

Running head: GESTURE AND SPATIAL PROBLEM SOLVING

The nature of gestures' beneficial role in spatial problem solving

Mingyuan Chu and Sotaro Kita

University of Birmingham

## Abstract

*Co-thought* gestures are hand movements produced in silent, non-communicative, problem-solving situations. The study investigated whether and how such gestures enhance performance in spatial visualization tasks such as a mental rotation task and a paper folding task. We found that participants gestured more often when they had difficulties solving mental rotation problems (Experiment 1). The gesture-encouraged group solved more mental rotation problems correctly than the gesture-allowed and gesture-prohibited groups (Experiment 2). Gestures produced by the gesture-encouraged group enhanced performance in the very trials in which they were produced (Experiments 2 &3). Furthermore, gesture frequency decreased as the participants in the gesture-encouraged group solved more problems (Experiments 2&3). In addition, the advantage of the gesture-encouraged group persisted into subsequent spatial visualization problems in which gesturing was prohibited: another mental rotation block (Experiment 2) and a newly introduced paper-folding task (Experiment 3). The results indicate that when people have difficulty in solving spatial visualization problems, they spontaneously produce gestures to help them and gestures can indeed improve performance. As they solve more problems, the spatial computation supported by gestures becomes internalized and the gesture frequency decreases. The benefit of gestures persists even in subsequent spatial visualization problems where gesture is prohibited. Moreover, the beneficial effect of gesturing can be generalized to a different spatial visualization task when two tasks require similar spatial transformation processes. We concluded that gestures enhance performance on spatial visualization tasks by improving the internal computation of spatial transformations.

*Keywords:* gesture, spatial problem solving, mental rotation, paper folding, embodied cognition

### The nature of gestures' beneficial role in spatial problem solving

Spatial ability plays a key role in reasoning and communication, and it is closely linked to our achievements in science, mathematics and engineering (Humphreys, Lubinski, & Yao, 1993). Therefore, how to improve one's spatial ability has become one of the central research topics in cognitive psychology and has wide implications for education.

Spatial ability is not a unitary construct. Several different spatial ability factors have been identified by psychometric studies (e.g., Carroll, 1993; Eliot & Smith, 1983; Lohman, 1979, 1988; McGee, 1979). For example, Lohman (1979) claimed that spatial ability is composed of three major factors, including: spatial visualization, spatial relations and spatial orientation. Spatial visualization is defined as the ability to mentally transform complex stimuli (e.g., three-dimensional object) in space. Spatial relation is an ability to rapidly recognize the identity of a simple item (e.g., a two-dimensional object) under rotation in a speeded task. Spatial orientation is an ability to imagine how stimuli will appear from another perspective. Among different factors, the strongest and most consistent factor that has been supported in the literature is spatial visualization. Hegarty and Waller (2006) defined spatial visualization as "the ability to mentally manipulate, rotate, twist, or invert objects without reference to one's self" (Hegarty & Waller, 2006, p. 127). The present study focused on the beneficial role of gesture in solving spatial visualization problems, such as a three-dimensional mental rotation task and a paper folding task, which are loaded on the spatial visualization factor (Hegarty & Waller, 2006).

Recent research has demonstrated that hand movements play a key role in solving three-dimensional mental rotation problems. A mental rotation task typically requires participants to judge whether a pair of asymmetric three-dimensional objects are identical or

mirror images (e.g., Shepard and Metzler, 1971). It has been found that when participants were required to manually turn a knob or joystick while solving a Shepard and Metzler (1971) type mental rotation task, they responded more quickly and accurately when the direction of hand movements were congruent with the direction of mental rotation than when they were incongruent (Wexler, Kosslyn, & Berthoz, 1998; Wohlschläger & Wohlschläger, 1998). In these studies, however, participants were forced to manually rotate an object in particular ways (e.g., turning a knob clockwise or anticlockwise) in every trial. Thus, it is not clear whether individuals would spontaneously produce hand movements, or gestures, in order to improve performance of mental rotation tasks.

The literature on speech-accompanying hand gestures (“co-speech” gestures) suggests that gesture and spatial visualization are closely linked. When individuals are required to provide verbal descriptions of how they solve spatial problems, they often spontaneously produce co-speech gestures (Chu & Kita, 2008; Ehrlich, Levine, & Goldin-Meadow, 2006; Garber & Goldin-Meadow, 2002; Schwartz & Black, 1996). Gestures are particularly frequent when people describe spatial transformations (e.g., to mentally change, move, create or remove objects) (Trafton et al., 2006). More importantly, individuals produce gestures more frequently when it is difficult to verbally describe spatial visualization (Hostetter, Alibali, & Kita, 2007). Therefore, co-speech gestures may be triggered in order to facilitate spatial visualization during spatial problem solving.

Most of the previous studies have investigated the role of gestures in spatial visualization by examining the co-speech gestures produced during the verbal explanation of solutions (e.g., Alibali, Kita & Young, 2000; Ehrlich, Levine, & Goldin-Meadow, 2006; Goldin-Meadow, Cook, & Mitchell, 2009; Kita & Davies, 2009; Melinger & Kita, 2007; Trafton et al., 2006); however, people also spontaneously produce gestures while

thinking silently (“co-thought” gestures) during spatial visualization problem solving (Chu & Kita, 2008, 2009; Hegarty et al, 2005; Schwartz & Black, 1996). To our knowledge, there have been very few studies that explicitly explore the role of co-thought gestures in spatial problem solving. In the present study we examine the functional role of co-thought gestures produced in spatial visualization tasks, in which participants were seated alone in a room without speaking.

Recent studies have shown that co-speech gestures can play a causal role in the learning process during mathematical problem solving. Children who were told to produce gestures relevant to mathematical equivalence problems on a whiteboard profited more from subsequent instruction about the problems compared to children told not to gesture: they solved more equivalence problems correctly in the post-instruction test (Broaders et al., 2007) and retained learned knowledge better up to four weeks later (Cook, Mitchell & Goldin-Meadow, 2008). However, neither of the two studies provided direct evidence that gesture itself improved mathematical problem solving: because both studies found that the told-to-gesture group and the told-not-to-gesture group solved comparable number of problems correctly during the gesture manipulation phase. Thus, gesture may activate relevant knowledge and make children more ready to learn, but it may not improve mathematical problem solving until children have received instructions that tell them how to use the relevant knowledge to solve the problems. In the present study, we directly tested whether co-thought gestures improve performance in spatial visualization tasks without any instruction.

The studies on gestural facilitation of learning of mathematical problems (Broaders et al., 2007; Cook et al., 2008; Goldin-Meadow et al. 2009) opened up a question as to whether gesture manipulation changed children's internal computation of mathematical problems. In

these studies, gesture manipulation was followed by an instruction and then a post-test. It is possible that gestures that highlighted useful information in the gesture manipulation phase were re-activated during the post-test. These studies did not report whether gestures were prohibited during the post-test and they did not provide information about participants' gesture behaviors during the post-test either<sup>1</sup>. Thus, it is possible that the combination of gesture manipulation and instruction taught children useful gestural strategies for future use, but they did not improve children's internal computation of mathematical problems. In the current study, we investigated whether gestures could improve internal spatial computation and benefit subsequent tasks in which gestures were not allowed.

To summarize, the previous research showed that people perform spatial visualization tasks better when they are instructed to move their hand in the way congruent with what they need to visualize (Wexler, Kosslyn, & Berthoz, 1998; Wohlschläger & Wohlschläger, 1998). However, the question remains whether people spontaneously produce useful hand movements to help them solve spatial visualization problems when they are not forced to move their hands. The previous research on the gestural benefit during problem solving has focused mainly on co-speech gestures produced in verbal descriptions of problem solving processes and how such co-speech gestures improve people's readiness to learn from subsequent instruction (Broaders et al., 2007; Goldin-Meadow et al. 2009). However, the question remains whether co-thought gestures (produced during silent thinking) can directly facilitate problem solving and improve relevant internal computation, such that the gestural facilitation extends to subsequent non-gesturing trials and to a new task that shares similar problem solving processes. The current study addressed these open questions by investigating the function of co-thought gestures produced during mental rotation, a classical spatial visualization task.

*Present Study*

We hypothesized that people would spontaneously produce co-thought gestures especially when they have difficulty in solving spatial visualization problems, such as mental rotation problems. Furthermore, we hypothesized that co-thought gestures would enhance performance of spatial visualization tasks by improving the internal computation of spatial transformations<sup>2</sup>. For example, by producing co-thought gestures that simulate manipulating, rotating or twisting an object, one taps into rich sensori-motor experiences of hand-object interaction and is, therefore, able to compute more accurate information on what the object would look like when transformed by the hand. Furthermore, by producing co-thought gestures in which the hand represents an object, one taps into the knowledge of how our body parts look when they are moved, and use that knowledge to compute more accurate information on how the object would look when the object moves. In addition, co-thought gestures could also improve performance of the spatial visualization tasks by facilitating spatial working memory. For example, one may offload the intermediate representations of spatial transformations to the gesturing hand in order to reduce the chance of forgetting the representations. If co-thought gestures only facilitate spatial visualization by offloading the intermediate representations, the beneficial effects should disappear in subsequent trials when gesture is no longer allowed.

We also hypothesized that the spatial computation supported by gestures would become internalized after one repeatedly solves similar spatial visualization problems (Chu & Kita, 2008). That is, when solving novel problems, people start with overt hand movements to help them explore information relevant for the problem solving process. The spatial computation supported by gestures consolidates gradually. This computation eventually becomes fully internalized and does not require the support from overt gestures. The internalized

computation can be either supported by imagined gestures or based on improved non-motoric visualization of object rotations. If internalization happens, the gesture rate should drop over the course of the experiment, and the gestural benefit should extend to subsequent trials in which gestures are not available. Furthermore, if the internalized computation is flexible enough to handle other similar types of spatial transformations, the benefit gained from gesturing during a mental rotation task should generalize to a new spatial visualization task that also requires similar spatial transformations. Lastly, although it has been shown that producing gestures in the primary task of a dual task experiment leads to better performance in a secondary digit span task (measuring verbal working memory) or a secondary visual pattern task (measuring visual working memory) (Wagner, Nusbaum, & Goldin-Meadow, 2004, we did not hypothesize that the gestural benefit gained in a mental rotation task can generalize to these two tasks (digit span and visual pattern tasks), as these working memory tasks do not require any spatial transformation processes.

To test these hypotheses, we first investigated whether difficulty in spatial visualization processes triggered spontaneous co-thought gestures, and whether co-thought gestures improved performance in spatial visualization tasks. We predicted that people would produce gestures more often in difficult spatial visualization problems than in easy ones (Experiment 1) and they would perform better when they were encouraged to gesture than when they were not encouraged to gesture or when they were prohibited from gesturing (Experiment 2).

Second, we investigated whether co-thought gestures enhanced performance in spatial visualization tasks through improved internal computation for spatial transformations. We predicted that the rate of overt gestures would drop over the course of the trials (Experiments 2 and 3) and that the benefit gained from gesturing would extend to subsequent spatial visualization problems where gestures were not available (Experiment 2).



Third, we investigated whether the beneficial effect of co-thought gestures was problem-specific or problem-general. We predicted that benefits of gesturing during mental rotation would generalize to a different spatial visualization task that requires similar spatial transformation processes (Experiment 3), but not to other working memory tasks that did not involve any spatial transformations (Experiments 2 & 3).

### Experiment 1

In Experiment 1, we investigated the spontaneous co-thought gestures produced in a non-communicative Shepard-Metzler (1971) mental rotation task. The main goal of Experiment 1 was to examine whether people produced more spontaneous co-thought gestures in difficult spatial visualization problems than in easy ones.

### Method

#### *Participants*

One hundred and thirty two students (98 females, mean age: 20.12, age range: 18-33) at the University of Birmingham took part in the study. All participants had normal or corrected-to-normal vision. They were awarded either course credits or £4 for their participation.

#### *Stimuli*

Shepard and Metzler (1971) type three-dimensional objects were used. They were created by “Blender”, which is free 3D graphics creation software (this software can be downloaded from [www.blender.org](http://www.blender.org)). 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. The edge length of each cube on the computer screen was 1 cm (see Figure 1).

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

Each stimulus consisted of two three-dimensional objects at the top of the screen and

one at the bottom of the 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 the horizontal axis, the vertical axis, or the axis pointing to depth. The lower object was rotated by four angles ( $60^\circ$ ,  $120^\circ$ ,  $240^\circ$  and  $300^\circ$ ) around the bisector which went through the object's centre between the horizontal and vertical axis (XY axis), the horizontal and in-depth axis (XZ axis), and the vertical and in-depth axis (YZ axis). The lower object was rotated from the upper left object in half of the trials and from the upper right object in the other half of the trials.

### *Apparatus*

Stimuli were presented centrally on a 15-in. (38.1-cm) LCD monitor. The participants' performance was captured by a hidden camera located on the participant's left and about 2 meters away. The video was recorded on a Sony DCR-HC19E PAL camcorder (at 25 frames per second).

### *Design*

There were 24 experimental trials (left vs. right  $\times$  4 angles  $\times$  3 axes). Stimuli were randomly presented.

### *Procedure*

The participants were tested individually. They were seated approximately 70 cm in front of the monitor. Each trial began with a white fixation cross in the center of the screen for 1,000 ms, followed by the stimulus. The task was to decide whether the lower object was the same as the upper left or right object by pressing the correspondent left or right foot pedal. The instruction did not mention gesture. When the response was given, the next trial started automatically. The participants 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

co-thought gestures were not suppressed due to time pressure. The participants solved the problems alone in a room and their gestures were captured by the hidden camera. There were no practice trials before the experimental trials. No feedback was given concerning the accuracy of the response. None of the participants reported awareness of the hidden camera and they were given the opportunity to request the recording to be erased.

### *Gesture Coding*

Gesture coding was carried out with video annotation software ELAN (European Distributed Corpora Project [EUDICO] Linguistic Annotator), developed by the Max Planck Institute for Psycholinguistics. Gestures were segmented into a series of gesture strokes (Kendon, 1980; McNeill, 1992). The segmentation was carried out following the procedure in Kita, van Gijn, and van der Hulst, (1998). Gesture strokes are performed more forcefully than other phases of gestures (e.g., preparation), and they express meanings of gestures.

Each gesture was coded according to the following classification system (developed on the basis of the classification system in McNeill, 1992). *Representational gestures* were the hand movements that either represented the interaction between the hand and objects, the perceptual and motion information of the objects themselves or pointing at the objects<sup>2</sup>objects<sup>3</sup>. For example, if a participant rotated her hand with the index finger and thumb opposed to each other as if to grasp the object in her hand and rotate it, or she drew circles with her right index finger as if to simulate the rotation of the object, or she simply pointed to a component of the object, these gestures were coded as representational gestures. *Unclear Gestures* were those that could not be clearly classified as representational gestures, including: (1) abandoned gestures, in which the gesturer interrupted a gesture before the stroke was completed and retracted it or moved on to the next different gesture; (2) those resembling emblem gestures, which conveyed some known meaning, such as “maybe” (e.g.,

a flat hand with the palm down, wavering).

To establish inter-coder reliability, 15% of all gestures coding ( $N = 119$ ) were randomly selected and coded by a second independent coder. The two coders matched on 98.32% of the selected gestures (Cohen's  $k = .87$ ,  $p < .001$ ).

### Results and Discussion

Participants solved on average 19.91 of the 24 problems correctly ( $SD = 3.51$ ). The average reaction time of each trial was 11.36 seconds ( $SD = 6.58$ ). Participants produced a total of 792 gestures of which 745 (94.07%) were representational gestures and 47 (5.93%) were unclear gestures (46 abandoned gestures and 1 gesture that resembled an emblem gesture). They produced representational gestures in 2.43 of the 24 trials on average ( $SD = 4.26$ ). In the analyses throughout the paper, trials without representational gestures are referred to as non-gesturing trials and participants who did not produce any representational gestures are referred to as non-gesturers.

#### *Mental Rotation Performance across Four Angles*

The purpose of the first set of analyses is to examine whether participants produced more co-thought gestures in difficult spatial visualization problems than in easy ones. The most robust findings in mental rotation studies have been that reaction times and error rates increase monotonically with the angle of rotation (Shepard & Cooper, 1982; Shepard & Metzler, 1971). Therefore, we first examined whether participants in our present study indeed had longer reaction times (RTs) and made more errors in 120° and 240° trials than in the 60° and 300° trials.

First, RTs were submitted to a one factor repeated measure analysis of variance (ANOVA) with rotation angles (60°, 120°, 240° and 300°) as the independent variable. There was a main effect of angles,  $F(3, 393) = 87.40$ ,  $MSE = 34.55$ ,  $p < .001$ ,  $\eta_p^2 = .40$  (See Figure

2 for the means and standard errors). Tukey posthoc tests showed that the RTs were significantly longer for 120° and 240° trials than for 60° and 300° trials ( $p < .01$ ), but there was no difference between 60° and 300° trials or between 120° and 240° trials.

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

Second, the error rates (proportion of trials with an erroneous response) were submitted to a one factor repeated measure ANOVA with rotation angles (60°, 120°, 240° and 300°) as the independent variable. There was a main effect of angles,  $F(3, 393) = 40.97$ ,  $MSE = 0.03$ ,  $p < .001$ ,  $\eta_p^2 = .24$  (See Figure 3 for the means and standard errors). Tukey posthoc tests showed that error rates were significantly higher in 120° and 240° trials than in 60° and 300° trials ( $p < .01$ ), but there was no difference between 60° and 300° trials or between 120° and 240° trials.

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

We also analyzed non-gesturing trials to examine how rotation angles affected mental rotation performance independently of possible effects of gestures. The aim was to establish the difficulty of the four angles independently of gesturing. We essentially replicated the above findings. There was a main effect of angles on RTs,  $F(3, 393) = 63.42$ ,  $MSE = 15.31$ ,  $p < .001$ ,  $\eta_p^2 = .33$  (60°:  $M = 7.79$  seconds,  $SD = 4.67$ ; 120°:  $M = 12.06$  seconds,  $SD = 6.97$ ; 240°:  $M = 12.36$  seconds,  $SD = 7.45$ ; 300°:  $M = 7.96$  seconds,  $SD = 4.86$ ). Tukey posthoc tests showed that the RTs were significantly longer in 120° and 240° trials than in 60° and 300° trials ( $p < .01$ ), but there was no difference between 60° and 300° trials or between 120° and 240° trials. There was a main effect of angles on error rates (proportion of trials with an erroneous response),  $F(3, 381) = 30.27$ ,  $MSE = 0.03$ ,  $p < .001$ ,  $\eta_p^2 = .19$  (60°:  $M = .10$ ,  $SD = 0.16$ ; 120°:  $M = .20$ ,  $SD = 0.22$ ; 240°:  $M = .26$ ,  $SD = 0.26$ ; 300°:  $M = .10$ ,  $SD = 0.15$ ). Tukey posthoc tests showed that error rates were significantly higher in 120° and 240° trials than in

60° and 300° trials ( $p < .01$ ), and the error rates were significantly higher in 240° trials than in 120° trials ( $p < .05$ ), but there was no difference between 60° and 300° trials.

Taken together, participants used longer RT and made more errors in 120° and 240° trials than in 60° and 300° trials. Thus, 120° and 240° trials were harder than 60° and 300° trials.

#### *Gesture Frequencies across Four Angles: Problem Difficulty Predicts Gesture Frequency*

If participants spontaneously seek help from gestures when they have difficulties in solving spatial visualization problems, they should produce co-thought gestures more often in 120° and 240° trials than in 60° and 300° trials. The rates of representational gestures (number of representational gestures per minute) were submitted to a one factor repeated measure ANOVA with rotation angles (60°, 120°, 240° and 300°) as the independent variable. There was a main effect of angles,  $F(3, 393) = 6.14$ ,  $MSE = 3.23$ ,  $p < .001$ ,  $\eta_p^2 = .05$  (See Figure 4 for the means and standard errors). Tukey posthoc tests showed that the gesture rates were significantly higher in 120° and 240° trials than in 60° ( $p < .01$ ) and 300° ( $p < .05$ ) trials, but there was no difference between 60° and 300° trials or between 120° and 240° trials.

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

The above results showed that people produced gestures more often in difficult spatial visualization problems than in easy ones and the difficulty was established independent of gestures. The results are consistent with the idea that people spontaneously seek help from gesture when solving difficult spatial visualization problems. However, the fact that participants produced more co-thought gestures in difficult spatial visualization problems does not mean that gestures can improve performance in spatial visualization problems. That is, even though individuals spontaneously produce gestures in order to help them solve spatial problems, these gestures do not necessarily lead to better performance. It is possible that

co-thought gestures are merely an external bodily reflection of individuals' spatial visualization process. The harder the spatial visualization process is, the more gestures are produced as by-products of the effortful spatial thinking process. Furthermore, it is even possible that gestures actually impede performance on a mental rotation task instead of improving performance. This issue was addressed in Experiment 2 in which we experimentally manipulate the availability of gestures in a mental rotation task.

### Experiment 2

The first goal of Experiment 2 is to directly examine whether co-thought gestures improve performance in spatial visualization problems. Participants were given two identical blocks of the same mental rotation task as used in Experiment 1, except that the length of each trial was fixed in order to eliminate a possible speed-accuracy trade-off. In the first block, participants were allocated to the following three groups: (1) the gesture-encouraged group in which participants were encouraged to gesture while solving mental rotation problems; (2) the gesture-prohibited group in which participants were required to sit on their hands to inhibit gesturing. If co-thought gestures improve performance in spatial visualization problems, the gesture-encouraged group should perform better than the gesture-prohibited group; (3) the gesture-allowed group in which gesture was not mentioned in the instruction although participants were still allowed to move their hands freely. The purpose of using a gesture-allowed group as another control group is to ensure that the better performance of the gesture-encouraged group compared to the gesture-prohibited group was not due to any detrimental effect caused by prohibiting gesture (e.g., participants felt uncomfortable when sitting on their hands). Thus, if gesture improved spatial visualization problem solving, the gesture-encouraged group should also perform better than the gesture-allowed group.

The second goal of Experiment 2 was to investigate how co-thought gestures improved

performance in spatial visualization problems. We first examined whether the beneficial effect of gesture could extend to the second mental rotation block where gesture was not available. If gestures improve solving spatial visualization problems only at the moment of gesturing, by offloading the intermediate representation of the stimulus object to the gesturing hand, then the three groups should perform equally well when they were not allowed to gesture. In contrast, if gestures improve the internal computation of spatial transformations, the gesture-encouraged group should perform better than the other two groups in the second mental rotation block, when gesture was prohibited.

We further examined whether the beneficial effect of gesture was only confined to spatial visualization problems or it could help any task, for example, through raised alertness. Participants were also given a digit span task, which is a typical verbal working memory task. If gesture only improves the internal computation of spatial transformation, the three groups should perform equally well in the digit span task.

## Method

### *Participants*

Sixty six students (57 females, mean age: 19.97, age range: 18 – 37) at the University of Birmingham took part in the study. None of the participants participated in Experiment 1. All participants had normal or corrected-to-normal vision. They were awarded either course credits or £3 for their participation.

### *Tasks*

**Mental Rotation Task.** The mental rotation task was the same as the one used in Experiment 1 except that the trial duration was fixed. In each trial, the stimulus stayed on the screen for 11.36 seconds (the same as the mean RT in Experiment 1), and the computer did not react to any response in this period. After 11.36 seconds, the stimulus was automatically



replaced by the word “Respond”. When a response was given, the next trial started automatically. If participants did not respond within 3 seconds after the word “Respond” appeared, the next trial started automatically and this trial was counted as an incorrect response.

Digit Span Task (adapted from Wechsler, 1991). This task measures verbal working memory. The stimulus consisted of a sequence of digits randomly selected from 0 – 9. Each digit was presented on the screen for 1000ms before the screen went blank for 500ms, and this procedure was repeated for other digits until the total number of digit in the sequence reached a certain number. Participants’ task was to immediately recall the digit sequence. Following two practice trials, there were 25 trials in the real experiment, starting from 5-digit sequence and steadily increasing to 9-digit sequence with 5 trials for each sequence-length. A response was considered correct only if all the digits were recalled in the correct order.

### *Procedure*

The participants were tested individually and they were assigned to either the gesture-encouraged group ( $N = 22$ ), the gesture-allowed group ( $N = 22$ ) or the gesture-prohibited group ( $N = 22$ ). The experiment consisted of three blocks. In the first block, the gesture-allowed group were given the same instruction as the one used in Experiment 1, except that participants were told that they had 11 seconds to work on the problem and that they needed to respond with the left or right foot pedal as fast as possible after they saw the word “Respond”. The other two groups were given the same instruction as the one given to the gesture-allowed group, except that the gesture-encouraged group was also encouraged to move their hands whenever they thought it might be helpful to solve the problems and the gesture-prohibited group was required to sit on their hands throughout the experiment. In the second and the third blocks, participants were either given the same mental

rotation task as used in the first block, although all three groups were required to sit on their hands while solving the problems or they were given the digit span task depending on the counterbalance order. There was no break in the middle of the experiment. Each mental rotation block took about 5 minutes to complete and the digit span task took about 10 minutes to complete. Participants were not required to sit on their hands during the digit span task. No feedback was given concerning the accuracy of participants' responses in all three blocks. The gestures in the first block were captured by the same hidden camera used in Experiment 1. None of the participants reported awareness of the hidden camera and they were given the opportunity to request erasing the recording.

### *Gesture Coding*

Gesture categories were the same as used in Experiment 1. 15% of all gesture coding ( $N = 70$ ) were randomly selected and coded by a second independent coder. The two coders matched on 98.57% of the selected gesture coding (Cohen's  $k = .85$ ,  $p < .001$ ).

### *Results and Discussion*

In the first block, the gesture-encouraged group produced a total of 400 gestures, of which 393 (98.25%) were representational gestures, and the gesture-allowed group produced a total of 63 gestures, of which 62 (98.41%) were representational gestures. Participants produced representational gestures in 8.86 ( $SD = 8.55$ ) out of 24 trials on average in the gesture-encouraged group, and in 1.64 ( $SD = 3.63$ ) out of 24 trials on average in the gesture-allowed group.

### *Gesture Rates Comparison between the Gesture-encouraged and Gesture-allowed Groups.*

In this analysis, we compared the rates of representational gestures (number of representational gestures per minute) produced by the gesture-encouraged group and the gesture-allowed group in order to ensure our experimental manipulation of gesture in the first

block was effective. The rates of representational gestures produced by the gesture-encouraged group ( $M = 3.93$ ,  $SD = 4.64$ ) were significantly higher than those of the gesture-allowed group ( $M = 0.62$ ,  $SD = 1.47$ ),  $t(42) = 3.19$ ,  $p < .01$ ,  $d = 0.96$ . Thus, the manipulation was successful in that the gesture-encouraged group produced more gestures than the gesture-allowed group did.

#### *Accuracy Comparison among Three Groups.*

To test whether co-thought gestures lead to better performance in the mental rotation task, the error rates (proportion of trials with an erroneous response) were submitted to a  $2 \times 3$  analysis of variance (ANOVA) with block (first mental rotation block vs. second mental rotation block) as a within-participants factor and group (gesture-encouraged group, gesture-allowed group, vs. gesture-prohibited group) as a between-participants factor (see Figure 5 for the means and standard errors). There was a main effect of block, such that all three groups made significantly less errors in the second mental rotation block than in the first mental rotation block,  $F(1, 63) = 89.08$ ,  $MSE = .38$ ,  $p < .01$ ,  $\eta_p^2 = 0.59$ . There was a main effect of group,  $F(2, 63) = 4.51$ ,  $MSE = 0.02$ ,  $p < .05$ ,  $\eta_p^2 = 0.13$ . The interaction between task and group was not significant,  $F(2, 63) = 0.43$ ,  $MSE < 0.01$ ,  $p = .65$ . As the comparisons of groups in the first and second mental rotation blocks address different theoretical questions, we carried out post-hoc comparisons of groups separately for the first and second blocks. Fisher LSD posthoc tests (as suggested by Howell (2007) for the comparison of three means) showed that the gesture-encouraged group solved significantly more problems correctly than the gesture-prohibited group ( $p < .01$ ) and the gesture-allowed group ( $p < .05$ ) in both the first and second mental rotation blocks. Furthermore, the performance of the gesture-allowed group and the gesture-prohibited group did not significantly differ from each other in the first and second mental rotation blocks. Thus,

encouraging participants to gesture not only improved the mental rotation performance in the first mental rotation block but also improved performance in the subsequent mental rotation block where gesture was not allowed.

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

#### *RT Comparison among Three Groups.*

In the current experiment, participants had a maximum of three seconds to respond after the word “Respond” replaced the stimulus on the screen. In order to make sure that the higher accuracy of the gesture-encouraged group was not because the gesture-encouraged group took longer to respond after the stimuli disappeared, the RTs were submitted to a  $2 \times 3$  analysis of variance (ANOVA) with block (first mental rotation block vs. second mental rotation block) as a within-participants factor and group (gesture-encouraged group, gesture-allowed group, vs. gesture-prohibited group) as a between-participants factor. There was a main effect of block, such that all three groups used less time to respond in the second mental rotation block than in the first mental rotation block,  $F(1, 63) = 20.06$ ,  $MSE = 0.01$ ,  $p < .01$ ,  $\eta_p^2 = 0.24$ . There was no main effect of group,  $F(2, 63) = 0.22$ ,  $MSE = 0.08$ ,  $p = .80$ . The interaction between block and group was not significant,  $F(2, 63) = 2.45$ ,  $MSE = 0.01$ ,  $p = .09$ . (In the first mental rotation block: for the gesture-encouraged group,  $M = 0.98$  seconds,  $SD = 0.23$ ; for the gesture-allowed group,  $M = 0.92$  seconds,  $SD = 0.18$ ; for the gesture-prohibited group,  $M = 0.95$  seconds,  $SD = 0.23$ ; In the second mental rotation block: for the gesture-encouraged group,  $M = 0.89$  seconds,  $SD = 0.25$ ; for the gesture-allowed group,  $M = 0.90$  seconds,  $SD = 0.19$ ; for the gesture-prohibited group,  $M = 0.83$  seconds,  $SD = 0.16$ ). Therefore, the lack of main effect of group indicates that there was no speed-accuracy trade-off, and the better performance of the gesture-encouraged group was not because they took longer to respond than the two control groups.

*Accuracy and Gesture Frequencies across Four Angles in the Gesture-encouraged group.*

In the gesture-encouraged group, we explicitly encouraged participants to gesture when they felt it necessary. Although participants still chose when and how to gesture freely, the participants were more conscious about gesture in the gesture-encouraged group than the participants who were not told any thing about gestures in Experiment 1. To ensure encouraging participants to gesture did not qualitatively change their gesture behaviour, we examined whether our findings from Experiment 1, that more difficult trials triggered more gestures, could be replicated in the gesture-encouraged group in the current experiment.

First, error rates were submitted to a one factor repeated measure ANOVA with rotation angles (60°, 120°, 240°, and 300°) as the independent variable. There was a main effect of angles,  $F(3, 63) = 10.17$ ,  $MSE = 0.02$ ,  $p < .01$ ,  $\eta_p^2 = 0.33$  (see Figure 6 for means and standard errors). Tukey posthoc tests showed that error rates were significantly higher for 120° and 240° trials than for 60° ( $p < .01$ ) and 300° ( $p < .05$ ) trials, but there was no difference between 60° and 300° trials or between 120° and 240° trials.

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

Second, the rates of representational gestures was submitted to a one factor repeated measure ANOVA with rotation angles (60°, 120°, 240° and 300°) as the independent variable. There was a main effect of angles,  $F(3,63) = 6.61$ ,  $MSE = 8.18$ ,  $p < .01$ ,  $\eta_p^2 = 0.24$  (see Figure 7 for means and standard errors). Tukey posthoc tests showed that gesture rates were higher for 240° trials than for 60° and 300° ( $p < .05$ ) trials, and gesture rates were higher for 120° trials than for 60° ( $p < .05$ ) ~~and 300° (not statistically significant) trials~~. There was no significant difference between 60° and 300° trials ~~or,~~ between 120° and 240° trials, or between 120° and 300° trials (thought there was a clear trend that the gesture rates were higher in 120° trials than in 300° trials). Participants made more errors in 120° and 240° trials

than in 60° and 300° trials and they produced gestures more often in 120° and 240° trials than in 60° and 300° trials. Thus, we replicated our findings in Experiment 1 that more difficult trials triggered more gestures. The results suggest that encouraging people to gesture did not qualitatively change participants' gesture behaviour in terms of the relationship between the difficulty of the problems and the gesture frequency.

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

#### *Analysis of the Digit Span Task*

To test whether the beneficial effect of gesture was specific to spatial visualization task or whether it could be generalized to any task, the error rates (proportions of trials with an erroneous response) of the digit span task were submitted to a one-way ANOVA with groups (gesture-encouraged group, gesture-allowed group, and gesture-prohibited group) as the independent variable. One participant was excluded from the analysis, as he was trying to use his mobile phone to record the digits as revealed by our hidden camera. There was no main effect of group,  $F(2, 62) = 0.52, p = .60$  (For the gesture-encouraged group:  $M = .56, SD = .22$ ; for the gesture-allowed group:  $M = .60, SD = .17$ ; for the gesture-prohibited group:  $M = .62, SD = .20$ ). Some participants used enumeration gestures during the digit span task. For example, some participants counted their fingers while they were trying to remember the digit sequences. In a further analysis, we performed the same one-way ANOVA only with participants who did not gesture during the digit span task (16 participants from the gesture-encouraged group, 16 from the gesture-allowed group and 17 from the gesture-prohibited groups were included in the analysis), there was no main effect of group,  $F(2, 46) = 0.21, p = .81$  (For the gesture-encouraged group:  $M = .55, SD = .19$ ; for the gesture-allowed group:  $M = .59, SD = .19$ ; for the gesture-prohibited group:  $M = .60, SD = .21$ ). Thus, it was not the case that the beneficial effect of gesture generalized to any task.

The results from Experiment 2 indicate that co-thought gestures can indeed improve performance in a mental rotation task, and the beneficial effects of gesture can extend to the subsequent mental rotation problems where gesture is not available. However, the beneficial effects did not extend to the verbal working memory task, which did not involve any spatial transformation. The results support our hypothesis that gesture facilitates the internal computation of spatial transformations during spatial visualization, and cannot be explained by the idea that gesture facilitates spatial visualization *only* by offloading the intermediate representation of the stimulus object to the gesturing hand.

### Experiment 3

The goal of Experiment 3 was to investigate whether the beneficial effect of gesture is problem-specific or problem-general. Although Experiment 2 showed that gesture has a lasting beneficial effect in solving mental rotation task, it is still unclear whether the beneficial effect of gesture can be generalized to new objects and new tasks that share similar spatial transformation processes. Therefore, in Experiment 3, we examined whether the benefit gained from gesture in the mental rotation task can be generalized to a subsequent non-gesturing paper folding task.

Performing a mental rotation task requires creating a mental representation of the object, rotating the object mentally until its orientation allows comparison to the object in the canonical position, comparing the mentally rotated object to the standard object, and finally making a judgement and a response (Johnson, 1990). Similarly, a paper folding task (Ekstrom et al., 1976) requires participants to first mentally form an image of a piece of paper, then mentally fold the piece of paper in the manner illustrated by the dotted lines and imagine a hole being punched through the folded paper in a location indicated by a circle, next mentally unfolding the piece of paper, and finally decide which of five figures corresponds to the

locations of the punched holes in the unfolded paper. It has been proposed that mental rotations of three-dimensional objects and paper folding test are typical spatial visualization tasks (Hegarty & Waller, 2006) and require similar spatial transformation processes (Wright et al., 2008). Furthermore, it has been found that individuals' performance on a mental rotation task significantly correlates with their performance on a paper folding test (Miyake et al., 2001). Thus, if the beneficial effect of co-thought gestures in spatial visualization tasks is problem-general, the participants who were encouraged to gesture in the first block with the mental rotation task should perform better on the subsequent (non-gesturing) paper folding block, compared to those who were not encouraged to gesture in the preceding mental rotation block.

Experiment 3 also investigated the limits of gestural benefit within the visuospatial domain. To this end, we replaced the digit span task in Experiment 2 with a visual pattern task (adapted from Della Sala, et al., 1997). The visual pattern task requires retaining a visual pattern for a short period of time. Crucially, this task does not involve any spatial transformations. If gesture improves internal computation for spatial transformations, performance of both groups should not significantly differ from each other in the (non-gesturing) visual pattern block.

Thus, Experiment 3 was the same as Experiment 2 except that: (1) only the gesture-encouraged group and the gesture-allowed group were included because the gesture-allowed group and the gesture-prohibited group showed very similar results in all analyses in Experiment 2; (2) we replaced the second mental rotation block in Experiment 2 by a block with a paper folding task, which took about 8 minutes to complete; (3) we replaced the digit span block in Experiment 2 by a visual pattern block, which took about 7 minutes to complete.



## Method

*Participants*

Thirty two students (27 females, mean age: 19.47, age range: 18 - 30) at the University of Birmingham took part in the study. None of them participated in Experiment [2-1](#) and [3-2](#). All participants had normal or corrected-to-normal vision. They were awarded either course credit or £3 for their participation.

*Tasks*

**Mental Rotation Task.** The mental rotation task used in Experiment 3 was the same as the one used in the first block of Experiment 2.

**Paper Folding Task.** Adapted from The Kit of Factor-Referenced Cognitive Tests (Ekstrom, French, Harman, & Derman, 1976), the test consisted of 20 problems which were randomly presented to participants. Each problem depicted a square piece of paper folded in the manners illustrated by the dotted lines and then hole-punched as indicated by a small circle (see Figure 8). Participants' task was to choose, among five choices, the correct pattern showing what the paper would look like when it was unfolded. In each trial, the stimulus stayed on the screen for 18.42 seconds (this was the mean RT of each trial calculated from 11 participants in a pilot study). After 18.42 seconds, the stimulus was automatically replaced by the word "Respond" and participants were required to say their answer out loud, into a tape recorder, as quickly as possible. The next trial started automatically 3 seconds after the word "Respond" showed up, and if participants did not respond within 3 seconds, this trial was counted as an incorrect response.

\*\*\* Figure 8 about here \*\*\*

**Visual Patterns Task.** Adapted from Della Sala et al. (1997), the stimulus consisted of a grid of blocks, in which half of the blocks were filled in black. The pattern was displayed for

3 seconds in each trial before being replaced by a blank grid with exactly the same dimensions as the one bearing the pattern but with different consonant letters in each block. Participants' task was to immediately recall the blocks that were filled in black by reading aloud the corresponding letters into a tape recorder. Following 2 practice trials, there were 25 trials in the experiment, starting from 7 black blocks and steadily increasing to 11 black blocks with 5 trials at each level. The patterns of each trial were never repeated. A trial was counted as correct only if all the black blocks were recalled.

### *Procedure*

The procedure was the same as used in Experiment 2, except that the second *non-gesturing* mental rotation block in Experiment 2 was replaced by the *non-gesturing* paper folding block, and the block with the digit span task in Experiment 2 was replaced by a block with the *non-gesturing* visual pattern task. The participants in both gesture-encouraged and gesture-allowed conditions were required to sit on their hand during the paper folding block and the visual pattern block.

### *Gesture Coding*

Gesture categories were the same as used in Experiment 1. 15% of all gesture coding ( $N = 35$ ) were randomly selected and coded by a second independent coder. The two coders matched on 97.14% of the selected gesture coding (Cohen's  $k = .65$ ,  $p < .001$ ).

### *Results and Discussion*

In the mental rotation task, the gesture-encouraged group produced a total of 206 gestures, of which 195 (94.66%) were representational gestures, and the gesture-allowed group produced a total of 19 gestures, all of which were representational gestures. The gesture-encouraged group produced representational gestures in 7.63 ( $SD = 6.11$ ) out of 24 trials on average, and the gesture-allowed group produced representational gestures in 1.25

( $SD = 1.65$ ) out of 24 trials on average.

*Gesture Rates Comparison between the Gesture-encouraged and Gesture-allowed Groups.*

We replicated our finding from the first mental rotation block of Experiment 2 that the rates of representational gestures (number of representational gestures per minute) produced by the gesture-encouraged group ( $M = 2.68$ ,  $SD = 2.75$ ) were significantly higher than those of the gesture-allowed group ( $M = 0.26$ ,  $SD = 0.72$ ),  $t(30) = 3.41$ ,  $p < .01$ ,  $d = 1.20$ . Thus, the manipulation to encourage the participants to gesture was successful.

*Error Rates Comparison between Gesture-encouraged group and Gesture-allowed group.*

We replicated our findings from the mental rotation task of Experiment 2. The error rates (proportions of trials with an erroneous response) for the mental rotation task were significantly lower in the gesture-encouraged group ( $M = .18$ ,  $SD = .13$ ) than in the gesture-allowed group ( $M = .27$ ,  $SD = .13$ ),  $t(30) = 2.03$ ,  $p = .05$ ,  $d = 0.69$ .

*Analysis of the Paper Folding Task*

*Accuracy Comparison between Two Groups.* To test whether the beneficial effect of gesture could be generalized to the paper folding task, we compared the error rates between the gesture-encouraged group and the gesture-allowed group. The error rates were significantly lower in the gesture-encouraged group than in the gesture-allowed group,  $t(30) = 2.57$ ,  $p < .05$ ,  $d = 0.91$  (see Figure 9) for means and standard errors). Thus, the benefit of gesturing can be generalized to a different spatial task that requires similar spatial transformations.

\*\*\* Figure 9 about here \*\*\*

*RT Comparison between Two Groups.* Again, we compared the RTs between the gesture-encouraged group and the gesture-allowed group in the paper folding block. The RTs were not significantly different between the gesture-encouraged group ( $M = 0.90$  seconds,  $SD$

= 0.22) and the gesture-allowed group ( $M = 0.87$  seconds,  $SD = 0.18$ ),  $t(30) = 0.48$ ,  $p = 0.63$ .

Therefore, there was no speed-accuracy trade-off.

#### *Analysis of the Visual Pattern Task*

To test whether the beneficial effect of gesture was specific to spatial transformations or whether it could be generalized to any visuospatial task, we compared the error rates of the visual pattern task between the gesture-encouraged group and the gesture-allowed group. The error rates (proportion of trials with an erroneous response) in the visual pattern block were not significantly different between the gesture-encouraged group ( $M = .73$ ,  $SD = .13$ ) and the gesture-allowed group ( $M = .76$ ,  $SD = .18$ ),  $t(30) = 0.53$ ,  $p = .60$ . Thus, the benefit of gesture in spatial problem solving can only be generalized when the tasks require similar spatial transformations, but not for any visuospatial task.

#### Further Analyses of the Pooled Data from Experiments 2 and 3

In this section, we report further analyses based on the data from Experiments 2 and 3. The two experiments used identical materials, task and procedure for the first mental rotation block in the gesture-encouraged and gesture-allowed conditions. Thus, we pooled the data from the two experiments to increase power and investigated the following two questions.

The first question concerned what the source of the advantage of the gesture-encouraged group over the gesture-allowed group was. Our theory predicted that gestures should enhance performance of the very trials in which they were produced. However, the analyses in Experiments 2 and 3 did not clarify whether gestures enhanced performance of the trials where they were produced or enhanced performance of non-gesturing trials, or both. We will clarify this issue.

The second question concerned whether the spatial computation supported by overt gestures became internalized over trials. According to our theory, individuals produce

co-thought gestures to help their internal computation of spatial transformations when they first solve spatial visualization problems. As internal computation improves over trials, at some point, overt gestures should become no longer needed. If this was the case, the number of gestures should decrease over the course of the first mental rotation block.

*Did gestures enhance performance in the very trials in which they were produced?*

To answer this question, we compared performance of the gesture-encouraged group and gesture-allowed group in an analysis that distinguished between the trials in which the gesture-encouraged group gestured but the gesture-allowed group did not and those in which neither group gestured. We predicted that the advantage of the gesture-encouraged group should be larger in the former type of trials than in the latter type of trials. In this analysis, the type of mental rotation problems (i.e., axes, angles and directions of rotation) was controlled for when comparing the performances of the two groups. Thus, any differences between the groups cannot be attributed to problem difficulty.

The data for this analysis was prepared in the following way. We first matched the participants in the gesture-encouraged group and those in the gesture-allowed group, according to their overall performance in the first mental rotation block. We ranked participants in each group separately, according to their overall error rates, and paired participants who had the same rank in their respective groups. The analysis focused on the gesturers ( $N = 33$ ) in the gesture-encouraged group and their paired participants in the gesture-allowed group. For each pair of participants, we paired trials with the same mental rotation problem in terms of the axis, angle, and direction of rotation. (Note that 24 distinct mental rotation problems were presented in 24 trials.) We divided the paired trials into gesturing trials and non-gesturing trials, depending on whether the participant in the gesture-encouraged group produced at least one gesture. We excluded the paired trials in

which the participant from the gesture-allowed group produced a gesture (26 trials out of 396 trials were excluded). This ensures that in the gesturing trials only the participant from the gesture-encouraged group produced a gesture and that in the non-gesturing trials neither participant produced a gesture. Note that pairing the trials ensured that we controlled for the type of problems (i.e. axes, angles and direction of rotation) solved when comparing the gesture-encouraged group and the gesture-allowed group. We also excluded paired participants who had only one gesturing or non-gesturing trial to reduce variability in the data by avoiding proportions where the denominator equaled one (4 pairs out of 33 pairs of participants were excluded).

Error rates (proportions of trials with an erroneous response) were submitted to a repeated measures  $2 \times 2$  analysis of variance (ANOVA) with trial type (gesturing trial vs. non-gesturing trial), and group (gesture-encouraged group, vs. gesture-allowed group) as independent variables. Not surprisingly, given the results from Experiments 2 and 3, there was a main effect of group, that is, the gesture-encouraged group made fewer errors than the gesture-allowed group,  $F(1, 28) = 32.02$ ,  $MSE = 0.01$ ,  $p < .01$ ,  $\eta_p^2 = 0.53$ . There was no main effect of trial type,  $F(1, 28) = 0.71$ ,  $MSE = 0.03$ ,  $p = .41$ . More importantly, the interaction between trial type and group was significant,  $F(1, 28) = 5.33$ ,  $MSE = 0.01$ ,  $p < .05$ ,  $\eta_p^2 = 0.16$ . (See Figure 10 for means and standard errors). Tukey posthoc tests showed that the error rates of the gesturing trials in the gesture-encouraged group were significantly lower than the error rate of the trials with the same mental rotation problems in the gesture-allowed group ( $p < .01$ ), but the error rates of the non-gesturing trials in the two groups were not significantly different from each other. In addition, the error rate differences between the gesturing and non-gesturing trials were not significant for either the gesture-encouraged group or the gesture-allowed group.

It is worth noting that in the current analysis the gesture-encouraged group made numerically fewer errors than the gesture-allowed group in the non-gesturing trials in the first mental rotation block. However, this difference was not statistically significant, unlike the finding in Experiments 2 and 3 that the gesture-encouraged group made significantly fewer errors than the gesture-allowed group in the subsequent non-gesturing blocks with spatial visualization problems. A question arises as to why the effect of gesturing was weaker in the non-gesturing trials in the first mental rotation block than in the subsequent non-gesturing blocks with spatial visualization problems. We speculate two reasons for this difference. First, in the first mental rotation block some of the non-gesturing trials occurred before participants started producing gestures, and therefore could not benefit from improved internal spatial computation due to preceding gestures. Second, it is possible that each gesture improves internal spatial computation only slightly and the beneficial effect of gesturing has to accumulate over a certain number of gestures (e.g., produced during the first mental rotation block) before it becomes detectable (e.g., in the second non-gesturing mental rotation block). Thus, we did not see a strong beneficial effect of gesturing in non-gesturing trials in the first mental rotation block.

\*\*\* Figure 10 about here \*\*\*

The result of the current analysis indicates that the gestures produced by the gesture-encouraged group enhanced performance in the very trials in which they were produced. In other words, the advantage of the gesture-encouraged group over the gesture-allowed group in the first mental rotation block reported in Experiments 2 and 3 was due to improved performance in the trials in which the gesture-encouraged group actually produced a gesture.

Finally, the finding that the gesture-encouraged group did not perform significantly

better in the gesturing trials than the non-gesturing trials might seem to contradict our conclusion that gesture improved performance in the gesturing trials. However, this is not surprising if we consider the fact that participants were more likely to gesture in the difficult trials than in the easy trials, as shown in Experiments 1 and 2. The gesturing trials may have simply included more difficult problems than the non-gesturing trials. To examine this possibility, we calculated the proportion of difficult problems (120° and 240° rotation angles) in the gesturing trials and the non-gesturing trials. The proportion of difficult problems was indeed marginally significantly higher in the gesturing trials ( $M = .56, SD = .25$ ) than in the non-gesturing trials ( $M = .44, SD = .16$ ),  $t(28) = 1.86, p = .07, d = 0.57$ . Thus, the gesturing trials were more difficult than the non-gesturing trials. Taken together, the key finding is clear: gestures improved mental rotation performance in the trials in which they were produced.

*Did gestures produced by the gesture-encouraged group internalize over the course of the task?*

In this analysis, we divided the 24 trials of the first mental rotation block into three sections (first 8 trials, second 8 trials, and last 8 trials). According to our hypotheses, gesture should gradually become internalized in the course of the experiment, so the rates of overt gestures should decrease over trials. The rates of representational gestures (number of representational gestures per minute) produced by the gesture-encouraged group in the first mental rotation block were submitted to a one factor repeated measures analysis of variance (ANOVA) with sections as the independent variable. There was a significant main effect of sections,  $F(2, 74) = 6.12, MSE = 4.37, p < .01, \eta_p^2 = 0.14$  (see Figure 11 for means and standard errors). Fisher LSD posthoc tests (as suggested by Howell (2007) for the comparison of three means) showed that the number of gestures was significantly higher in the first section than in the second ( $p < .05$ ) and third sections ( $p < .01$ ). Although the difference



between the number of gestures in the second and third sections was not statistically significant, there was a clear trend that participants gestured more often in the second section than in the third section. Furthermore, we examined the error rate across the three sections. We found a main effect of section,  $F(2, 74) = 3.74$ ,  $MSE = 0.02$ ,  $p < .05$ ,  $\eta_p^2 = 0.09$  (first section:  $M = .21$ ,  $SD = 0.16$ ; second section:  $M = .20$ ,  $SD = 0.15$ , third section;  $M = .14$ ,  $SD = 0.14$ ). Namely, the error rate was significantly lower in the third section than in the first and second sections ( $p < .05$ ). Taken together, the results suggest that as participants became more experienced in the task, the spatial computation supported by gestures became internalized and overt hand movements were no longer required. The drop of gesture rates was not due to fatigue or lack of concentration towards the end of the block, as the performance improved over trials.

\*\*\* Figure 11 about here \*\*\*

### General Discussion

The purpose of the present study was to investigate the functional role of silent co-thought gestures produced during spatial visualization problem solving. Our results showed that: (1) participants spontaneously produced gestures more often when they had difficulties in solving mental rotation problems (Experiment 1); (2) the gesture-encouraged group correctly solved more mental rotation problems than the gesture-allowed group or the gesture-prohibited group (Experiment 2); (3) performance of the gesture-encouraged group in gesturing trials was better than performance in the equivalent trials of the gesture-allowed group in the first mental rotation block (Experiment 2 & 3); (4) the rates of representational gestures produced by the gesture-encouraged group decreased over the course of the first mental rotation block (Experiment 2 & 3); (5) the group who were encouraged to gesture in the first mental rotation block solved more problems correctly in the subsequent

non-gesturing mental rotation block, compared to the gesture-allowed or the gesture-prohibited groups (Experiment 2); (6) the group who were encouraged to gesture in the mental rotation block solved more problems correctly in the subsequent non-gesturing paper folding block, compared to the gesture-allowed group (Experiment 3); (7) the group who were encouraged to gesture in the first mental rotation block and the two control groups solved a comparable number of problems correctly in the subsequent verbal working memory and visual working memory tasks (Experiment 2 & 3). Therefore, the results indicate that difficulty in a spatial visualization task triggers spontaneous co-thought gestures and these gestures indeed improve performance in the spatial visualization task at the moment of gesture production. As participants become used to the task, the spatial computation supported by gestures become internalized and gesture frequency decreases. Furthermore, the benefit of co-thought gestures can extend to subsequent trials in which gesture is prohibited. This gestural benefit can be generalized to a different spatial visualization task that requires similar spatial transformation, but not to verbal and visual working memory tasks that do not involve spatial transformation. The following subsections will discuss the findings in more details.

#### *Difficulty in Spatial Visualization Triggers Spontaneous Gestures*

Previous studies have shown that people are more likely to produce co-speech gestures when they talk about spatial information than when they talk about non-spatial information (Alibali, Heath, & Myers, 2001; Krauss, 1998; Rauscher, Krauss, & Chen, 1996; Trafton et al., 2006), and that people gesture more often when they have difficulty organizing spatial information for speaking (Alibali, Kita, & Young, 2000; Hostetter, Alibali, & Kita, 2007; Kita & Davies, 2009; Melinger & Kita, 2007). In the present study, we have shown that when people solve spatial visualization problems without speaking or communication, they still

spontaneously produce silent co-thought gestures in order to facilitate problem solving. To our knowledge, the present study is the first study to demonstrate that individuals spontaneously produce hand movements, or co-thought gestures, in order to improve performance in spatial visualization tasks especially when the tasks are difficult.

However, the participants in Experiment 1 produced gestures only in 2.43 out of 24 trials. If participants produced gestures in order to improve performance in spatial visualization tasks, why did they only gesture in so few trials? We have two speculations. First, producing gestures might require extra resources from the cognitive system, so gesture might be triggered only when the spatial visualization is very difficult and could not be accomplished purely mentally. In Experiment 1, participants on average solved about 20 out of 24 trials correctly, so most trials may not have been sufficiently difficult to trigger co-thought gestures. Secondly, participants may not have had a meta-awareness that gesture could help them, and therefore they did not produce many gestures. When we explicitly encouraged participants to use their hands to help them solve the problem, the average number of gesturing trials increased to 8.86 (Experiment 2) and 7.63 (Experiment 3) out of 24 trials.

### *Gesture Improves Spatial Problem Solving*

Action has often been found to facilitate solving spatial visualization problems. For example, manually rotating a knob or a joystick could enhance mental rotation performance if the manual and mental rotation direction were congruent with each other (Wexler, et al. 1998; Wohlschlager & Wohlschlager, 1998). In these studies, however, participants were instructed to move their hands in a specific way that was hypothesized to be helpful in solving the problems. In Experiment 2 of the present study, participants in the gesture-encouraged group were only told that they could use their hands to facilitate problem solving when they felt it

necessary, so they could freely choose when and how to move their hands. Therefore, the better performance in the gesture-encouraged group than in the two control groups indicates that participants were able to use appropriate gestural movements at appropriate times (when there was a need) to improve performance on spatial visualization tasks, and these gestures indeed lead to better performance.

Our results are also compatible with the findings that gesture plays an active role in children's learning process (Broaders et al., 2007; Cook, Mitchell, & Goldin-Meadow, 2008; Goldin-Meadow, Cook & Mitchell, 2009). For example, children who were encouraged to gesture during their explanation of mathematical problems were more likely to profit from subsequent instruction and solved more problems correctly in the post-instruction test, compared to those who were prohibited from gesturing (Broaders et al., 2007). Furthermore, children, who were required to produce gestures that contained a correct strategy for solving mathematical problems, learned more than children, who were required to produce gestures that contained a partially correct strategy (Goldin-Meadow, Cook & Mitchell, 2009). However, encouraging children to gesture did not directly improve their performance in mathematical problem solving at the moment when gestures were produced. The gesture-encouraged group and the control group performed equally well at the manipulation phase prior to the instruction phase. The gesture-encouraged group only performed better than the control group after both groups received further instruction on how to solve the problems correctly (Broaders et al., 2007; Cook, Mitchell & Goldin-Meadow, 2008; Goldin-Meadow, Cook & Mitchell, 2009). In contrast, the present study showed that the gesture-encouraged group solved more mental rotations problems correctly than the control groups did, though no instruction was given to any of the groups. Solving mathematical equation problems may require abstract mathematical knowledge that can readily be provided

by instruction but not by gestures. Thus, merely encouraging children to gesture was not sufficient to make children understand the problems. In addition, these studies only examined children who did not know how to solve the mathematical problems, and therefore verbal instruction may have been an essential part of the children's learning process. In contrast, adults in the present study knew how to solve mental rotation problems, and therefore encouraging them to gesture was sufficient to lead to better performance.

It should be noted that in Experiment 1 and the first mental rotation block in Experiment 2, co-thought gestures were associated with trials with more errors and longer reaction times. From this, we concluded that difficulty in problem solving triggered gestures. However, a question can be raised as to why performance was worse in the trials with gestures if gesture facilitated spatial computation. We speculate that this is because the benefit gained from gestures was not big enough to completely overcome the difficulty caused, for example, by larger angles of rotation. It is also a logical possibility that co-thought gestures actually hurt performance in spatial visualization tasks. However, this idea was ruled out by the finding from the analysis on the pooled data from Experiments 2 and 3. We found that performance of the gesture-encouraged group in gesturing trials was significantly better than performance of the gesture-allowed group in the corresponding trials with the same problems but without any gesturing. Thus, we maintain that gesture improves spatial problem solving.

#### *How Gesture Improves Solving Spatial Visualization Tasks*

Little is known in the literature about how gestures improve performance in spatial visualization tasks. Solving spatial visualization tasks requires both the execution of multiple steps of mental spatial transformations and the temporary maintenance of the intermediate representations of the stimulus object (Carpenter & Just, 1986; Hegarty & Sims, 1994; Salthouse et al., 1990; Shah and Miyake, 1996). Gesture may either improve performance in

spatial visualization tasks by helping spatial working memory or by facilitating the internal computation of spatial transformations, but of course gesture may benefit both of these.

Previous studies on co-speech gestures have shown that gesture can reduce working memory load during speaking (Goldin-Meadow et al., 2001; Wagner, Nusbaum and Goldin-Meadow, 2004). It has been found that when children and adults were asked to remember verbal and visuospatial items while explaining mathematical problems, they remembered significantly more items when they gestured than when they did not gesture during their verbal explanation of math problems (Wagner, Nusbaum and Goldin-Meadow, 2004). Therefore, it is possible that co-thought gestures in the current study facilitated spatial working memory in the spatial visualization tasks by offloading the intermediate representations of spatial transformations to the gesturing hand to reduce the chance of forgetting the representations.

However, the fact that the beneficial effect of co-thought gestures could extend to the subsequent non-gesturing mental rotation and paper folding tasks cannot be explained by the idea that gesture facilitated the maintenance of intermediate representations in spatial working memory. If co-thought gestures only helped gesturers offload intermediate mental representations to physical (gestural) representations, the beneficial effect should have disappeared when gestures were no longer available. Therefore, our study provides evidence that co-thought gestures improved the internal computation of spatial transformations in visualization tasks. It is important to note that our study did not refute the offloading hypothesis. Rather, it demonstrated that the off-loading hypothesis alone cannot account for all the gestural benefits in spatial visualization tasks<sup>4</sup>.

In the first mental rotation block in Experiment 2 and Experiment 3, participants who were encouraged to use their hand movements had more chances to use gestures to explore and construct more efficient solution strategies to complete the spatial transformations

required by the mental rotation task. Gesture, as a simulated action (Hostetter & Alibali, 2008), can provide a rich sensori-motor representation of the physical world and pick up organization of information that is less readily available to visuo-spatial processes (Kita, 2000). For example, using gestures to simulate grasping an object and rotating it might allow gesturers to use their rich experience of hand-object interaction to compute more accurate information regarding how the object could be rotated by hand for different axes, angles, and directions. In fact, evidence has shown that adults' knowledge about a physical event can be constructed through imagined actions on the physical object (e.g., Gibbs, 2006; Shapiro, 2007). Individuals can make correct physical inferences about water in a glass when imagining acting on the glass with a hand (Schwartz & Black, 1999). In addition to using their hands to simulate manipulating an object, participants might also use their hands to represent an object itself and rotate their hands in the air to simulate the rotation of the object. This can help participants to capture more spatial characteristics of the object, such as relative location of parts and the orientation of planes during rotation. By gesturing, participants were more able to keep track of the changes in the appearance of an object under rotation, and consequently became more able to predict how the object would look when it was rotated on a given axis for given degrees and directions.

As the gesture-encouraged group became better at solving the problem through the help of gestures, participants did not need to rely on external hand movements any more towards the end of the first mental rotation block. We suggest that the decline in gesture frequency reflects internalization of the computation of spatial transformations supported by gestures (Chu & Kita, 2008). Thus, even when gesture was prohibited in the subsequent mental rotation block or paper folding block, participants in the gesture-encouraged group were still able to use the benefit gained from gesturing in the first mental rotation block, as compared to

the control groups. Here, we speculate two possibilities as to how gestures could improve performance on subsequent non-gesturing blocks. The first possibility is that when participants in the gesture-encouraged group were prohibited from gesturing, they used imagined hand movements to facilitate spatial visualization and consequently performed better than the control groups. For example, they might imagine using their hands to rotate the three-dimensional object, or to fold and unfold the paper around different axes and to different directions. Or they might imagine rotating their hands to represent the rotation of the three-dimensional object in the mental rotation task or the paper in the paper folding task. The second possibility is that in the non-gesturing mental rotation block and paper folding block, participants might solve the problems without imagining hand movements at all, as they became more proficient at predicting the consequences of rotation of an object along certain axes with the help of gesture in the first mental rotation block.

*What kind of tasks do gestures facilitate?*

The current study showed that gestures can facilitate spatial visualization tasks during which gestures are produced (Experiments 2, 3). This does not rule out the possibility that the production of gestures can facilitate other types of concurrent tasks. In fact, it has been shown that producing gestures in the primary task of a dual task experiment leads to better retention of verbal or visual information in secondary working memory tasks (Wagner, Nusbaum, & Goldin-Meadow, 2004). It was claimed that gesture reduces demand on mental resources, which leaves extra resources available to a variety of concurrent tasks.

However, the current study showed that gestures can only benefit the subsequent mental rotation and paper folding tasks, but not in the digit span or visual pattern tasks. In other words, the gestural benefit was specific to tasks that involved spatial transformations. Note that the digit span and visual pattern tasks in the current study were essentially the same



as the secondary tasks in Wagner and colleague's (2004) study, which were facilitated by gesturing in the concurrent primary task. Thus, gestures benefit a wide range of concurrent tasks, but only specific types of subsequent (non-gesturing) tasks that require mental computation directly improved by gestures such as spatial transformations.

### *The Relationship between Co-speech and Co-thought Gestures*

Although in the current study we focused on co-thought gestures, produced during silent thinking process, it is interesting to investigate the relationship between co-thought and co-speech gestures that represent aspects of spatial visualization problems. If these two types of gestures were completely distinctive behaviors, it would be difficult to generalize our findings on co-thought gestures to co-speech gestures. However, a systematic relationship between co-thought and co-speech gestures has been demonstrated in three ways. First, in Chu and Kita (2008), participants were either required to verbally explain their solution of a mental rotation task to the experimenter or to silently solve the same mental rotation task as we used in Experiment 1. Over the course of the experiment, the rates of both co-thought and co-speech gestures decreased, and the representational contents of both co-thought and co-speech gestures changed in the same manner. Second, both co-speech and co-thought gestures are triggered by task difficulty. The current study showed that people produced more co-thought gestures in more difficult mental rotation problems. Similarly, people produce more co-speech representational gestures when it is more difficult to verbalize spatial information (Alibali, Kita & Young, 2000; Melinger & Kita, 2007; Hostetter, Kita, & Alibali, 2007; Kita & Davies, 2009). Third, in Kita and Chu (2009), the participants who produced more co-thought gestures in a non-communicative mental rotation task also produced more co-speech gestures in an animation description task. In other words, the rates of co-thought gestures and co-speech gestures were positively correlated. These findings suggest that

co-thought and co-speech gestures are likely to share some underlying processes. Thus, like co-thought gestures, co-speech gestures could benefit spatial visualization problem solving as well. This would be an interesting topic for future studies.

### *Conclusion*

To summarize, the present study investigated the functional role of co-thought gestures in solving spatial visualization problems. Difficult spatial visualization problems triggered more co-thought gestures, and co-thought gestures improved performance in spatial visualization problems. As participants solved a particular type of problem repeatedly, spatial computation supported by gesturing became internalized and gesture frequency decreased. The beneficial effect of co-thought gestures extended to subsequent problems in which gesturing was prohibited and the beneficial effect also generalized to a different spatial visualization task that required similar spatial transformations. Our findings indicate that people spontaneously deploy gestures to facilitate spatial problem solving, and gesture improves the internal computation of spatial transformation in a problem-general way.

References

- 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 & Language, 44*, 169-188.
- 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.
- Broaders, S. C., Cook, S. W., Mitchell, Z. A. & Goldin-Meadow, S. (2007). Making children gesture brings out implicit knowledge and leads to learning. *Journal of Experimental Psychology: General, 136*, 539-550.
- Carpenter, P. A. & Just, M. A. (1986). Spatial ability: an information processing approach to psychometrics. In R. J. Sternberg (Eds.), *Advances in the psychology of human intelligence* (Vol. 3), (pp. 221-253). Hillsdale NJ: Erlbaum.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytical studies*. New York : Cambridge University Press.
- 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.
- Chu, M., & Kita, S. (2009). Co-speech gestures do not originate from speech production processes: Evidence from the relationship between co-thought and co-speech gestures. In *Proceedings of the 31st Annual Conference of the Cognitive Science Society*.
- Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2008). Gesture makes learning last. *Cognition, 106*, 1047 - 1058.
- Della Sala, S., Gray, C., Baddeley, A., Wilson, L. (1997) *The Visual Pattern Test*. Thames Valley Test Company.

- Ehrlich, S. B., Levine, S., & Goldin-Meadow, S. (2006). The importance of gesture in children's spatial reasoning. *Developmental Psychology, 42*, 1259-1268.
- Ekstrom, R. B., French, J. W., Harman, H., & Derman, D. (1976). *Kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service.
- Eliot, J., & Smith, I. M. (1983). *An international directory of spatial tests*. Windsor, England: NFER-Nelson.
- Garber, P., & Goldin-Meadow, S. (2002). Gesture offers insight into problem-solving in adults and children. *Cognitive Science, 26*, 817-831.
- Gibbs, R. W., Jr. (2006). *Embodiment and cognitive science*. Cambridge: Cambridge University Press.
- Goldin-Meadow, S. (2003). *Hearing gesture: How our hands help us think*. Cambridge, MA: Harvard University Press.
- Goldin-Meadow, S., Cook, S. W., & Mitchell, Z. A. (2009) Gesturing gives children new ideas about math. *Psychological Science, 20*, 267-272.
- Goldin-Meadow, S., Nusbaum, H., Kelly, S., & Wagner, S. (2001). Explaining math: Gesturing lightens the load. *Psychological Science, 12*, 516-522.
- Hegarty, M., Mayer, S., Kriz, S., & Keehner, M. (2005). The role of gestures in mental animation. *Spatial Cognition and Computation, 5*, 333-356.
- Hegarty, M. & Sims, V. K. (1994). Individual differences in mental animation during mechanical reasoning. *Memory & Cognition, 22*, 411-430.
- Hegarty, M. & Waller, D. (2006). Individual differences in spatial abilities. In P. Shah & A. Miyake (Eds.). *Handbook of Visuospatial Thinking* (pp. 121-169). Cambridge University Press.
- Hostetter, A. B., & Alibali, M. W. (2008). Visible embodiment: Gestures as simulated action.

*Psychonomic Bulletin and Review*, 15, 495-514.

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.

Howell, D. C. (2007). *Statistical methods for psychology* (6th edition). Belmont CA: Duxbury.

Humphreys, L. G., Lubinski, D., & Yao, G. (1993). Utility of predicting group membership: Exemplified by the role of spatial visualization for becoming an engineer, physical scientist, or artist. *Journal of Applied Psychology*, 78, 250-261.

Johnson A. M. (1990). Speed of mental rotation as a function of problem solving strategies. *Perceptual and Motor Skills*, 71, 803-806.

Kendon, A. (1980). Gesticulation and speech: two aspects of the process of utterance. In M. R. Key (Eds.), *The Relationship of verbal and nonverbal communication* (pp. 207-227). The Hague: Mouton and Co.

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., & Davies, T. S. (2009). Competing conceptual representations trigger co-speech representational gestures. *Language and Cognitive Processes*, 24, 761 – 775.

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* (pp. 23-35). Berlin: Springer.

Krauss, R. M. (1998). Why do we gesture when we speak? *Current Directions in Psychological Science*, 7, 54-59.

Lohman, D. F. (1979). *Spatial ability: A review and reanalysis of the correlational literature*.

Technical Report No. 8, Aptitude Research Project, School of Education, Stanford University.

Lohman, D.F. (1988). Spatial abilities as traits, processes and knowledge. In R. J. Sternberg (Eds.), *Advances in the psychology of human intelligence* (Vol. 4), (pp. 81-248).

Hillsdale NJ: Erlbaum.

McGee, M.G. (1979). Human spatial abilities: Psychometric studies and environmental, genetic, hormonal, and neurological influences. *Psychological Bulletin*, 86, 889-918.

McNeill, D. (1992). *Hand and Mind*. Chicago: University of Chicago Press.

Melinger, A., & Kita, S. (2007). Conceptualisation load triggers gesture production.

*Language and Cognitive Processes*, 22, 473-500.

Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent variable analysis. *Journal of Experimental Psychology: General*, 130, 621-640.

Rauscher, F. B., Krauss, R. M., & Chen, Y. (1996). Gesture, speech and lexical access: The role of lexical movements in speech production. *Psychological Science*, 7, 226-231.

Salthouse, T.A., Babcock, R.L., Mitchell, D.R., Skovronek, E. and Palmon, R. (1990). Age and Experience Effects in Spatial Visualization, *Developmental Psychology*, 26, 128-136.

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.

- Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General*, *125*, 4–27.
- Shapiro, L. (2007). The embodied cognition research programme. *Philosophy Compass*, *2*, 1-9.
- Shepard, R., & Cooper, L. (1982). *Mental Images and Their Transformations*. Cambridge, MA: MIT Press.
- 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.
- Wagner, S. M., Nusbaum, H. C., & Goldin-Meadow, S. (2004). Probing the mental representation of gesture: Is handwaving spatial? *Journal of Memory and Language*, *50*, 395-407.
- Wechsler, D. (1991). *WISC-III Manual*. New York: The Psychological Corporation, Harcourt Brace Jovanovich, 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.
- Wright, R., Thompson, W. L., Ganis, G., Newcombe, N. S. & Kosslyn, S. M. (2008). Training generalized spatial skills. *Psychonomic Bulletin and Review*, *15*, 763-771.

Footnote

1. Although the post-test was a paper and pencil test, children could still use the pencil to point or trace the numbers in the mathematic equations.
2. In the present study, “spatial transformations” refers to object-based spatial transformations, which are defined as “to imagine the results of changing the positions of objects in the environment, while maintaining one’s current orientation in the environment” (Hegarty & Waller, 2004, p. 188). Another type of spatial transformations is egocentric spatial transformations, which are defined as “to imagine the results of changing one’s egocentric frame of reference with respect to the environment” (Hegarty & Waller, 2004, p. 188).)
3. We believe pointing gestures can also facilitate the internal computation of spatial transformation by helping participants create a more accurate internal representation of the object before and after rotation. However, even when we excluded the deictic gestures from the representational gesture category, the results essentially remained the same.
4. It is possible that the gesture-encouraged group used imagined gestures to offload the intermediate representations of the stimulus object in the second non-gesturing mental rotation block, and therefore performed better than the gesture-allowed group. However, this is unlikely because participants had to offload a representation in working memory to an imagined hand that also required working memory.



Author Note

We would like to thank Zeshu Shao for her help in the reliability check of gesture coding and Beatrice Hannah, Katerina Kantartzis, and Katherine Mumford for their proof reading of our manuscript. We would also like to thank Prof. Karen Pine and Prof. Alan Wing for their insightful comments on earlier versions of the manuscript. We also greatly benefited from the comments by Susan Goldin-Meadow, Greg Trafton and an anonymous reviewer.

Correspondence concerning this paper should be addressed to Mingyuan Chu, School of Psychology, University of Birmingham, Birmingham, B15 2TT, United Kingdom. [m.chu.1@bham.ac.uk](mailto:m.chu.1@bham.ac.uk).

Figure Captions

*Figure 1.* An example stimulus of the mental rotation task in Experiment 1 (Lower object was rotated from the upper left object 240 degrees about the bisector of horizontal and vertical axes).

*Figure 2.* Mean reaction times in the four rotation angles in Experiment 1. The error bars represent standard errors.

*Figure 3.* Mean error rates (proportions of trials with an erroneous response) across four rotation angles in Experiment 1. The error bars represent standard errors.

*Figure 4.* Mean representational gesture rates across four rotation angles in Experiment 1. The error bars represent standard errors.

*Figure 5.* Mean error rates (proportions of trials with an erroneous response) of the gesture-encouraged group, the gesture-allowed group and the gesture-prohibited group in the first and second mental rotation tasks in Experiment 2. The error bars represent standard errors.

*Figure 6.* Mean error rates (proportions of trials with an erroneous response) across four rotation angles in the gesture-encouraged group in Experiment 2. The error bars represent standard errors.

*Figure 7.* Mean representational gesture rates across four rotation angles in the

gesture-encouraged group in Experiment 2. The error bars represent standard errors.

*Figure 8.* An example stimulus of the paper folding task in Experiment 3.

*Figure 9.* Mean error rates (proportions of trials with an erroneous response) of the gesture-encouraged group and the gesture-allowed group in the paper folding task in Experiment 3. The error bars represent standard errors.

*Figure 10.* Mean error rates (proportions of trials with an erroneous response) of gesturing and non-gesturing trials of the gesture-encouraged group and the corresponding non-gesturing trials of the gesture-allowed group from an analysis of pooled data from Experiments 2 and 3. The error bars represent standard errors.

*Figure 11.* Mean representational gesture rates across the three sections in the gesture-encouraged group from an analysis of pooled data from Experiments 2 and 3. The error bars represent standard errors.

Figure 1.

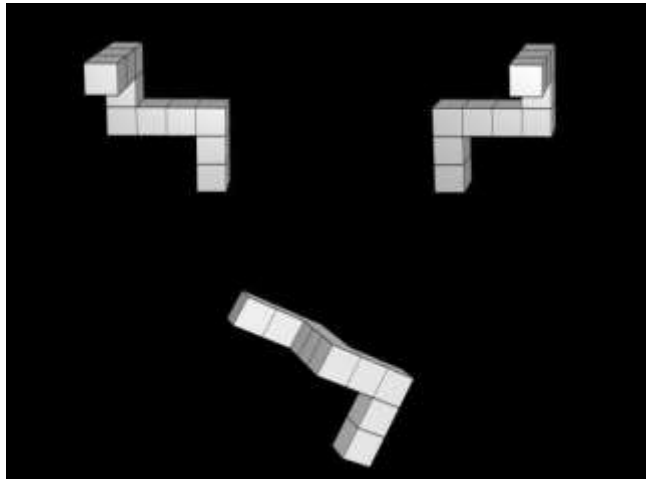


Figure 2.

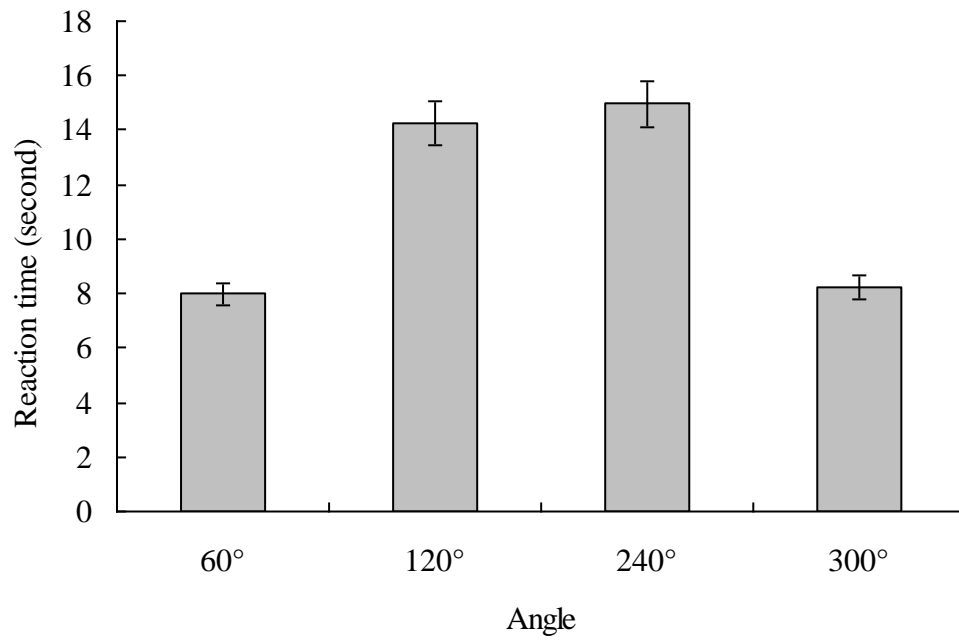


Figure 3.

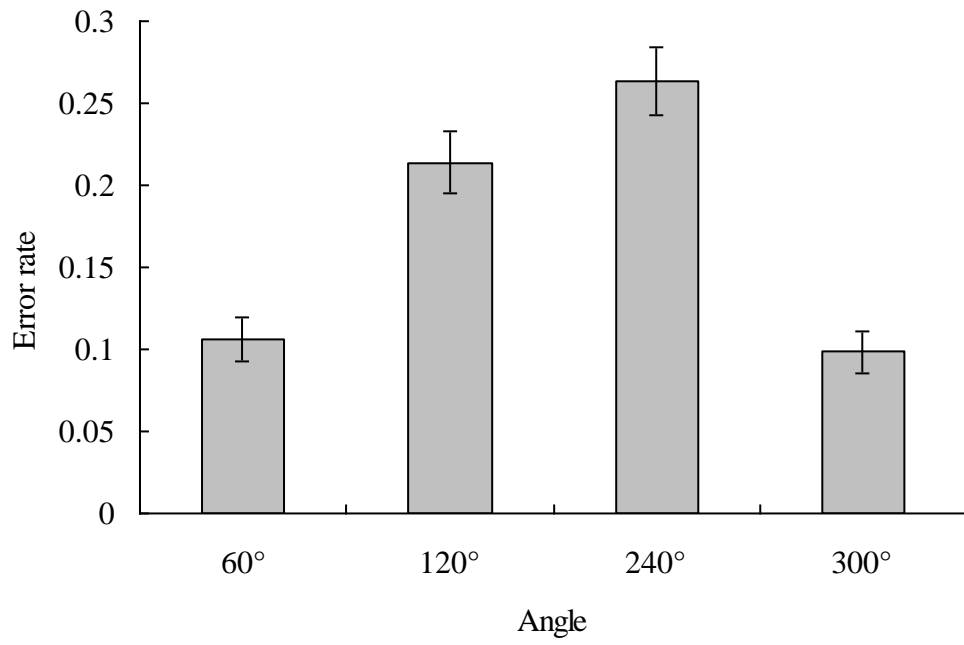


Figure 4.

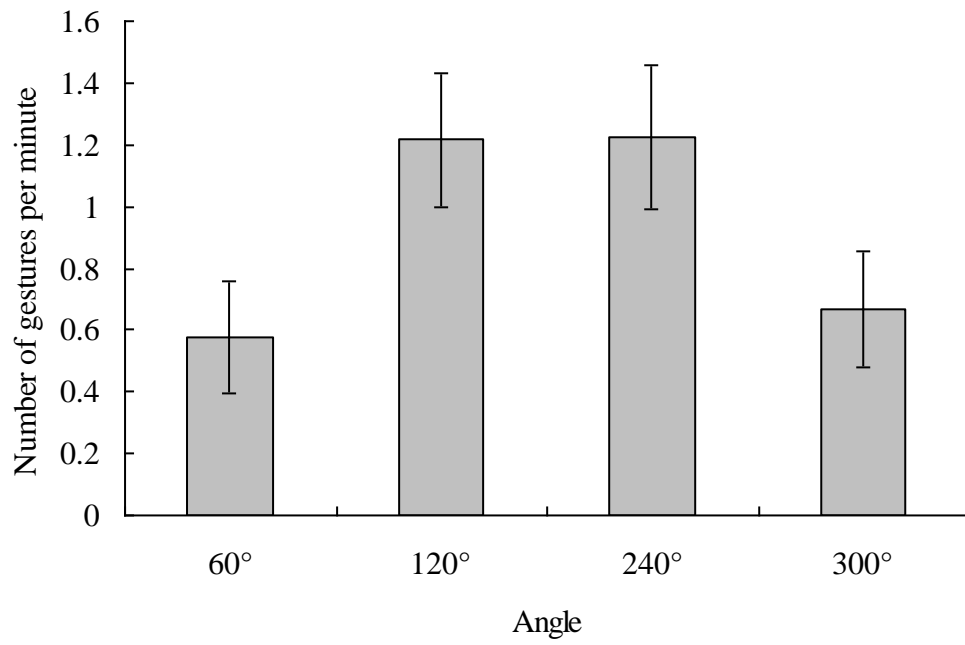


Figure 5.

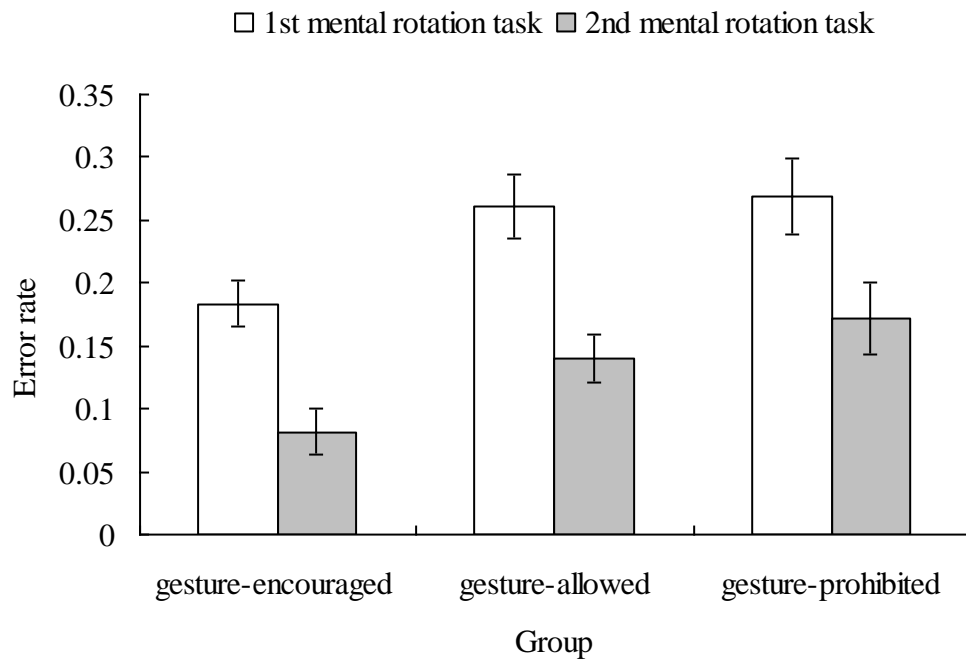




Figure 6.

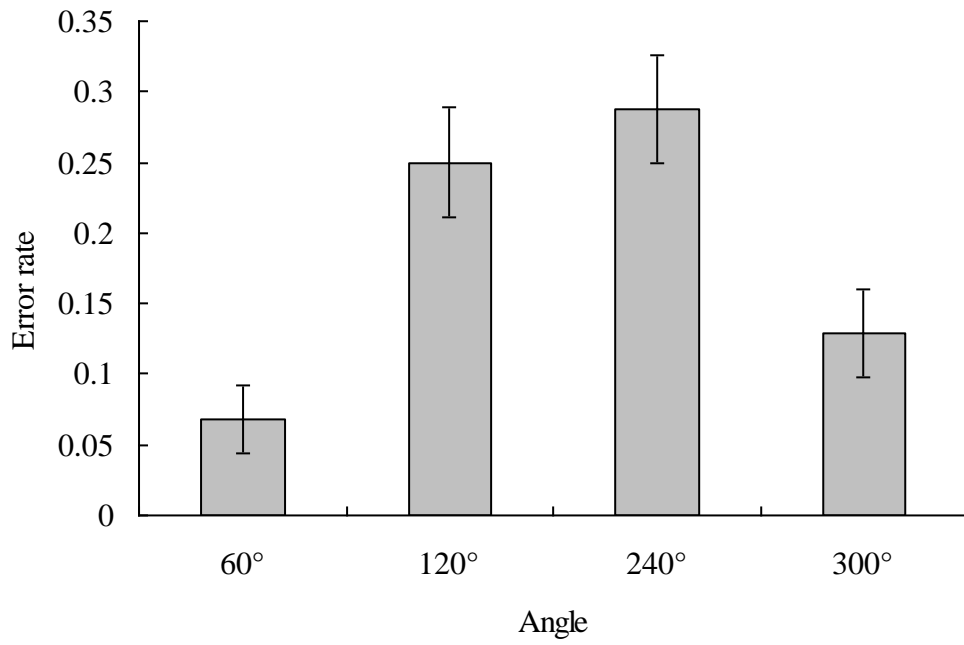


Figure 7.

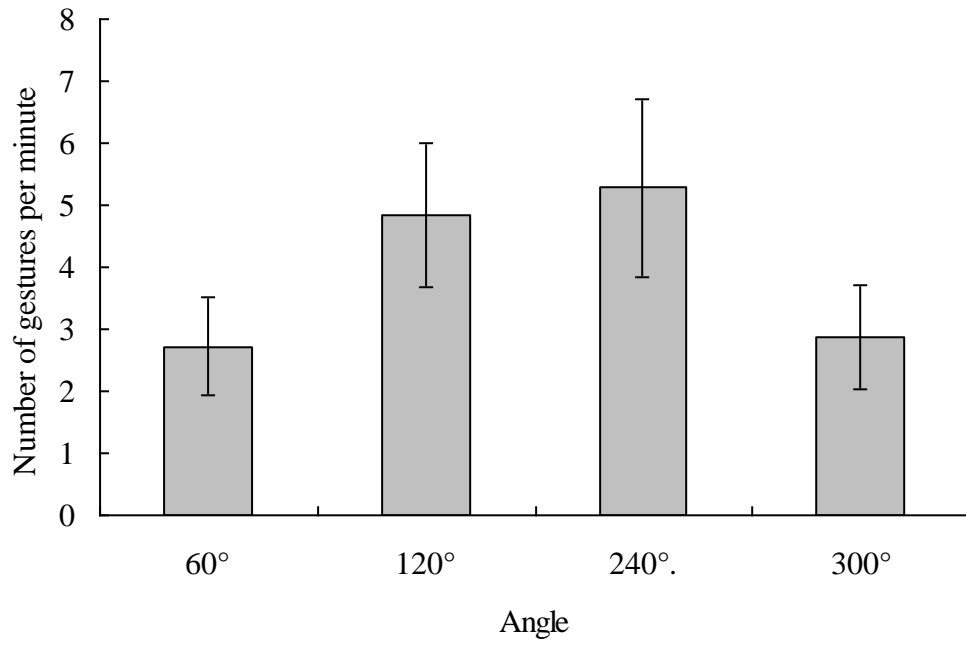


Figure 8.

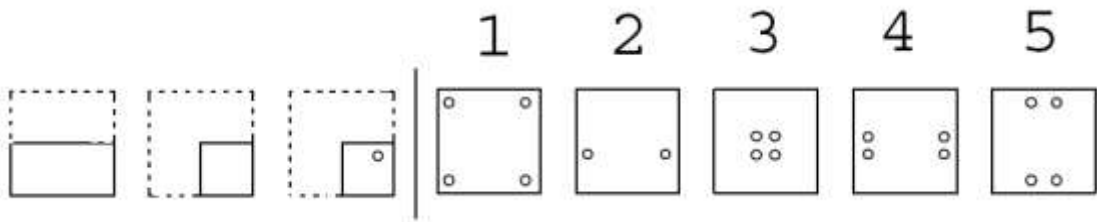


Figure 9.

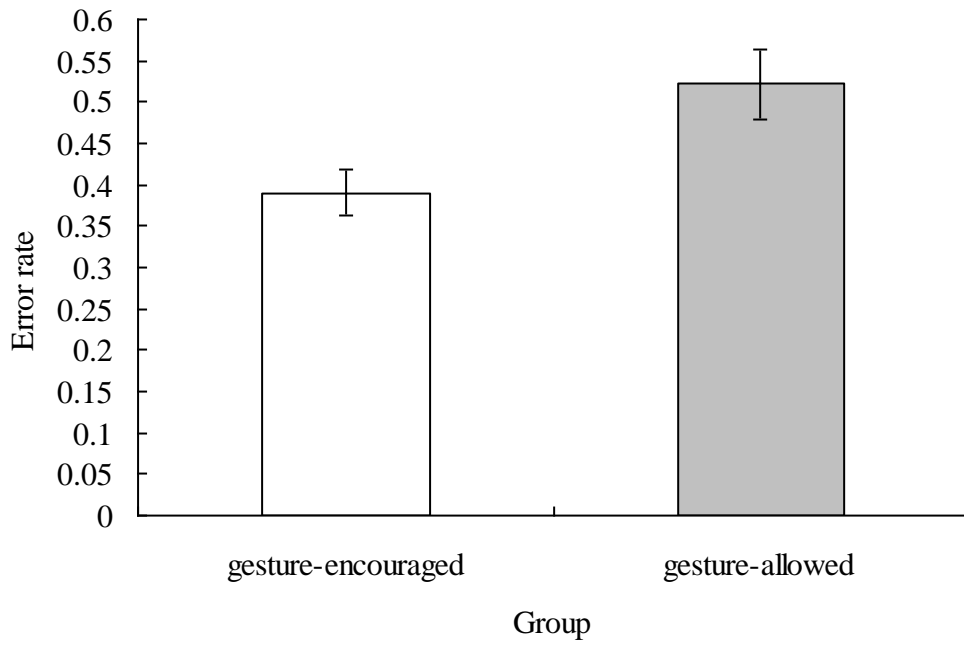


Figure 10.

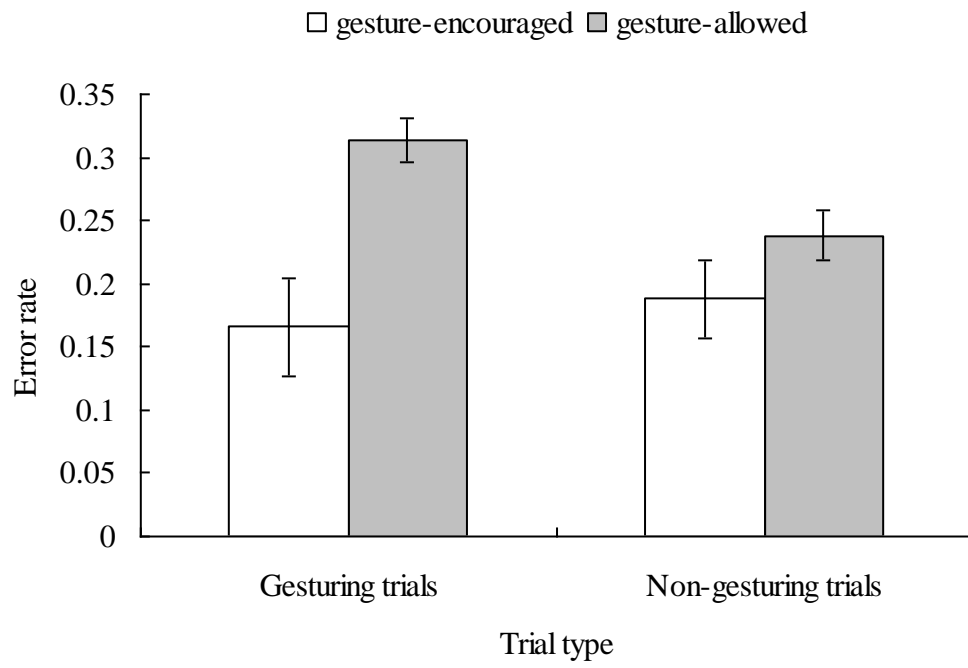


Figure 11.

