COGNITIVE PSYCHOLOGY **30,** 154–219 (1996)
ARTICLE NO. 0006

Analog Imagery in Mental Model Reasoning: Depictive Models

DANIEL L. SCHWARTZ

Peabody College, Vanderbilt University

AND

JOHN B. BLACK

Teachers College, Columbia University

We investigated whether people can use analog imagery to model the behavior of a simple mechanical interaction. Subjects saw a static computer display of two touching gears that had different diameters. Their task was to determine whether marks on each gear would meet if the gears rotated inward. This task added a problem of coordination to the typical analog rotation task in that the gears had a physical interdependency; the angular velocity of one gear depended on the angular velocity of the other gear. In the first experiment, we found the linear relationship between response time and angular disparity that indicates analog imagery. In the second experiment, we found that people can also solve the problem through a non-analog, visual comparison. We also found that people of varying spatial ability could switch between analog and non-analog solutions if instructed to do so. In the third experiment, we examined whether the elicitation of physical knowledge would influence solution strategies. To do so, we manipulated the visual realism of the gear display. Subjects who saw the most realistic gears coordinated their transformations by using the surfaces of the gears, as though they were relying on the friction connecting the surfaces. Subjects who saw more schematic displays relied on analytic strategies, such as comparing the ratios made by the angles and/or diameters of the two gears. To explain the relationship between spatial and physical knowledge found in the experiments, we constructed a computer simulation of what we call *depictive modeling.* In a depictive model, general spatial knowledge and context-sensitive physical knowledge have the same ontology. This is different from prior simulations in which a non-analog representation would be needed to coordinate the analog behaviors of physical objects. In our simulation, the inference that coordinates the gear motions emerges from the analog rotations themselves. We suggest that mental depictions create a bridge between imagery and mental model research by positing the referent as the primary conceptual entity.

This research was supported in part by a Ben D. Wood Fellowship to the first author from Teachers College, Columbia University. Acknowledgments to David Krantz and Kathryn Spoehr who helped in many kind ways. Correspondence may be addressed to Daniel L. Schwartz, Department of Psychology and Human Development, Box 512 GPC, Vanderbilt University, Nashville, TN 37203. E-mail: SchwarDL@ctrvax.vanderbilt.edu.

0010-0285/96 $18.00

In everyday activity, people can infer simple physical behaviors. With varying degrees of accuracy, people can infer behavior for situations as various as colliding jellyfish, a water balloon hitting a bat, and the flexing of a candy bar. In the current paper, we develop evidence that for a simple mechanical interaction—the meshing of two gears—these inferences can take place through spatial imagery. At the same time, we gather evidence indicating the role of physical knowledge in determining the specific spatial solutions that subjects deploy. This raises the question of how people integrate general spatial knowledge and processes with more domain-specific, or context-sensitive, knowledge. To describe our theory of how this integration occurs, we present the construct of a *depictive model* in the form of an object-oriented computer model.

Although there is disagreement about the role of spatial imagery in simple mechanical inferences, there is some agreement within the literature about their non-spatial aspects. One point of agreement is that these inferences are local in that they are used to reason about simple physical interactions, perhaps embedded within a larger system. A second point of agreement is that they attempt to model the temporal structure of events. For example, Hegarty (1992, 1993) examined the sub-goals people constructed to understand pulley systems. In addition to the sub-goal of mapping the larger structure of a system, she found that people frequently focused on and inferred the local behavior of a specific component embedded within the larger system. For example, people might infer the behavior of a pulley as a rope pulls through it. Implying that these inferences have a temporal structure that maps on to the physical world, Hegarty found that inferences about the forward flow of events were drawn more quickly than inferences about antecedent events. Hegarty labeled these local inferences with the term ‘‘mental animation,’’ suggesting a spatial imagery that mimics the flow of visual events.

Many cognitive scientists working in the field of qualitative physics have also characterized physical inferences as dependent on local, simulative reasoning (for a review see Forbus, 1988). However, they have not found it necessary to invoke an analog of perceptual space to infer local behaviors. For example, de Kleer and Brown (1984) developed a model of qualitative inference in which a computer program first maps the connections within a complex device (e.g., a hose has a valve, the valve is above a bucket, there is a hole in the bucket). The program then determines the function of the device by propagating the results of local inferences throughout the device topology. For example, when the valve opens water falls out, when water falls from the valve it goes into the bucket, and so forth. Unlike Hegarty’s (1993) suggestion that these local inferences involve animation, de Kleer and Brown employed a qualitative calculus to derive local behavior. A qualitative calculus captures directional changes in relationships, as in the case when a bucket switches from gaining water to losing water. de Kleer and Brown did not find it necessary to include representations that had the precision of a metric space, but instead could used qualitative, or partially ordered, spatial and temporal relations (e.g., above, below, before, after). Thus, although we find that researchers agree in principle on the local and temporally ordered nature of mechanical inference, we find that there is a lack of concordance about the representations behind these inferences.

The disagreement as to whether everyday, physical inferences are based in spatial imagery or a qualitative calculus may be, to a large extent, the result of different research traditions. The majority of research into vision and imagery has looked towards invariant sources of spatial knowledge. Because spatial properties, like foreshortening, hold universally across multiple contexts, the knowledge and processes that work over space can be characterized in a generally applicable fashion. For example, some researchers have posited a special working memory structure, like a two-dimensional array, that supports all spatial reasoning (e.g., Forbus, Nielsen, & Faltings, 1990; Kosslyn, 1980). In contrast, researchers who work under the general umbrella of mental models have emphasized the variability of human reasoning in different domains (e.g., Bassock, 1990; Greeno, 1983; Johnson-Laird, Legrenzi, & Legrenzi, 1972; Norman, 1983). In the context of qualitative physics these domains range from fluid dynamics to electronics to classical mechanics (e.g., Hayes, 1979; Gentner & Gentner, 1983; McCloskey, Caramazza, & Green, 1980). For the task of characterizing learned knowledge and explaining why it is accessed in some contexts but not others, a more flexible representation of world knowledge has been necessary (e.g., a production network; Anderson, 1983).

There are tasks, however, that require both context-specific information about physical interactions and detailed spatial information. For example, consider the task in Fig. 1. The gear on the left is called the *power gear*, and the gear on the right is called the *driven gear*. The question is whether the knob and groove will mesh when the power gear rotates clockwise. Like most physical problems, there are several solution strategies. For example, one might use a *measurement* strategy by comparing the lengths of the subtended arcs. Or, one might employ an *analytic* strategy by comparing the inverse ratios of the angles and diameters. For our immediate interests, the most relevant inferences are the ones that model the temporal behavior of the gears. Because the task requires spatial precision, it seems reasonable that if people model the problem, they would require a metric representation of space. A qualitative calculus, designed to capture changes to the sign of a derivative, cannot model the smooth metric transformation of the gears. At the same time, because the gears are physically connected, it seems reasonable that people would need qualitative, non-geometric knowledge about the behaviors of connected rigid bodies. For example, the forces that join the gears into a coordinated relationship are invisible and could not be modeled with purely visual or spatial representations.

There is good evidence that people can transform mental images so they

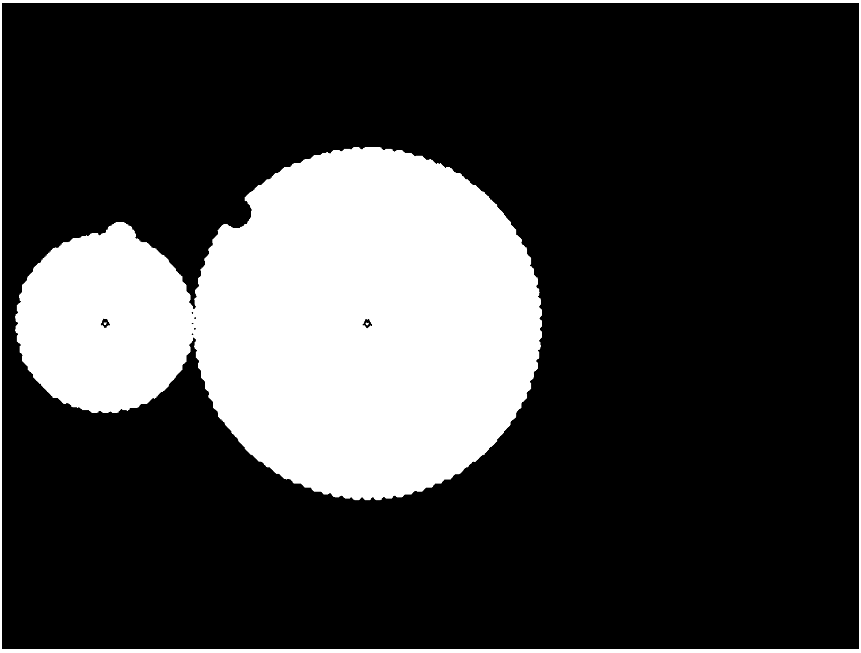

FIG. 1. A task with a problem of coordination—Will the knob and groove mesh if rotated inward? If one mentally animates the gears, it is necessary to coordinate their relative velocities.

mimic perceived spatial transformations (e.g., Kosslyn, 1980; Pinker & Finke, 1980; Shepard & Cooper, 1986a). For example, in the analog imagery paradigm (Shepard, 1968/1986), subjects judge whether two figures of different orientation are identical or mirrored versions of one another. Judgment latencies are generally in a positive linear relationship with the angular disparities between the two figures, just as they would be if one actually witnessed the two figures turning into alignment prior to making a judgment. Analog imagery research has used a variety of stimuli including alphanumeric characters (Hochberg & Gellman, 1977), alphabet blocks (Just & Carpenter, 1985), Cheng blocks (Shepard & Metzler, 1971), Atteneave figures (Cooper & Podgorny, 1976), ice cream cones (Marmor & Zaback, 1976), and miscellaneous utensils hung from strings within a box (Pinker & Finke, 1980). A characteristic of all these tasks, however, is that they do not require knowledge about physical interactions between objects or forces. In contrast, the gear task adds a *problem of coordination* to the traditional analog imagery task. In the gear task, the transformation rate of one gear must be coordinated with the transformation rate of the second gear; as long as the gears are of different sizes, they should have different angular velocities.

This small methodological addition to the traditional analog imagery paradigm may provide some new leverage for understanding how people control their imagery transformations. In many prior tasks, spatial information and spatial processes can sufficiently prescribe and predict the trajectory of a mental transformation. For example, imagine a letter rotated from the upright position. How do people decide which way to rotate the letter to make it upright—clockwise, counter-clockwise, or through the visual plane? When detectable, people choose the shorter direction of rotation (Shepard & Metzler, 1971). Thus, the explanation is that people choose a specific image transformation according to spatial properties. This type of geometric explanation, sometimes coupled with capacity considerations, has also been extended to transformations that appear to map onto physically constrained transformations (e.g., Leyton, 1989; Proffitt, Kaiser, & Whelan, 1990). For example, using an implied motion paradigm, Foster (1975) alternately flashed two lines that were at different coordinates and orientations. The task was to position a third line between the two lines to create the smoothest animation. The results showed that subjects positioned their lines such that the animation looked as though the tail end of a longer line were pivoting around a remote axis point. Shepard and Cooper (1986b) argued that people applied the pivoting transformation because it was spatially less complex than separate rotation and translation transformations (the alternative animation possibility). Thus, there was no need to ascribe a role for knowledge about a physical domain in determining the image transformation.

For the gear task, however, there are two possible analog solutions which, if both evinced, mean that spatial properties alone cannot predict the transformation subjects will choose. These strategies correspond to Johnson-Laird's (1983) distinction between kinematic and dynamic mental models. In the kinematic solution, the image mimics the spatial behavior of the gears. First, one determines the relative rotation rates of the two gears by comparing size ratios. For example, if the driven gear is twice as large as the power gear, it should turn half as fast. Subsequently, one transforms the images of the gears ensuring the correct relative angular velocities. We will call this a *global analog strategy,* because the control of the relative rotation rates is external to the images themselves. In the dynamic solution, the image mimics the spatial *and* physical behavior of the gears. In this scenario, the turning surface of the power gear ''forces'' the driven gear to rotate. Thus, rather than globally computing and controlling relative angular velocities, the angular velocities emerge from the local action of ''forcing'' the gear surfaces through the point of contact at the same rate. We will call this a *local analog strategy,* because the coordination of the rotations results from the transformation process itself. If the following experiments show that people can employ both analog strategies, then it suggests that people are not relying solely on general spatial principles to determine the behavior of the gears.

A growing body of research suggests that purely spatial explanations cannot sufficiently account for people's spatial reasoning about a physical world.

For example, the original research using a change of focus paradigm found that the Euclidean distance between memorized objects was an excellent predictor of retrieval times (Kosslyn, 1976; Thorndyke, 1981). If a subject thinks of one object, then the amount of time to recall another object is a function of the imagined distance between the two objects, as though the subject were scanning from one object to another across a spatial representation. However, researchers have also found that physical and social boundaries can influence retrieval times over and above spatial distances (Garnham, 1981; Glenberg, Meyer, & Lindem, 1987; McNamara, 1986). Even though two objects separated by a wall may be quite close spatially, subjects exhibit retrieval latencies that imply that their imaginations traveled around a boundary wall (Morrow, Greenspan, & Bower, 1987).

Research on imagined displacements also suggests a role for physical knowledge in people's spatial transformations. For example, Freyd, Pantzer, and Cheng (1988) found that some subjects spontaneously transformed their memories of a diagrammed scene when an implied physical constraint, such as a hook holding a pot, was removed from the diagram. Parsons (1987, 1994) conducted research in which subjects determined whether a left or right hand was displayed on a computer screen. Parsons found that people imagined moving their hands from the computer keyboard into the orientation of the hand on the screen. Interestingly, he could best predict response latencies by hypothesizing a trajectory of image transformations that avoided awkward limb movements. Shiffrar and Freyd (1990) also showed that people imagined body transformations following physical constraints. In an implied motion paradigm, subjects saw alternating pictures of a woman with her arm in front of and behind her body. If the inter-stimulus interval was sufficient, subjects reported seeing the arm animate around the woman's torso, rather than passing through her body. These experiments show that non-spatial constraints such as pivot joints and object solidity can regulate the transformation of a spatial image. However, in each case, it is not clear whether these constraints informed inferences about objects. For example, subjects in Parson's experiments may have been replaying pre-formulated motor plans for moving their hands into various orientations. Consequently, we cannot determine whether people can make an inference to solve the problem of how to coordinate different motions, especially those of distal, non-biological objects (cf. Shiffrar, 1994).

The purpose of the following experiments is to demonstrate that people can combine physical knowledge and analog imagery in the course of drawing an inference about a simple physical mechanism. We call this form of reasoning depictive modeling. To demonstrate the psychological plausibility of depictive models, we must show that people can use imagery to complete a task involving a problem of coordination. But, we must also show that neither spatial processing factors nor direct memories of spatial changes provide sufficient explanations for how people resolve the problem of coordination inherent to the gear task. Alternatively, we must demonstrate the central role of contextual, physical knowledge in how people accomplish their spatial transformations. To accomplish these goals we use the analog imagery paradigm. The primary evidence implicating an analog rotation is a positive linear relationship between angular disparity and response latency. The logic is that if people rotate an image through a string of intermediate states, then more intermediate states should produce longer response times.

EXPERIMENT 1

The purpose of the first experiment was to investigate whether people can conduct a mental rotation to solve the gear problems, and whether all subjects do so. If subjects mentally rotate the gears, latencies should increase as the marks on the gears are placed further from the meshing point. If there is a linear relationship between angular disparity and latency, then we can examine whether the disparity of the power or driven gear best predicts the response times.

In this experiment, the power gear had a constant radius of 20 mm. The *knob* on the power gear was placed at 10° intervals from 30° to 110° counterclockwise from the point of contact between the gears. The driven gear had a radius of either 40 or 60 mm. The *groove* in the driven gear was positioned as a function of the relative sizes of the gears and whether the problem was a positive or negative instance. For a positive, or *mesh,* instance the correct groove placement was determined with the equation: *Groove angle* = *Knob angle* × (*Power gear radius*/*Driven gear radius*), where the groove angle is clockwise from the point of contact between the gears. For example, with a knob disparity of 90°, the groove meshes when placed at 30° on the 60 mm driven gear, and 45° on the 40 mm gear. A negative, or *no-mesh,* instance was constructed by displacing the groove 20% of the mesh angle in either direction. For a *groove-early* problem, the groove was displaced towards the point of contact so it would arrive before the knob. For a *groove-late* problem, the groove was displaced away from the point of contact so it would arrive after the knob. With a knob disparity of 90°, the groove-early and groove-late positions were 24° and 36°, respectively, for the 60 mm driven gear. For the 40 mm driven gear, the corresponding groove positions were 36° and 54°. We used a proportional displacement of 20%, as opposed to a constant difference (e.g., 5°), to provide reasonably equal discriminability across all angles in accordance with Weber’s Law (Woodworth & Schlosberg, 1965).

Method

Subjects. Fifteen adult volunteers (9 females and 6 males) from the Columbia University community were recruited. Subjects were run separately.

Apparatus. Subjects sat at a comfortable distance from an IBM PS/2 with VGA graphics displayed on an IBM 8513 color monitor (640 × 480 mode). The ‘‘/’’ and ‘‘z’’ keys were labeled ‘‘yes’’ and ‘‘no,’’ respectively.

Stimuli. The stimuli consisted of two horizontally touching, solid white circles (henceforth, gears) against a black background (see Fig. 1). The gears had ragged perimeters that resembled sprockets but that were neither distinct nor reliable enough to count. Each gear had a small mark in the center representing its axis. The power (left-hand) gear had a constant radius of 20 mm. A 3.4 mm semicircular knob was positioned on its perimeter. The driven gear had either a 40 mm or 60 mm radius. On its perimeter was a semicircular groove into which the knob fit. The two gears touched at the same screen location in all the experiments.

Design. Two within-subject factors were crossed. One factor was the size of the driven gear using the levels of 40 mm and 60 mm radii. The second factor was the angular disparity of the knob from the point where the two gears touched. These disparities ranged from 30° to 110° in 10° intervals. We systematically varied the placement of the groove in the driven gear to construct negative problems. For each angle at each size of the driven gear, there were two mesh, one groove-early, and one groove-late problems. For a mesh problem, the angular disparity of the groove equaled the knob's angular disparity multiplied by the ratio of the power and driven gears' radii. For a groove-early problem, the angular disparity of the groove was reduced by 20% of the correct mesh angle. For a groove-late problem, the angular disparity was increased by 20% of the correct mesh angle. All told, there were a total of 72 problems (2 sizes $\times$ 9 angles $\times$ 4 mesh/no-mesh) of which half would mesh. Problem order was randomized for each subject in this and subsequent experiments.

Procedure. Subjects saw a computer demonstration of four different gear combinations rotating through two complete revolutions of the power gear. A practice problem then appeared on the screen. To remove potential effects of memorization, the demonstration and practice gears were 5 mm larger and smaller than the gears used in the recorded trials. The angular disparities were also 5° larger than those of the recorded trials. The experimenter explained that the ‘‘power gear’’ on the left turned the ‘‘driven gear’’ on the right. Subjects were directed to press the ‘‘yes’’ or ‘‘no’’ keys depending on whether they thought the knob would meet the groove. They were told that both speed and accuracy were important. Subjects subsequently received 10 practice problems with textually displayed accuracy feedback. For the no-mesh problems, feedback also included a green circle that marked the groove's true, mesh starting position. After the practice, subjects pressed any key to start the recorded trials. For the recorded trials there was no feedback, and the screen went blank after each response. Subjects pressed any key to initiate the next trial. Halfway through the trials an automated message directed the subjects to take a short break. Subjects were debriefed after the experiment.

Results

We focus primarily on the correctly answered mesh problems (hits), because they provide the most stable latency estimates owing to the larger

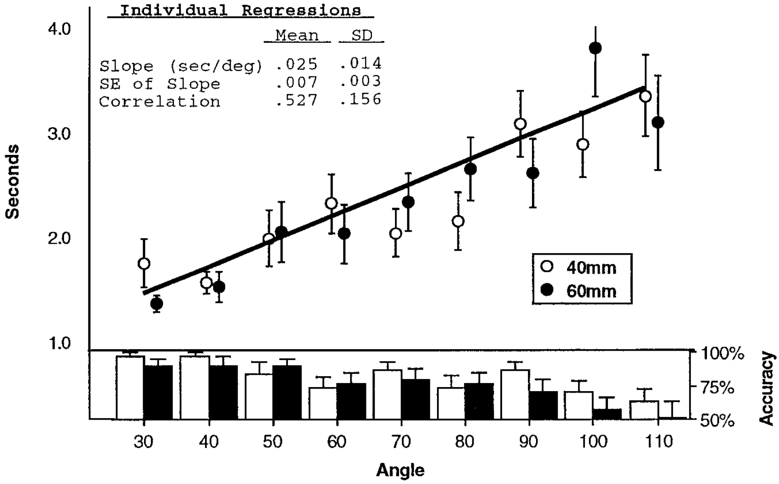

FIG. 2. Correct answer latencies and percentage correct for the mesh problems (Experiment 1). The data points represent group averages. The regression slope and tabled statistics represent the mean of the individuals' regression coefficients and the variability between subjects. The standard error bars are across trials for the latency data and across subjects for the accuracy data.

frequency of mesh problems. Figure 2 demonstrates the linear relationship between the angular disparity of the knob and the latencies for the correctly answered (79%) mesh problems.[1] Although accuracy rates are low for the larger angular disparities, the latencies provide first evidence that people can solve the gear problem using a mental rotation.

In Figure 2 and all following figures, each data point represents the average across correct trials. The slope, reflecting the statistical analyses, is the mean slope found by averaging the individuals' separate regression statistics. For example, in Fig. 2, the slope of .025 s/deg represents the average slope across individuals. The standard deviation of .014 represents the between-subjects variability in slopes. Throughout the paper, we conducted the statistical analyses using subjects' individual regression coefficients in multivariate analyses of variance. Given a normal distribution of coefficients, this style of analysis satisfies the requirements for generalizing to the members of a population (Lorch & Myers, 1990).

To show that the linear pattern reliably holds across subjects and to explore

[1] In this and the subsequent experiments, outliers greater than three standard deviations from the subject's mean were removed (approximately 1.6% of all responses). All within-subject analyses in the experiments were multivariate; however, we reported the simpler univariate results. All assumptions for using the univariate results were met or adjusted for in significance calculations.

the effect of the different sizes of the driven gear, we computed two regression slopes for each subject; one per driven gear. Each subject's resulting two slope coefficients can be thought of as representing the rotation rate of the power gear when it is coupled with either driven gear. The two coefficients were data points in a repeated measures analysis with driven gear size as a within-subject factor. The slopes were reliably greater than zero indicating that latencies increased with angular disparity across the subjects; $F(1,14) = 44.3$ $p < .01$. The effect of the driven gear's size on the rate of rotation was also significant; $F(1,14) = 5.45$, $p < .05$. The 60 mm driven gear led to somewhat slower rotations than the 40 mm driven gear; .03 s/deg, $SD = .018$ and .021 s/deg, $SD = .015$, respectively. However, as may be seen in Fig. 2, the effect of the driven gear's size was quite small compared to the overall effect of the power gear's angular disparity. For the mesh problems, the best determinant of latency was the angular disparity of the power gear's knob; mean $r = .53$, $SD = .16$. Further indicating that the disparity of the power gear's knob was the best predictor of correct response time, an exploratory stepwise regression on the hit trials (not nested within-subject) showed that the angular disparity of the power gear entered the equation first; $r = .41$, $B = .023$, $SEB = .002$, $t = 8.99$, and blocked the entry of other potentially relevant stimuli characteristics; all t's < 1. This included the angular disparity of the groove; partial $r = -.031$, the initial distance between the knob and groove; partial $r = .004$, the difference between the initial motion vectors of the marks; in X coordinate, partial $r = .018$, in Y coordinate, partial $r = .042$, the differences between the cord lengths connecting each gear's mark with the point of contact; partial $r = .040$, as well as interactions between these and other stimuli characteristics.

Figure 2 indicates a positive correlation between error rates and latencies across angular disparities; $r = .22$. This suggests that the longest latencies were due to problem difficulty and not a speed/accuracy tradeoff. In general, the gear problems yielded less accurate and more variable responses than other analog imagery tasks. Although some of this variability may be attributed to strategy switches and a lack of practice, a large part of the variability may be the result of the need to make a metric comparison rather than a configural comparison (e.g., are two shapes equivalent). This may yield less response stability for two reasons. First, metric comparisons probably require the maintenance of more precise information than configural comparisons. For example, when comparing two letter R's one may only need to represent the major axes with precision (Hochberg & Gellman, 1977); size information is less critical. Second, the mesh judgment is a matter of degree. Because people cannot make perfectly coordinated rotations, they must decide if slight misses should really count as hits. A similar judgment about error tolerances is not as central for configural comparisons. In the design of the false problems, we tried to make the error margins large enough that subjects could be confident that a close-call was really a hit. However, for the 30° groove-early problems

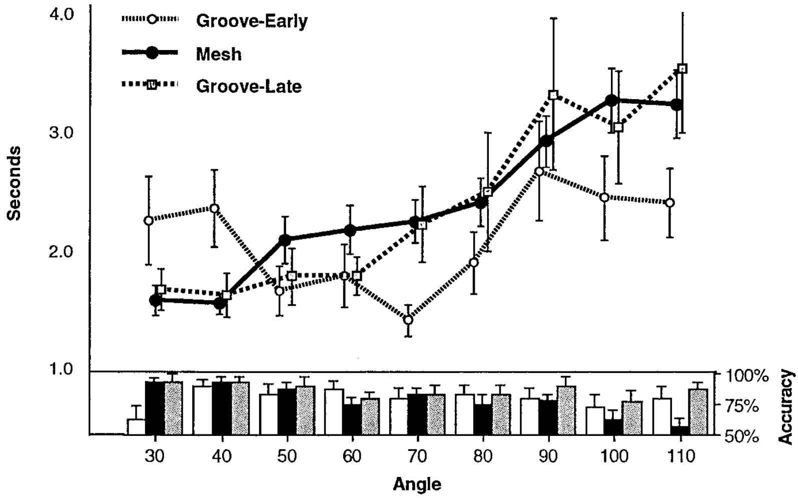

FIG. 3. Correct answer latencies and percentage correct for all problems (Experiment 1).

and the 100° and 110° mesh problems it appears that the error tolerances were insufficient given the subjects' minimal practice time. Consequently, as judgments became less sure, response bias could play a larger role. For example, at the 100° and 110° disparities, one may see that the subjects had a bias towards no-mesh responses. In the following experiment, subjects received more practice. This removed the near chance performances and rejection bias that make it difficult to interpret latencies.

Figure 3 presents the accuracy data and correct answer latencies for all three types of problem: groove-early, mesh, and groove-late. For the small angles of the groove-early problems, there was no evidence of the positive linear trend that implicates an analog rotation. Perhaps the problems invited the subjects to make a perceptual judgment of static distances, because the groove was so close to the point of contact. For example, when the knob on the power gear was at 30°, the groove was only 8° from the point of contact for the 60 mm driven gear. The next experiment provides better evidence on this issue.

The latencies for the larger angles of the three problem types provides a second source of support that subjects used analog rotations in this task. If subjects mentally rotated the gears, one would expect them to terminate their rotations as soon as either the knob or groove entered the point of contact. Notice that the mesh and the groove-late problems were answered at nearly identical speeds. One possible explanation is that the subjects terminated their rotations as soon as they noticed the knob entering the point of contact. Because the knob was positioned identically for positive and negative trials of the same angle, a consistent set of analog rotations would always put it at the point of contact at the same time. Thus, for the groove-late and mesh problems, the first evidence indicating the correct answer occurred when the knob entered the point of contact; the groove entered later in the case of a groove-late problem. Also notice that for the larger angles, the groove-early problems were answered the most quickly. This follows if subjects terminated their rotations when the groove entered the point of contact prior to the knob. Thus, the results for the larger angles make sense if one supposes that the knob and groove enter into the point of contact during an analog rotation.

The possibility that the smaller angles of the groove-early problems were answered without recourse to analog imagery suggests that some individuals could have applied non-analog strategies to solve the mesh and groove-late problems as well. As a first step, we compared individuals' slopes for the correctly answered mesh problems. Again, we chose to examine the hit latencies because they provided the most stable data. Averaging the slopes of the three fastest subjects yielded a rotation rate of 112° per second, whereas the slowest three averaged to 22° per second. We thought that these inter-subject differences might reflect different strategies as opposed to different rotation rates. Using the debriefing as a clue to strategic differences, we found that thirteen of the subjects either described a rotation strategy or were unable to articulate their method of solution. Two subjects, however, explicitly mentioned that they compared the arc lengths subtended by the marks on each gear. We will call this non-analog, visual comparison the *measurement strategy*. We explore these two subjects' data as a prelude to the next experiment in which subjects are trained to use the measurement strategy.

If the two subjects did compare arcs, latencies should be most similar for the smaller angles. These problems would be easy to compare at a glance and would not require the steadily increasing processing times of a longer rotation. At larger angles, subjects might switch to an analog strategy as the arc lengths become too great to handle in a glance. Moreover, at larger angles there is a decreasing correlation between the length of a cord and its corresponding arc—both within and across the gears. This is important because the cord, which is the straight line that connects the two ends of an arc, may provide a simpler yard stick for comparing the two arc lengths. This heuristic use of the cord may be effective at smaller angles but as the correlation becomes weaker at the larger angles, this strategy might be abandoned for a rotation strategy.

Figure 4 shows that the latencies of the two self-described measurers were relatively constant until about 90° where they increased. As a source of contrast, the figure includes the plots of the ten remaining subjects whose overall accuracy was above 75%. (We plot these subjects to show that the linear increase in latencies holds when removing subjects with near-chance performances.) In addition to the slopes shown in the figure, which were fitted post-hoc, a regression of latency on angular disparity and its quadratic indicated a non-linear component for the two measurers'; quadratic partial r

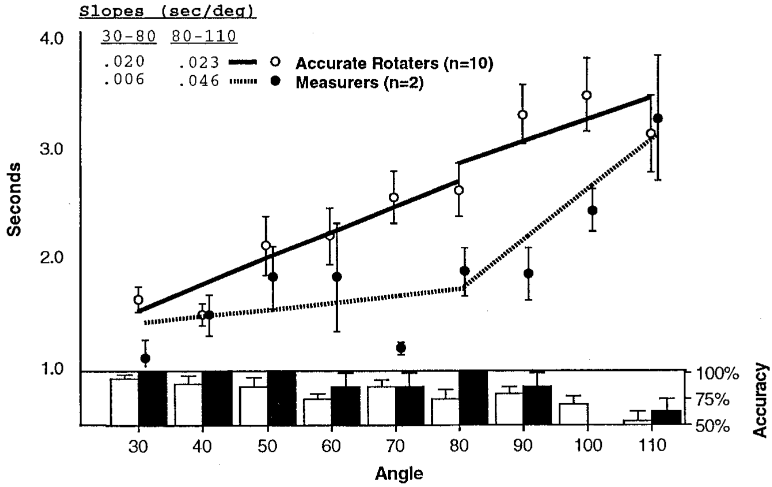

FIG. 4. Comparison between the two subjects who mentioned a measuring strategy and ten other subjects who had at least 75% overall accuracy (Experiment 1). The three subjects who performed close to chance are excluded from the plot to show that their removal did not remove the linear pattern.

= .21, but not the 10 other subjects; partial r = .004. Further intimating different underlying processes, the figure indicates that the two measurers were more accurate than the other subjects. Experiment 2 addresses the question of whether these differences were the result of strategy selection, ability, or a combination of both.

Discussion

Our interpretation of the positive correlation between angular disparity and hit latency is that people can coordinate the imagined behaviors of physically interacting objects through analog imagery. Although subjects did not exhibit a linear increase for the smallest angles of the groove-early problems, the interpretation of an analog rotation was corroborated by the correct rejection latencies for the larger angles. The groove-late and mesh problems had similar latency patterns, and the groove-early problems were solved more quickly. This makes sense if one supposes that subjects mentally rotated the gears and made their judgments as soon as the knob or groove entered the point of contact. An alternative to the analog interpretation might be that subjects implicitly counted sprockets. This would have caused longer latencies at larger disparities as there were more sprockets to count. However, the sprockets could not be reliably counted because they were indistinct and there was no direct relationship between the sprockets of the two gears. Moreover, to anticipate Experiment 3, subjects exhibit a linear relationship when the gears have no sprockets.

The best predictor of the hit latencies was the angular disparity of the knob, implying that the driven gears rotated at different speeds depending on their size. One explanation for this result is that subjects rotated the power gear at a constant velocity and used this as a baseline to determine the rotation rates for the two driven gears. There are several reasons that the power gear might have been animated with fixed angular velocity. One is that its size was constant, making it a suitable anchor from which to derive the rates of the two driven gears. Another is that power gear would always have to rotate the most quickly and consequently may have represented the upper limit of how fast subjects could rotate the gears. Another possible factor is the label ''power gear,'' implying that it should control the rate of rotation. Finally, the feedback focused on the driven gear, suggesting that it was the one to be regulated.

Another possibility is that the different angular velocities of the driven gears were a side effect of the subjects' efforts to make the circumlinear velocities equivalent. If subjects used the local analog strategy, they coordinated the motions of the two gears by driving their circumferences through the point of contact at the same rate. Using circumlinear velocity as the measure of transformation speed, the power and driven gears move the same number of arc length units per second. Experiment 3 addresses the question of whether we should conceive of subjects' transformation rates as being regulated by angular or circumlinear velocity.

A third explanation of why the driven gear had a variable rotation rate relies solely on the processing requirements of a spatial representation. It might be argued that the ''larger'' an internal image, the more slowly it can be transformed (Kosslyn, 1980; Shepard & Feng, 1972). The larger images of the larger gear may have slowed the rotation rates in direct proportion to the fewer number of degrees the gears needed to traverse. If this were the case, the driven and power gears may have coincidentally rotated at appropriate rates independently of one another. Consequently, there would be limited grounds for postulating the deployment of physical knowledge to resolve the problem of coordination. We do not have an a priori reason to suspect that the slowing effect of a larger gear's area would be in a close inverse relationship with the fewer number of degrees it must traverse. Nonetheless, the statistical results suggest that the larger driven gear may have slowed the overall rate of rotation. Perhaps the driven gear slowed the rotation to a greater extent than the compensating effect of having fewer degrees through which to rotate. Before drawing firm conclusions, it is worth re-testing the influence of the size of the driven gear on the rates of rotation.

The analysis of individual differences showed that two subjects solved the mesh problems under 90° without a stable increase in latency. This implies that these subjects did not use an analog strategy for the smaller angles. Figure 4 shows that these two subjects were faster. Moreover, both subjects were in the top third of the sample with respect to accuracy. A natural question is why the other subjects' results did not also follow this more optimal pattern. One possibility is spatial ability differences (Egan & Grimes-Farrow, 1982). For example, Just and Carpenter (1985) found that higher spatial ability yielded more efficient letter block rotations, because subjects used rotation trajectories that did not follow the natural axes of the block. Spatial ability may have similarly interacted with how subjects made their comparisons of the two gears. Another possibility is that the two measuring subjects had access to explicit knowledge about the task. For example, they may have inferred or known that the arcs should be identical lengths. A third possibility is an interaction between spatial ability and knowledge of the specific task. Experiment 2 examines these three possibilities by collecting measures of spatial ability and by providing explicit training in rotation and measurement strategies. If training causes people to use one or the other strategy and spatial ability does not, there will be support for the argument that the spatial resolution of the problem is knowledge dependent.

EXPERIMENT 2

The purposes of Experiment 2 were fourfold. The first was to replicate the results of the first experiment. The second was to show that linear increases in response latency are not a necessary outcome of the task regardless of strategy. The third was to examine the effect of the size of the driven gear more closely. The fourth was to show that individual differences in linear response patterns are primarily a function of strategy selection rather than spatial aptitude.

Each subject completed three treatments. They all began with the *no-training* treatment which was a replication of Experiment 1 with more practice to increase accuracy, and three driven gear sizes to further examine the effect of the driven gear's size. In the remaining two treatments, subjects were explicitly trained in different methods of solution. In the *rotation* treatment subjects were directed to use the strategy of ''pulling'' the gears through the point of contact. In the *measurement* treatment subjects were told to compare the arcs from each mark to the point of contact. The order of training was counter-balanced across subjects.

This experimental design can develop convergent and divergent evidence that people spontaneously use mental rotations to solve this problem of coordination. If a linear response pattern in the no-training treatment parallels the pattern in the rotation treatment, this would provide convergent evidence that subjects rotated in the no-training treatment. If the measurement treatment yields latencies that do not increase with the smaller angles, then there would be evidence that the linearity in the other two treatments was not a necessary outcome of the task, regardless of strategy.

To explore the influence of spatial aptitude, subjects took two spatial ability tests that measure transformational ability and visual memory. It was expected that the test of transformational ability, Paper Folding, would be a good predictor of performance in this rotation task, whereas the Shape Memory test would not. However, we did not expect that spatial ability would be a determinant of whether the subject could or would use an analog solution. Instead, our prediction was that strategy training would overwhelm any effect of spatial ability.

Method

Subjects. The 13 females and 3 males who responded to an advertisement posted in the Columbia University community represented a broad spectrum of education, age, and ethnicity. They were paid a flat fee after completing all three conditions.

Design. Subjects completed three treatments. All subjects completed the *no-training* treatment first to provide a sample of spontaneous solutions and to replicate Experiment 1. Subjects then completed the *rotation* and *measurement* treatments with completion order counter-balanced between subjects to preclude order confounds. In all treatments, the radius of the power gear was 20 mm, and the driven gear had radii of 40, 50, and 60 mm. The angular disparity of the knob ranged from 40° to 120° in increments of 10°. This gave an equal sample of data points about 80° which was suspected to be the point at which subjects would switch from a measurement to an analog strategy. This yielded a total of 108 problems for each treatment (3 sizes $\times$ 9 angles $\times$ 4 mesh/no-mesh). Spatial ability measures were the Paper Folding and Shape Memory tests (Ekstrom, French, & Harman, 1976) which test transformational ability and visual memory, respectively. Scores on Paper Folding were used to balance spatial ability across the two orders of training.

Procedure and stimuli. To improve accuracy and because this experiment involved training, subjects completed 60 training problems for each treatment. For each treatment the procedure was as follows: (1) see the rotating demonstration gears; (2) receive directions as to task; (3) practice 48 problems; (4) take a 5 min break; (5) see rotating gears again; (6) do 12 more practice problems; (7) do 54 recorded trials; (8) take a short break; (9) complete the recorded trials; (10) debrief; and, (11) return in 3 or 4 days for the next treatment if applicable. After the no-training treatment, subjects completed the spatial ability tests. In the no-training treatment, feedback was the same as in Experiment 1. In the measurement treatment, subjects were told to compare the lengths of the arcs that start at the point of contact and terminate at the marks on each gear. Feedback took the form of red arcs that indicated the correct lengths for a mesh problem. In the rotation treatment, subjects were told to imagine the gears pulling through the point of contact. To receive feedback, subjects pressed the space bar several times in succession. Each press animated the gears one step inward.

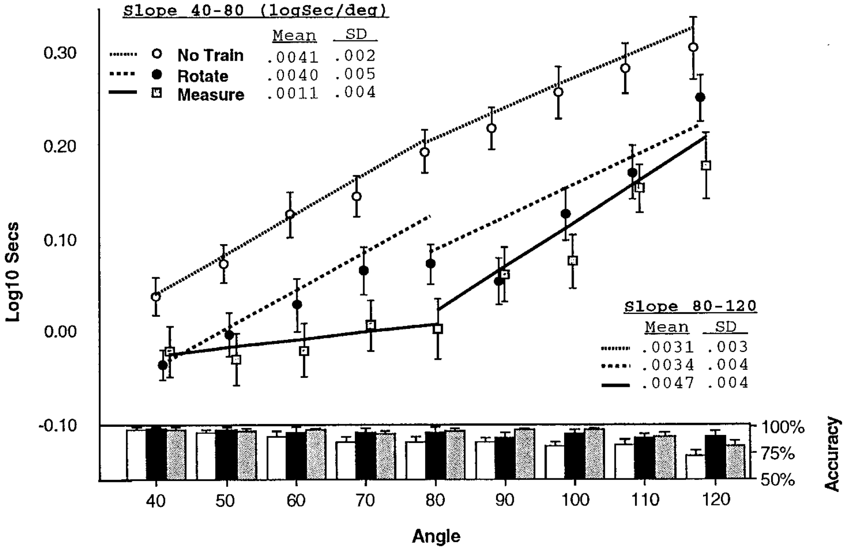

FIG. 5. Log transformed correct answer latencies and percentage correct for the mesh problems (Experiment 2). The slopes were computed separately for the smaller and larger angles to demonstrate the flat latency pattern for the smaller angles of the measurement condition.

Results

The Effects of Strategy Training on Latencies

To make the variances of each treatment comparable and to control the noise associated with the explicit training and the repeated measure of each subject over several days, hit latencies were transformed into base 10 logarithms. Figure 5 shows that the no-training and rotation treatments led to increasing latencies across the angular disparities but that the measurement treatment led to relatively flat response times for the smaller angles. (Figure 7 displays the data without the log transform.) Importantly, the accuracy rates for the measurement and rotation treatments were comparable for the smaller angles. This indicates that the effect of the measurement training was not merely a speeded response at the smallest angles; if it were, accuracy should have diminished.

To compare the effects of the three treatments, separate slope coefficients of log10 response time on angular disparity were calculated for each subject for each treatment for the *smaller angles* of 40° to 80° and for the *larger angles* of 80° to 120°. The two slope coefficients by the three treatments served as within-subject factors and the order of training was a between-subjects factor. The Paper Folding and Shape Memory scores served as covariates. Differences between the pairs of slopes within a treatment indicate nonlinearity. We used two separate slopes rather than a quadratic term so that we could capitalize on the results of the first experiment that showed the linear break would occur around 80°.

The slopes did not differ between the rotation and no-training treatments, $F(1,14) = .05$, $p > .8$. Importantly, the slopes for the larger and smaller angles did not differ in the two combined treatments; $F(1,14) = .002$, $p > .9$. In contrast, there was a significant difference between the slopes for the smaller and larger angles in the measurement treatment compared to the other two combined treatments; $F(1,14) = 5.89$, $p < .05$. This effect was primarily located in the 40° to 80° slope of the measurement treatment compared to the other two combined treatments; $F(2,28) = 3.95$, $p < .05$. The negligible slope in the measurement treatment suggests that subjects did not perform analog rotations for the smaller angles.

There was a main effect of training order; $F(1,14) = 7.15$, $p < .05$. However, there were no significant interactions involving order. This makes it difficult to interpret the effect. For example, the lack of an interaction implies that the order effect held for the no-training condition even though all subjects went through this condition first. Importantly, the interaction between order, treatment, and the two slopes was not significant; $F(2,28) = 1.78$, $p > .15$. This indicates that subjects could switch between strategies after training.

Effects of Spatial Ability on Strategy Application

The scores for the Paper Folding (PF) and Shape Memory (SM) measures were comparable to the means and variances of population norms and were minimally correlated; $R = -.07$. PF scores reliably covaried with the regression slopes in the preceding analyses; $t(12) = 2.28$, $p < .05$. Higher PF scores were associated with steeper slopes (i.e., slower rotations), perhaps due to more careful rotations. The SM scores did not reliably covary with the regression slopes; $t(12) = 1.77$, $p > .1$. There were no interactions involving the ability scores; all t's < 1. Table 1 shows the correlation of the ability scores with selected within-subject time and accuracy measures. We confine our discussion to the PF scores, because they were expected to tap into a subject's ability to transform images. We do not currently have an interpretation of the SM correlations.

The primary question addressed by these data is whether spatial ability can account for whether an individual implements a measurement or rotation strategy. There are too few subjects to draw firm conclusions about the effect size of spatial ability in this task, and thus we do not examine the magnitudes of correlations for significant effects. Instead, we look for correlation patterns across the measures, treatments and angular disparities. In particular, we focus on the correlations between the PF scores and the slopes ($\log_{10}$ s/deg), found in the middle rows of Table 1. These correlations should vary if spatial ability influences strategy application because the slopes of the smaller angles discriminate the measurement and rotation strategies. We first consider the two training treatments because they may provide guidance for how to dis-

TABLE 1
Correlations between Spatial Ability and Selected Measures (Experiment 1)

	Training treatment (N = 16)					
	None		Rotate		Measure	
Dependent measure	Paper folding	Shape memory	Paper folding	Shape memory	Paper folding	Shape memory
$\log_{10}$ RT Average for Hit						
40–80°	−.39	.49	−.36	.25	−.33	.19
80–120°	−.18	.56[a]	−.23	.40	−.27	.21
Slope[b]						
40–80°	.41	.39	−.20	.17	.28	.44
80–120°	.41	.47	.14	−.04	−.04	−.41
% Correct[c]						
40–80°	.25	.01	.38	−.08	.53[a]	.07
80–120°	−.10	−.44	.41	−.32	.23	.15

[a] $p < .05$, two-tailed.
[b] Individuals' coefficients from regressions of $\log_{10}$ response time on angular disparity for true positives.
[c] Percentage of correct answers across positive and negative trials.

criminate whether spatial ability influenced strategy selection in the no-training condition. In the measurement treatment, higher PF scores predicted steeper slopes for the smaller angles ($R = .28$), whereas in the rotation treatment they predicted shallower slopes ($R = -.20$). Assuming these correlations are reliable, one explanation is that high PF subjects were inclined to imagine the gears rotating in the measurement treatment, thus they had steeper slopes than subjects who made the static comparison. In the rotation treatment, the high PF subjects may have completed their rotations quickly yielding shallower slopes, whereas the low PF subjects, who were not allowed to measure the arcs, took more time carrying out their rotations. This explanation is highly speculative, however, in that the larger angles, which presumably involved rotations in both treatments, showed little evidence that higher PF ability led to faster rotations.

The key question is whether spatial ability determined how subjects spontaneously solved the task. If it did, then there is less reason to posit a tight relationship between context-sensitive knowledge and spatial imagery in the solution of problems of coordination. For the smaller angles, six subjects in the no-training treatment had slopes that were within a half standard deviation of the mean slope for the measurement treatment. Thus, it seems possible that some subjects used a measurement strategy at times during the no-training treatment. If spatial ability determined strategies for the smaller angles in the no-training treatment, then there should be a minimal correlation between PF ability and these slopes. This is because low PF ability would presumably be the cause of a measurement strategy which would have shallow slopes, and high PF ability would be the cause of a rotation strategy which would also have shallow slopes (subjects could make faster rotations). The results did not support this implication because the correlation between PF ability and the small-angle slopes was large compared to the two trained conditions ($R = .41$). Moreover, if low transformational ability led some subjects to use a measurement strategy for the smaller angles in the no-training treatment, then one would expect the correlations between PF and slope to vary for the smaller and larger angles. This is because lower PF ability would predict a shallow slope for the smaller angles (due to a measurement strategy) and a steep slope for the larger angles (presumably, all subjects rotate at the larger angles and low ability subjects would take longer completing the rotations.) However, the correlations between the PF scores and the two sets of slopes were the same (R's $= .41$). Consequently, PF scores are not differentiating the signatures of the two strategies. Finally, in the no-training treatment higher PF scores predicted slower rotation rates. This undermines the speculation that higher PF subjects used a rotation strategy in the no-training condition because it was faster and easier for them, compared to the low PF subjects who presumably found a different strategy. Given these three lines of evidence, it is difficult to build a story that explains how PF ability influenced strategy selection consistently across the treatments.

The most consistent finding was that larger PF scores predicted briefer latencies, as shown by the negative correlations in the top two rows of Table 1. The bottom rows of Table 1 also show that higher PF scores correlated with greater accuracy in the two training treatments. However, for the no-training treatment, PF scores had correlations with accuracy in opposite directions for the smaller and larger angles. Overall, this suggests that PF ability did not determine how subjects solved the gear problem, but it did determine overall speed and how effectively they performed the task once taught a specific strategy.

The Effect of Gear Size on Rotation Rates

Figure 6 shows the effect of the driven gear sizes in the combined no-training and rotation treatments. Because our interest is in the effect of gear size on analog solutions, we only analyze the no-training and rotation treatments. A within-subject analysis of subjects' regression slopes ($\log_{10}$ hit latency on angle) for each gear combination showed that the main effect of the driven gear size was significant; $F(2,28) = 3.69$, $p < .05$. One interpretation is that the size of the driven gear influenced the speed at which subjects could conduct their rotations. However, Fig. 6 shows that the effect of gear size was primarily confined to the largest gear at the largest angles. It also shows that for these largest problems the accuracy levels dropped off. This suggests

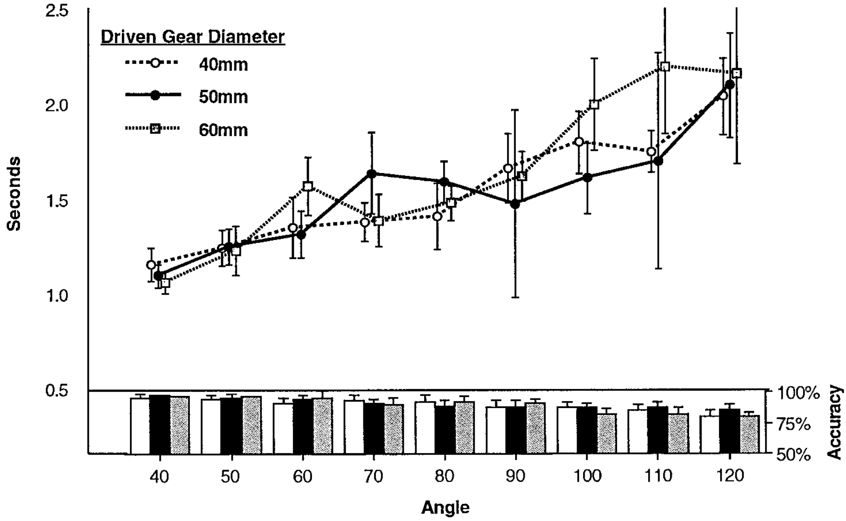

FIG. 6. Correct answer latency and percentage correct for the mesh problems broken down by the size of the driven gear for the combined No-Training and Rotation Training treatments (Experiment 2).

that the longer latencies did not reflect slower rotations but rather may have resulted from a non-linear increase in problem difficulty. Experiment 3 investigates this possibility.

The Effect of the Mesh and No-Mesh Problems on Latency

Figure 7 shows the correct response latencies in seconds, the average slope (s/deg) and accuracy for the groove-early, mesh, and groove-late problems. Perhaps due to the increased practice, this experiment improved the picture of what happens with the smaller angles of the groove-early problems. Latencies remained relatively flat. Our interpretation is that subjects used a measurement strategy for the smallest angles of the groove-early problems. Because the groove was very close to the point of contact for the smaller angles, subjects did not find it necessary to rotate the groove into the point of contact to make their judgments. Moreover, the margin of error built into the groove-early stimuli for the smaller angles may have been large enough that subjects could detect a difference without rotating. Because detecting a difference does not require the precision of determining a match, responses were quicker. For the larger angular disparities the differences were not easily detected, and consequently subjects used a rotation strategy. This explanation is the same as the one used to explain the latency shift in the measurement treatment. As the angular disparities became larger and the curvatures of the two gears became increasingly

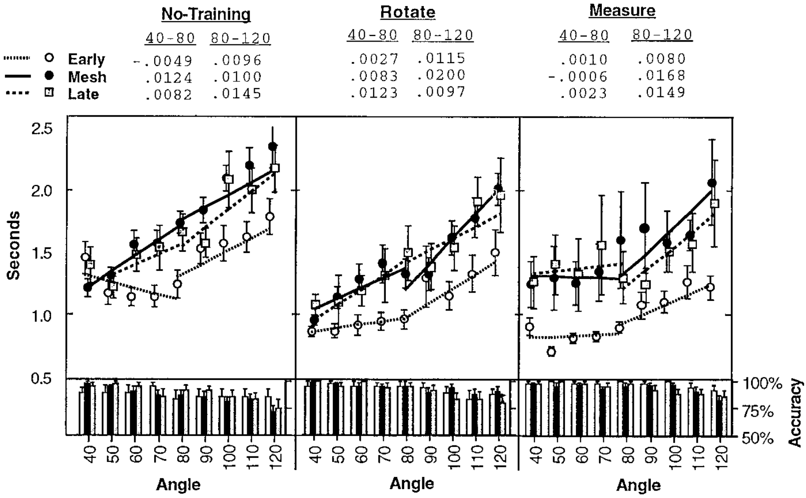

FIG. 7. Correct answer latency and percentage correct for positive and negative trials (Experiment 2).

disparate, subjects could no longer compare arc lengths in a glance and needed to switch strategies. This switching of strategies within a given treatment reflects the flexibility people may often bring to problems of coordination. Even in the explicitly trained rotation treatment, one may interpret the shallow slope of the groove-early problems as an attempt by some subjects to apply a measurement strategy for the smaller angles.

For the larger angles in all three treatments, it appears that subjects made their judgments as soon as the knob or groove entered the point of contact. The groove-late and mesh problems exhibited overlapping linear patterns, and the groove-early problems, which also showed a linear increase in latency, had consistently shorter latencies. This makes sense if one notes that the knob would be the first indicator of a match or mismatch for the mesh and groove-late problems. Because the knob is at the same initial angle for both types of problem, a subject would need to rotate the gears the same amount before the evidence of a match or mismatch would arise. However, for the groove-early problems, evidence of a mismatch would arise sooner, because the groove would enter the point of contact prior to the knob. An alternative interpretation is that the groove-early problems were easier to solve than the mesh and groove-late problems, thereby yielding faster response times. However, if this were the case, one would expect the accuracy levels to be greater for the groove-early problems. Figure 7 shows that the groove-early problems did not yield a systematic accuracy advantage, thereby making it difficult to ascribe the latency results to problem difficulty.

Discussion

The results from the measurement treatment support the argument that people can solve mechanical problems through analog imagery. The direct intervention of measurement training showed that these gear problems can yield a non-linear pattern of reaction times. This defeats the common argument against analog imagery that the linear results are due to unspecified task characteristics. Coupled with the convergent evidence that the no-training and the explicitly trained rotation treatments showed highly similar results, the case that subjects were using an analog rotation in the no-training treatment gains validity.

One might argue, however, that subjects used tacit knowledge rather than imagery in the no-training condition. This led them to hold off responding according to their implicit estimations of how long such a rotation would take, an estimation that would be aided by the initial demonstration of the rotating gears. However, it seems unlikely that subjects would have tacitly determined that for the larger angles they should have responded to the groove-late and mesh problems at the same speed, and responded to the groove-early problems more quickly. Alternatively, one might claim that the training interventions cued the subjects to the desired results. This account seems unlikely given that the subjects were not familiar with analog imagery research; debriefing revealed that the subjects had no idea of the predicted patterns of reaction times. Nonetheless, it is desirable to achieve a similar divergent and convergent pattern of evidence through less intrusive methodologies. The approach in Experiment 3 is to induce different strategies by changing the visual character of the gears. Moreover, we removed the initial demonstration of rotating gears. This less directive manipulation of strategies should counter objections about experimenter effects.

The current experiment suggests that spatial ability cannot provide a complete account for how people solve the gear task. Although the statistical power is too small to disallow an effect of spatial ability, the results showed that compared to strategy training, which used the same small sample, spatial ability had limited effects. While there was evidence that higher spatial ability led to faster responses and greater profit from instruction, the training intervention had a large influence on how subjects solved the problem. This allocates a role for context-sensitive knowledge in imagery about connected objects. In the measurement treatment, subjects used knowledge about the gears' static properties—they should have equal arc lengths. In the other treatments, subjects used knowledge of the gear's dynamic properties—either that their angular velocities are proportional to their sizes or that they move through the point of contact at an equal rate.

An alternative to a knowledge-based explanation of how subjects solved the problem of coordination is that the size of an image determines its rate of rotation. Thus, one might argue that the driven and power gears rotate at an individual's maximum velocity, it just so happens that a gears' maximum velocity is in direct proportion to its size. Supporting this argument is the significant result that a larger driven gear yielded a steeper slope than a smaller one.

A different hypothesis for the effect of the driven gear is that the longer latencies were the result of a non-linear increase in the difficulty of the problems as they approach and exceed 90°. Figure 6 shows that the slowing effect of the largest driven gear took hold around 90°. This angle is close to the point where the subjects abandoned a measurement strategy, although this may be a coincidence of the gear sizes used in this experiment. With respect to the measurement strategy, the knob and groove may become too far apart around 90° to be easily incorporated into the same image with sufficient resolution for a simple length comparison (Kosslyn, 1980). This divergence, as well as the divergence in the arc curvatures, may have led subjects to switch to a rotation strategy. With respect to a rotation strategy, the angular disparity itself, in addition to absolute distances, may have added difficulties. At 90° the power and the driven gears no longer share a common direction of motion; the horizontal and vertical motion vectors of the two marks are opposed between the gears. This may make the coordination of the respective motions more difficult, assuming that people are using the marks to coordinate or evaluate their rotation rates. The compounding difficulties of the opposing motions and the large absolute distances between the marks may have reduced the subjects' margin of error. Consequently, they might have re-rotated, slowed down their rotations, or spent more time encoding the gears. Notably the slowing effect occurred mostly for the 60 mm driven gear past 90° which is where the divergences in the direction of motion and the absolute distance between the knob and groove are most pronounced. The interpretation of a non-linear increase in difficulty is further supported by the decline in accuracy around 90°. If problems become more error prone because of a change in the processing requirements, Weber's Law of equal discriminability may not hold across the 90° point. In follow-up work, we found that increasing the margin of error by 2% past 90° was sufficient to bring the error rates back into line. In Experiment 3 we used this adjustment so that all problems past 90° had a 22% margin of error. If our hypothesis of a non-linear increase in problem difficulty is correct and our adjustment sufficient, the small effect of the driven gear size should disappear.

EXPERIMENT 3

Experiment 3 attempted to isolate another set of strategies without relying on explicit instruction. We distinguish three reasoning methods: analytic strategies for calculating an answer; a global analog strategy that depends on the angular disparity of the power gear; and a local analog strategy that depends on the circumlinear disparity of the gears. To help determine whether subjects used the angles or surfaces of the gears to coordinate their mental rotations, we employed power gears with 20 mm and 30 mm radii. If subjects' latencies for a given angular disparity are the same regardless of power gear size, then we can conclude that subjects are controlling the angular velocities of the gears. However, if their latencies are 1.5 times longer for the 30 mm gear, then subjects are using the surfaces of the gears. This is because the knob on the 30 mm gear subtends an arc that is 1.5 times longer than the 20 mm gear for a given angle.

To induce the different strategies, we manipulated the surface features of the gears displayed on the computer. Inducing different reasoning strategies through the manipulation of surface features has worked in domains like algebra (Hinsley, Hayes, & Simon; 1977) and statistics (Ajzen, 1977). In the current case, we changed perceptual characteristics rather than problem semantics. Research paradigms including implied motion (Shiffrar & Freyd, 1990), analog imagery (Schwartz, 1995) visual recall and recognition (Freyd, Pantzer, & Cheng, 1988; Mandler & Parker, 1976; Mandler & Stein, 1974), rapid identification (Biederman, Mezzanotte, & Rabinowitz, 1982), and problem solving (DeLoach & Marzolf, 1992; Moore & Schwartz, 1994) have shown effects depending on whether a display corresponds to a possible, perceived world. Our hypothesis is that as the features of a picture approximate the perceived physical world (e.g., a photograph), people gravitate towards depictive reasoning strategies. But, as a picture approximates an abstracted and symbolic world (e.g., a schematic diagram), people gravitate towards analytic strategies (cf. Rumelhart, 1989). If this hypothesis is correct and different strategies result from changes to a visual display, then there is evidence that differences in strategy and performance are not simply due to covert experimenter effects.

Prior evidence suggests how pictorial realism will influence problem solving strategies in the gear task. Schwartz (1995) asked subjects to judge whether marks on two legs of a giant hinge would meet if the hinge closed. Subjects saw either realistic or line drawings of the hinge. The opening of the hinge was varied to give the range of angles necessary for a test of analog imagery. The subjects who saw the realistic drawings tended to imagine the hinge closing, as though gravity were pulling the upper part of the hinge around the pivot point connecting the two legs of the hinge. Evidence to this effect was a consistent linear increase in latencies over the angles of the hinge opening. In contrast, the subjects who saw the line drawings did not initially simulate the physical behaviors of hinges, but instead they tried to analyze and compare static features such as line lengths and angles. Evidence to this effect included the lack of a consistent linear increase over the angles of the hinge opening. In the current experiment, we expected a similar effect whereby the most realistic drawing of the gears pulls for a simulation of physical behaviors and the most abstract drawings of the gears pull for a more analytic approach. Figure 8 demonstrates the feature combinations that make this experiment's five between-subject conditions. The conditions are

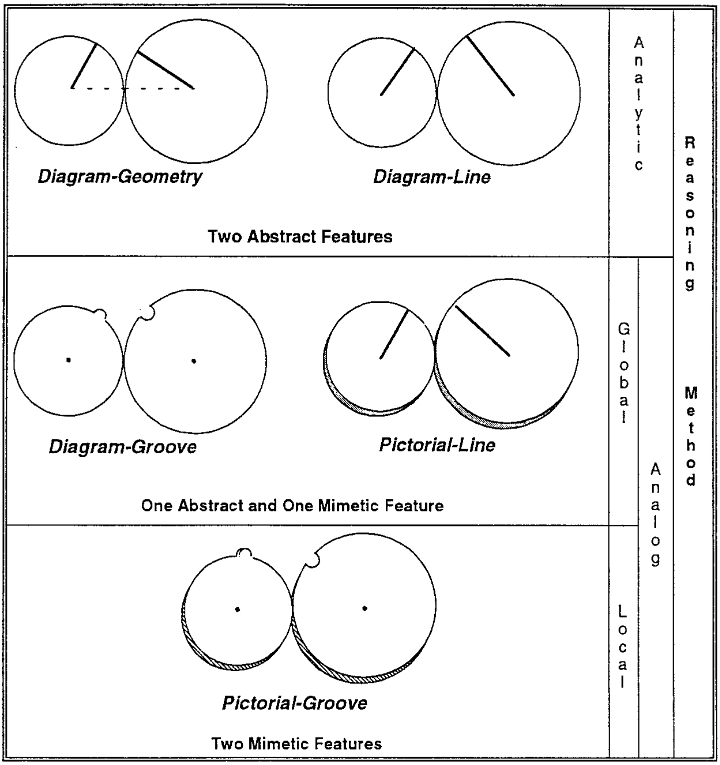

FIG. 8. A schematic of the five gear displays of Experiment 3. On the right side of each diagram is the expected strategy for processing the problem.

organized into three levels of pictorial realism. We call visual features mimetic or abstract corresponding to whether they mimic or schematize the characteristics of physical gears.

For the drawing with two mimetic features, *pictorial–groove,* we thought subjects might employ the local analog strategy. As in the case of the hinge study, the mimetic cues may cause subjects to think in terms of physical gears and their behaviors. In particular, subjects might represent the physical forces that connect the gears at their surfaces, much as subjects may have represented the physical constraints of a pivot point in the hinge task. This representation could help solve the problem of coordination because the forces that act upon the gears' surfaces cause them to move point-for-point at their touching surfaces. As subjects imagine the turning of one gear, they might imagine its surface forcing the movement of the second gear's surface. If subjects do resolve the problem of coordination at the gear surfaces, response times should be a function of circumlinear disparity rather than angular disparity.

Similar to the hinge study, we thought that subjects would exhibit more analytic approaches for the two displays with two abstract cues, *diagram–line* and *diagram–geometry*. Although the problem is framed in terms of physical behaviors, the perceptual features are more akin to a geometry problem. Consequently, subjects might be inclined to compare static features. For example, subjects could judge whether the ratio of the gear sizes is the inverse of the ratio of the two angular disparities. Evidence that subjects do not consistently follow an analog strategy in this condition would be relatively constant latencies for the angular disparities past 85°. It should be noted that we predicted flat latencies for the smaller angles when using a measurement strategy, whereas we are predicting flat latencies for the larger angles when using the analytic strategy. In prior research (Black, Schwartz, & Marks, 1988), we found that 85° marked the point at which subjects' latencies no longer increased. This makes sense if subjects exclusively employ analytic approaches at the larger angles; the algorithm of an analytic strategy should be deployed at the same speed regardless of the quantities involved (i.e., the sizes of the circles and angles). A number of possibilities may have led subjects to use exclusively analytic strategies past 85° in both our previous and current research. One possibility may be the high salience of the differences in the gears' angles at the larger angles. Another possibility is that subjects may default to an analog solution, but for the larger angles the rotation takes long enough for more reflective strategies (e.g., computing proportions) to come into play. As a manipulation check that subjects use an analytic approach at the larger angles in the diagram–line condition, we included the diagram–geometry condition in which dotted lines connected the centers of the two circles. The resulting wedges make the problem reminiscent of textbook geometry problems. If subjects in the diagram–geometry condition show the same flattening of latencies past 85° as in the diagram–line condition, there is reasonable evidence that subjects pursued an analytic strategy.

We also included two conditions, *diagram–groove* and *pictorial–line*, in which there were one mimetic cue and one abstract cue. One possible prediction is that subjects will either choose an analytic or local analog strategy depending on the salience of either pictorial feature. For example, subjects in the pictorial–line condition might focus on the line and use an analytic strategy. On the other hand, they might focus on the three-dimensional surfaces of the gear and use a local analog strategy. A second possibility, the one that motivated the inclusion of these two conditions, is that subjects might employ a global analog strategy. Our reasoning was that the presence of a mimetic feature and the presentation of a problem about gear motions would yield an analog solution. However, the abstract feature might also cause the subjects to solve the problem of coordination on the basis of geometric properties rather than physical forces. In the case of the pictorial–line display, the line may prime geometric knowledge to determine the relative rates of rotation. In the diagram–groove condition, the lack of a three-dimensional surface may make it less likely that subjects will represent surface forces. As a result they might also rely on global relationships between the gear sizes to resolve the problem of coordination. Using the ratio of the two gear sizes, subjects in both conditions could compute appropriate relative angular velocities for each gear and use this derivation to coordinate the rates of rotation. If subjects do control the angular velocities of the gears, then latencies should be a function of angular disparity and not circumlinear disparity.

Secondary lines of evidence may also be available for distinguishing between the strategies subjects use in response to each graphical display. Consider the difference between the global and local analog strategies. With the global strategy, it is necessary to calculate the appropriate ratio of angular velocities for the two gears. In contrast, the local strategy does not require this calculation, because the motion of one gear directly coordinates the motion of the second through the gear surfaces. Thus, for a global strategy there is a period of computation preceding the analog rotations that would not occur for the local strategy. This would be reflected in the constant portion of the regression equation that relates latency and angular disparity. One would also expect the global strategy to be more error prone and less efficient than the local strategy. For the global strategy one needs to make a rough estimation of relative gear sizes, and one needs to control two separate angular velocities. In the local strategy, the size relationship between the gears is irrelevant, and both gears are moving at the same rate with respect to their surfaces.

Similar lines of evidence may separate the analytic strategy from the two analog strategies. For an analytic strategy, a large part of one's time is spent estimating extent and computing size and angle ratios, while very little time, if any at all, is spent transforming the images. Reflecting this period of computation, the constant portion of a regression of latency on angular disparity should be greatest for the analytic strategy. One might also expect analytic strategies to be quite variable and errorful in this type of task (Schwartz, 1995). The problem does not suggest a single analytic solution and subjects may try a variety of heuristics. Subjects may also commit compounding errors as they make the necessary size estimations.

Method

Subjects. Fifty men and women, fulfilling course requirements at Columbia University, were randomly separated into five groups with eight or nine females in each.

Stimuli. There were two attributes that determined a gear's appearance: The formal features of the display, and the type of mark used for each angle. A gear was drawn either as a *diagram* or a *picture.* A diagrammatic gear was simply an unfilled white circle drawn against a magenta background. The center of the circle was marked. The pictorial gears were solid white, with accurate but somewhat irregular teeth, and three-dimensional shading. The

TABLE 2

Average Latency (s), Slope Coefficients of Latency by Angle (s/deg), and Percentage Correct Broken Down by the Larger and Smaller Angular Disparities in the Five Display Conditions (Experiment 3)

	Diagram line	Diagram geometry	Pictorial line	Diagram groove	Pictorial groove
Hit RT[a]					
40–85°	2.77 (2.24)	2.48 (1.85)	2.20 (1.35)	1.76 (1.13)	1.16 (.56)
85–120°	4.07 (2.79)	3.52 (2.36)	2.74 (1.45)	2.10 (1.05)	1.47 (.71)
Slope[b]					
40–85°	.044 (.036)	.040 (.040)	.017 (.025)	.021 (.036)	.007 (.011)
85–120°	.000 (.053)	−.002 (.044)	.018 (.030)	.021 (.018)	.012 (.012)
Correct[c]					
40–85°	71%	76%	77%	90%	95%
85–120°	59%	71%	67%	82%	93%

[a] Seconds averaged across subjects.
[b] Average individual slope of seconds regressed on the angular disparity of the power gear.
[c] Percentage correct over positive and negative trials.

angle marks were either the usual knob and *groove* (appropriately shaded or diagrammatic), or a radial *line* drawn from the center of the gear to the perimeter. In one diagrammatic display, *geometry–line*, a dotted line connected the centers of each circle. For the displays that used lines (*diagram–line*, *geometry–line* and *pictorial–line*), training feedback was a second, radial line indicating the correct mesh angle on the driven gear. For the two groove displays (*diagram–groove* and *pictorial–groove*) feedback was the same small green circle as in Experiment 1. In the current experiment, the percent used to determine the no-mesh disparities was increased to 22% past 90°.

Design and procedure. There were five between-subjects conditions: Diagram–geometry, diagram–line, diagram–groove, pictorial–line, and pictorial–groove. In each condition there were three different gear combinations (power::driven): 20::40 mm, 20::60 mm, and 30::60 mm. Within each gear combination there were 11 different angles: 40, 50, 60, 70, 80, 85, 90, 95, 100, 110, and 120°. The 5° increments around 90° created a balance of data points about 85°. Within each angle there were two mesh and two no-mesh problems, yielding a total of 132 recorded trials. The procedural changes to the no-training treatment of Experiment 2 were that subjects did not see the initial demonstration of rotating gears, and there were 72 practice trials.

Results

Separating Analog and Analytic Strategies

Table 2 summarizes the differences in overall performance. These results may be ranked left to right in a rough continuum from fewest to most mimetic

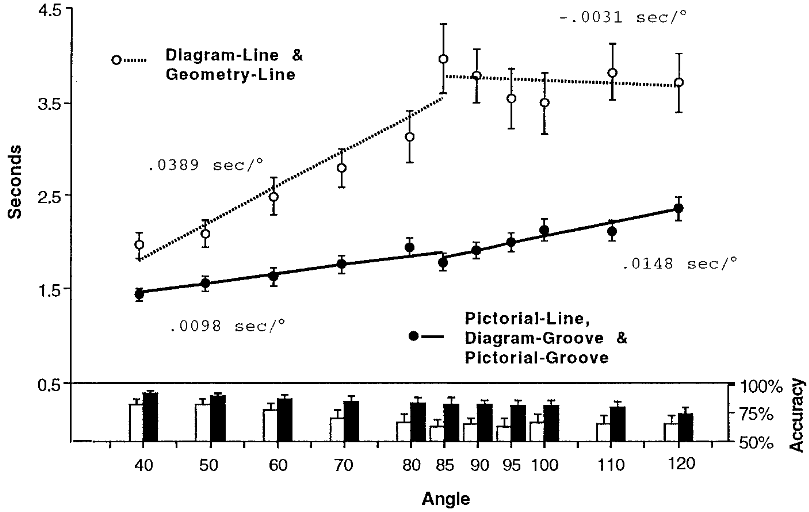

FIG. 9. Correct answer latency and percentage correct for mesh problems broken down by whether the display included a mimetic feature (Experiment 3).

features. The pictorial–groove condition, which had two mimetic features, had responses that were the fastest, least varied, and most accurate. The diagram–line and diagram–geometry conditions, which had no mimetic features, had the slowest, most variable, and least accurate responses. The two conditions with a single mimetic feature, diagram–groove and pictorial–line, fell in between the other conditions.

Table 2 and Fig. 9 show that the latencies from the diagram–geometry and diagram–line conditions did not increase past 85°, indicating a nonanalog, analytic approach at the larger angles. Relying on prior evidence that the latencies for the abstract gears flatten past 85° (Black, Schwartz, & Marks, 1988), separate regression slope coefficients of hit latency on the power angle were found for each subject for the 40° to 85° angles and the 85° to 120° angles. Table 2 shows the mean slopes and between-subjects standard deviations for each condition. The two slope coefficients were a within-subject factor, and the type of gear display was a between-subjects factor. Four planned, orthogonal contrasts were conducted. The first contrast showed that the slopes of the diagram–line and diagram–geometry conditions did not differ significantly from each other; $t(45) = -1.04$, $p > .3$. The second contrast showed that the difference between the slopes of the smaller and larger angles in the combined diagram–line and geometry–diagram conditions were significantly different from the other three conditions combined; $t(45) = -3.89$, $p < .01$. The third contrast showed that the slopes from the pictorial–line condition were not significantly different from the two groove conditions combined;

TABLE 3
The Effect of the Size of the Power and Driven Gears on the Slope of Latency by Angular Disparity (s/deg) for the Five Display Conditions (Experiment 3)

Power::driven	Diagram line	Diagram geometry	Pictorial line	Diagram groove	Pictorial groove
20 mm::40 mm	.030 (.036)[a]	.018 (.026)	.026 (.024)	.012 (.008)	.009 (.012)
20 mm::60 mm	.026 (.024)	.032 (.020)	.016 (.017)	.049 (.101)[b]	.010 (.007)
30 mm::60 mm	.025 (.045)	.021 (.023)	.018 (.013)	.017 (.013)	.013 (.011)[c]

[a] The mean individual slope coefficient of latency on angular disparity (and the standard deviation between subjects).

[b] When removing an outlier, the average 20::60 slope for diagram–groove is .017 (.012).

[c] Contrast of 30::60 vs (20::40 and 20::60), $p < .05$.

$t(45) = .85$, $p > .4$. This shows that it was not solely the line feature that caused the flattening of responses for the larger angular disparities. Lastly, the slopes for the two groove conditions were not significantly different from each other; $t(45) = -.29$, $p > .7$. Thus, both the line and the diagrammatic character of the display were necessary to elicit the flat response times for the larger angles.

Separating Global and Local Analog Strategies

Table 3 shows the average of the subjects' slopes that result from regressions of hit latency on angular disparity for each gear combination. A contrast of the 20::40 mm gear slopes with the 20::60 mm gear slopes tests whether the size of the driven gear had an effect on rotation rates. A contrast of the two 20 mm power gear combinations with the 30::60 mm combination tests whether latencies were dependent on the angular or circumlinear disparity. The three slope coefficients for the three gear combinations created a within-subject design. Because the variances of the five conditions were unacceptably heterogeneous when the data were partitioned this way, we ran separate analyses for each condition. The comparisons of the slopes for the two driven gear sizes showed no reliable effects in any of the conditions; diagram–line $t(9) = .33$; diagram–geometry $t(9) = -1.79$; pictorial–line $t(9) = -.041$; diagram–groove $t(9) = -1.2$; pictorial–groove $t(9) = -.48$. Adding 2% to the no-mesh problems greater than 90° removed the effect of the driven gear's size.

When contrasting the slopes of the two sizes of the power gear, only the pictorial–groove showed a significant effect for the size of the power gear; $t(9) = 2.79$, $p < .05$.; all other conditions, $t < 1$. Further analysis supported the hypothesis that circumlinear disparity would better model the data for the pictorial–groove condition. In the pictorial–groove condition, the average within-subject difference between the slopes for the 20 and 30 mm power

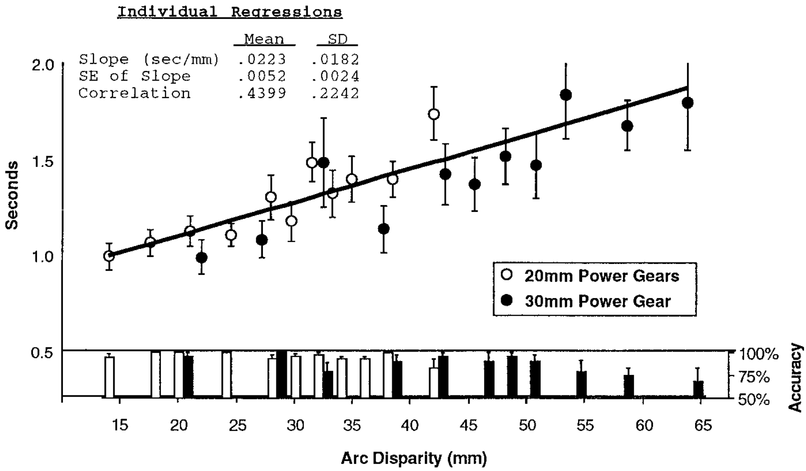

FIG. 10. Hit latencies for the pictorial–groove condition plotted against the circumlinear disparity of the knob (Experiment 3).

gear was 1.54. This is close to the 3:2 ratio one would ideally expect. In contrast, when removing three of the subjects who contributed undue variance to the diagram–groove and pictorial–line conditions, the difference between the 20 and 30 mm slopes was .93 and .94, respectively. Figure 10 shows the pictorial–groove's hit latencies plotted by circumlinear disparity. The regression slope was found by regressing the hit latencies on circumlinear disparity. For the pictorial–groove condition, a linear velocity of .02 s/mm is a more appropriate description of mental transformation rate than an angular velocity measure. This linear metric describes a feature of the display and not mental or retinal distances.

Separating the Strategies on the Basis of Initial Calculation Time

If we are correct that the proposed strategies require different amounts of calculation prior to any rotations, this should influence the intercepts in each condition. To give the best estimate of the constant portion of the hit latencies, we used the regression model most appropriate to each condition. For the diagram–line and diagram–geometry conditions, this consisted of two separate vectors, one representing the angles from 40° to 85° and another representing the angles from 85° to 120°. This non-linear model is warranted because subjects' latencies yielded statistically different patterns for each set of angles. Regressing across hit trials without blocking by subject, the intercepts were 1.45 s (*SE* = .37) for the diagram–line condition and 2.00 s (*SE* = .29) for the diagram–geometry condition. For the pictorial–groove condition, the model consisted of a separate vector for each size of the power gear. This model is preferable because the two gears yielded a significant difference.

The intercept for this condition was .63 s (*SE* = .09). For the pictorial–line and diagram–groove conditions, we only used the power angle vector. This model is warranted by the repeated finding that this angle was a reliable predictor of response time and no contrasts indicated the need for a more complex model. The intercepts were 1.37 s (*SE* = .21) for the pictorial–line condition and 1.08 s (*SE* = .17) for the diagram–groove condition.

To compare the intercepts between-conditions, we found each subject's intercept using the appropriate regression model. These intercepts served as data points in an analysis of variance. Because the variances could not be pooled due to excessive heterogeneity, and because this analysis required contrasts between condition means, we used separate estimates of variance to conduct the contrasts (Cliff, 1987). The SPSS[x] Oneway program calculated the degrees of freedom for the error terms. Four a priori contrasts tested whether the intercepts fell in the predicted order. The average intercepts for the diagram–line (1.72 s, SD = 1.73) and diagram–geometry (2.80 s, SD = 1.80) conditions were not significantly different from each other; $t(18.0) = 1.37$, $p > .15$. Similarly, the average intercepts for the diagram–groove (.98 s, SD = .83) and pictorial–line (1.43 s, SD = 1.00) conditions were not significantly different from each other; $t(17.4) = 1.09$, $p > .25$. However, the combined diagram–line and diagram–geometry conditions had intercepts reliably greater than the combined pictorial–line and diagram–groove conditions; $t(27.0) = 2.36$, $p < .05$; and the combined pictorial–line and diagram–groove conditions had intercepts reliably greater than the average intercept for the pictorial-groove (.50 s, SD = .57) condition; $t(25.4) = 2.58$, $p < .05$.

Although we believe these intercept differences reflect up-front computation, an alternative interpretation is that they reflect differences at the end of the problem-solving process. The most likely scenario for end-of-process differences involves judgment bias. A large rejection bias could cause subjects to perseverate on mesh problems until they built enough confidence or evidence to overcome their bias. Evidence in favor of this interpretation would be a large rejection bias accompanied by longer hit latencies. The diagram–groove condition demonstrates this pattern in Fig. 11. Subjects exhibited a bias towards rejection (i.e., more accurate on negative problems) and had longer hit latencies. However, the total picture did not support this interpretation, because the other conditions did not show this pattern. Moreover, the pictorial–groove latencies followed the same pattern even though the error data did not indicate a response bias. In lieu of a strong alternative for why tail-end processes would affect the intercepts, we interpret the intercept differences to be the result of processes that occurred prior to any analog transformations.

Discussion

The formal features of the gear displays in Experiment 3 produced three distinct data signatures without using direct experimenter intervention. Like

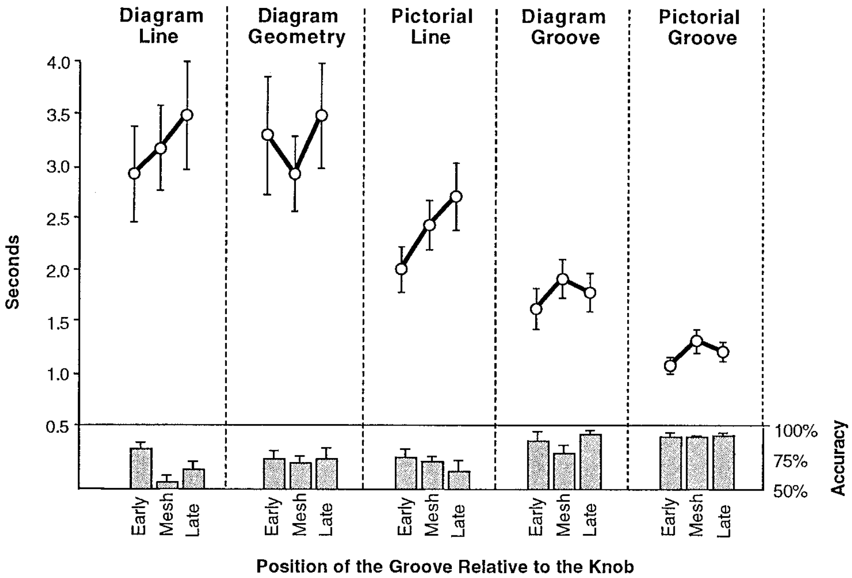

FIG. 11. Correct response times and percentage correct collapsed across angular disparity for each problem type in each condition.

Experiment 2, the diagram–line and diagram–geometry conditions showed that linear response times are not a necessary outcome of the task. This divergent evidence helps to isolate the association between linearity and analog rotation. The interpretation of linearity as a reflection of an analog strategy gains further support from a consideration of the conditions in which each of the data signatures occurred. In terms of their analytic and analog qualities, the different strategies intuitively correspond with the abstract and mimetic features of the display. The mental representation engendered by a mimetic feature could be expected to emphasize the experienced properties of actual gears. And of course, one of the key features of any perceived gear system is the rotary motion. Thus modeling the gear behaviors through an analog rotation is a reasonable approach to completing the task. Moreover, the most realistic display, pictorial–groove, led to a depiction that incorporated the dynamics of the gear movements themselves. The analog transformations were not only kinetically faithful to the motions of the gears, as in the diagram–groove and pictorial–line conditions, they were also dynamically faithful. As demonstrated by the relationship between circumlinear disparity and latency, subjects relied on the surfaces of the gears to coordinate their rotation rates.

In contrast, one can consider the effect of the abstract features. In the two diagram–geometry and diagram–groove conditions, the latencies were not linear across the full range of angles. This is what one should expect if subjects applied an analytic strategy of calculating size and angle ratios. Our interpretation is that the schematic, geometric appearance of the problems elicited an emphasis on the static, measurable attributes of the gears rather than their dynamic behaviors. Because the resulting representation emphasized circles and angles, it de-emphasized gear behaviors. Consequently, subjects tried to determine how angle and circle sizes should be related.

A second type of evidence also separates the analytic and analog approaches to the problem. Although anecdotal, the diagram–line and diagram–geometry conditions led to a striking effect in subject debriefing. Through all the experiments, subjects typically had extreme difficulty articulating their processes. In contrast, subjects in these two conditions were quite detailed in their descriptions of the heuristics and thumbnail calculations used to solve the problems. This makes sense in that the knowledge used to conduct an analog solution is largely non-verbal, whereas analytic knowledge would be of a more verbal character.

Our explanation of the differences among the five conditions is that the displays elicited different forms of knowledge which led to different solution strategies. An alternative explanation that does not involve knowledge elicitation is that the pictorial–groove display included more information that could support an accurate solution. The accurate sprockets of the pictorial displays, although irregular in shape and very small (i.e., the same size as in Fig. 1), may have helped guide the transformation. Because the number of sprockets on the power and driven gears matched for a mesh problem (unlike the prior experiments), subjects may have used these to gauge how accurately they were coordinating the two motions. This would have led to a more efficient transformation that was controlled at the gear surfaces. It was not the mimetic features of the pictorial–groove display, per se, which led to the effects. Rather, it was the sprockets specifically.

A defense against this increased information argument is that the pictorial–line condition, which had identical sprocket detail, did not show similar effects to the pictorial–groove condition. This line of defense cannot be definitive, however. For example, one might argue that lines are inherently more difficult to rotate than knobs and grooves, and thereby masked the use of the gear surfaces. As a plausible compromise prior to further empirical research, we suggest that subjects capitalized on whatever the stimulus, the experimenter, and their ability, experience, and knowledge had to offer. In all the strategies that we have identified, there have been signs of other strategies also at play. The clearest cases of multiple strategies were found for the measurement condition, the diagram–line and diagram–geometry conditions, and the groove-early problems, where it was found that subjects conducted analog rotations for some angles, and calculated or measured for others. There were no pure cases of strategy application across trials. Thus, we suggest that the mimetic features of the gears elicited knowledge about the forces that coordinated the gear motions, and the accurate rendition of the sprockets further helped to regulate the course of the transformations.

GENERAL DISCUSSION

Empirical Summary

Through three experiments, we demonstrated that people can reason about a simple physical interaction using imagery. Implicating analog imagery we found that true-positive latencies were linearly related to the angular displacement of the gears from their target orientation. We also found supporting evidence when we additionally considered the correct response latencies for the larger angles of the negative problems. The latencies for the mesh, groove-late, and groove-early problems followed the pattern one would expect if subjects made an analog rotation and reached their judgments as soon as the groove or knob entered the point of contact.

In Experiments 2 and 3, we showed that people adaptively used analog imagery under certain conditions but not others. The analog rotation pattern of results occurred when subjects were explicitly told to rotate the gears, or when the graphical display of the gears presumably suggested the use of a physical simulation. On the other hand, evidence of non-analog solutions was found when subjects were directed to use a visual comparison strategy, or when they saw a computer display that presumably suggested an analytic strategy. For these two latter cases, latencies did not increase linearly across all the angular disparities.

We developed a third line of evidence to argue that imagery processing for interacting objects cannot be accounted for solely in terms of the spatial abilities of a person or the geometric properties of their representations. To exclude the alternative, purely spatial processing accounts of task performance, we showed that spatial ability, as measured by two tests, bore a relatively weak relationship to the variation in rotation rates in comparison to the effects of strategy training. We also showed that the mental coordination of the gears' motions was not a coincidental side effect of larger images requiring longer transformation times. Finally, because subjects in Experiment 3 all had the same practice period, we can eliminate practice, or direct memory of gears turning, as a sufficient explanation of the different performances and data signatures that occurred across conditions.

In favor of using the knowledge a subject brings to bear to the problem as an account for different response patterns, we showed that giving the subject knowledge or eliciting knowledge with different graphical displays determined how the problem was solved. Even within the analog solutions, we found that a graphical display could influence the type of knowledge that was employed to solve the problem of coordination. When the subjects worked with the most realistically drawn gears, they coordinated their rotations by using the gear surfaces, just as it occurs with real gears. However, if the gears were drawn more schematically, subjects appeared to coordinate their gear images by pre-determining their appropriate angular velocities. The use of different spatial transformations for spatially isomorphic tasks is important because it makes it difficult to account for these differences solely on the basis of the problems' geometric or spatial properties. Instead, it implicates some integration of spatial and physical knowledge in the course of one's imaginations about the world. Considering all the evidence, we have constructed a reasonable case that people can reason about physical events by transforming an analog mental model of the world.

A Computer Model of Mental Depiction

In the remaining sections of the paper we develop a representational model to account for subject performance in the final experiment. We call the representation that supports and coordinates analog transformations of corresponding physical events a depictive model. Because some of the assumptions behind our account of depictive models cannot or do not yet have sufficient empirical support, we developed a computer model to test whether these assumptions produce a workable account of human cognition. To make this test, we compare the results of the computer simulations and the human performances for the pictorial–groove and the diagram–groove conditions.

The Constraint That Depictive Models Reveal Behavior

The highest level constraint on our implementation is that a depictive model must reveal physical inferences through the course of its behavior, much as an action on the world reveals behavior. To bring this constraint into relief and to see its implications vis-a-vis the integration of physical and spatial knowledge, consider how subjects in the pictorial–groove condition made the physical inference that yielded the surface coordination of the gear motions. Did this inference emerge from the modeling itself, or was it inferred beforehand and then applied to the analog rotations?

First consider the possibility that subjects made the physical inference prior to their analog rotations. For example, they may have reasoned that because both gears have rough surfaces and the surfaces are touching, friction should keep the surfaces moving in tandem. In this scenario, subjects did not model the gears to infer their physical behaviors. Instead, they derived the appropriate physical relationship and used this to make sure the image would behave accordingly. One consequence of this scenario is that spatial and physical inferences take different forms. Physical inferences occur by describing behaviors, whereas spatial inferences occur by portraying behaviors. Next consider the possibility that the physical inference emerged through the imagined gear motions. For example, the representation of the gears may have included analog, mental forces as well as an analog, mental space. As people turned one gear, the other gear naturally followed due to ''mental friction.'' In this scenario, the physical and spatial inferences take similar forms in that both resulted from imagined transformations. It is this type of representation that satisfies our primary constraint on depictive modeling. It allows physical inferences to emerge from model behavior, rather than requiring them to be been drawn before running the model.

We have chosen to develop a computer implementation of depictive models under this constraint for three related reasons. It requires fleshing out a commitment to referent-based reasoning found in both the imagery and mental model literatures. It creates an alternative to dual code models of physical inference. It explains why people use models in some situations but not others.

Depictive models are referent based. A common theme within the imagery and mental model literatures is that people's reasoning can mimic the structure of experience. For example, the analog imagery and qualitative physics literatures point to how people draw inferences by simulating the temporal flow of events. An important aspect of physical experience is that changes occur to referents. Accordingly, the imagery and mental model literatures (e.g., Finke, 1980; Holland, Holyoak, Nisbett, & Thagard, 1986; Johnson-Laird, 1981; Perky, 1910) depend on representations that are referent based in the sense that inferences arise from the form of, or changes to, object representations. An image is a rather direct perceptual representation of a referent, and a mental model is generally a model of some ''thing.'' The commitment to referent-based reasoning is not shared by all accounts of inference. For example, a dominant model of inference is that reasoning occurs through rule application (e.g., Forbus, 1988; Hegarty, Just, & Morrison, 1988; Rips, 1986). In the domain of physical inference, rule-based reasoning allows for inferences that violate a commitment to referent-based representations. Rules can operate over attributes that have been extracted from their objects. An example of this comes from a subject in the diagram–line condition who stated that he was comparing the ratio of the gears' diameters to the ratio of their angular disparities. In this case, he was comparing relations between abstracted attributes rather than reasoning by mimicking the behavior of the world. Reasoning over abstract relationships can be done without a referential model altogether (e.g., using numbers).

Depictive models employ a single ontology. Our simulation is an attempt to evaluate whether referent-based models of reasoning are sufficient to draw inferences about physical interactions, or whether it is necessary to employ rules that operate in a manner unlike the structure of perceptual experience. Previous characterizations of physical inference that included referent-based models have employed rules to draw the physical inferences that guide the model's behavior (e.g., Forbus, 1988; Kosslyn & Shwartz, 1981; however, Funt, 1980). These characterizations assume dual codes in physical inference. Kosslyn's (1980) work on imagery and memory provides the computational prototype for dual code models of inference. In his computer simulation, spatial relations are represented by placing imagined objects in a two-dimensional array (i.e., an analog of perceived space). For example, to draw an inference about the relative distance between objects, the program scans across the cells of the array from object to object, much as one would scan across an actual scene. In this respect, the program is referent based. However, the information that determines where to place objects in the array is represented separately as a list of propositions. For example, to draw an inference about where a glove should be placed relative to an arm, one would search through a list of propositions to infer the appropriate relationship. Kosslyn's (1980) dual code framework can be extrapolated to physical inferences whereby the system combines propositions to infer physical behaviors, and then maps the spatial consequences into an analog representation. The description of the global analog strategy provides one example of this form of reasoning, as does the possibility that people infer the point-for-point behavior of the gears prior to imagining their rotations in the local analog strategy.

Despite the articulation of and evidence for dual codes in memory and inference (e.g., Pavio, 1971), a problem with dual code models is that only spatial relations are represented in analog form, whereas all other forms of relation are represented in a more descriptive code. We see no a priori reason why physical knowledge must be propositional, whereas spatial knowledge can be analog. For the same reasons that one would propose an analog spatial representation based on visual imagery, one should be able to posit an analog physical representation based on haptic imagery (e.g., Clement, 1994). The computer implementation of depictive models provides an alternative formulation of physical inference that entertains this possibility.

Depictive models handle novelty. Enforcing the constraint that depictive models reveal behavior provides a potential explanation for why people construct models to reason about novel situations. Namely, modeling may be the only way that the physical knowledge embedded in a referent-based representation can emerge. There is some evidence that people employ models in novel situations. For example, we examined people's hand gestures as they solved problems involving parity relations among gears (Schwartz & Black, in press). We found that people relied on a model (i.e., depictive hand motions) when rule-based knowledge was unavailable (i.e., a parity rule). However, when people had a rule they did not model the gear behaviors. Text comprehension researchers have also posited the use of mental models to explain how people comprehend descriptions of novel situations that do not have the stereotypy necessary for a script or schema representation (Black & Bower, 1980; Collins, Brown, & Larkin, 1980; van Dijk & Kintsch, 1983). Our proposal is that people use models to draw inferences about novel situations because the imagined transformations provide access to knowledge and inferential procedures that may be otherwise unavailable.

Schwartz and colleagues (Schwartz & T. Black, 1996; Schwartz & Hegarty, in press) developed evidence that some physical inferences are best drawn depictively. In these studies, subjects were shown two glasses of identical height but of different diameter. The glasses had lines indicating identical levels of pretend water. In the ‘‘descriptive’’ conditions, subjects saw the two glasses side-by-side and decided whether they would start pouring at the same angle of tilt, and if not, which glass required a lesser tilt. The subjects were rarely correct. In the ‘‘depictive’’ conditions, subjects closed their eyes and tilted (or imagined tilting) each glass in turn, until they thought the water would just reach the lip of the glass. As in the descriptive conditions, there was no actual water in the glasses. The subjects’ tilts were almost always right; they correctly tilted the narrower glass further than the wider glass. These results suggest forms of knowledge that can only be deployed through depiction. The computer model describes how this could occur.

Additional Constraints on Depictive Models

We designed the implementation to meet three additional constraints. One constraint was that the program needed to show how depictive models could be constructed on case-by-case basis. Although we have no data bearing on this constraint and our implementation was designed post-hoc for Experiment 3, we needed an implementation that could conceivably be scaled to encompass a range of novel situations. Second, the implementation needed to provide a reasonable explanation for the effect of visual realism found in Experiment 3. The most realistic display, pictorial–groove, led to responses indicating that subjects were depictively coordinating the gear rotations at the surfaces, whereas the other conditions did not. Consequently, it is incumbent upon us to explain the relationship between pictorial realism and depictive modeling. Finally, the program needed to produce data that would determine whether it had a reasonable fit to human performance. To meet the two latter goals, we modeled psychological processes for the diagram–groove and pictorial–groove conditions. These conditions provided the most stable empirical data, as well as an effect of pictorial realism. In our simulation, we show how the local analog strategy develops from a depictive model based on the pictorial–groove stimuli, and we use the global analog strategy, based on the diagram–groove stimuli, to demonstrate a solution that requires a rule-based inference prior to the simulation.

An Object-Oriented Framework for Simulating Mental Depiction

A referent-based representation. The program was written with the Common Lisp Object System (CLOS), an object-oriented programming language (see Keene, 1989). At the highest level, object-oriented languages map onto referent-based reasoning because the primary data structure is an object and not a relationship (e.g., an entailment). All the data and procedures relevant to an object are *encapsulated* in that object. For example, our depictive simulation does not use a list that specifies the sizes and orientations of the gears. Instead, this information is part of the gear objects. This corresponds to the referent-based nature of experience in which size information is integral to the objects we see or imagine. The encapsulation of data and procedures within an object has strong implications for how one can compute relationships between objects. For example, because the program cannot represent information independently of an object, it is difficult to compute a ratio between the sizes of two gears—where would this ratio information be stored, in one gear or the other? Similarly, because procedures cannot be represented independently of objects, it is difficult to represent a procedure that could direct the rotations of both gears according to this ratio. Consequently, encapsulating data and procedures into objects helps enforce an implementation in which reasoning cannot operate ''outside'' the referent model.

Instead of representing relations independently of objects, the object-oriented architecture provides a mechanism called *message passing* for creating local relationships between objects. Assume an object passes a message. If another object has a *method* for handling this message, it invokes its own procedures in response to the message. As a simplified example, as the power gear turns, it broadcasts a message that it is creating a force. The driven gear, which has a *force method* for responding to a *force message,* catches the message and computes an appropriate response (e.g., it turns). An important property of message passing is that it exemplifies why some physical inferences might only arise through imagined object interactions. For example, consider a scenario in which only the driven gear is represented. In isolation the driven gear still has a force method that can respond to force messages. However, the physical knowledge embodied in the force method is inert because there are no other objects to generate a force message. This scheme of message-activated methods provides an analogy for understanding why physical inferences might reveal themselves through imagined object interactions, but not otherwise.

Representing space through objects. Given that relations and properties cannot be represented ''outside'' of objects, then space cannot be represented outside of objects either. Consequently, there should not be an abstract spatial structure that stands over and above objects (cf. Freyd, 1987). The lack of an abstract spatial data structure runs counter to models like Kosslyn's (1980) in which there is a static two-dimensional array that exists and has functional properties that are independent of any object filling that space. In our model, spatial properties like dimensions only become available through objects. There is some reason to accept that objects are necessary for representing visual space. For example, Gibson showed that when there are no objects filling an external space, people do not perceive a space with any functional properties (e.g., dimensions, Gibson & Waddell, 1952). According to our account of imagery and modeling, the referent is the primary representational entity rather than the spatial data structure.

To create space for an object, we defined each object locally within its own 2-$\frac{1}{2}$ dimensional (D) array (cf. Marr, 1982). In a 2-$\frac{1}{2}$ D array, vectors in the cells of a two-dimensional array indicate the orientation of an object surface. To represent a sphere as seen from a given perspective, it is like

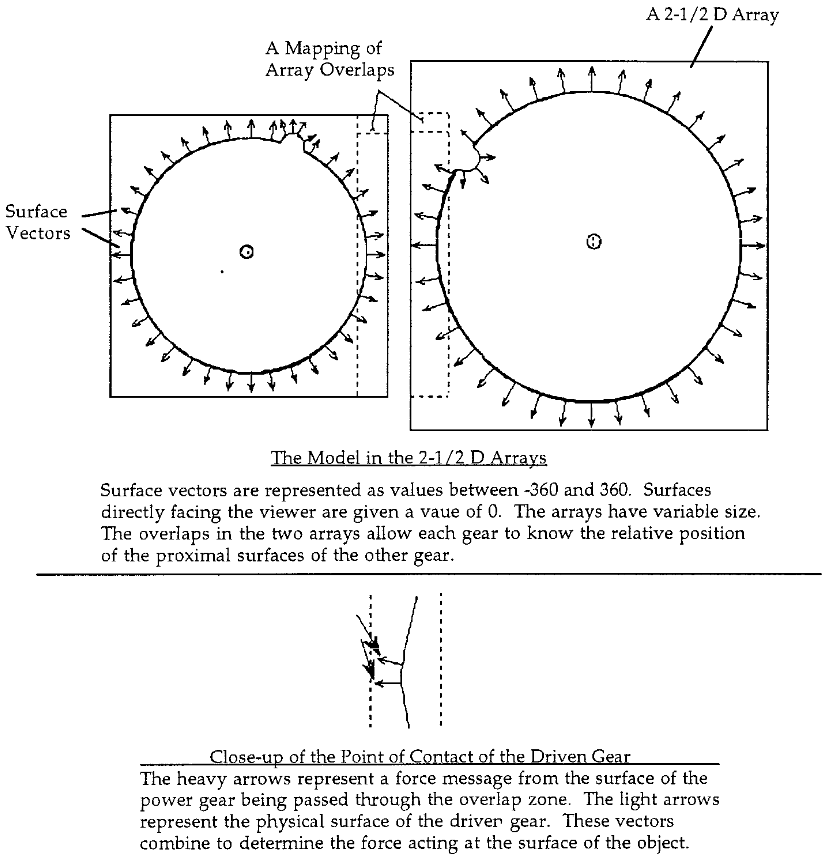

FIG. 12. A schematic of how two spatial objects communicate spatial and force information. The overlap zone in the upper half of the figure allows two object-centered representations to map their relative positions to one another. The bottom half of the figure shows that the gears communicate, or pass messages, when a change to one gear falls into the overlap zone. In the current model, these changes are represented by motion and/or force vectors that indicate the direction of the change.

writing into a grid the angles of a thousand pins stuck into the visible side of a baseball. If one represented a flat disk, each vector would point in the same direction and all the grid entries would be the same. The top half of Fig. 12 shows a graphic rendition of two 2-$\frac{1}{2}$ D arrays. The vectors pointing from the lateral surfaces of the gears indicate their orientation. If the gears were simply circles, as with the diagram–groove display, there would be no lateral surfaces and these arrows would not be included. To reduce clutter, we have omitted the face vectors.

Given that each gear has a local spatial coordinate system, we need a way that the two objects can be in a spatial relation. To enable objects in separate coordinate systems to relate spatial information to one another, we use *overlap zones.* The top half of Fig. 12 provides a visual schematic. The overlap zones are *not* defined with respect to a larger coordinate system (e.g., the page). Instead, each gear defines the position of the other gear relative to itself. For example, the power gear represents the fact that there is something within its ''right'' overlap zone, while the driven gear represents something within its ''left'' overlap zone. Consequently, if an action (message) from the left gear falls within the right gear's overlap zone, the right gear can position the message relative to itself.

Representing force between objects. The 2-$\frac{1}{2}$ D array is important for our depictive simulation because surfaces, broadly construed, are needed to model impinging forces. However, to make spatial surfaces capable of modeling force exchange, their physical properties need to be specified. For example, a purely spatial representation of a surface does not indicate whether the surface deforms, displaces, or breaks in response to a force. There are numerous ways to include physical information about a surface in a 2-$\frac{1}{2}$ D array. For example, multiple vectors within a cell could capture various surface properties such as resilience and directional surface resistance (e.g., a surface that can be easily brushed one way but not another). For the gear task, because the surfaces have uniform physical properties, we simply allow the program to use the spatial surface vectors to represent a rigid physical surface with high friction.

In addition to specifying physical properties, the surfaces need to generate forces during the movement of a gear. This way, when the model of the power gear turns, it can create a force vector indicating the direction of movement along its surface. In our object-oriented implementation, these vectors take the form of messages which can be caught by the neighboring gear if they fall within its overlap zone. The bottom half of Fig. 12 provides a visual rendition of how forces messages pass from one gear to the other in the overlap zone. Once the driven gear catches these force messages, it uses its own surface vectors to determine how it should respond. This process is described more fully in the *Narrative* section below.

Constructing novel and domain-specific models. CLOS provides a mechanism for dynamically constructing specific objects. In our program, this corresponds to the construction of a novel, domain-specific model. The model construction routines make the most specific object representation possible given the available information. This is important because the more determinate (Mani & Johnson-Laird, 1982) an object representation, the more it constrains the possible behaviors of the model. For example, if one represents a glass as brittle, one can infer its breakage upon falling to a slate floor, whereas not representing its brittleness leaves open a range of possible inferences. So, when our simulation receives information specifying physical surfaces, as in the case of the pictorial–groove input, it constructs an object that includes a representation designed to hold surface information. However, when the input does not specify surfaces, as in the case of the diagram–groove input, the program cannot construct an object that has a representation of surfaces. Consequently, the object is indeterminate with respect to its physical perimeter and cannot model the effects of interacting surfaces.

Figure 13 presents the hierarchy we use to layout the different forms of representation that can be combined into a model. Each box indicates the classes of representation that are designed to model specific types of information. For each class there are two entries. The Information column indicates the domain of information that a given class can represent. The Methods column indicates the *attached* procedures that can manipulate the information. The information and method entries are ''attached'' in that the information is inert without the methods, and the methods can only operate over that domain of information. As an analogy, numbers are only meaningful if one has methods like adding, and arithmetic methods are only meaningful if one has numbers. So, in our program, when the information to the program specifies a fixed point, the program constructs an object with fixed point information (e.g., allows turning around point X) as well as the specific procedures for computing and responding to forces around the fixed point (e.g., fixed point + force = torque). To simplify exposition, we refer to the ''class of representation'' to indicate the information and methods the program uses at a given point.

In CLOS, the program constructs a model by traversing downward through the hierarchy collecting classes of representation to fit the specific input. Our hierarchy, however, is not intended to reflect a temporal, psychological process. Although there is some evidence that spatial relations are resolved prior to physical ones (Shiffrar & Freyd, 1990), we do not intend any process ordering, and thus we do not include arrows in the hierarchy. Instead, the hierarchy is designed on the grounds of informational dependency and specificity. Classes lower in the hierarchy depend on classes higher in the hierarchy, but not vice versa. For example, our simulation can model rotations without Pivot class representations, but it cannot model pivots without Visual class representations. This is because a pivot is specific form of rotation—it requires more specificity than a rotation (e.g., when a pivot point is at the edge of an object), yet it is still requires the ability to rotate. This scheme of informational dependencies is intended to correspond to our experience of the world in which some forms of information are more prevalent and serve as prerequisites to more domain specific forms of information. For example, one needs surfaces to model rigidity, but one does not need rigidity to model surfaces. It also reflects the spirit of geometric kinematic accounts of spatial imagery (Shepard, 1994) in which a few general spatial transformations are the basis of, and constraints on, more specific transformations. It should be stressed, however, that this hierarchy of informational dependencies and these specific classes of information are not empirical claims, although it seems

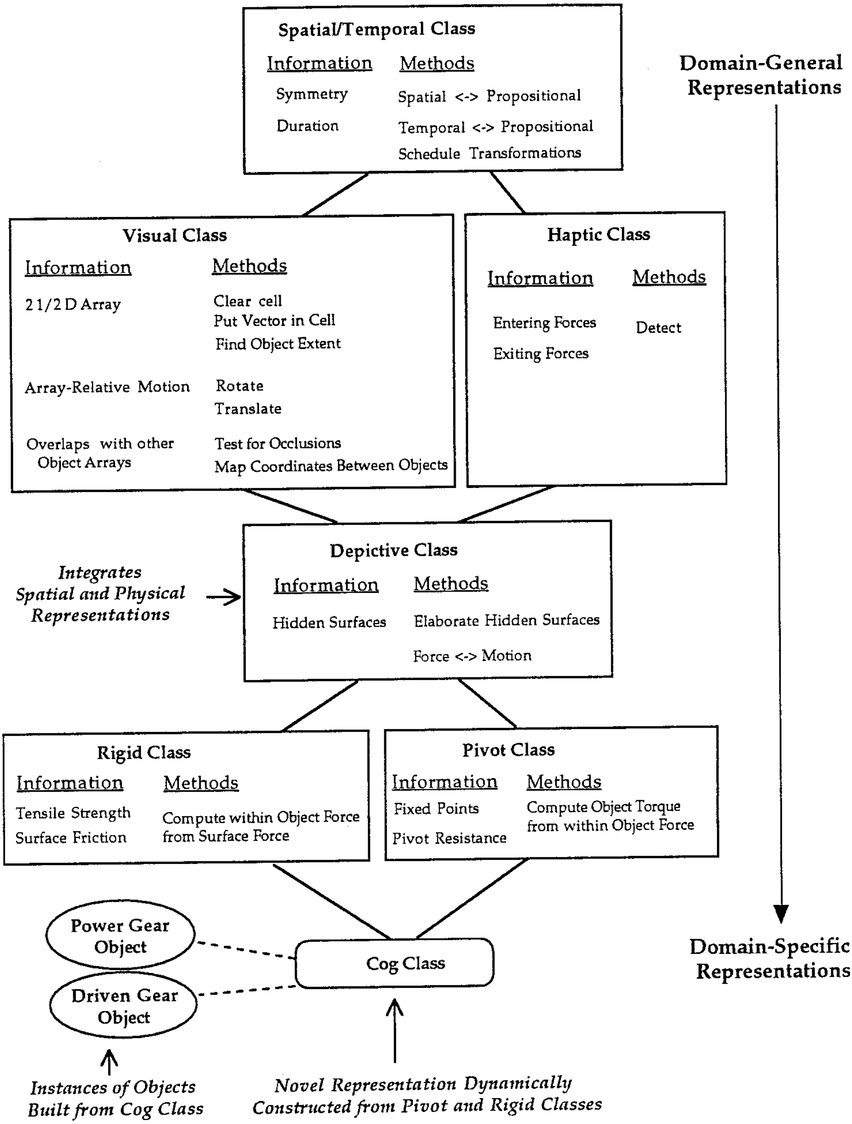

FIG. 13. The hierarchy of classes that construct a depictive model of a gear. The hierarchy reflects the different classes of representation that can be included in an object. In each box, the Information column indicates what type of information is available to the modeling process. The Methods column indicates the transformations that are attached and available to that information type. The top of the hierarchy shows the most generally available types of information and the most generally applicable methods. The bottom of the hierarchy has more situation-specific information and methods that depend on more general information. This dependency is why they are lower in the hierarchy.

possible to approach the issue empirically (e.g., diSessa, 1993). Instead they exemplify how the construction of a referent-based model could allow for spatial representation alone, as well as the integration of general spatial knowledge and specific physical knowledge.

To emphasize the distinction between, and the integration of, spatial and physical representations we made Visual, Haptic, and Depictive classes. The Visual class specifies analog spatial representations, and the Haptic class specifies the potential for receiving and sending analog forces. The Depictive class is the first level of specificity where the relationship between spatial and physical information becomes articulated. In the current scheme, the Depictive class combines and adds to the representations of the Visual and Haptic classes so that spatial motions at a surface can become physical forces and vice versa. We associate the Haptic class with forces because we presume that the projection of haptic experience to distal objects is the intuitive basis for reasoning about object forces. For the current purposes, we could have embedded the representations of the Haptic class into the Depictive class, because all forced-based reasoning in the simulation begins with surface forces. More generally, however, there may be forms of physical and spatial integration that do not involve well-defined surfaces (e.g., proprioception), and consequently, it should be possible to combine physical and spatial properties into a class besides the Depictive class.

One advantage of this form of hierarchy is that it allows the program to construct novel objects from different classes of representation, if the classes occur at the same level within the hierarchy. For example, in the simulation using the pictorial–groove input, the program dynamically constructs a new class of representation which we label the Cog class for presentation purposes. It does this because it has information specifying both a rigid object and a pivoting object. To accommodate both of these properties into a single object, it constructs a novel object class, of which the power and driven gear objects become instances. If the input specified a non-rigid, pivoting object (e.g., a jellyfish nailed to a board), theoretically the program would construct a different, new class of non-rigid, pivoting representation. Although we did not design a Non-rigid class for our hierarchy, we made the Pivot class sufficiently modular that it could combine with a Non-rigid class (or other classes at the same level in the hierarchy) as long as the Non-rigid class has access to the representations of the Depictive class. The program's combinatorial capability demonstrates how input can be constructed into a novel object representation with a novel combination of properties, rather than requiring the input to be interpreted into a pre-existing data structure (cf. Bransford, Barclay, & Franks, 1972).

Narrative of the Simulation

The simulation took inputs corresponding to the pictorial–groove and diagram–groove conditions of Experiment 3. This included the three gear combinations at all of their angular disparities. In the model construction phase for both conditions, the program received propositions that the goal was to determine whether the marks on the two rigid, high friction, pivoting gears would meet. The program also received a 2-$\frac{1}{2}$ D sketch that represented a plausible cognitive parsing of the screen display (Marr, 1982). The program separates the two gears in the sketch into separate objects by tracing their closed circumferences. The program then builds an object description for each gear. After constructing the objects, the program begins the transformation of the model. During this phase, it is sometimes necessary to violate the constraint that all behaviors result from model transformations. For example, the display of the gears is static and without implied force; therefore the rotations need to be initiated from outside the model. Similarly, the program needs a description of when to stop. To identify violations like these, we refer to the *global level* of the program. To permit the global level limited access to the model, we gave the Spatial/Temporal class the ability to translate between global level representations (i.e., propositions) and Visual class representations (i.e., objects). For example, the Spatial/Temporal class translates the propositional input, ‘‘rotate the power gear clockwise,’’ into a message sent downward to the Visual class which can perform the rotation.

A Depictive Model for the Pictorial–Groove Input

Model construction. Appendix A provides an annotated description of a small gear based on the pictorial–groove input (in practice a 2-$\frac{1}{2}$ D array had at least 1024×1024 cells). Because of the secondary importance, we do not describe how the input information is matched to the most specific classes of representation, nor how the program builds the overlap zones for each object. Instead, we focus on the unanticipated finding that the program needs to represent hidden surfaces.

In our implementation, the Depictive class relies on surfaces to catch and communicate forces. However, recall that each object is built from a 2-$\frac{1}{2}$ D sketch which shows an object from a single view point. Consequently, due to occlusion, it is impossible to ‘‘see’’ both surfaces of the gears near the point of contact. This meant that the original program did not represent both touching surfaces simultaneously, and therefore could not model interactions between the surfaces. We had to augment object construction in the Depictive class so that it could enrich the available visual information to extend a surface obscured by a second object. The possibility that mental images can represent hidden surfaces with some degree of accuracy has received indirect empirical support (Cooper, 1989) and seems reasonable given that we can imagine a ball bouncing off the back side of an open door.

Transforming the depictive model. In this sub-section we sketch the model transformations. Our intent is to fill in the picture of how a physical inference might be made using referent-based transformations without dwelling on the specific algorithms which are of minimal psychological relevance. Appendix B provides a fuller, but still simplified, annotated trace of the first steps of the transformation phase of the program. The following narrative matches the more detailed Appendix.

The transformations begin when the Spatial/Temporal class translates a global level command into a clockwise message for the Visual class of the power gear. The Visual class responds to this message by invoking its method for rotating the gear. This method shifts the surface vectors in the 2-$\frac{1}{2}$ D array one cell clockwise and updates their orientation (i.e., the knob is one step closer to the point of contact and pointing in a new direction). This spatial movement causes the Depictive class to generate force messages along the surface of the power gear. These messages indicate the direction and location of force vectors on the power gear's surface. Except for the leading edge of the knob, where the direction of motion and surface orientation are congruent, the force vectors are normal to the surface of the rotating gear. Those messages that fall into the overlap zone are detected by the Haptic and Visual classes of the driven gear. The Depictive class of the driven gear receives these messages from the Haptic and Visual class and determines whether and where the force impinges relative to its own surface. For those forces that do impinge, the Depictive class passes force messages to the Rigid class. The Rigid class computes a deflectance vector that takes into account the orientation of the affected surface and the surface's physical properties. The deflectance vector is then combined with the original force vector sent by the Depictive class to determine the direction of force on the overall object, rather than just its surface. So, because the surface is rigid and non-slippery, the object force vector indicates that the force has a ‘‘size’’ and that it is pushing downward on the left side of the gear. This vector becomes a message for the Pivot class. (If there were no Pivot class, the Rigid class would need to determine the object's response, e.g., translate, crack.) The Pivot class receives the force on the object and computes how this force interacts with the pivot point of the object. This results in a message from the Pivot class that indicates a counter-clockwise torque. The Depictive class translates this counter-clockwise torque into a counter-clockwise rotation message for the Visual class. The Visual class then makes the surface vectors of the driven gear take a single step in the counter-clockwise direction, moving the groove closer to the point of contact. As with the power gear, this rotation generates force messages along the surface of the driven gear. Under one version of the simulation, these force messages caused the power gear to turn a step, yielding a ‘‘self-propelled’’ string of transformations. Here, we simply allow these messages to serve as a ‘‘handshake’’ to the power gear. Upon receiving the handshake, the Spatial/Temporal class of the power gear initiates another rotation step.

The process of communicating forces and precipitating spatial changes to the gears continues until one of two termination conditions is reached. In one case, the gears jam when the knob hits the surface of the driven gear. The jamming arises from the same depictive routines that determine the gear motions. When a force vector from the knob enters into contact with the surface of the driven gear, it is at an angle that drives into the driven gear's surface towards the fixed point of the gear. As a result, the Pivot class cannot compute a torque. Moreover, the fixed point prevents the gear from translating, and the rigid frictional surface prevents it from deforming or slipping. Consequently, the driven gear does not respond to the knob's force vector with a movement. This lack of movement is sent back as a message to the power gear and indicates that the knob cannot move further. This ''negative'' handshake reaches the Spatial/Temporal class which determines that the transformations have halted. It subsequently sends a message to the global level which translates this result into a ''no'' response. In the other termination case, the Visual class of the driven gear tests every few steps whether the groove is close to the point of contact. In the case where it is, it sends a message to the power gear to determine the distance of the knob from the point of contact. The results of the visual inspections become messages passed to the global level. The global level then decides whether the distances are within error tolerances and produces a ''yes'' or ''no'' response accordingly.

A Non-Depictive Model for the Diagram–Groove Input

Model construction. The diagram–groove input was identical to the pictorial–groove input except that the 2-$\frac{1}{2}$ D sketch did not include surface vectors. Instead, it consisted of 0's and Nil's. A Nil indicates that there is nothing at a location, and a 0 indicates the presence of the line in the original display. This yields 2-$\frac{1}{2}$ D arrays that simply have circles of 0's. The lack of surface information has a large impact on the objects the program constructs. Because there are no surfaces that can catch or generate forces, the program is unable to make Depictive representations. Consequently, it is unable to include the Rigid and Pivot classes because they depend on surface forces for their operation. As a result, the program cannot model the physical interactions between the gears. The bottom half of Appendix A, from the Visual Class downward, provides an example of the object description (note that the array would only have 0's and Nil's).

Our assumption in the computer model was that the construction process should be constrained by the initially available information. In the diagram–groove case, the lack of specific surface information prevents the program from making an object description that can support force-based interactions between objects. This assumption is not meant to imply that people cannot provide elaboration, as is clearly the case when constructing a model from text. Perhaps verbal coaxing would have led subjects to construct a model that included surfaces. Additionally, it seems probable that the inclusion of abstract features primed the construction of alternative types of representation. For example, although it included surface information, the pictorial–line condition led to responses that were not dependent on the circumlinear velocities of the surfaces. Perhaps the lines that marked out the angles in this condition led people to construct a representation that eliminated or ignored surface information. Given these caveats, the visual features of a problem in general (Schwartz, 1995) and the surface information in particular (Farah, Rochlin, & Klein, 1994) may still play a large role in determining a problem representation. We propose that one contributing factor to the effect of visual realism found in Experiment 3 is that subjects in the diagram–groove condition did not receive sufficient surface information to model the effects of between-gear forces, whereas subjects in the pictorial–groove condition did receive this information.

Transforming the Non-Depictive Model. Because the model for the diagram–groove input does not include Depictive class representations to model forces, the simulation requires global level inferences. Appendix C provides an annotated trace of the simulation of the non-depictive model. As before, the simulation begins with a global directive to turn the power gear. And as before, the power gear rotates one step in the array. However, in this case, the rotation does not generate a set of force vectors along the object's perimeter. Instead, it only generates a set of spatial motion vectors that are without physical influence. Consequently, the driven gear cannot respond to the motion of the power gear and sends a no-motion ''handshake'' to the power gear. Similarly, the power gear does not have any way to respond to a lack of movement by the driven gear (i.e., it cannot jam). The no-motion message reaches the Spatial/Temporal class of the power gear which attempts to check whether the program has reached the termination conditions. As the marks on either gear have not reached the point of contact, the Spatial/Temporal class sends a message to the global level that it has unresolved circumstances. This precipitates a global level solution to the problem of coordination.

Reflecting our interpretation that subjects in the diagram–groove condition used a global analog strategy, we gave the global level a set of instructions for determining the relative angular velocities of the gears. The global level first determines which direction the driven gear should move given a clockwise rotation of the power gear. Here, it simply checks an answer built into the global level. This answer may be treated either as a memory or the result of an inference based on a parity rule. The global level then passes a message to the Spatial/Temporal class of each gear specifying the respective directions of rotation. The global level then asks the Spatial/Temporal class of each gear to find the gear's size. This is accomplished in two steps. The Visual class times how long it takes to scan, or ''walk,'' across the gear in the array. And, the Spatial/Temporal class turns this temporal information into a quantitative size estimate. This information is sent to the global level. Using the estimates from each object, the global level derives the appropriate relative angular velocities by comparing the sizes of the two gears. It requests the Spatial/Temporal classes to set up schedules of transformation that correspond to the derivation of the relative angular velocities. The Spatial/Temporal classes create this schedule by sending messages to the Visual representations

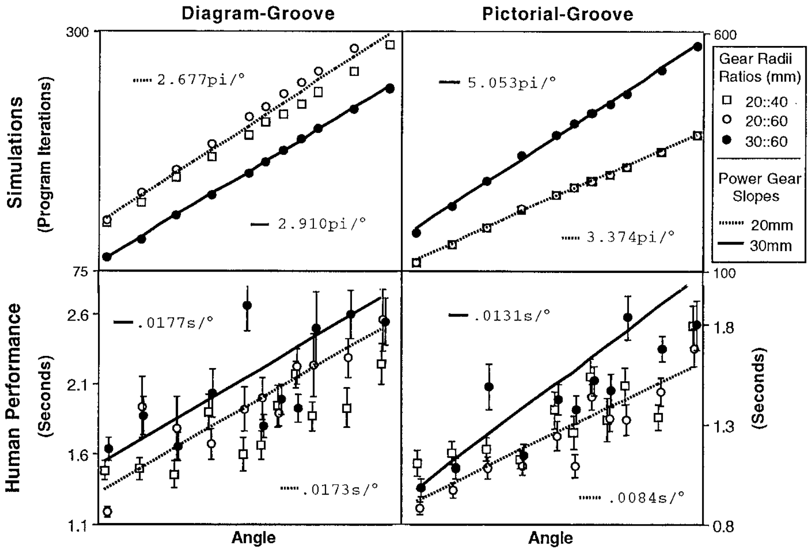

FIG. 14. Rates of rotation for hit trials from the computer simulations and human performances for the diagram–groove and pictorial–groove stimuli of Experiment 3. To facilitate comparison of the slopes, the range of the Y-axis has been scaled differently for each plot. Program iterations refer to the number of times the power gear rotated one step. The error bars represent a half standard error.

to turn a specific angular distance at regular intervals. For example, the 20 mm power gear might turn 2° for every 1° of the 40 mm driven gear. The program terminates when the Visual classes detect the presence of the knob or groove in the overlap zone and pass this information through the classes to the global level.

Results and Discussion of the Computer Simulations

The fit of the computer data to human performance. The pictorial–groove input led to depictive reasoning in the sense that the physical inferences coordinating the gear rotations arose from the transformation of a referential model. The diagram–groove input, because it did not include surface information, led to non-depictive reasoning in the sense that the problem of coordination had to be solved before the analog transformations of the gears. In this latter case, we gave the program knowledge of the global analog strategy in which the relative sizes of the gears determines their relative angular velocities.

The top of Fig. 14 plots the results of five computer runs over the full set of gear sizes and angular disparities for the diagram–groove and pictorial–groove inputs each. The Y-axis represents the number of times the power gear moved spatially, or program iterations. We varied the Y-axes to emphasize the differences in the slopes from the two simulations. As may be seen, the pictorial–groove input led to a ''rotation rate'' that was sensitive to the differences in the circumlinear disparities of the two power gears, whereas the diagram–groove input did not. The figure does not reflect the extra time the diagram–groove simulation took to solve the problem of coordination at the global level. This required ''looking up'' the counter-clockwise motion of the driven gear, estimating gear size, and deriving relative angular velocities. We have no way of interpreting these computations relative to the rotation steps. Similarly, we cannot psychologically interpret the fact that the pictorial–groove input required more program iterations per degree than the diagram–groove, as shown by the two Y-axes in the figure. This difference reflects the fact that our program took larger ''steps'' when coordinating the gear motions by their angles instead of their circumferences. This does not imply that angular solutions should be faster. In our implementation, angle-based rotations required many intermediate computations, not required of the surface-based rotations, to make the gears rotate with sufficient angular precision to yield an equivalent level of accuracy.

The bottom of Fig. 14 shows the empirical data gathered from Experiment 3. As in the plot of the computer data, we have varied the Y-axis to remove the intercept from the plots and to accentuate the slope differences. As may be seen, the results of the pictorial–groove simulations map onto the regression slopes found for the pictorial–groove subjects, and the diagram–groove simulations map onto the diagram–groove subjects. (For the diagram–groove display, the inversion of the 20 and 30 mm data between the computer model and the subjects was an artifact of how we mapped the pre-computed angles of rotation onto rectangular coordinate systems.) For the diagram–groove display, the ratio between the 20 and 30 mm slopes was 1.09 for the computer data and 1.02 for the subjects. For the pictorial–groove input, the slope ratio was 1.50 for the computer data and 1.56 for the subjects.

Given the minimal unexplained variance of the simulations (R's $> .95$),[2] an appropriate test of the fit between the computer and subject data is whether the predictions made by the computer data correlate with the human data. Because the computer simulation was not intended to capture individual differences, and because we wanted to focus on the effects of angular disparity and gear size on latency, we subtracted each individual's mean latency from his or her responses. We then correlated the predictions from both the pictorial–groove and diagram–groove simulations with the hit latencies for each group of subjects.

[2] We did not attempt to model error data extensively with the program. We experimented with some stochastic elements that caused the program to make mistakes. Interestingly, one type of stochastic variation that did not make a difference was the size of each rotational step. We found that if the rotational steps had sufficient resolution to differentiate the 20% that separated mesh and no-mesh problems, the law of large numbers (i.e., lots of little steps) removed the effect of any jitter in the rotations.

The data from the pictorial–groove subjects fit the pictorial–groove simulation ($R = .44$) better than the diagram–groove simulation ($R = .35$). The data from the diagram–groove subjects had a minimally better fit to the diagram–groove simulation ($R = .34$) than the pictorial–groove simulation ($R = .31$). Given the reasonable fit of the pictorial–groove simulation with the pictorial–groove subjects, the referent-based reasoning of the depictive simulation may provide a feasible account of simple mechanical inference.

Mental depiction for qualitative inferences. The depictive simulation suggests how people might solve a larger class of mechanical inferences, including those that only require simple qualitative inferences. As the simulation resolved the problem of coordination, the counter-clockwise behavior of the driven gear emerged from the simulation, as did the jamming behavior of the knob bumping into the groove. The gear objects in the program did not include parity or jamming rules for connected gears. Instead, the counter-clockwise and jamming behavior issued from lower level knowledge about friction, rigidity, pivot points and so forth. Our proposal is that this lower level knowledge comes together in a depictive model to generate an inference. The only way this lower level knowledge can reveal itself is through the construction and transformation of a referential representation. In other words, people model the things of the world to see what they will do.

Possible capacity constraints for future implementations. The computer model was designed to model a specific, veridical human performance, on a specific task. Although our implementation was constrained by theoretical assumptions and the empirical data, there were many implementation decisions that were unconstrained. Prior to using the simulation to guide research in a broader range of tasks, it is worth investigating what degrees of freedom can be removed from future implementations. A reasonable place to begin is to consider resource constraints. Mental depiction apparently models relatively local interactions. For example, people do not mentally animate pulley systems whole scale. Rather they animate one component at a time (Hegarty, 1993). Although we do not currently have a satisfactory operationalization of complexity, the empirical results showing a non-linear increase in difficulty past 90° for the power gear provide a clue. As a starting point, we suggest that people are only able to focus on a single confluence of causal and spatial change. This would explain the non-linear increase in problem difficulty past 90°. The motions of the two marks are not converging at the start of a rotation, because the knob is moving ‘‘up’’ and the groove is moving ‘‘down.’’ Drawing an implication of our proposed constraint of spatial and causal confluences, we would predict that rotating the marks away from the point of contact is more difficult than rotating them towards the point of contact. This is because the motions of the marks do not converge on the location of causal interaction.

Implementing a processing constraint that insists that spatial and causal changes share proximal locations and relatively congruent vectors could be done by manipulating the distance a ‘‘message’’ can travel. This way a force message at one location would not be able to cause spatial changes at a remote location. Such a constraint might provide some explanation of naive physics errors. For example, when people are shown a static diagram of a marble that is about to exit a spiral tube, they often predict that the marble will mark out a curvilinear trajectory rather than a straight one (McCloskey, 1983). One possible explanation of this error is that people are unable to depict the prolonged trajectory of the marble, because its spatial movement is away from the original force vectors that held it within the tube. Consequently, subjects may apply naive, explicit theories of object behavior, or simply recall its behavior within the tube. Perhaps if there were a way to make the forces and spatial movements in the problem more congruent and proximal, people could use a depictive model to draw a more accurate inference.

Linking depictive and non-depictive knowledge. Although people may have a fairly limited capacity for complex mechanical inferences, these limited inferences may still reflect the perceptual basis of our qualitative knowledge about the world. If the computer model is a reasonable account of the effects of visual realism, then the current results may have some practical implications. In particular, it begins to suggest a method for moving people between analytic and depictive understandings of a problem. The behaviors of novices solving physics problems are quite similar to the behavior of the subjects who saw the abstract displays of the gears. According to their debriefings, they approached the problems according to the most salient surface features (Chi, Feltovich, & Glaser, 1981). They also jumped to analytic solutions without considering the underlying forces (Larkin, McDermott, Simon, & Simon, 1980). Our computer model suggests that this type of behavior may result from a lack of visual information that can serve to constrain and guide an individual's intuitive reasoning. Although the complexity and non-experiential basis of many physics problems deny the immediate utility of mental depiction (McCloskey, Caramazza, & Green, 1980), it is possible that more realistic diagrams of problems could have some beneficial effects. In particular, a more concrete representation of a problem, like that of a picture, could engender the qualitative understanding of a depictive model. This, in turn, could anchor the articulation and refinement of connections between experiential knowledge and formal algorithms.

CONCLUSIONS

The empirical contribution of this paper was to demonstrate that people can make mechanical inferences using an imagistic simulation. The theoretical contribution was to develop a computer representation of mental depiction that may provide a bridge between the assumptions that unite mental models and imagery. This is important because both perspectives are necessary to provide an account of the current set of results. An account that relied solely on qualitative, mental model semantics could not provide a psychological description of the spatial processing in the task. On the other hand, a purely spatial description of subject performance could not explain why people solved problems that had similar spatial properties in such diverse ways.

We find that the uniting assumption of many imagery and mental model theories is that people construct representations that reveal, rather than describe, the behavior of the world. This assumption holds whether one is arguing from the point of view of perceptual equivalence in imagery research (Finke, 1980), or from the point of view of a semantic, set theoretic in mental model research (Fauconnier, 1985; Johnson-Laird, 1983). However, there is an implicit divide between the two traditions that involves claims about the universality of underlying knowledge. Imagery research has attempted to identify universal properties of our spatial world and knowledge. In contrast, mental model research has often shown that a universal, syntactic description of human reasoning cannot apply across multiple domains. In the theory of mental depiction, we suggest that the primary conceptual entity is the referent. This move can bridge the gap between some aspects of the two traditions, because spatial and domain-specific knowledge have the same referent-based ontology, and both are contingent on the type of object that is being considered.

REFERENCES

Ajzen, I. (1977). Intuitive theories of events and the effects of base-rate information on prediction. *Journal of Personality and Social Psychology,* **35,** 303–314.

Anderson, J. R. (1983). *The architecture of cognition.* Cambridge, MA: Cambridge University Press.

Bassock, M. (1990). Transfer of domain-specific problem-solving procedures. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **16,** 522–533.

Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology,* **14,** 143–177.

Black, J. B., & Bower, G. H. (1980). Story understanding as problem solving. *Poetics,* **9,** 223–250.

Black, J. B., Schwartz, D. L., & Marks, L. (1988). *What does it mean to know how something works?* Proceedings of the Annual Meeting of the Psychonomic Society, Chicago, IL.

Bransford, J. D., Barclay, J. R., & Franks, J. J. (1972). Sentence memory: A constructive versus interpretive approach. *Cognitive Psychology,* **2,** 193–209.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science,* **5,** 121–152.

Clement, J. (1994). Use of physical intuition and imagistic simulation in expert problem solving. In D. Tirosh (Ed.), *Implicit and explicit knowledge* (pp. 204–242). Hillsdale, NJ: Erlbaum.

Cliff, N. (1987). *Analyzing multivariate data.* New York: Harcourt Brace Jovanovich.

Collins, A., Brown, J. S., & Larkin, K. M. (1980). Inference in text understanding. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension: Perspectives from cognitive psychology, linguistics, artificial intelligence, and education* (pp. 385–407). Hillsdale, NJ: Erlbaum.

Cooper, L. A. (1989). Mental models of the structure of visual objects. In Soledad & D. Schacter (Eds.), *Object perception: Structure and process* (pp. 91–119). Hillsdale, NJ: Erlbaum.

Cooper, L. A., & Podgorny, P. (1976). Mental transformations and visual comparison process: Effects of complexity and similarity. *Journal of Experimental Psychology: Human Perception and Performance,* **2,** 503–514.

de Kleer, J., & Brown, J. S. (1984). A qualitative physics based on confluences. *Artificial Intelligence,* **24,** 7–83.

DeLoache, J. S., & Marzolf, D. P. (1992). When a picture is not worth a thousand words: Young children's understanding of pictures and models. *Cognitive Development,* **7,** 317–329.

diSessa, A. A. (1993). Toward an epistemology of physics. *Cognition and Instruction,* **10,** 105–225.

Egan, D. E., & Grimes-Farrow, D. D. (1982). Differences in mental representations spontaneously adopted for reasoning. *Memory & Cognition,* **10,** 297–307.

Ekstrom, R. B., French, J. W., & Harman, H. H. (1976). *Kit of factor-referenced tests.* Princeton, NJ: Educational Testing Service.

Farah, M. J., Rochlin, R., & Klein, K. L. (1994). Orientation invariance and geometric primitives in shape recognition. *Cognitive Science,* **18,** 325–344.

Fauconnier, G. (1985). *Mental spaces: Aspects of meaning construction in natural language.* Cambridge, MA: The MIT Press.

Finke, R. A. (1980). Levels of equivalence in imagery and perception. *Psychological Review,* **87,** 113–132.

Forbus, K. D. (1988). Qualitative physics: Past, present and future. In H. Shrobe (Ed.), *Exploring artificial intelligence* (pp. 239–296). San Mateo, CA: Morgan Kaufman.

Forbus, K. D., Nielsen, P., & Faltings, B. (1990). Qualitative kinematics: A framework. In D. S. Weld and J. de Kleer (Eds.), *Readings in qualitative reasoning about physical systems* (pp. 562–567). San Mateo, CA: Morgan Kaufman.

Foster, D. H. (1975). Visual apparent motion and some preferred paths in rotation group SO(3). *Biological Cybernetics,* **18,** 81–89.

Freyd, J. J. (1987). Dynamic mental representations. *Psychological Review,* **94,** 427–438.

Freyd, J. J., Pantzer, T. M., & Cheng, J. L. (1988). Representing statics as forces in equilibrium. *Journal of Experimental Psychology: General,* **117,** 395–407.

Funt, B. V. (1980). Problem-solving with diagrammatic representations. *Artificial Intelligence,* **13,** 201–230.

Garnham, A. (1981). Mental models as representations of text. *Memory & Cognition,* **9,** 560–565.

Gentner, D., & Gentner, D. R. (1983). Flowing waters or teeming crowds: Mental models of electricity. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 99–130). Hillsdale, NJ: Erlbaum.

Gibson, J. J., & Waddell, D. (1952). Homogeneous retinal stimulation and visual perception. *American Journal of Psychology,* **65,** 263–270.

Glenberg, A. M., Meyer, M., & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and Language,* **26,** 69–83.

Greeno, J. (1983). Conceptual entities. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 227–252). Hillsdale, NJ: Erlbaum.

Hayes, P. J. (1979). The naive physics manifesto. In D. Michie (Ed.), *Expert systems in the micro-electronic age.* Edinburgh, Scotland: Edinburgh University Press.

Hegarty, M. (1992). Mental animation: Inferring motion from static displays of mechanical systems. *Journal of Experimental Psychology: Learning, Memory, & Cognition,* **18,** 1084–1102.

Hegarty, M. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory and Language,* **32,** 717–742.

Hegarty, M., Just, M. A., & Morrison, I. R. (1988). Mental models of mechanical systems: Individual differences in qualitative and quantitative reasoning. *Cognitive Psychology,* **20,** 191–236.

Hinsley, D., Hayes, J. R., & Simon, H. A. (1977). From words to equations. In P. Carpenter and M. Just (Eds.), *Cognitive processes in comprehension.* Hillsdale, NJ: Erlbaum.

Hochberg, J., & Gellman, L. (1977). The effect of landmark features on mental rotation times. *Memory & Cognition,* **5,** 23–26.

Holland, J. H., Holyoak, K. J., Nisbett, R. E., & Thagard, P. R. (1986). *Induction: Processes of inference, learning and discovery.* Cambridge, MA: The MIT Press.

Johnson-Laird, P. N. (1981). Mental models of meaning. In A. K. Joshi, B. L. Webber, & I. A. Sagg (Eds.), *Elements of discourse understanding* (pp. 106–126). Cambridge, MA: Cambridge University Press.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness.* Cambridge, MA: Cambridge University Press.

Johnson-Laird, P. N., Legrenzi, P., & Legrenzi, M. S. (1972). Reasoning and a sense of reality. *British Journal of Psychology,* **63,** 305–400.

Just, M. A., & Carpenter, P. A. (1985). Cognitive coordinate systems: Accounts of mental rotation and individual differences in spatial ability. *Psychological Review,* **92,** 137–172.

Keene, S. E. (1989). *Object-oriented programming in Common Lisp: A programmer's guide to CLOS.* Reading, MA: Addison-Wesley.

Kosslyn, S. M. (1976). Can imagery be distinguished from other forms of internal representation? Evidence from studies of information retrieval times. *Memory & Cognition,* **4,** 291–297.

Kosslyn, S. M. (1980). *Image and mind.* Cambridge, MA: Harvard University Press.

Kosslyn, S. M., & Shwartz, S. P. (1981). Empirical constraints on theories of visual mental imagery. In J. Long & A. Baddeley (Eds.), *Attention and performance IX* (pp. 241–260). Hillsdale, NJ: Erlbaum.

Larkin, J. H., McDermott, J., Simon, D., & Simon, H. A. (1980). Expert and novice performance in solving physics problems. *Science* **208,** 1335–1342.

Leyton, M. (1989). Inferring causal history from shape. *Cognitive Science,* **13,** 357–387.

Lorch, R. F., Jr., & Myers, J. L. (1990). Regression analyses of repeated measures data in cognitive research. *Journal of Experimental Psychology: Learning, Memory & Cognition,* **16,** 149–157.

Mandler, J. M., & Parker, R. E. (1976). Memory for descriptive and spatial information in complex pictures. *Journal of Experimental Psychology: Human Learning and Memory,* **2,** 38–48.

Mandler, J. M., & Stein, N. L. (1974). Recall and recognition of pictures by children as a function of organization and distractor similarity. *Journal of Experimental Psychology,* **102,** 657–669.

Mani, K., & Johnson-Laird, P. N. (1982). The mental representation of spatial descriptions. *Memory & Cognition,* **10,** 181–187.

Marmor, G. S., & Zaback, L. A. (1976). Mental rotation by the blind: Does mental rotation depend of visual imagery? *Journal of Experimental Psychology: Human Perception and Performance,* **2,** 515–521.

Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information.* San Francisco, CA: W. H. Freeman & Co.

McCloskey, M. (1983). Naive theories of motion. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 299–324). Hillsdale, NJ: Erlbaum.

McCloskey, M., Caramazza, A., & Green, B. (1980). Curvilinear motion in the absence of external forces: Naive beliefs about the motion of objects. *Science,* **210,** 1139–1141.

McNamara, T. P. (1986). Mental representations of spatial representations. *Cognitive Psychology,* **18,** 87–121.

Moore, J. L., & Schwartz, D. L. (1994). Mental models for proportional reasoning. In A. Ram & K. Eiselt (Eds.), *Proceedings of the Annual Meeting of the Cognitive Science Society* (pp. 640–645). Hillsdale, NJ: Erlbaum.

Morrow, D. G., Greenspan, S. L., & Bower, G. H. (1987). Accessibility and situation models in narrative comprehension. *Journal of Memory and Language,* **26,** 165–187.

Norman, D. (1983). Some observations on mental models. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 7–14). Hillsdale, NJ: Erlbaum.

Parsons, L. M. (1987). Imagined spatial transformations of one's hand and feet. *Cognitive Psychology,* **19,** 178–241.

Parsons, L. M. (1994). Temporal and kinematic properties of motor behavior reflected in mentally simulated action. *Journal of Experimental Psychology: Human Perception and Performance,* **20,** 709–730.

Pavio, A. (1971). *Imagery and verbal processes.* New York: Holt, Rinehart and Winston.

Perky, C. W. (1910). An experimental study of imagination. *American Journal of Psychology,* **21,** 422–452.

Pinker, S., & Finke, R. A. (1980). Emergence of two-dimensional patterns in images rotated in depth. *Journal of Experimental Psychology: Human Perception and Performance,* **6,** 244–264.

Proffitt, D. R., Kaiser, M. K., & Whelan, S. M. (1990). Understanding wheel dynamics. *Cognitive Psychology,* **22,** 342–373.

Rips, L. J. (1986). Mental muddles. In H. Brand & R. Hernish (Eds.), *The representation of knowledge and belief* (pp. 258–286). Tucson, AZ: University of Arizona Press.

Rumelhart, D. E. (1989). Toward a microstructural account of human reasoning. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 298–312). Cambridge, MA: Cambridge University Press.

Schwartz, D. L. (1995). Reasoning about the referents of a picture versus reasoning about the picture as the referent: An effect of visual realism. *Memory & Cognition,* **23,** 709–722.

Schwartz, D. L., & Black, J. B. Shuttling between depictive models and rules: Induction and fallback. *Cognitive Science,* in press.

Schwartz, D. L., & Black, T. (1996). *Situated physical inference: Depictive knowing versus descriptive beliefs.* Manuscript submitted for publication.

Schwartz, D. L., & Hegarty, M. (in press). Coordinating multiple representations for reasoning about mechanical devices. To appear in *Cognitive and computational models of spatial representation: Papers for the 1996 Symposium.* Stanford, CA: AAAI Press.

Shepard, R. N. (1968/1986). On turning something over in one's mind. In R. Shepard & L. A. Cooper (Eds.), *Mental images and their transformation* (pp. 19–23). Cambridge, MA: The MIT Press.

Shepard, R. N. (1994). Perceptual-cognitive universals as reflections of the world. *Psychonomic Bulletin & Review,* **1,** 2–28.

Shepard, R. N., & Cooper, L. A. (Eds.) (1986a). *Mental images and their transformations.* Cambridge, MA: The MIT Press.

Shepard, R. N., & Cooper, L. A. (1986b). Apparent motion. In R. N. Shepard & L. A. Cooper (Eds.), *Mental images and their transformations* (pp. 303–324). Cambridge, MA: The MIT Press.

Shepard, R. N., & Feng, C. (1972). Chronometric study of mental paper folding. *Cognitive Psychology,* **3,** 228–243.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three dimensional objects. *Science,* **171,** 701–703.

Shiffrar, M. (1994). When what meets where. *Current Directions in Psychological Science,* **3,** 96–100.

Shiffrar, M., & Freyd, J. J. (1990). Apparent motion of the human body. *Psychological Science,* **1,** 257–264.

Thorndyke, P. W. (1981). Distance estimation from cognitive maps. *Cognitive Psychology,* **13,** 526–550.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension.* New York: Academic Press.

Woodworth, R. S., & Schlosberg, H. (1965). *Experimental Psychology* (pp. 194–195). New York: Holt, Rinehart & Winston.

(Received December 4, 1992)
(Revision received January 9, 1995)

APPENDIX A

Object Information Structure

This print-out represents a small depictive model of a gear. One can determine the equivalent non-depictive model by reading from the Visual Class downward and replacing the vector entries with 0's. In the current implementation, we used slots to represent motion, friction and tensile strength. When the gear turns, these slots determine the types of vectors that are generated along the gear surface. In a larger scale implementation that worked over less homogeneous objects, we would represent motion, friction and rigidity with additional, permanent vectors in each cell of the 2-$\frac{1}{2}$ D array. We do not include the methods that are attached to each class of information for the sake of clarity.

Computer Print-Out of a Depictive, Power Gear Object	Added Comments
#<COG #x5F27E1>	Object Handle
Class: #<STANDARD-CLASS COG>	Class made by program from Rigid and Pivot Classes
Instance slots	
	Rigid Class Information
`TENSILE: STRONG`	Holds surface→object vector
`FRICTION: HIGH`	
`SURFACEFORCES: NIL`	
	Pivot Class Information
`PIVOTPOINT: (5 6)`	Holds object torque
`RESISTANCE: NONE`	
`AXISFORCES: NIL`	
	Depictive Class Information
`HITSPOT: NIL`	Position of force
	Haptic Class Information
`HITVEC: NIL`	Holds entering/exiting force

Computer Print-Out of a Depictive, Power Gear Object	Added Comments
The description of a Non-Depictive Object only includes the information below.	
	Visual Class Information
OVERLAPS: (((11 10 9) (0 11)) ((3 2 1) (0 11)))	Mappings into OVERLAPOBJ
OVERLAPOBJ: #<COG #x5F2871>	
ARRAY: #2A ((NIL NIL NIL NIL NIL NIL NIL NIL NIL NIL NIL NIL)	A small 2-1/2 D Sketch
(NIL NIL NIL NIL NIL 360 NIL NIL NIL NIL NIL NIL)	Blip surface pointing ‘‘right’’
(NIL NIL NIL 119 112 97 82 68 NIL NIL NIL NIL)	
(NIL NIL 141 128 0 0 0 61 51 NIL NIL NIL)	Numbers indicate vectors on
(NIL 158 151 0 0 0 0 0 38 28 NIL NIL)	360 degrees. This corresponds
(NIL 178 0 0 0 0 0 0 0 22 NIL NIL)	to surface direction. Numbers
(NIL 187 0 0 0 0 0 0 0 360 NIL NIL)	are rounded for presentation.
(NIL 202 0 0 0 0 0 0 0 353 NIL NIL)	
(NIL 209 219 0 0 0 0 0 331 338 NIL NIL)	Non-Depictive Object array would
(NIL NIL 231 241 0 0 0 308 321 NIL NIL NIL)	only hold 0’s and NIL’s.
(NIL NIL NIL 248 263 277 292 299 NIL NIL NIL NIL)	
(NIL NIL NIL NIL NIL NIL NIL NIL NIL NIL NIL NIL))	
MOTION: NIL	Holds current object motion
	Spatial/Temporal Class Information
SYMMETRY: RADIAL	
DURATION: TO-ALIGNMENT	Terminal goal state
SCHEDULE: NIL	Holds temporal sequences

APPENDIX B

An Annotated and Simplified Trace of the First Steps in a Depictive Rotation

The InfoClass column indicates which level within the class hierarchy is sending the message or information. The Message column indicates the information that is being output by the current object. For example, in the third row, a Visual class method outputs a message indicating that a motion vector with an orientation of 270° is located at point 918, 460 on the power gear. The ToClass column indicates which InfoClass catches the message. The program always attempts to have the lowest, or most informationally complete, level in the class hierarchy catch the message. Thus, in the third row, although the Visual class could catch the message, the Depict class is more informationally specific—it operates with motions that may involve forces. If there were no Depict class methods, the message would move up to the next level of generality. The ToMethod column specifies which method (or transformation) within the ToClass catches the message. The results of the method in the ToMethod column become the messages shown in the next row.

In the preceding example that used the third row, all the messages were being passed within a single object. In the trace shown below, messages cross between objects as well. For example, in the fifth row the driven gear catches a message from the power gear. In the simulation, this occurs when a message from one object falls within the purview (overlap zone) of another object, and it has not been previously caught by its own methods.

InfoClass	Message	ToClass	ToMethod	Comments
Global	Clockwise (power) Fit (mark (power, driven))	Spatial	Scheduler	Global level tells the scheduler how to start the simulation and how to terminate the simulation.
Spatial	Clockwise (power)	Visual	Mover	Spatial class converts the clockwise assertion into a spatial command that is sent to the motion methods of the power gear.
Visual	270° at X = 918, Y = 460	Depict	Motion	The depictive method converts the motion into a force that will be meaningful to the driven gear. For simplicity we present the direction and location of a single point of motion. In the model, there are numerous motion messages occurring in parallel.
Depict	Force (270°) force (X = 918, Y = 461)	Haptic Depict	InForce LocateForce	The depictive methods send the direction and location of the surface force into the overlap zone. The driven gear's haptic method ''feels'' the direction of the force and a depictive method locates the surface position.
Haptic	Force (270°)	Depict	TakeForce	The haptic sense of the force becomes spatially defined when passed to the depictive methods.
Depict	Force (270°, X = 2, Y = 461)	Rigid	Force	The force and location are mapped to the driven gear's coordinate system before being passed to the rigid methods.
Rigid	ForceIn (270°) Deflection (180.2°)	Rigid	SurfaceForce	The rigid methods turn the surface orientation of the contact point into a deflection potential. It passes this and original force in to another rigid class method that determines the resulting force on the object.

Rigid	ObjForce (315°, X = 2, Y = 461)	Pivot	Force	The force on the object and its location go to the pivot class methods
Pivot	ObjForce (315°) FixPointForce (−180.4°)	Pivot	TorqueForce	The pivot method uses the vector from the pivot point to the surface point to compute object torque. The program only determines the direction of the torque. There is currently no provision for forces of differing strength.
Pivot	Force (Counter)	Depict	InsideForce	The counter-clock torque within the object goes to the depictive methods.
Depict	Counter	Visual	Mover	The depictive method converts torque into a motion for the visual methods.
Visual	270° at X = 2, Y = 461	Depict	Motion	The resulting surface motion is translated into a force in the overlap zone.
Depict	Force (270°) Force (X = 2, Y = 461)	Depict	Confirm	The motion in the overlap zone serves as a handshake to the original power gear depictive method that sent the motion message. Had the driven gear been unable to turn, as in the case when the blip bumps into its surface, it would send a message that it could not move. This would cause the depictive method of the power gear to interpret this as a jam.
Depict	OK	Spatial	Scheduler	The handshake is sent on to the scheduler. At this point the scheduler could request a test from the visual class to see if the marks on the objects are in alignment. This occurs at random intervals.
Spatial	Clockwise (power)	Visual	Mover	In this instance, the scheduler precipitates the next step in the animation.

APPENDIX C

An Annotated and Simplified Trace of the First Steps in a Non-Depictive Rotation

InfoClass	Message	ToClass	ToMethod	Comments
Global	Cockwise (power) Fit (mark (power, driven))	Spatial	Scheduler	Global level tells the scheduler how to start the simulation and how to terminate the simulation.
Spatial	Clockwise (power)	Visual	Mover	Spatial class converts the clockwise assertion into a spatial command that is sent to the motion methods of the power gear.
Visual	270° at X = 918, Y = 460	Visual	InMotion	The visual method of the driven gear detects the motion in the overlap zone.
Visual	No-motion	Spatial	Confirmation	The driven gear has no way to respond. The No-Motion message is caught by the scheduler. Without the depictive methods, the power gear, like the driven gear, cannot act on motion messages from other objects.
Spatial	Test-fit (power, driven)	Visual (power) Visual (drive)	MarkInZone MarkInZone	The scheduler requests each object to test whether their marks have entered into the overlap zone as a preliminary to seeing if it is a mesh or no-mesh. Because neither mark is in the overlap zone, the spatial class will issue a help message that is picked up by the global level.
Spatial	Help	Global	Breakdown	The breakdown message initiates a propositional attempt to resolve the problem of coordination.
Global	Look up gear motions	Memory	Search	The global level recalls the turning direction of each gear from the training period. In the current implementation, this is simply a global variable.
Global	Clockwise (power) Counterclock (driven)	Spatial Spatial	Scheduler Scheduler	Tell the scheduler the direction of each gear.

Global	GetSize (power)	Spatial	GetMeasure	The global level requests the size of the power gear from the spatial class.
Spatial	GetExtent	Visual	ScanExtent	The spatial class requests a scan of the power gear's extent.
Visual	150 steps	Spatial	Measure	The visual method returns an indicator of size. In this case, it is the number of steps it took to cross the object. This could just as well be a time-based measure (i.e., how long it takes the visual class to return a message of having completed the scan).
Spatial	150 distance units	Global	StoreMeasure	The spatial class converts the analog measure into an attribute measurement. This value is stored by the global level.
. . . . . .	. . . . . .	. . . . .	. . . . .	The same process is repeated for the driven gear. In this case, it returns a
Global	300 distance units	. . . . .	StoreMeasure	value of 300 distance units.
Global	Compute 150/300	Global	StoreRatio	The global level computes and stores the size ratio of 1:2.
Global	Driven angle = 1/2 power angle	Spatial	Scheduler	The global level tells the scheduler to make the driven gear turn at 1/2 the rate of the power gear.
Spatial	1 angular unit (power gear)	Visual	Mover	The scheduler sets up a schedule of 1 angular unit for the power gear and 1/2 angular unit for the driven gear.
Spatial	1/2 angular unit (driven gear)	Visual	Mover	This process continues until termination conditions are met.