

Running head: GESTURES AND MENTAL ANIMATION

The Role of Gestures in Mental Animation

Mary Hegarty, Sarah Mayer, Sarah Kriz & Madeleine Keehner

Department of Psychology

University of California, Santa Barbara

Address for Correspondence: Mary Hegarty

Department of Psychology

University of California, Santa Barbara

Santa Barbara, CA 93106

hegarty@psych.ucsb.edu

Abstract

We examined the use of hand gestures while people solved spatial reasoning problems in which they had to infer how components of a mechanical device will move from a static diagram of the device (mental animation problems). In Experiment 1, participants were asked to think aloud while solving mental animation problems. They gestured on more than 90% of problems, and most gestures expressed information about the component motions that was not stated in words. Two further experiments examined whether the gestures functioned in the mechanical inference process, or whether they merely served functions of expressing or communicating the results of this process. In these experiments, we examined the effects of instructions to think aloud, restricting participants' hand motions, and secondary tasks on mental animation performance. Although participants who were instructed to think aloud gestured more than control groups, some gestures occurred even in control conditions. A concurrent spatial tapping task impaired performance on mechanical reasoning, whereas a simple tapping task and restricting hand motions did not. These results indicate that gestures are a natural way of expressing the results of mental animation processes and suggest that spatial working memory and premotor representations are involved in mental animation. They provide no direct evidence that gestures are functional in the thought process itself, but do not rule out a role for overt gestures in this type of spatial thinking.

Use of Gestures in Mental Animation

When people are asked to think aloud while solving problems, they often make hand movements, or gestures, and these gestures often communicate information about their thought processes that is not communicated in speech (e.g., Alibali, Bassok, Solomon, Syc, & Goldin-Meadow, 1999; Garber & Goldin-Meadow, 2002; Schwartz & Black, 1996a; Emmorey & Casey, 2001). Gestures are particularly evident in problem solving with spatial content, such as puzzles in which people have to place blocks such that they fill a puzzle grid (Emmorey & Casey, 2001), and mechanical reasoning problems in which people have to predict the motion of mechanical components (Schwartz & Black, 1996a). In these cases, people's gestures often simulate the object motions that they have to imagine in solving the problem.

What are the functions of gestures in problem solving situations? Much of the research on gestures and problem solving has focused on the ways in which gestures reflect people's mental representations and problem solving processes. This research suggests that gestures may be a natural way of expressing or communicating the results of internal thought processes and may reveal alternative solution processes that participants do not describe verbally. Research in semiotics has shown that gestures can also be used to overtly express or reinforce information presented in the speech stream. In this sense, gestures can also function to highlight or communicate information about problem solving process. An additional possibility is that gestures may not only reflect thought but be functional in the thought processes themselves. That is, gestures might aid mental computations, and not merely reflect the results of these computations.

In this article we examine the roles of gestures in a type of spatial inference problem in which people have to infer motion from static diagrams. We refer to these as mental animation

problems (Hegarty, 1992). Following Alibali (2005), we define a gesture as any movement of the hand or body that accompanies problem solving, regardless of whether the gesture is accompanied by speech. We examine how people naturally use gestures when solving spatial inference problems, and how preventing people from gesturing affects performance.

Gestures as reflections of thought. There is mounting evidence that gestures reflect people's internal representations while they reason and solve problems. One source of evidence comes from studies of gesture-speech mismatch (Goldin-Meadow, Alibali & Church, 1993; Garber & Goldin-Meadow; 2002; Alibali et al., 1999). In these studies, investigators examine mismatches between the representation of a problem or its solution that are expressed in speech, and the representations and solution processes revealed by gesture. For example in development of the concept of conservation, children sometimes describe one level of understanding in speech while communicating a more advanced level in gesture and those who show such gesture-speech mismatches are more responsive to instruction about conservation (Goldin-Meadow et al., 1993). Gesture-speech mismatches also predict adults' solution processes, for example in mathematics problem solving (Alibali et al., 1999) and in solving the Tower of Hanoi problem (Garber & Goldin-Meadow, 2002). One interpretation of these results is that gestures reveal secondary concepts or solution processes that are never expressed in the speech stream.

In other problem solving situations, gestures have been found to match the content found in co-occurring language, such that gestures can provide extra information not specified in the language alone. For example, when explaining to an addressee how to solve a puzzle in which blocks of different shapes had to be arranged so that they fit a puzzle grid, Emmorey & Casey (2001) found that people often produced deictic anaphoric constructions (e.g., "turn it this way") in which the manner of turning the block was communicated in gesture only, and the speech

referred to the gesture. This use of gesture appeared to be intentional in that it occurred only when the addressee was visible (in other conditions of the experiment, in which the addressee was not visible, people still gestured, but did not use as many deictic anaphoric constructions). Schwartz & Black (1996a) observed similar constructions in which speech referred to gesture (which they called exophoric references) when people “thought aloud” while solving mental animation problems. One interpretation of these results is that details of motion are often not easily communicated in language, and that gesture is a more natural means of communication in this case, because the hands can be used to mimic the details of the motion to be communicated (Emmorey & Casey, 2001).

Functions of gestures in thought. In the examples presented above, people were asked to think aloud or describe to an addressee how they solved problems, so they were required to express their thoughts verbally. In these situations, gestures can be interpreted as reflecting thought, in the sense that the thought or internal representation comes first, and the gestures express or communicate the thought process. In addition to this function of gesture, an additional possibility is that gestures are functional in the thought processes themselves. This view is supported by preliminary evidence that gestures often precede speech (Emmorey & Casey, 2001), that speakers use gestures in communicating the solution of a problem, even when the addressee is not visible (Emmorey & Casey, 2001) and even when they are not asked to think aloud (Schwartz & Black, 1996a). Note that this is not an alternative view to the idea that gestures communicate thought processes, but rather it suggests an additional role for gesture, reflecting McNeill’s (1992) view that gesture is both an act of communication and an act of thought.

There are at least two ways in which gestures might be functional in solving spatial problems such as puzzles and mental animation problems. First, one's hand can be used to represent an object that must be mentally transformed, such that moving the hand reveals something about the motion or its result. For example, in solving a puzzle that involved imagining moving an L-shaped block, participants made an L shape with their hands and moved their hands to consider possible positions of the block (Emmorey & Casey, 2001). Similarly in a problem that involved imagining how interlocking gears would move, Schwartz & Black (1996a) found that some participants used their hands to represent the gears, interlocked their fingers, and observed that when one hand moved clockwise the other had to move counterclockwise. In these cases, hands, which were free to move, represented objects that could not be physically moved while solving the problem.

Another way in which gestures might be functional in thought is in representing the results of intermediate computations. For example, in mentally animating mechanical systems, Hegarty (1992) found that participants infer the motion of components one by one, in order of the causal chain of events in the system's operation. This strategy involves maintaining information about the motions of earlier components in the causal chain, which is necessary to infer the motion of later components. However, internal maintenance of this information may compete with processing demands, as proposed by theories of working memory (Baddeley, 1986; Logie, 1995). In this situation, information maintenance might be offloaded on the motor system. For example, if a person has inferred that one gear in a gear system turns clockwise, she could make a clockwise motion with her hand to maintain this information, thus freeing up working memory resources for computing how the next component moves.

In somewhat related research, there is mounting evidence that imagining spatial transformations may be dependent on the motor system, in the sense that mental transformations involve high-level mental representations and brain mechanisms that are also used in action planning (Kosslyn, 1994; Logie 1995). Brain imaging studies have shown that areas of premotor cortex are active when people imagine spatial transformations of objects (Wraga, Shepherd, Church, Inati & Kosslyn, in press). However, evidence that spatial transformation processes involve premotor activation does not necessarily imply that they will be accompanied by overt gestures.

Other research has found that spatial transformations such as mental rotation are facilitated by hand motions that are congruent with the motions to be imagined, and impaired by incongruent hand motions (Wexler, Kosslyn & Berthoz, 1998; Wohlschlagel & Wolschlagel, 1998). This result generalizes to motions that would produce the imagined motion, when using a tool, even if the hand motion is not congruent with the object motion to be imagined (Schwartz & Holton, 2000). If congruent hand motions facilitate spatial thinking, it is plausible that people might spontaneously make congruent hand motions during spatial thinking. However, using hand motions strategically in this way may also depend on metacognitive knowledge of their effects on spatial thinking, so these results do not necessarily imply that people will spontaneously gesture while solving spatial problems.

The Present Study. In the present study, we examined hand motions that people made while solving mental animation problems about mechanical systems. In these problems, people are shown a static diagram of a mechanical system, given information about how one component of the system is moving, and must infer how another component moves (see the example in Figure 1). These problems are highly spatial in that they involve analyzing the spatial

arrangement of components of a mechanical system and inferring motion, which is a spatial property. Reaction time and eye-fixation data suggest that people solve mental animation problems piecemeal, inferring the motion of components one at a time in order of the causal chain of events (Hegarty, 1992). Several sources of evidence suggest that for novel mental animation problems, the inference process is often one of mentally simulating the motions of components (Hegarty, 2004). For example, mental animation performance is highly correlated with spatial ability (Hegarty & Sims, 1994) and mental animation interferes more with a spatial working memory load than a verbal working memory load (Sims & Hegarty, 1997). Furthermore when imagining the motion of two interlocking gears, time to infer whether a knob on one gear will fit into a groove on the other is related to the angle of rotation, suggesting an analog imagery process (Schwartz & Black, 1996b).

Schwartz & Black (1996a) first documented the use of hand gestures in solving mental animation problems. They asked people to imagine gear chains of different lengths (e.g., 5, 6 or 7 gears) on the basis of a verbal description. Participants, who worked either alone or in dyads told which direction the gear on one end was moving, and were asked to infer how the gear on the other end of the chain would move. They solved several problems of this type. When they began solving these problems, participants frequently gestured, using their hands to simulate the gears. As they solved more problems, their gestures became less frequent and they induced a rule that interlocking gears move in opposite directions so that every other gear in the chain will turn in the same direction. Schwartz & Black, interpreted their gestures as indicating a mental simulation process, which was replaced by a rule-based reasoning process. While early, simulation-based reasoning was accompanied by gestures, they acknowledged that they could

not offer any proof on the basis of their data that hand gestures were functional in the inference process.

In the present study, the participants solved several different types of mental animation problems, including gear, pulley, lever problems and problems showing mechanical systems that were made up of several different types of components. We used problems about a variety of different mechanical systems, to minimize the development of rules of mechanical reasoning, as documented by Schwartz & Black (1996a), which occurs if people solve the same type of problem (e.g. gear problems) over and over. However, we cannot be sure that some participants did not either possess rules of mechanical reasoning or develop them in the course of the experiment.

In a pilot study, 12 participants were asked to think aloud and were videotaped while they solved 22 mechanical reasoning problems of the type shown in Figure 1. Participants were shown the problems, which were printed on paper, one at a time. In previous studies with similar problems, participants had difficulty finding names to distinguish the different mechanical components (the upper pulley on the right, etc.) and this naming process clearly introduced additional cognitive processes that were not part of the mental animation process itself. Therefore, participants were given a pencil and were instructed that they could use this to point to the mechanical component in the diagram that they were referring to as they described their thinking. That is, they were instructed to use the pointer for deictic purposes. We discovered that rather than just pointing, participants frequently moved the pencil while reporting their thought processes, and we coded these movements as gestures. On average, participants gestured during their explanations of 20.1 (91%) of the problems. The most common type of gesture was to trace the motion of a component on the diagram with the pencil, e.g., if the participant inferred that a

gear was turning clockwise, he or she would trace the outside of the gear in a clockwise motion. Thus, participants produced what can be called “imitative” gestures (i.e., gestures that imitate a path or direction of motion), and not just deictic gestures.

We examined the expressive and communicative functions of gestures more closely in Experiment 1. In Experiments 2 and 3, we examined how both gesturing and mental animation performance are influenced by several manipulations, including instructions to think aloud, incongruent hand motions, and preventing gestures. We considered three possibilities regarding the function of gestures in mental animation. The first possibility is that gestures are functional in verbally expressing or communicating the results of mental animation processes only, and motor processes have no role in the mental animation process itself. If this is true, people should only gesture when asked to think aloud and incongruent hand motions or preventing gestures should not affect performance. The second possibility is that gesture is functional in both verbal expression and inference. If this is true, gestures should be observed in both think-aloud and control situations and both incongruent hand motions and preventing gestures should interfere with mental animation. The third possibility is that mental animation depends on premotor processes but not on overt gestures, and that overt gestures are used only in verbal expression of the results of mental animation processes. If this is true, we should observe gestures when people are asked to think aloud, but not in situations in which people do not have to report their thought processes, and incongruent motions should interfere with mental animation, but preventing gestures should not interfere.

Experiment 1

The goal of Experiment 1 was to examine the role of gestures in expressing the results of mental animation processes. Participants were asked to think aloud while solving mental

animation problems. As in standard protocol instructions, they were asked to say whatever came to their mind as they were solving the problems. Although verbal protocol instructions do not include instructions to communicate to the experimenter, participants knew they were being videotaped and audiotaped, and the experimenter was present, so that they so that they were in a communicative situation, at least implicitly. They were shown static diagrams of mechanical systems and asked to infer how parts of the systems moved. We examined how they described the component motions, both in words and in gestures.

If the internal representation of how a mechanical component moves is spatial in nature, it should be more readily expressed in gestures than in words, whereas if it is verbal in nature, it should be more readily expressed in words. There are also ways in which gestures might be used in conjunction with speech to express how components move. One possibility is that gestures help package thoughts into units that can then be formulated into speech. In this case, gestures are used only in the service of speech and should always be accompanied by verbal descriptions of the motion of components. Another possibility is that gestures co-occur with speech such that they provide precise spatial information not expressed in the language alone, in which case the gestures should not be redundant with speech.

Method

Participants. Ten undergraduate students from the University of California Santa Barbara took part in the experiment as part of a course requirement.

Materials. The materials consisted of 20 mechanical reasoning problems adapted from tests of mechanical reasoning such as the Bennett Mechanical Reasoning test (Bennett, 1969) and the mechanical reasoning test of the Differential Aptitudes test (Bennett, et al., 1981). Eight of these problems were classified as *mental animation* problems in that they presented a diagram

of a mechanical system with several interacting components, indicated how one component in the system was moving and asked participants to infer the motion of another component (see the examples in Figure 1). The other 12 problems examined different aspects of mechanical reasoning (such as comparisons of forces) that are part of a broader research program and are not of interest in the present research. Therefore only the eight mental animation problems will be analyzed in the present paper.

Procedure. Participants were tested individually. First, they were given instructions for the mechanical reasoning task, including instructions to think aloud as they solved the problems. They were handed a pointer and asked to use this pointer to indicate which component in the diagram they were referring to at a given time. The problems were presented one at a time on a 17-inch computer monitor. Participants were allowed as long as they needed to solve each problem. While they solved the problems, a video camera was aimed on the computer screen to record any hand motions (gestures) they made, and their verbalizations were recorded by means of a microphone.

Coding. We transcribed participants' verbalizations while solving the problems and three raters together identified all verb phrases that described the motion of a mechanical component. We classified the description of motion in each verb phrase as *lexical* if the motion was described completely in words (e.g. "it turns clockwise), *exophoric* (cf. Schwartz & Black, 1996a) if the verbal description referred to a gesture (e.g., "it turns this way"), *anaphoric* if it referred to a previously described motion (e.g., "it turns in the opposite direction") or as not specifying the direction of motion (e.g. "it turns"). Three raters independently coded 25% of the trials, All 3 agreed on 94% of the codes and resolved the discrepancies by consensus. The remainder of the trials were coded by one of these raters.

We then coded each gesture accompanying a verb phrase as *tracing* if the participant moved the pointer or their finger on a component of the diagram to indicate how it would move, *deictic* if the participant pointed to a component, or *other* if the participant made some hand movement that could not be classified as one of these types. Finally we noted when a verb phrase was not accompanied by a gesture. Three raters independently coded the gestures on 25% of the trials and agreed on 85% of cases, and again discrepancies were resolved by consensus. The remainder of the trials were coded by one of these raters.

Results and Discussion

Participants solved on average 5.9 of the eight problems correctly ($SD = 1.6$). Across the eight problems, participants described the motion of a mechanical component an average of 28.3 times ($SD = 17.8$). Table 1 presents data on how the motion was described in each case and the types of gestures that accompanied the verbal description. The majority (59%) of verb phrases describing motion ($M = 16.7$, $SD = 11.14$) were classified as exophoric, such that the verbal description (e.g. “this way”) depended on a simultaneous gesture whereas an additional 13% ($M = 3.7$, $SD = 4.7$) of verb phrases did not include an adverbial phrase specifying motion direction of motion. Almost all of these cases were accompanied by a tracing gesture that demonstrated the direction of motion. In the remainder of cases, the direction of motion was described either lexically or using an anaphor. These verb phrases were also typically accompanied by tracing gestures. In fact, 98.6% of verb phrases were accompanied by a gesture, so that gestures accompanied almost all verb phrases regardless of whether the words expressed how the component moved. Finally, 92.5% of all gestures were tracing gestures in which the participant traced the direction of motion on a component of the static diagram¹.

To illustrate these points, Table 2 presents the transcripts of two participants solving a gear problem that showed 4 interlocking gears (the problem was similar to the one shown in Figure 1 but cannot be reproduced here for copyright reasons). It is clear that gesture was central to Participant A's description of how the mechanical components moved, as it is not possible to tell what he was thinking from the verbal information alone. In contrast, Participant B described the motion of components lexically. Note that although this participant's speech fully expressed the direction of motion of the components, it was also accompanied by tracing gestures.

It is clear from this and the pilot experiment that gesture is important in expressing how components of mechanical systems move, that is, in externalizing the results of internal mental animation processes. In both of these experiments, participants were required to think aloud, and tracing gestures were pervasive in their "think aloud" protocols. In a minority of cases gestures were accompanied by speech expressing the same information as the gesture, so it is possible that these gestures were produced in the service of speech production. However, in most cases, the information given in the gesture was not expressed verbally. One possibility is that these gestures expressed precise spatial information that is difficult to translate to a verbal form, i.e. that their function is to communicate to the observer. But another possibility is that these gestures are functional in thinking itself, that is, they facilitate the inference process. We investigated this possibility in two further experiments.

Experiment 2

In Experiment 2 we examined the effects of think aloud instructions and a spatial secondary task on mental animation performance. There were three groups in this experiment, a control group, a verbal protocol group who were instructed to think aloud, and a dual task group, who were required to tap a spatial pattern while they solved the problems.

We predicted that the think aloud group would gesture while solving most problems, as in Experiment 1. If gestures are used for verbal expression purposes alone, then we should observe gesturing only in the verbal protocol group. On the other hand if gestures are functional in both inference and verbal expression, we should observe gesturing in both the verbal protocol and control groups.

We predicted that spatial tapping would impair performance on the mental animation problems. There are several levels at which this task might interfere with performance. First, spatial tapping prevents participants from gesturing, so that if gestures are functional in the inference process, interference might occur at this level. Second spatial tapping requires the production of hand motions that are incongruent with the motions to be inferred, so that the interference might occur at the level of premotor representations, if these are involved in inferring motion. Third, in studies of working memory, spatial tapping is often used as a secondary task that depends on spatial working memory resources (Baddeley, 1986) and previous research has suggested that mental animation is also dependent on spatial working memory (Sims & Hegarty, 1997), so that spatial tapping should impair mental animation performance.

Method

Participants and design. Forty-five undergraduate students from the University of California, Santa Barbara took part in the experiment in return for course credit. Fifteen participants served in the verbal protocol group, fifteen in the spatial tapping group, and fifteen in the control group.

Materials and Apparatus. The materials consisted of a Mental Animation scale, consisting of 20 mechanical reasoning problems and a spatial tapping sheet. The mechanical

problems included the eight problems analyzed in Experiment 1 along with an additional 12 mental animation problems. They were presented one at a time on a 17 inch computer monitor using a Gateway computer running Superlab software. The internal consistency of the mental animation scale (consisting of the 20 problems), as measured by Chronbach's Alpha, was .71.

For the spatial tapping group, we used a piece of paper, divided into 4 squares, with each square labeled with the numbers 1, 2, 3 or 4 (see Figure 2). Participants were asked to tap these squares throughout each problem, according to the numerical sequence. A metronome was used to give them a time signal for their tapping responses.

Procedure. Participants were randomly assigned to an experimental condition and were tested individually. First, they were then given instructions for the mechanical reasoning task, and were given three sample problems to view and answer. If they answered any of the sample problems incorrectly, they were corrected and the experimenter explained how to solve the problem. A final set of instructions was then read to them, asking them to verbally state the answer (either A, B or C) aloud for each problem. At this point, participants in the control group began the 20 mechanical reasoning trials. Each participant viewed the problems in the same order.

Participants in the verbal protocol group were given the same initial instructions as the control group. Then they were given additional instructions to think aloud as they completed the task. They were handed a pencil and told that they could use it to point to the part of the diagram that they were thinking about, but were not instructed to gesture. After these additional instructions, they began the 20 mechanical reasoning problems.

Participants in the spatial tapping group were given the same initial instructions as the control group. Then they were instructed to tap the squares on the piece of paper in the order "1,

2, 3, 4” to the tick of a metronome. The metronome ticked at a rate of 108 ticks per minute. They practiced spatial tapping alone for 1 minute. Upon conclusion of the practice session, they were given final instructions for the mechanical reasoning task, which involved solving the mechanical problems and spatial tapping at the same time. Participants were instructed to commence spatial tapping immediately prior to the presentation of each problem, and to continue tapping throughout, until they had announced the solution. They were allowed to rest between problems.

In all conditions of the experiment, the experimenter pressed a response key when the participant announced his or her response to the problem, and this was used as an approximate measure of solution time. Then a blank screen appeared and the participant informed the experimenter when to press a key to advance to the next problem. Participants were allowed up to 30 seconds to solve each problem. The experimenter recorded their answers (A, B or C) on a scoring sheet. She also observed the participants while they were solving the mechanical reasoning problems and noted whether they made one or more gestures on each problem.

Results and Discussion.

Use of Gestures. We predicted that participants in the think-aloud condition would gesture on most problems. If gestures function only in the communication process, we should observe gestures only in this condition, and not in the control condition. There was a large significant difference in gesturing between the groups, $F(2,42) = 55.07$, $MSE = 23.36$, $p < .001$, *partial* $\eta^2 = .72$. Post-hoc (Scheffe) comparisons ($\text{Alpha} = .05$) indicated that the verbal protocol group gestured significantly more than the control and spatial tapping groups, and the difference between the control and spatial tapping group was marginally significant ($p = .05$)².

As predicted, participants in the verbal protocol group gestured on nearly all of the 20 problems ($M = 18.53$, $SD = 1.51$). Participants in the control condition also gestured on some problems ($M = 5.20$, $SD = 8.06$). There was no strong evidence that particular problems were more or less likely to evoke gestures (the rate of gestures by control participants ranged from .13 to .40 for the different problems). However, there were large individual differences in gesturing among the control group, with 4 of the 15 participants in this group gesturing on 17 or more of the 20 problems, and the other 11 of the 15 participants gesturing on 3 or fewer problems. Rate of gesturing was not significantly correlated with performance on the mental animation problems for either the verbal protocol group ($r = .01$) or the control group ($r = -.20$). The spatial tapping condition was largely successful in preventing participants from gesturing, such that participants in this group gestured on less than one of the 20 problems on average (0.73 , $SD = 1.67$) – these participants occasionally attempted to gesture with their left hand while tapping with the right hand.

The result that participants in the verbal protocol group gestured more than control participants suggests that gestures function primarily in expressing the results of mental animation processes process. However, the result that some participants in the control group also gestured, suggests that verbal expression may not be the only function of gestures in solving these types of problems, i.e., gestures may be important in the inference process also, at least for some individuals.

Mechanical Reasoning Performance. If gestures function in only in expressing or communicating the results of mental animation processes, spatial tapping, which prevents gesturing, should not have an effect on mechanical reasoning. On the other hand, if they function in the inference process, spatial tapping should impair performance because the action of tapping

precludes simultaneous gesturing. As shown in Figure 3, the spatial tapping group had poorer performance on the mechanical reasoning problems than the other two groups. Problem accuracy was significantly affected by experimental condition, $F(2,42) = 5.21$, $MSE = 10.63$, $p = .01$, $partial \eta^2 = .20$. Post-hoc (Scheffe) comparisons with Alpha level set at .05 indicated that the control group was significantly more accurate than the spatial tapping group, but the difference between the control group and the verbal protocol group was not statistically significant, nor was the difference between the verbal protocol group and the spatial tapping group. The experimental groups also differed in response times, shown in Table 3, $F(2,42) = 55.07$, $MSE = 23.36$, $p < .001$, $partial \eta^2 = .22$. Post hoc (Scheffe) comparisons with an alpha level of .05 indicated that the control group had a significantly lower response times than both the spatial tapping and verbal protocol groups, which did not differ significantly from each other.

The seemingly poorer performance of the verbal protocol group, relative to the control group, might have been an experimental artifact, due to the fact that participants were given a time limit for each problem, and it took additional time for participants in the verbal protocol group to explain their answers. Participants in the verbal protocol group were somewhat more likely to time out while solving a problem (on 3.7 problems on average) than the control group (who timed out on 2.2 problems on average). Both the accuracy and reaction time measures indicated that spatial tapping significantly impaired mental animation performance. One possible explanation of this result is that spatial tapping prevented participants from gesturing, which is functional in the inference process. However, the spatial tapping task did more than merely prevent gesturing. It also engaged premotor and spatial representations. If either or both of these functions are required by the mental animation task, it could account for the interference that we observed. We conducted a third experiment to disambiguate these possibilities.

Experiment 3

There are three possible explanations of interference of spatial tapping with mental animation in Experiment 3. First spatial tapping precludes overt gesturing, second, it engages premotor and motor processes and third, it depends on spatial working memory. Thus, the interference that we observed might have occurred at any, or all of these levels. To disambiguate these possibilities, in Experiment 3, we contrasted the effects of spatial tapping with two other conditions. First, we examined the effects of simple tapping, a simpler secondary task that involved tapping a single key to the beat of a metronome. This task engaged premotor and motor processes, and prevented gestures, but made minimal demand on spatial processing resources. Second, we examined the effects of restricting participants' hand movements while solving the mechanical reasoning problems. This task prevented gesturing, but did not engage either the motor system or spatial processing functions. If overt gestures are functional in the inference process, all three of these experimental conditions should impair mental animation performance. If premotor representations and processes are involved in mental animation, mental animation should be impaired by spatial tapping and simple tapping but not by hand restriction. If mental animation involves spatial working memory resources but does not depend on the motor system, spatial tapping should impair performance but simple tapping and hand restriction should not.

In this experiment we also corrected for some limitations of Experiment 2. First we measured not just mental animation performance, but also performance on the secondary tasks (spatial tapping and simple tapping) to ensure participants were actually carrying out the secondary task properly. Second, we measured reaction time more precisely, by requiring participants to press a key when announcing their answer to the problem

Method

Participants and design. Sixty right-handed undergraduate students from the University of California, Santa Barbara took part in the experiment, in return for course credit. Fifteen participants served in the control group, fifteen in the spatial tapping group, fifteen in the simple tapping group and fifteen in the hands-restricted group.

Materials and Apparatus. The materials included the same 20 mechanical reasoning problems as used in Experiment 2. The problems were presented on a 17 inch computer monitor. Responses were recorded by means of a Cedrus RB 834 button box, with 4 large buttons in the center and four smaller buttons on the side (see Figure 4). As in Experiment 2, a metronome was used to give participants a time signal for their tapping responses.

Procedure. Participants were randomly assigned to an experimental condition and tested individually. First, they were then given instructions for the mechanical reasoning task and solved three sample problems, printed on paper, as in Experiment 2. Then they were given additional instructions that during the experimental trials, they should press the “response” key on the button box with their left index finger as they verbally stated their answer (A, B or C) for each problem. They were given two more sample problems to view and answer on the computer, in order to practice pressing the “response” key while answering verbally.

Then participants in the control group and hands restricted group completed the 20 experimental trials. While completing the trials, participants in the hands-restricted condition were instructed to hold the button box with their two hands, so that their left index finger was positioned above the response box. Their hands were held in position by elastic bands, thus restricting any movements.

Participants in the spatial tapping group were given the same instructions as in Experiment 1 and practiced spatial tapping for 30 seconds to the tick of a metronome, before

commencing the twenty experimental trials. They tapped the spatial pattern with their right hand and tapped the response key with their left hand to indicate when they made their response.

Participants in the simple tapping group were instructed to tap a single button (the top left large button) with their right index finger to the beat of a metronome during the experimental trials.

They practiced this tapping response for 30 seconds alone before commencing the experimental trials. For both tapping tasks, the metronome ticked at a rate of 108 ticks per minute.

Results and Discussion.

Gestures. Three of the 15 participants in the simple tapping condition gestured during one or two of the 20 problems and no hand gestures were observed in either the hands restricted or spatial tapping conditions, indicating that the experimental manipulations were successful in preventing hand gestures. On average, participants in the control group gestured on 4.33 of the 20 problems (SD = 6.38). Although the amount of gesturing was somewhat lower in this experiment, there were again large individual differences, such that four of the 15 participants in this condition gestured on 10 or more of the problems, whereas ten of the participants gestured on fewer than 3 problems. As in Experiment 2, this difference in rate of gesturing was not significantly related to performance on the mental animation problems ($r = .20$, $p = .47$).

Mechanical Reasoning Performance. If gestures are functional in inferring motion from static diagrams, then performance of the spatial tapping, simple tapping and hands restricted conditions should all be impaired relative to the control group. If the effects of spatial tapping observed in Experiment 1 were due to interference at the level of premotor representations, spatial tapping and simple tapping, but not hand restriction, should impair performance. Finally, if the interference is at the level of spatial working memory resources, only spatial tapping should impair performance

As Figure 5 shows, there was a significant difference between the groups, $F(3, 56) = 5.152$, $MSe = .3728$, $p = .003$. Post-hoc (Sheffe) comparisons with an alpha level of .05 indicated that the hands restricted ($M = 13.87$, $SD = 3.09$) and control groups ($M = 13.67$, $SD = 2.5$), had superior performance to the spatial tapping group ($M = 10.4$, $SD = 2.38$). The performance of the simple tapping group did not differ significantly from any of the other groups. This experiment replicated the impairment effect of spatial tapping that was observed in Experiment 2. The lack of an effect of hand restriction and the non-significant effect of simple tapping are more consistent with the spatial working memory hypothesis than with the idea that gestures or motor representations are functional in the process of inferring motion from static diagrams. Note that these non effects cannot be interpreted as ceiling effects, as participants were far from ceiling (20 problems correct) on this task. Table 3 shows the means and standard deviations of response times for the four groups There were no significant effects of experimental condition on reaction times in this experiment, $F(3, 56) = .562$, $MSE = 12.28$, $p = .643$.

Secondary Task Performance. In dual-task experiments, it is important to examine both primary and secondary task performance, to ensure that non-effects of the secondary task on primary task performance were not due to the fact that participants were simply not carrying out the secondary task. To examine the effects of the mental animation task on the secondary tasks of spatial and simple tapping, we compared tapping performance during the initial practice period (single task condition), in which participants performed the tapping task alone for 30 seconds, to their performance when tapping was paired with mental animation (dual task condition). Participants in the spatial tapping single task condition never tapped a key out of order, whereas 3.4% of spatial taps ($SD = 3.0$) in the dual task condition were out of order, and this rate was significantly different from zero, $t(14) = 4.48$, $p = .001$. Participants in the simple

tapping condition never tapped the wrong key in either the single task or dual task conditions.

Thus spatial tapping performance was impaired by mental animation, suggesting that these tasks shared resources. In contrast, simple tapping was not impaired, so that the non-significant effect of simple tapping on the primary task cannot be attributed to the idea that participants were not carrying out the simple tapping task correctly.

Since participants were required to tap to the tick of a metronome, variability in rate of tapping is also an indication of impaired performance on the secondary tasks. We also measured participants' rate of tapping (i.e., mean time between taps) and how regularly they tapped, (assessed by computing the standard deviation of times between taps). The metronome sounded at a rate of 108 ticks per second, so perfect performance on the tapping measures would be a tapping rate of 555 milliseconds and a standard deviation of 0. Table 4 presents data on rate of tapping and regularity of tapping in the single and dual task conditions. Although there were trends for participants to tap less frequently and less regularly in the dual-task conditions, these differences were not significant.

General Discussion

In this series of experiments we have examined gesture production during mental animation of static mechanical diagrams in order to better understand how gestures co-occur with linguistic descriptions of mechanical movement as well as how gestures may be produced to help in facilitating the inference of motion from static diagrams. It is clear from Experiments 1 and 2 that gestures serve an important function in expressing how machine components move. In both experiments, people who were asked to think aloud, gestured on almost all of the problems. In Experiment 1, most gestures expressed information that was not stated in words. In Experiment 2, people who were asked to think aloud gestured much more than those in a control

group. It appears therefore that when people have to communicate how parts of a machine move, gestures are naturally evoked. A likely explanation of these results is that the internal representations in mental animation tasks are spatial in nature, and that these representations can be expressed more easily and more completely in gesture, because the hands can be used to mimic the manner of motion (Emmorey & Casey, 2001). Another possibility is that the gestures are used in the service of verbal communication, such that they help package the information in units for speaking (cf. Kita, 2000). However, in the think aloud conditions, most gestures did not lead to a verbal description of motion (i.e., motion was expressed in gestures alone) and some participants gestured in control conditions where they were not required to speak. Therefore, this cannot have been the only function of gestures in this study.

It is possible that aspects of our experimental procedures influenced the rate of gesturing in this study. First, in all of the “think aloud” conditions, participants were given a pencil and asked to point to components as they referred to them in their verbal protocol. This pointer was introduced so that participants would not have to find names for the different mechanical components, but it might have caused them to make more gestures than they would have made naturally. However it is important to note that participants were instructed to use this pointer for deictic purposes only, and the vast majority of their gestures (tracing gestures) were imitative and not purely deictic. Nevertheless, it is important to replicate these studies in a situation where people are not given a pointer or explicitly instructed to point to components. Another factor that influences gesture production is the presence of another individual who can see the gestures. In the present study, an experimenter was present in all conditions, so differential amounts of gesturing across the experimental conditions can not be attributed to presence of another. However, to fully investigate the functions of gestures in thinking as opposed to communication,

it is also important to repeat these experiments in a situation where an experimenter is not present. Finally, it will be important in future studies to directly compare the types of gestures that accompany speech to those used in control conditions where participants are not asked to think aloud.

With respect to the function of gestures in thinking, the results of these studies provide no direct evidence that overt gestures are functional in thinking. Although some people gestured while solving the problems, even in control conditions in which they were not required to communicate, preventing gestures by restricting people's hand movements or by simple tapping did not significantly impair performance. Also, there was no evidence that individual differences in gesturing were related to problem solving performance. On the other hand, in both Experiments 2 and 3, concurrent spatial tapping did impair performance. In addition to preventing gesture and engaging the motor system, spatial tapping has been found to rely on spatial working memory (Logie, 1995; Smyth, Pearson & Pendleton, 1988; Smyth & Pendleton, 1989). The selective effects of spatial tapping on performance therefore suggest that mental animation relies strongly on spatial working memory resources, a result that is consistent with previous research (Sims & Hegarty, 1987).

Our results suggest that mental animation is a process that both depends on spatial working memory resources and is naturally expressed in gestures. This prompts us to ask if there is a specific relation between spatial working memory representations and gesture. In support of this view, gestures are associated not just with the solution of spatial problems, but there is considerable evidence in the language production literature that gestures are particularly associated with the communication of spatial information more generally (Alibali, Heath & Myers, 2001; Kita & Oezyurek, 2003; Rauscher, Krauss & Chen, 1996). One reason why spatial

content may be naturally expressed in gesture, is that the spatial working memory system may in fact be highly related to the premotor and motor systems. Logie (1995) argues that similar neuroanatomical structures and pathways underlie spatial and motor representations, concluding that “the retention of spatial information, the retention of movements, and the production of movements all rel[y] on overlapping cognitive resources” (p. 116). This conclusion, which is related to theories of embodied cognition (Wilson, 2002), is supported by evidence that concurrent spatial movements impair performance on a range of spatial and imagery tasks (Farmer, Berman & Fletcher, 1986; Quinn & Ralston, 1985, Smyth, Pearson & Pendleton, 1988). It is also consistent with the results of neuroimaging studies, showing premotor activation in spatial imagery tasks (Wraga et al., in press).

Although we found no evidence in this study that gestures are functional in thinking, it would be premature to conclude this on the basis of this study alone. We found striking individual differences in the spontaneous use of gestures among participants in control conditions, with a minority of students gesturing on most of the problems, even when not asked to think aloud. Furthermore, a small number of participants attempted to gesture, even in the spatial tapping and simple tapping conditions, and we also observed that in the hands-restricted condition, some participants gestured with their heads, behavior that was also reported by St. John (1992). Thus it is possible that gestures are more functional in thought for some individuals than others.

It is important to investigate individual differences in gesture in future studies. These studies could assess the degree of impairment by secondary tasks on an individual level, using a within-subjects design in which participants perform mental animation in both dual-task and control conditions. For example, in the present study, there was a non-significant trend for

simple tapping to impair performance, raising the question of whether simple tapping impaired performance for some participants but not others, and a significant effect would be observed with a more powerful experimental design. It would also be illuminating to investigate whether the use of gestures is correlated with other individual differences measures, such as expertise, spatial ability and cognitive style.

In summary, we considered three possible accounts of the role of gestures and motor processes in mental animation performance. The first account is that gestures are functional only in expressing the results of mental animation and there is no involvement of the motor system in mental animation. The second is that overt gestures are functional in both the mental animation process and expression. The third is that overt gestures are functional only in expression, but that mental animation depends on premotor representations and spatial working memory resources that are involved in planning movements. The results presented in this paper are most consistent with the third account. However they have revealed important individual differences in the use of gesture in both communication and inference, suggesting that one possible account of the function of gestures in mental animation and other problem solving situations may not fit all individuals. As we continue to investigate the role of gestures in spatial cognition, it is important that we take account of individual differences and not assume that one model will necessarily account for the performance of all individuals.

References

- Alibali, M. W. (2005). Gesture in spatial cognition: Expressing, communicating and thinking about spatial information. *Spatial Cognition and Computation*.
- Alibali, M.W., Bassok, M, Solomon, K.O., Syc, S.E., & Goldin-Meadow, S. (1999). Illuminating mental representations through speech and gesture. *Psychological Science, 10*, 327-333.
- Alibali, M. W; Heath, D. C & Myers, H. J. (2001). Effects of visibility between speaker and listener on gesture production: Some gestures are meant to be seen. *Journal of Memory & Language, 44*, 169-188.
- Baddeley, A.D. (1986). *Working Memory*. Oxford: Oxford University Press.
- Bennett, C. K. (1969). *Bennett Mechanical Comprehension Test*. San Antonio: The Psychological Corporation.
- Bennett, G. K., Seashore, H. G., & Wesman, A. G. (1981). *Differential aptitude tests*. Sidcup, UK: The Psychological Corporation.
- Emmorey, K., & Casey, S. (2001). Gesture, thought and spatial language. *Gesture, 1*, 35-50.
- Farmer, E. W., Berman, J. V. F., & Fletcher, Y. L. (1986). Evidence for a visuo-spatial scratch-pad in working memory. *Quarterly Journal of Experimental Psychology, 38A*, 675-688.
- Garber P. & Goldin-Meadow, S. (2002). Gesture offers insight into problem solving in adults and children. *Cognitive Science, 26*, 817-831.
- Goldin-Meadow, S., Alibali, M., & Church, R. (1993). Transitions in concept acquisition: Using the hand to read the mind. *Psychological Review, 100*, 279-297.
- Hegarty, M. (1992). Mental animation: Inferring motion from static diagrams of mechanical systems. *Journal of Experimental Psychology: Learning, Memory and Cognition, 18(5)* 1084-1102.

- Hegarty, M. (2004). Mechanical reasoning as mental simulation. *TRENDS in Cognitive Sciences*, 8, 280-285.
- Hegarty, M. & Sims, V. K. (1994). Individual differences in mental animation during mechanical reasoning. *Memory & Cognition*, 22, 411-430.
- Kita, S. (2000). How representational gestures help speaking. In D. McNeill (Ed.) *Language and Gesture* (pp. 162-165). Cambridge, UK: Cambridge University Press.
- Kita, S. & Oezyurek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal? Evidence for an interface representation of spatial thinking and speaking *Journal of Memory & Language*. 48, 16-32.
- Kosslyn, S. M. (1994). *Image and Brain: The Resolution of the Imagery Debate*. Cambridge, MA: MIT Press.
- Logie, R. H. (1995). *Visuo-spatial working memory*. Hove: Lawrence Erlbaum Associates.
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago, IL: University of Chicago Press.
- Quinn, J. G., & Ralston, G. E. (1986). Movement and attention in visual working memory. *Quarterly Journal of Experimental Psychology*, 38A, 689-703.
- Rauscher, F. H., Krauss, R. M., & Chen, Y. (1996). Gesture, speech, and lexical access: The role of lexical movements in speech production. *Psychological Science*, 7, 226-231.
- Schwartz, D. L. & Black, J. B. (1996a). Shuttling between depictive models and abstract rules: Induction and fall-back. *Cognitive Science* 20, 457-497.
- Schwartz, D. L. & Black, J. B. (1996b). Analog imagery in mental model reasoning: Depictive models. *Cognitive Psychology* 30, 154-219.

Schwartz D. L. & Holton, D. L. (2000). Tool use and the effect of action on the imagination.

Journal of Experimental Psychology: Learning Memory and Cognition, 26, 1655-1665.

Sims, V.K. & Hegarty, M. (1997). Mental Animation in the visuospatial sketchpad: Evidence

from dual-task studies. *Memory and Cognition*, 25, 321-332.

Smyth, M. M., Pearson, N. A., & Pendleton, L. R. (1988). Movement and working memory:

Patterns and positions in space. *Quarterly Journal of Experimental Psychology*, 40A,

497-514.

Smyth, M. M., & Pendleton, L. R. (1989). Working memory for movements. *The Quarterly*

Journal of Experimental Psychology, 41A, 235-250.

St. John, M. (1992). Learning language in the service of a task. *Proceedings of the 14th annual*

conference of the Cognitive Science Society. Hillsdale, NJ: Lawrence Erlbaum Associates.

Wexler, M., Kosslyn, S.M., & Berthoz, A. (1998). Motor processes in mental rotation.

Cognition, 68, 77-94.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin and Review*, 9,

625-636.

Wohlschlager, A. & Wohlschlager, A. (1998). Mental and manual rotation. *Journal of*

Experimental Psychology, 24, 397-412.

Wraga, M.J., Shepherd, J. M., Church, J. A. Inati, S. & Kosslyn, S. M. (in press). Imagined

rotations of self versus objects: An fMRI study. *Neuropsychologia*.

Author Notes

This research was supported in part by grant N00014-03-1-0119 from the Office of Naval Research. We are grateful to Autumn Elling and Pamela Bernaba for help with coding of the data in Experiment 1.

Footnote

¹As indicated by the standard deviations in Table 1, there were large individual differences in frequency of gestures. Accuracy was correlated .53 ($p = .11$) with frequency of tracing gestures and .33 (= .36) with frequency of deictic gestures. It would be premature to interpret these trends given their lack of statistical significance and the small number of participants, but they might have merely reflected more verbalizations and gestures by participants who were more confident in their answers.

²We did not record the gestures in Experiments 2 and 3 so cannot formally compare the types of gestures used in the verbal protocol and control conditions. However, the experimenter could detect no differences between the types of gestures used in these two conditions, and reported that as in Experiment 1 they were primarily tracing gestures.

Table 1

Mean number of gestures of different types used by participants in conjunction with verb phrases describing the motion of mechanical components across the 8 experimental trials. The verb phrases are classified by how they expressed the manner of motion of a component. Standard deviations are in parentheses.

Type of Verbal Description	Type of Gesture			
	Tracing	Deictic	Other	None
Lexical (e.g. “clockwise”)	5.2 (6.5)	0.4 (0.7)	0.1 (0.3)	0.1 (0.3)
Exophoric (e.g., “this way”)	16.1 (10.4)	0.4 (0.7)	0.2 (0.4)	0.0 (0.0)
Anaphoric (e.g., “opposite”)	1.5 (2.0)	0.5 (1.0)	0.1 (0.3)	0.0 (0.0)
No Adverb (e.g., “it’s turning”)	3.0 (3.6)	0.3 (0.7)	0.1 (0.3)	0.3 (0.7)
Total	25.8 (15.7)	1.6 (2.0)	0.5 (1.7)	0.4 (0.7)

Table 2

Verbal protocols (with simultaneous gestures in parentheses) of two participants solving a problem that showed four interlocking gears. An arrow indicated that the leftmost gear was turning counterclockwise, and the participant had to choose whether the rightmost gear would turn counterclockwise (direction A) or clockwise (direction B). We refer to the gears as Gear 1, 2, 3 and 4 from left to right. Participant A uses primarily exophoric expressions whereas Participant B describes the motions lexically.

Participant A

S (Reading Question): When the left-hand gear moves in the direction shown, which way will the right one, right-hand gear turn? *(no gestures)*

(Solving): So it's going this way, *(tracing pointer counterclockwise on Gear 1)*

which means this one moves this way, *(tracing pointer clockwise on Gear 2)*

this one moves this way *(tracing pointer counterclockwise on Gear 3)*

this one moves this way *(tracing pointer clockwise on Gear 4)*

Participant B

S (Reading Question): When the left, O.K. when the left-hand gear moves in the direction shown, which way will the right one, right-hand gear turn? *(no gestures)*

(Solving): Um, O.K.

so this is turning counterclockwise *(tracing pencil counterclockwise on Gear 1)*

so this would be turning clockwise *(tracing pencil clockwise on Gear 2)*

so this be counterclockwise *(tracing pencil clockwise on Gear 3)*

so this would be clockwise *(tracing pencil clockwise on Gear 4)*

so it's B

Table 3

Mean Response Time and Standard Deviation per problem by Condition in Experiments 2 and 3.

		Response Time (in sec)	
	Group	Mean	Standard Deviation
Experiment 2	Control	17.94	4.15
	Verbal Protocol	21.42	3.09
	Spatial Tapping	21.58	2.48
Experiment 3	Control	17.32	3.71
	Spatial Tapping	18.69	3.42
	Simple Tapping	18.25	4.02
	Hands Restricted	17.35	2.73

Table 4

Mean rate of tapping and regularity of tapping (measured in seconds) in single task and dual task conditions of Experiment 3 for participants in the spatial tapping and simple tapping conditions (standard deviation in parentheses).

	Tapping Rate (Mean time between taps)	Tapping Regularity (Standard deviation of time between taps)
Spatial tapping alone	0.587 (.044)	.098 (.082)
Spatial tapping (dual task)	0.618 (.077)	.142 (.132)
Simple tapping alone	0.597 (.039)	.064 (.037)
Simple tapping (dual task)	0.594 (.036)	.079 (.037)

Figure Captions

Figure 1. Sample mental animation problem

Figure 2. Diagrams printed on paper that participants tapped in the spatial tapping condition of Experiment 2. Participants tapped the numbers in order.

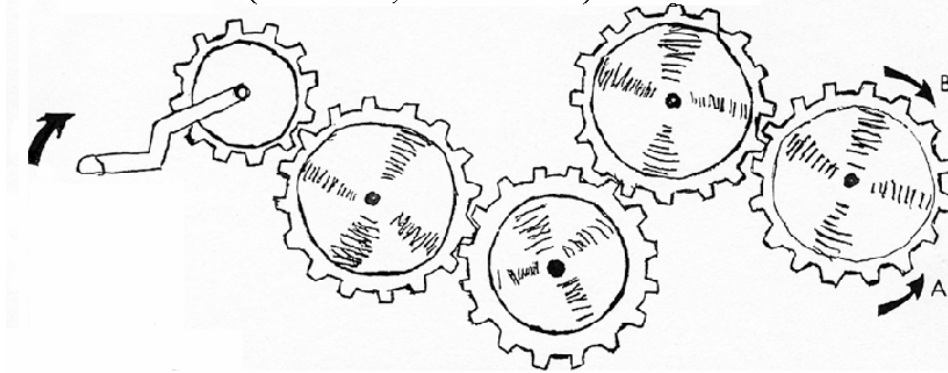
Figure 3. Mean accuracy in the different experimental conditions of Experiment 2. Error bars show standard error of the mean.

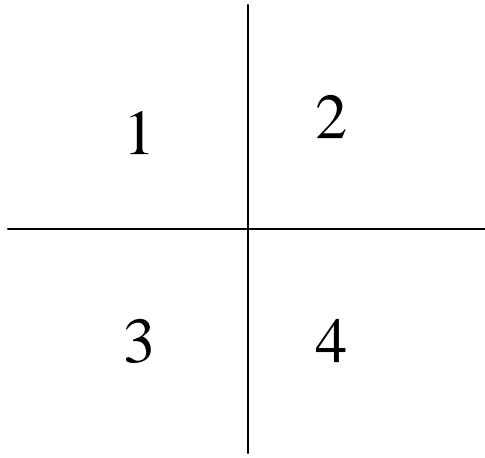
Figure 4. Diagram of the button box used in Experiment 3. Participants in the spatial tapping condition tapped the number keys in order. Participants in the simple tapping condition tapped the key marked “1” repeatedly.

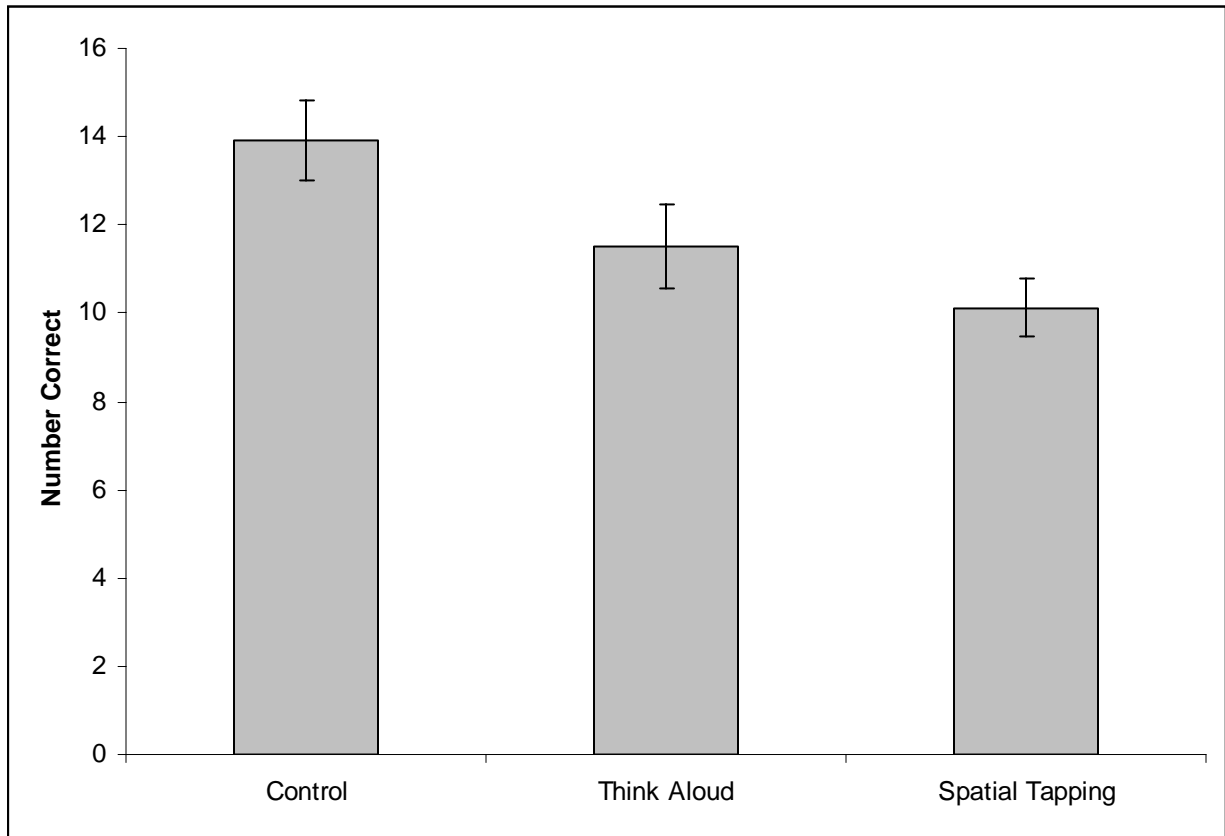
Figure 5. Mean accuracy in the different experimental conditions of Experiment 3. Error bars show standard error of the mean.

When the handle is turned in the direction shown,
which direction will the final gear turn?

(If either, answer C.)







Response Key

