allowed.

\*\*\* Table 3 about here \*\*\*

*Discussion.* The main purpose of the above speech analyses was to examine whether gesture played a causal role in the change of motor strategy. We hypothesized that gesture facilitates deagentivization of the motor strategy. We found that the verbal description of rotation in the three trials overall indicated more deagentivized strategies in the gesture-allowed condition than in the gesture-prohibited condition. Note that this result might, at first glance, seem to contradict the finding from Experiment 2 that verbal description modes in the non-gesturers were more deagentivized than those in the gesturers. However, these results are compatible with each other. One possible interpretation of the spontaneous non-gesturers in Experiment 2 is that they had gone through deagentivization and internalization processes, and could directly use internalized (and thus deagentivized) motor strategies to solve the problem from the first trial. We suggest that this is why in Experiment 2 verbal description modes in the non-gesturers were more deagentivized than those in the gesturers in Experiment 2. In the gesture-prohibited condition in Experiment 4, the participants were forced to use internal strategies to solve the problem, even if they had

not gone through the natural progression from deagentivization to internalization. In other words, some participants in the gesture-prohibited condition were forced to prematurely internalize their motor strategy. Without the help of gestures, those participants, who would have produced gestures in the gesture-allowed condition, were less likely to deagentivize their motor strategies. Thus, the overall verbal description modes were more deagentivized in the gesture-allowed condition than in the gesture-prohibited condition.

In a further analysis, we found that the participants were more likely to use an agent-implicit or agent-less description in the first trial in the gesture-allowed condition than in the gesture-prohibited condition. This suggested that gesture facilitated deagentivization within the first trial even before the verbal description started. This is consistent with our finding in Experiment 2 that the participants were more likely to use more deagentivized description modes when they initiated a gesture before their verbal description than when they gestured after their verbal response.

In the last analysis, we showed that those participants who used agent-explicit description in the first trial were more likely to deagentivize their descriptions in the following two trials in the gesture-allowed condition than in the gesture-prohibited condition. Taken together, we conclude that gesture plays a causal role in strategy change. More specifically, gesturing facilitates deagentivization of the motor strategy.

Because the nature of the gesture prohibition manipulation, we could not, in principle, rule out the alternative explanation that consequences of sitting on ones' hands other than lack of gesturing (e.g., discomfort, distraction) might inhibit or interferes with the deagentivization process. In the current experiment, however, it is difficult to imagine why discomfort or distraction in the gesture-prohibited condition should prevent the deagentivization process. It is reasonable to assume that discomfort or distraction leads to

easier descriptions, and that the agent-less mode (i.e., thirty degrees to the right) is easier than the agent-explicit mode (i.e., I would rotate it thirty degrees to the right). Then, one would predict that gesture-prohibition group should use more descriptions in the agent-less mode and fewer in the agent-explicit mode, as compared to the gesture-allowed group. However, we found the opposite pattern of results, namely, that participants in the gesture-prohibited group were more likely to use the agent-explicit description mode than the gesture-allowed group.

### General Discussion

Two main findings of the study concerned spontaneous gestures that were produced while two different types of mental rotation tasks involving the Shepard-Metzler (1971) style figures. First, the type, frequency, and location of these gestures changed over the course of the experiment. This change was found in three different time-scales: within a single trial (Experiment 1 and 3), within the first three trials (Experiments 1 and 2), and over the whole experiment (Experiment 1 and 3). Patterns of change were always compatible with the idea that the motor strategy becomes less and less constrained by the external physical world over the course of the experiment. Second, the motor strategy expressed in the verbal response was in a more deagentivized form in the gesture-allowed condition than in the gesture-prohibited condition (Experiment 4). This supports the idea that gesturing facilitates deagentivization of the motor strategy. Furthermore, this facilitation can happen even before the verbal response starts if the gesture was initiated before the onset of the verbal response (Experiments 2 and 4). The following subsections will discuss these findings in more details.

#### *Deagentivization and Internalization of the Motor Strategy*

Participants were more likely to produce hand-object-interaction gestures (representing an agent manipulating the object) before object-movement gestures

(representing a moving object), and this appearance order could be observed both across trials and within a single trial. Meanwhile, over the course of the whole experiment, the rates of both types of gestures decreased. In addition, at the beginning of the descriptive mental rotation task, the rate of hand-object-interaction gestures decreased whereas the rate of object-movement gestures increased.

Furthermore, the two types of gestures differed in terms of location at which they were performed. Hand-object-interaction gestures were more likely to be performed near the stimulus object on the computer screen than object-movement gestures, in general, which confirms our interpretation that hand-object interaction gestures (but not object-movement gestures) are representationally anchored to the object. Moreover, location of hand-object-interaction gestures became further away from the stimulus object in the second half of the experiment.

This set of findings is in line with the idea that manual and mental rotation share a processing mechanism (Wohlschläger & Wohlschläger, 1998), and participants use motoric simulation to solve the mental rotation task (Schwartz & Holton, 2000; Wexler, et al. 1998; see also Hegarty, 2004). Furthermore, change in gesture type, frequency and location indicates the following time course of strategy change. The external motor strategy starts out in a form of hand-object interaction, as if participants try to use their hands to manipulate the stimulus object. Then, it gradually becomes more self-contained (i.e., gesturing hand itself representing the object). This is the deagentivization process, in which the agent of an action becomes less and less salient, eventually leaving just the movement of the object in the representation. The deagentivization process is compatible with the idea that people schematize their strategies over repeated trials in problem solving (Schwartz & Black, 1996). That is, people throw out the irrelevant information during the schematization process. In the

deagentivization process, the information about the agent, which is not logically necessary for the solution, gradually drop out of the gestural representation.

Within a longer time-span, gestures are produced further away from the referent object, and are eventually internalized presumably because no overt gestural simulation of rotation is needed. The external motor strategy is replaced by more efficient internal strategies. This internalization process can explain Wexler et al.'s (1998)'s finding that overt rotary movement by the hand facilitated mental rotation performance only in the first half but not the second half of the experiment.

The external motor strategy, in the form of spontaneous gestures, thus, gradually becomes more liberated from constraints of the physical world. The deagentivization process separates the object in the problem from the agent, removing constraints stemming from hand-object interaction. Deagentivized gestural simulation, however, is still constrained by anatomical restrictions of the gesturing hand. The internalization process then further reduces these constraints stemming from the execution of gestures, though it may not completely remove such constraints (Sekiyama, 1982). Consequently, once the motor strategy goes through both deagentivization and internalization, it becomes much freer from the constraints of the physical world. This change should make the problem solving strategy more efficient and flexible.

The micro-development of gestural simulation is reminiscent of cognitive and symbolic development in young children. Piaget (1968) proposed that young children learn about the physical world through bodily interaction with it, and after the repeated experience, a certain feature of the physical world becomes internalized as a schema. This schema can be used in cognitive processing efficiently because it is free from the constraints of the physical world. Werner and Kaplan (1963) proposed that young children's use of symbols does not

clearly differentiate the referent and the form (i.e., the "vehicle") of a symbol, but gradually the referent and the form become independent from each other both physically and representationally. In other words, the "symbolic distance" increases. Through this process, symbols become self-contained and available to be used freely in thought without the need for anchoring to external referents. The results from the present study suggest that these mechanisms may be at work even in adults, albeit within a shorter time span, when they solve novel problems regarding the physical world.

This conclusion is also compatible with the findings from a qualitative study on gestures in instructional settings by LeBaron and Streeck (2000). They analyzed gestures produced by a professor who commented on a cardboard model of a building in an architecture class. The professor first produced gestures that indicated the shape of the model by tracing the curved shape on the object with his index finger. Then later in his comment, he expressed the same concept of the curved shape with similar gestures that were more detached from the object and performed in mid-air.

Note that it is not possible to explain all the changes in gesture behaviours discussed above in terms of Gricean pragmatics or common grounds that built up between the participant and the experimenter over the course of the experiment. This is because the changes in gesture type, frequency and location were observed not only in the description tasks (Experiments 1 and 2), but also in a non-communicative task (Experiment 3). In Experiment 3, the participants performed the mental rotation task alone in the room, while being recorded by a hidden video camera.

The current study also examined the verbal description of rotation in the first three trials in order to provide converging evidence for the deagentivization process. The participants who used the agent-explicit description mode, which expressed an agent acting

on the object (an active transitive verb) produced hand-object-interaction gestures more often than those who used other description modes. Moreover, gesture behavior and verbal description mode changed in the same direction in the first three trials. Thus, both gestural and verbal representations of rotation reflected the same underlying motor strategy. This allowed us to investigate the causal role of gestures in strategy change by investigating how verbal description of rotation changed as a function of availability of gestures.

#### *Gestural Facilitation of Deagentivization of the Motor Strategy*

The current study investigated the function of spontaneous gestures in the deagentivization process of the motor strategy by prohibiting participants from gesturing. When gestures were allowed, people who initiated their gestures before the onset of their verbal description of rotation were more likely to use more deagentivized description modes than those who initiated their gestures after the onset of their verbal description. Moreover, the verbal descriptions of rotation overall were more deagentivized in the gesture-allowed condition than in the gesture-prohibited condition. Participants were more likely to use more deagentivized description (passive transitive verbs or no transitive verbs) in the first trial in the gesture-allowed condition than in the gesture-prohibited condition. Those participants who used agent-explicit description (active transitive verbs) in the first trial were more likely to deagentivize the description mode in the following two trials in the gesture-allowed condition than in the gesture-prohibited condition.

In summary, gesture facilitates deagentivization of the motor strategy. This is compatible with the idea that action can play an important role in problem solving in adults (Alibali, Spencer, & Kita, in preparation; Schwartz & Black, 1999) and gesture influences conceptualization processes that underlie speaking (Alibali & Kita, under review; Hostetter,

Alibali, & Kita, 2007; Kita, 2000; Melinger & Kita, 2007). The question arises as to what the mechanism of gestural facilitation of deagentivization is.

We conjecture two possible mechanisms that underlie this effect. First, gestures may enrich people's motoric experience. They provide a vivid first hand experience of the nature of a problem, and explore a more appropriate way to solve a problem (Kita, 2000). Second, inherent instability of motor execution may serve as a reservoir for different possible strategies. The gestural simulation with the grasping handshape may sometimes be performed, by chance, with a more lax flat handshape. This may provide an "insight" that the gesturing hand does not have to represent a manipulating hand, but it could represent the object itself. Such a "chance discovery" may prompt the shift to object-movement gestures, namely the deagentivization process. These two conjectures are both in line with the claim of the embodied nature of cognition, namely, action is the foundation of cognition, and the simulations of real-world actions and perceptions are the foundation of cognition (Barsalou, 1999; Glenberg, 1997).

#### *Parallelism between Co-speech Gestures and "Co-thought" Gestures*

The patterns of the gesture behaviour were similar between the description task (Experiments 1 and 2) and the non-communicative (non-linguistic) task (Experiment 3), and this parallelism has implications for theories of gesture production. The parallelism suggests that co-speech gestures and "co-thought" gestures (in a non-linguistic task) may be generated from the same mechanism. This is not compatible with the theories in which co-speech gesture production is intrinsically linked to speaking. For example, it has been proposed that co-speech gestures may be generated from one of the stages of the speech production process (Butterworth & Hadar, 1989; de Ruiter 2000). Co-speech gestures may also be generated from a "growth point" consisting of a combination of an image and a linguistic category,



which serve as the seed representations for a gesture and an utterance (McNeill, 1992). The above mentioned parallelism, rather, suggests that co-speech gestures are generated from an action generation mechanism that is highly coordinated with, but independent from, the speech production system (Kita, 2000; Kita & Özyürek, 2003).

### *Conclusion*

To summarize, the current study investigated gestural and verbal expression of rotation during mental rotation tasks. Gestures provided an insight into micro-development of the motor strategy for mental rotation tasks. The external motor strategy initially took the form of hand-object interaction as if an agent manipulated the stimulus object. It, then, became more self-contained and lost the representation of the agent, and eventually became fully internalized. At this point, the motor strategy was liberated from many of the constraints of the physical world, and thus it was more efficient and flexible. In other words, when confronted with a new problem from the physical world, adults go through developmental processes such as internalization (Piaget, 1968) and symbolic distancing (Werner & Kaplan, 1963), just like young children, albeit within a much shorter time span. Gestures also facilitated deagentivization of the motor strategy (i.e., the removal of agent from the representation of rotation). When participants produced gestures, they were more likely to deagentivize their motor strategy (as inferred from their verbal response) than when they were prohibited from gesturing. Thus, gestures are not only a mere reflection of mental representations used in problem solving, but they also play an active causal role in problem solving.

## References

- Alibali, M. W., Bassok, M., Olseth Solomon, K., Syc, S. E., & Goldin-Meadow, S. (1999). Illuminating mental representations through speech and gesture. *Psychological Science, 10*, 327-333.
- Alibali, M. W., Heath, D.C., & Myers, H.J. (2001). Effects of visibility between speaker and listener on gesture production: Some gestures are meant to be seen. *Journal of Memory and Language, 44*, 169-188.
- Alibali, M. W., & Kita, S. (under review). On the role of gesture in thinking and speaking: Prohibiting gesture alters children's problem explanations.
- Alibali, M. W., Kita, S., & Young, A. J. (2000). Gesture and the process of speech production: We think, therefore we gesture. *Language and Cognitive Processes, 15*, 593-613.
- Alibali, M. W., Spencer, R. C., & Kita, S. (in preparation). Spontaneous gestures influence strategy choices in problem solving.
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*, 577-660.
- Butterworth, B., & Hadar, U. (1989). Gesture, speech, and computational stages: a reply to McNeill. *Psychological Review, 96*, 168-174.
- Church, R. B., & Goldin-Meadow, S. (1986). The mismatch between gesture and speech as an index of transitional knowledge. *Cognition, 23*, 43-71.
- Cohen, A. A. (1977). The communicative function of hand gestures. *Journal of Communication, 27*, 54-63.

- Cooper, L., & Shepard, R. (1973). Chronometric studies of the rotation of mental images. In W. Chase (Eds.), *Visual Information Processing* (pp. 135-142). New York, NY: Academic Press.
- de Ruiter, J. P. (2000). The production of gesture and speech. In D. McNeill (Eds.), *Language and Gesture* (pp. 284-311). Cambridge, UK: Cambridge University Press.
- Garber, P., & Goldin-Meadow, S. (2002). Gesture offers insight into problem-solving in adults and children. *Cognitive Science*, 26, 817-831.
- Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences*, 20, 1-55.
- Grice, H. P. (1975). Logic and conversation. In P. Cole, & J. Morgan, (Eds.) *Syntax and Semantics*, (pp. 41-58). New York, NY: Academic Press..
- Hegarty, M. (2004). Mechanical reasoning by mental simulation. *Trends in Cognitive Sciences*, 8, 280-285.
- Hostetter, A.B., Alibali, M. W., & Kita, S. (2007). I see it in my hands' eye: Representational gestures reflect conceptual demands. *Language and Cognitive Processes*, 22, 313-336.
- Kail, R. (1986). The impact of extended practice on rate of mental rotation. *Journal of Experimental Child Psychology*, 42, 378-391.
- Kendon, A. (1980). Gesticulation and speech: Two aspects of the process of utterance. In M. R. Kay (Eds.), *The relation between verbal and nonverbal communication* (pp. 207-227). The Hague: Mouton.
- Kita, S. (2000). How representational gestures help speaking. In D. McNeill (Eds.), *Language and Gesture* (pp. 162-185). Cambridge, UK: Cambridge University Press.

- Kita, S. & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal? Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language*, 48, 16-32.
- Kita, S., van Gijn, I., & van der Hulst, H. (1998). Movement Phases in signs and co-speech gestures, and their transcription by human coders. In I. Wachsmuth & M. Fröhlich (Eds.), *Gesture and sign language in human-computer interaction, International Gesture Workshop Bielefeld, Germany, September 17-19, 1997, Proceedings. Lecture Notes in Artificial Intelligence, Volume 1371*, (pp. 23-35). Berlin: Springer-Verlag.
- Krauss, R. M., Chen, Y., & Gottesman, R. F. (2000). Lexical gestures and lexical access: A process model. In D. McNeill (Eds.), *Language and Gesture* (pp. 261-283). Cambridge, UK: Cambridge University Press.
- LeBaron, C. D., & Streeck, J. (2000). Gestures, Knowledge, and the World. In D. McNeill (Eds.), *Language and Gesture* (pp. 118-138). Cambridge, UK: Cambridge University Press.
- McNeill, D. (1992). *Hand and Mind*. University of Chicago Press, Chicago.
- Melinger, A., & Kita, S. (2007). Conceptualisation load triggers gesture production. *Language and Cognitive Processes*, 22, 473-500.
- Parsons, L. M (1987) Imagined spatial transformations of one's hands and feet. *Cognitive Psychology*, 19, 178-241.
- Perry, M., Church, R. B., & Goldinmeadow, S. (1988). Transitional knowledge in the acquisition of concepts. *Cognitive Development*, 3, 359-400.
- Piaget, J. (1968). *Six psychological studies*. NY: Random House.
- Sekiyama, K. (1982). Kinesthetic aspects of mental representations in the identification of left and right hands. *Perception & Psychophysics*, 32, 89-95.

- Schwartz, D. L., & Black, J. B. (1996). Shuttling between depictive models and abstract rules: Induction and fallback. *Cognitive Science*, 20, 457-497
- Schwartz, D. L., & Black, T. (1999). Inferences through imagined actions: knowing by simulated doing. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 25, 116-136.
- Schwartz, D. L., & Holton, D. L. (2000). Tool use and the effect of action on the imagination. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 26, 1655-1665.
- Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, 171, 701-703.
- Trafton, J. G., Trickett, S. B., Stitzlein, C. A., Saner, L., Schunn, C. D., & Kirschenbaum, S. S. (2006). The relationship between spatial transformations and iconic gestures. *Spatial Cognition and Computation*, 6, 1-29.
- Vygotsky, L. S. (1981). The genesis of higher mental functions. In J.W. Wertsch (Eds.). *The concept of activity in Soviet psychology* (pp. 162). New York, NY: Sharp.
- Werner, H., & Kaplan, B. (1963). *Symbolic formation. An organismic-developmental approach to language and the expression of thought*. New York: John Wiley & Sons, Inc.
- Wexler, M., Kosslyn, S. M., & Berthoz, A. (1998). Motor processes in mental rotation. *Cognition*, 68, 77-94.
- Wohlschläger, A., & Wohlschläger, A. (1998). Mental and manual rotation. *Journal of Experimental Psychology-Human Perception and Performance*, 24, 397-412.

### Author Note

This research was supported by a grant from the University of Bristol and the University of Birmingham to Mingyuan Chu.

We would like to thank Flora Wilson, Katherine Donneley and Paula Robinson for their help in speech transcription and gesture coding. We also thank Katerina Kantartzis for her proof reading of our manuscript.

Correspondence concerning this paper should be addressed to Mingyuan Chu, School of Psychology, University of Birmingham, Birmingham, B15 2TT, United Kingdom.  
mxc668@bham.ac.uk.

Footnotes

<sup>1</sup> Vygotsky (1981) had a related but different conception of internalization. He focused on the importance of communication and social interaction in development: “any higher mental function was external because it was social at some point before becoming an internal, truly mental function” (pp. 162).

<sup>2</sup> One of the reviewers suggested that a significant interaction between gesture types and trial halves would have supported our deagentivization and internalization claim as the rate of hand-object-interaction gestures should decrease more than the rate of object-movement gestures. However, our theory does not necessarily predict such an interaction. According to our theory, the rate of object-movement gestures should increase first and then decrease, but our theory does not specify by how much the rate of object-movement gestures should increase and decrease in the first half of the experiment, which would influence whether the interaction would be significant or not. In addition, our theory does not specify how long the deagentivization process would last. If most of the deagentivization process happened in the first three practice trials, internalization process would be the main source for the decrease of both types of gestures in the experimental trials, thus a significant interaction between trial halves and gesture types would not be likely.

<sup>3</sup> In Experiment 3, we did not include tracing gestures which comprised 39.24% of all gestures in our analysis. One might argue that tracing gestures could potentially be conceived as a part of hand-object-interaction gestures in the sense that these gestures were anchored to the object and represented an agent tracing the outlines of the stimulus object, though they did not indicate the axes, direction and degrees of the rotation. However, including tracing gestures into hand-object-interaction gestures did not change any of our findings in Experiments 1, 2, and 3.

## Appendix A: Sample excerpts of three verbal description modes of rotation in Experiment 2

### 1) Agent-explicit mode

Em, rotate it through to the left about a central axis, em, about a hundred and twenty degrees.

You want to, em, turn it, say a hundred and thirty degrees, em, anticlockwise, away from me.

### 2) Agent-implicit

Em, it needs to, be sort of made level by tilting downwards towards my left by about, em, forty five degrees maybe.

It will be rotated towards me upward and about forty, thirty five, forty degrees.

### 3) Agent-less

Em it's, it's a rotation sort of clockwise, but through the horizontal plane, em, by around a hundred degrees.

About, em, about eighty degrees to the right.



### Figure Captions

*Figure 1.* Presentation of a stimulus in Experiment 1 (Left, 60 degrees x-axis rotation; Right, the object in the canonical position).

*Figure 2.* Mean hand-object-interaction and object-movement gesture rates (per minute) in the first and second halves of Experiment 1. The error bars represent standard error.

*Figure 3.* Mean hand-object-interaction and object-movement gesture rates (per minute) in the first, second and third practice trials of Experiment 1. The error bars represent standard error.

*Figure 4.* Mean proportion of near-screen hand-object-interaction and object-movement gestures in the first and second half of Experiment 1. The error bars represent standard error.

*Figure 5.* Mean hand-object-interaction and object-movement gesture rates (per minute) in the first, second and third trials of Experiment 2. The error bars represent standard error.

*Figure 6.* Presentation of a stimulus in Experiment 3 (Lower object, 60 degrees on bisector of x-axis and y-axis rotation; Upper left and right objects in the canonical position).

*Figure 7.* Mean hand-object-interaction and object-movement gesture rates (per minute) in the first and second halves of Experiment 3. The error bars represent standard error.

Figure 1.

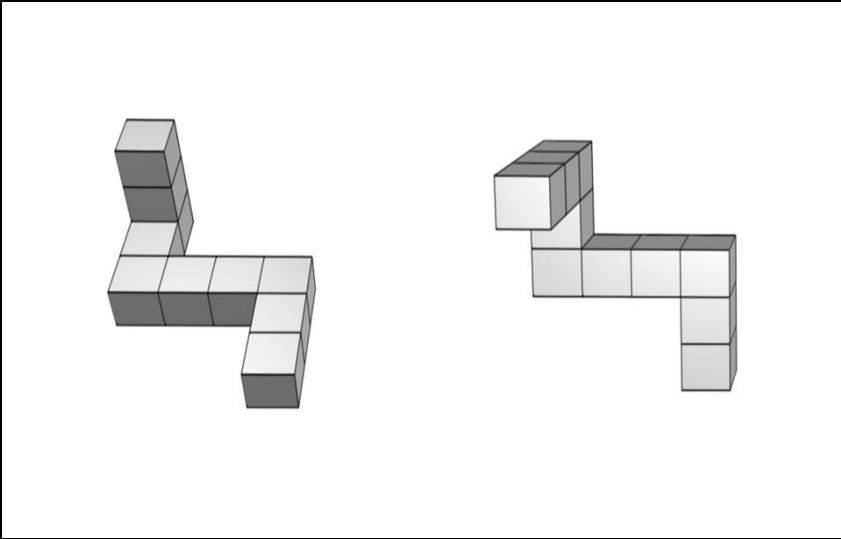


Figure 2.

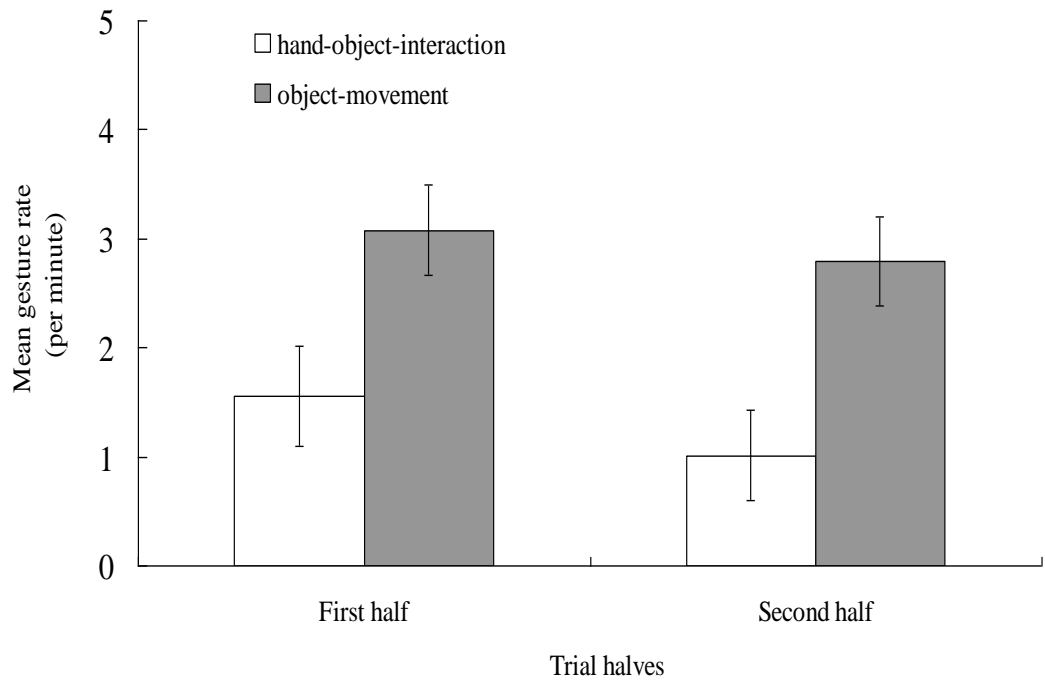


Figure 3.

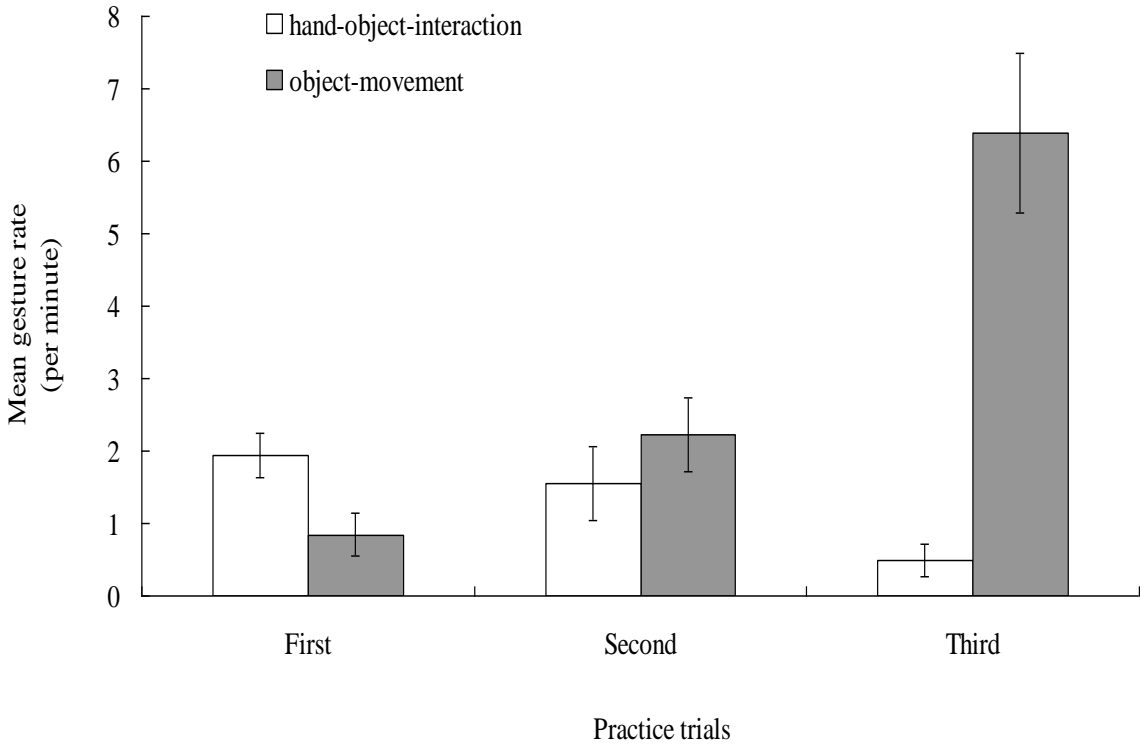


Figure 4.

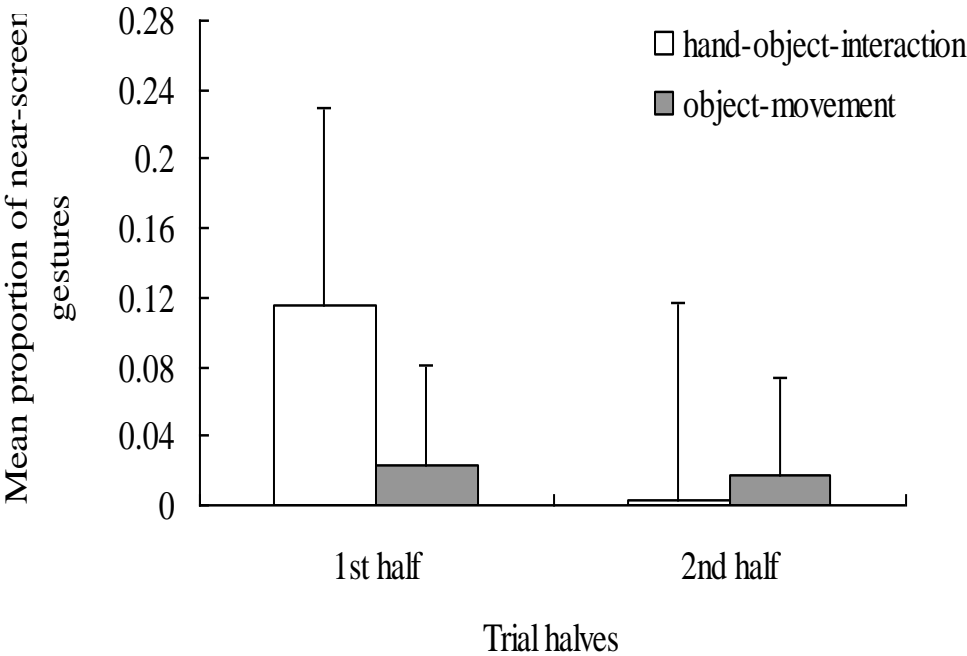


Figure 5.

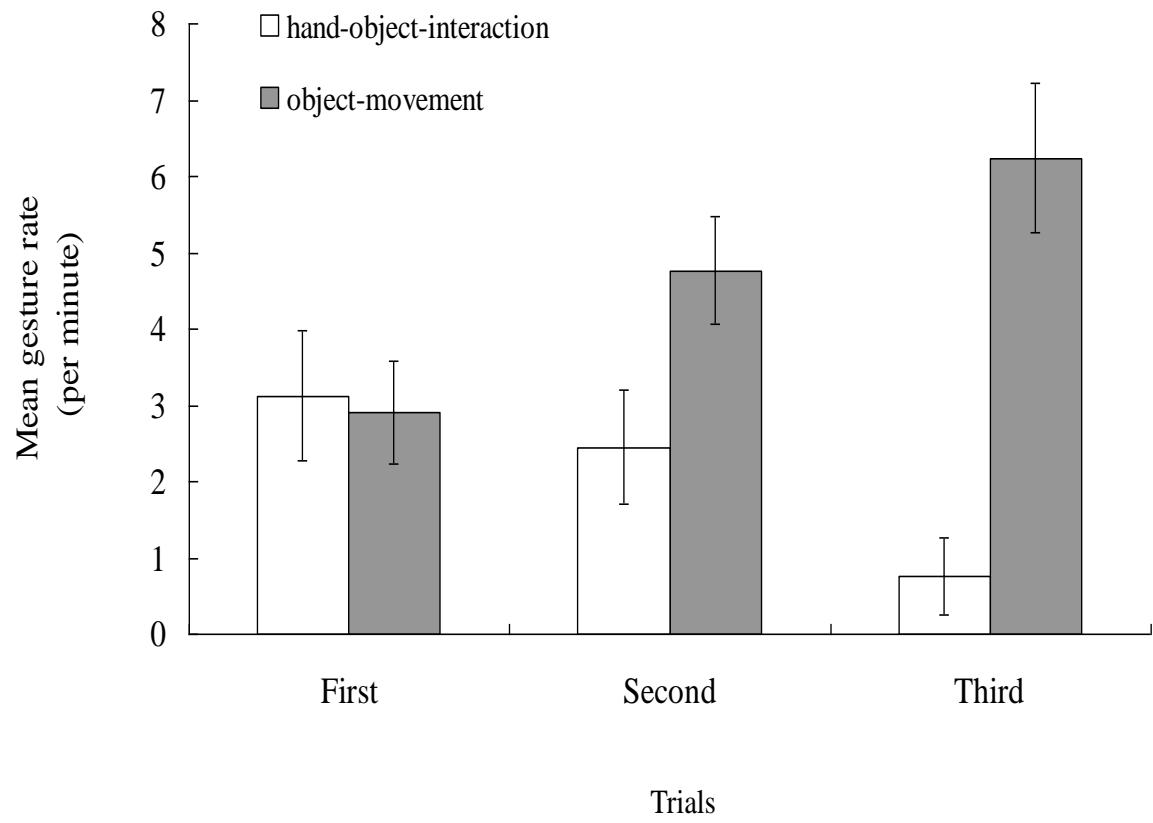


Figure 6.

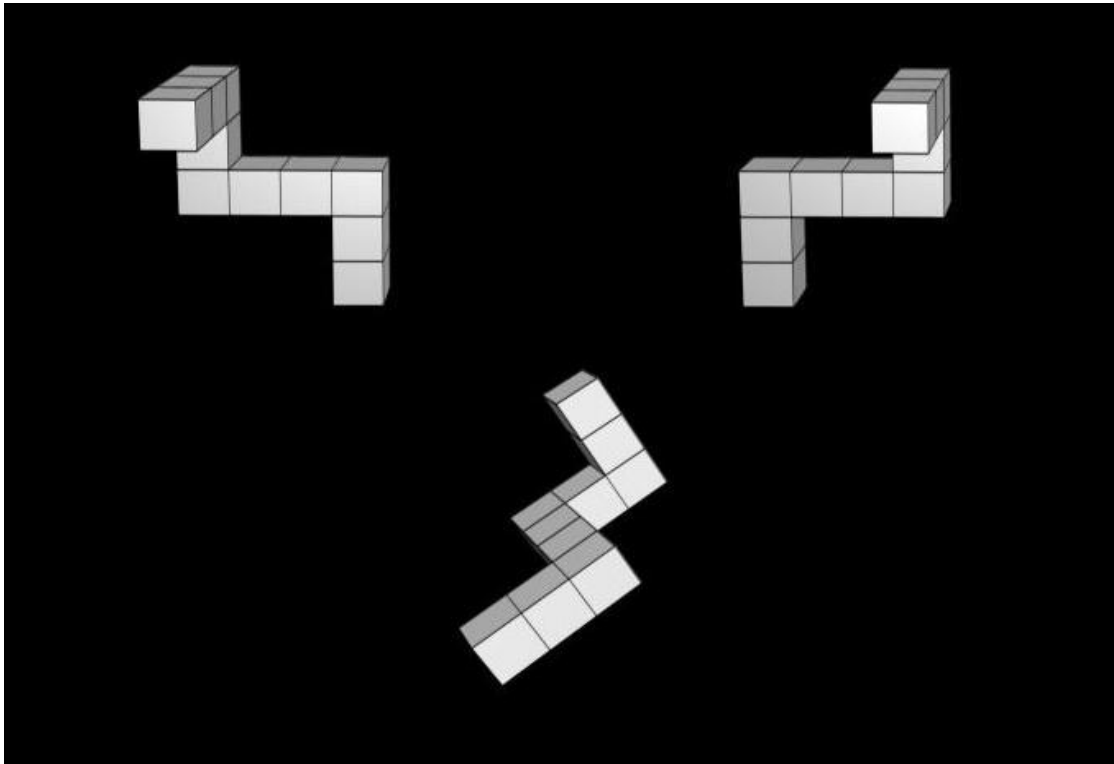


Figure 7.

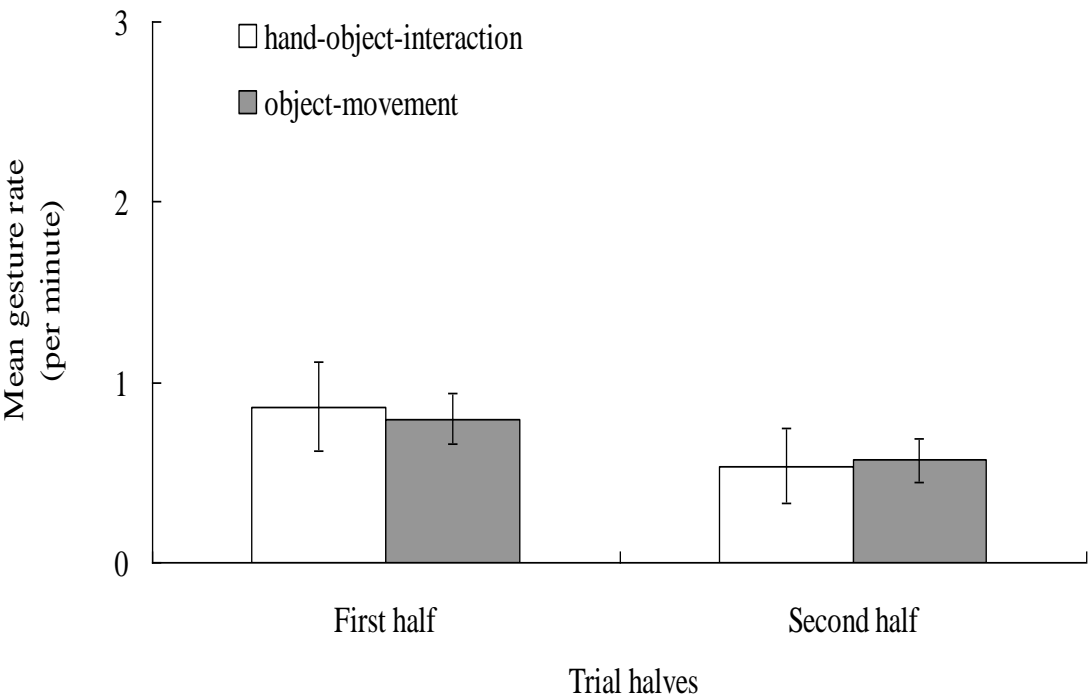




Table 1

*Number of participants in the four groups created by the gesture-based and speech-based criterion for deagentivization. This table includes only the gesturers whose linguistic description mode changed over the three trials.*

Gesture	Speech	
	Unequivocal Deagentivization	Not
Unequivocal Deagentivization	4	2
Not	1	9

Table 2

*Number of participants in the four groups based on whether using agent-explicit or non-agent-explicit description in the first trial and whether gestures were allowed or prohibited.*

Condition	Speech mode in the first trial	
	Agent-explicit	Non-agent-explicit
Gesture allowed	8	17
Gesture prohibited	15	9

Table 3

*Number of participants in the four groups based on whether description modes deagentivized or not in the second or the third trial and whether gestures were allowed or prohibited.*

Condition	Speech mode in the second or the third trial	
	Deagentivized	Not
Gesture allowed	3	5
Gesture prohibited	0	15