

Running Head: SPATIAL REASONING WITH EXTERNAL VISUALIZATIONS

Spatial Reasoning with External Visualizations:

What Matters is What You See, not Whether You Interact

Madeleine Keehner<sup>1,2</sup>, Mary Hegarty<sup>1</sup>, Cheryl Cohen<sup>1</sup>, Peter Khooshabeh<sup>1</sup>,  
and Daniel R. Montello<sup>1</sup>

<sup>1</sup>University of California, Santa Barbara

<sup>2</sup>University of Dundee

Address for correspondence:

Dr Madeleine Keehner

School of Psychology

University of Dundee

Dundee

DD1 4HN

UK

Email: [M.M.Keehner@Dundee.ac.uk](mailto:M.M.Keehner@Dundee.ac.uk)

Tel: +44 (0)1382 344683; Fax: +44 (0)1382 229993

## Abstract

In three experiments, we examined effects of interactive visualizations and spatial ability on a task requiring participants to infer and draw cross sections of a three dimensional (3-D) object. We manipulated whether participants could interactively control a virtual 3-D visualization of the object while performing the task, and compared participants who were allowed interactive control of the visualization to those who were not allowed control. In Experiment 1, interactivity produced better performance than passive viewing, but the advantage of interactivity disappeared in Experiment 2 when we equalized visual input for the two conditions in a yoked design. In Experiments 2 and 3, differences in how interactive participants manipulated the visualization were large and related to performance. In Experiment 3, non-interactive participants who watched optimal movements of the display performed as well as interactive participants who manipulated the visualization effectively and better than interactive participants who manipulated the visualization ineffectively. Spatial ability made an independent contribution to performance on the spatial reasoning task, but did not predict patterns of interactive behavior. These experiments indicate that providing participants with active control of a computer visualization does not necessarily enhance task performance, whereas seeing the most task-relevant information does, and this is true regardless of whether the task-relevant information is obtained actively or passively.

Keywords: psychology, distributed cognition, situated cognition, human-computer interaction, human experimentation, metacognition, individual differences, interactivity, active passive, spatial ability.

Spatial Reasoning with External Visualizations: What Matters is What You See, not  
Whether You Interact

Imagine that a scientist examining multidimensional data creates a three-dimensional (3-D) plot of the data on a computer screen and rotates this plot to reveal patterns in the data set. Imagine that a schoolchild learning the geography of the United States “flies” across the country using Google Earth software. Imagine that a surgeon who is about to remove a tumor from a patient’s liver interacts with a 3-D computer visualization of the abdomen in order to plan how to navigate to the site of the tumor. In each of these cases, a person is using an external visualization (a visual-spatial representation presented on a computer screen) to aid in some internal computation, such as comprehension, learning or planning.

With developments in computer graphics and widespread availability of computers, computerized representations of 3-D structures have become prevalent in fields such as engineering, architecture, science, and medicine. Although these three-dimensional models are projected on the two dimensions of a computer screen, rotating these visualizations provides powerful depth cues, creating the impression of viewing a 3-D object. When they are controlled by an intuitive interface, the user has the capability to match what is viewed on the screen to his or her momentary intentions so that the interface actions and the results of these actions are tightly coupled with internal cognitive processes. As a result, much has been made of the potential of these visualizations to augment cognition and to allow effortful cognitive processes to be offloaded onto less effortful perceptual-motor processes of interacting with the external visualization and observing the result (Card, MacKinlay, & Shneiderman, 1999; Gordin & Pea, 1995; Kirsh, 1997). But are all individuals equally able

to use external visualizations to augment internal cognition? If a person does not use an interactive visualization effectively, might a non-interactive visualization, showing them the most task relevant information, be more effective?

*Previous findings on interactivity.* The results of previous controlled studies of the effects of interactivity in visual-spatial tasks are mixed. Some studies, on visual object recognition (Harman, Humphrey, & Goodale, 1999; James et al., 2002), learning to tie knots (Schwan & Riempp, 2004), and acquiring spatial knowledge of a virtual environment (Peruch, Vercher, & Gauthier, 1995) have found significant advantages of interactivity. By contrast, other studies on inferring structure in three-dimensional data (Marchak & Marchak, 1991) and navigating desktop and immersive virtual environments (Foreman, Sandamas, & Newson, 2004; Melanson, Kelso, & Bowman, 2002; Wilson, 1999; Wilson, Foreman, Gillett, & Stanton, 1997) have found no difference between active and passive participants. Finally, in studies on searching for structure in three-dimensional data (Marchak & Zulager, 1992) and tactile maze learning (Richardson, Wuillemin, & MacKintosh, 1981), participants who were given active control were found to perform worse than passive participants. Even within a single study different comparisons of active and passive conditions have sometimes produced apparently contradictory results (Attree et al., 1996; Christou & Bühlhoff, 1999; Wilson & Peruch, 2002).

There are several possible reasons for the inconsistent results in previous research. First, in many previous studies the provision of interactive control has been confounded with the specific visual information that a user receives. Because interactive users can manipulate the system but non-interactive users cannot, the visual information available to interactive and non-interactive participants is not equivalent, unless a yoked design is used

to match the information presented in the two conditions. Moreover, any one interactive participant may receive quite different information from any other, because what they see varies according to how they manipulate the representation. By contrast, in many studies there is no such variability in the information available to non-interactive participants.

Second, provision of an interactive visualization does not guarantee that users will discover the *most effective* way to manipulate it in order to accomplish a task. Importantly, the *quality* of the information that users gain depends not just on *whether* they are permitted to interact with an external visualization, but on *how* they interact with it, and this presumably varies among different individuals and in different studies. Using an external visualization effectively depends on understanding of the task and of how the visualization can be used productively to accomplish that task, which we will define as metacognitive understanding. Previous research suggests that such metacognitive understanding is often lacking (Hegarty, 2004; Lowe, 1999, 2004; Rieber, Tzeng, & Tribble, 2004). Thus Betrancourt (2005) has hypothesized that only more experienced learners are likely to benefit from certain types of interactivity. Interestingly, variability in how people use external representations has received little attention in the literature to date.

Third, different studies have used systems that vary in the type or level of interactivity available. Krygier, Reeves, Cupp, and Di Biase (1997) distinguish between resources that are *static* (e.g., images, maps, diagrams, graphs), *animated* (express change or motion when activated), *sequential* (present information in a predetermined linear sequence), *hierarchical* (allow a non-linear exploration of embedded information or nested concepts), and *conditional* (which respond directly to the user's manipulations).

Betrancourt (2005) distinguishes between low-level control over pace and direction of the

presentation (play, pause, rewind, etc), and more sophisticated capabilities such as altering parameters in a simulation, or changing the viewpoint to allow exploration from different perspectives. These different levels of interactivity may be more or less effective depending on the characteristics of the task and the user.

*Theoretical accounts of interactivity.* Current theories view interactive behavior within a theoretical framework of distributed cognition (Hutchins, 1995; Zhang & Norman, 1994), which argues that cognitive processes occur both *internally* (in the mind) and *externally* (in the world, within some external medium). In this view, external representations are not merely peripheral aids to cognition; they intersect with internal representations to form a distributed representational space for solving a problem. External representations are an obligatory component of the representational space of a distributed task, and their inherent properties affect how we interact with them. They anchor and structure cognitive behavior within an “action space” that constrains the range of possible behaviors, and they can change the very nature of a task, as different external representations can mean that more or less of the task load is carried out internally (Zhang and Norman, 1994).

In distributed tasks, performance involves a tradeoff between use of internal cognitive resources such as working memory, and perceptual-motor processes involved in interacting with external displays and observing the results. Some models of embodied cognition, which we refer to as minimum memory models, assume that when possible, people minimize the reliance on internal memory or other internal cognitive processes and offload cognitive processes onto perceptual-motor processes (e.g., Kirsh & Maglio, 1994; Ballard, Hayhoe, Pook & Rao, 1997; Wilson, 2002; Zhang & Norman, 1994). Other

models, known as soft-constraints models, argue that perceptual-motor processes are not necessarily preferred over cognitive processes because “the human information processing system is indifferent to the source of its information” (Gray, Sims, Fu and Schoelles, 2006, p.18). In this view, the cognitive-perceptual-motor system integrates “knowledge in the head” with “knowledge in the world” for the most efficient cost-benefit tradeoffs, and interactive routines are selected to achieve benefits with minimum costs. Cognitive, perceptual, and motor resources are allocated flexibly depending on relative utility, which is typically measured in time (Gray & Fu, 2004; Gray, Sims, Fu and Schoelles, 2006).

In this paper we examine the effects of different external visualizations on a spatial inference task in which using the external representation effectively depends on relatively complex cognition. In our task, interacting effectively with the external visualization involves reasoning to infer what is the most task-relevant information, as well as planning and executing the motor actions that reveal that information.

*Experimental Task.* The task studied in our experiments involved inferring and drawing a cross-section of an unfamiliar 3-D object with a complex internal structure consisting of ducts that branch in different directions (see Figure 1a and b). A superimposed vertical or horizontal line on printed images indicated where participants should imagine the object had been sliced, and an arrow showed the orientation from which participants were to imagine the cross section (see Figure 1a). The participants’ task was to infer and draw what the cross section would look like from the viewpoint of the arrow. The correct cross section for one trial is shown in Figure 1c (with scoring grid and dotted lines superimposed), examples of participants’ drawings for that trial are shown in Figure 1d-f, and further examples of cross-sections used in the studies are shown in Figure 1g.

An informal task analysis suggests that to perform this task, a person must (1) construct an internal representation of the three-dimensional structure from the information in the diagram (encode the object), (2) imagine slicing the object and removing the section between the viewpoint arrow and the cut plane, (3) imagine changing his or her perspective to the view indicated by the arrow, (4) infer what the cross section will look like, and (5) draw the imagined cross section from this perspective<sup>1</sup>. While they performed the cross-section task, our participants had access to a 3-D computer visualization of the object. In different experimental conditions we varied whether they could interact with the external visualization to select a specific view of the object, or whether it played a pre-recorded animation of the object rotating in depth, which could not be controlled.

*Effects of Interactive Visualizations.* There are several ways in which a 3-D computer visualization might be used to help in our task. Any rotation of the 3-D visualization, regardless of whether its level of interactivity is *animated* (merely playable) or *conditional* (responding to user manipulations; Krygier et al., 1997), provides motion-based depth cues such as motion parallax, accretion, and deletion, whereas a static 2-D view provides only pictorial depth cues<sup>2</sup>. This additional depth information, which is available in both the interactive and non-interactive versions of the external representation, should aid in the initial step of constructing an accurate internal representation of the 3-D structure for all individuals.

One potential benefit of interactivity is that it allows the visualization to be rotated in order to view the object from different perspectives. In the context of our task, this level of interactivity allows the participant to rotate the object in the visualization to the orientation from which he or she must imagine the cut plane (i.e., the orientation indicated

by the arrow in Figure 1a). Such an action would reduce the discrepancy between the view shown on the computer monitor and the view that must be imagined, and would reveal information about the structure that is highly relevant to completing the task. Rotating the visualization in this way corresponds to what Kirsh (1997) refers to as a *complementary* action, which is performed in the world and relieves the individual of the need to perform an internal computation. Kirsh and Maglio (1994) found that experienced players of Tetris often rotated the Tetris pieces on the computer screen rather than mentally rotating them. In our task, participants can rotate the external visualization to the arrow view instead of performing a mental rotation or imagined perspective shift, thus offloading cognition onto the perceptual-motor system (Wilson, 2002). A related benefit of interactivity is that the user can rotate the external visualization in his or her own time, e.g., once the stimulus object has been encoded, so that the actions performed on the external visualization can be tightly coupled with his or her internal cognitive processes (cf. Hollan, Hutchins & Kirsh, 2000). In contrast, a non-interactive visualization may show the more task-relevant views of the object when the user is not ready to benefit from seeing these views.

Another possible advantage of interactivity may be the correspondence between the motor commands made to control the visualization and the resulting movements observed. Monitoring motor commands may provide especially strong cues about spatial properties (e.g., Christou & Bühlhoff, 1999; Feldman & Acredolo, 1979; Philbeck, Klatzky, Behrmann, Loomis, & Goodridge, 2001; Wang & Simons, 1999), particularly with a naturalistic interface designed such that manipulations made by users are exactly mirrored in the resulting movements of the visualization (cf. Schneiderman 1983). Recent research indicates that congruent hand motions can facilitate and incongruent hand motions can

impede mental rotation (Wexler, Kosslyn & Berthoz, 1998; Wohlschlagel & Wohlschlagel, 1998; but see Schwarz & Holton, 2000).

Given these potential benefits of interactive control it is perhaps surprising that previous studies have not shown consistent advantages of interactive over non-interactive visualizations. In this research we propose that interactive visualizations are not always more effective because people do not always use them effectively. Distributed cognition theorists have focused on how the design of external tools can affect cognitive performance by providing affordances for adaptive control, and we agree that manipulation is partly driven by the properties of external tools (Kirsh, 2004). However we believe that for complex and novel tasks, understanding of how to use external tools, and use of these tools, may be cognitively demanding and subject to individual differences among users. Thus we propose that *both* individual differences among users *and* differences in external tools will affect performance.

These observations allow us to formulate two alternative hypotheses regarding the effects of interactivity in our spatial inference task. The first hypothesis is that interactive visualizations always lead to improved performance compared to non-interactive visualizations, because they allow users to rotate the 3-D model in their own time to the most task-relevant views. This hypothesis presumes that people are generally able to use the visualization effectively and find this less effortful than using internal visualization processes. The alternative hypothesis is that the effects of interactivity will depend on whether people discover how to best use the external visualization, and perhaps how they perceive the relative costs and benefits of manipulating the external visualization versus using internal visualization processes. According to this hypothesis, there will be individual

differences in use of the visualization and these differences will be related to task performance, so that there will be an advantage of an interactive visualization only for individuals who use it to reveal the most task-relevant information. A corollary of the second hypothesis is that a non-interactive visualization that shows the most task-relevant information will be more effective than an interactive visualization for those who do not use it effectively.

*Effects of Spatial Ability.* An additional goal of our experiments was to examine how both performance on the cross section task and use of interactive visualizations are modulated by spatial ability. Spatial ability can be interpreted as the ability to mentally store and manipulate visual-spatial representations accurately (Hegarty & Waller, 2005). Because inferring a cross-section of a 3-D object involves mentally manipulating a representation of its internal structure, we hypothesize that spatial ability will predict performance in the cross section task itself, regardless of whether the individual has interactive control. In studies within the domain of geology, Kali and Orion (1996) found that high spatial participants were able to deduce the internal properties of structures while low-spatial participants were unable to mentally penetrate the structure and depended on patterns visible on external faces. They referred to the relevant ability as *visual penetration ability*. Spatial ability also predicts performance in inferring cross-sections of simple solids (Lord, 1985).

Spatial ability may also modulate the effects of external visualizations. There are at least three possible ways in which interactive visualizations may affect individuals of different spatial abilities (Hegarty, 2004). First, such visualizations might “augment” performance equally for high-spatial and low-spatial individuals, so that users of all

abilities are helped equally. Second, provision of an interactive visualization may be particularly effective for low-spatial individuals who have poor internal visualization abilities. In this case, an external visualization may act as a cognitive “prosthetic” for low-spatial individuals. Third, spatial ability might be a necessary prerequisite for using an external visualization effectively. For example, spatial ability may be important for inferring which views of the structure are most useful for performing the cross-section task and planning and executing rotations to those views without becoming disoriented. Alternatively spatial ability might be necessary for making sense of the information contained within the external visualization and effectively integrating that information with one’s own internal representation. In either of these cases, external visualizations should provide the most benefit to participants with stronger spatial abilities, magnifying performance differences between high- and low-spatial individuals.

*Overview of the Experiments.* In three experiments we examined the effects of interactivity and spatial ability on the task of inferring and drawing a cross section of a 3-D object. In all experiments, we compared an interactive condition to a non-interactive condition. Within the interactive conditions we varied the level of control available by allowing either unconstrained object rotation or rotation constrained to cardinal axes of the object. In Experiment 1, we contrasted performance with an interactive visualization to that with a constantly rotating non-interactive visualization. In Experiment 2, we used a yoked design to observe effects of interactivity while controlling for visual input in interactive and non-interactive conditions. In Experiment 3 we contrasted an interactive visualization with a non-interactive visualization that was designed to model the visual information accessed by the most successful interactive participants in earlier experiments. In addition to

examining the effects of interactive versus non-interactive conditions on performance, we also examined how the interactive visualizations were manipulated and how this, in turn, affected performance.

### Experiment 1

In our first experiment, participants performed the cross-section task while viewing a computer visualization of the 3-D object. Half of the participants could interactively manipulate the visualization, while the other half could only view the visualization passively. We also tested participants' spatial ability with a psychometric test, predicting that spatial ability would correlate with overall performance on our cross-section task. However, because previous research has been so mixed with respect to the effects of interactivity on performance, because our cross-section task is novel, and because we allowed for individual differences in use of the interactive visualization, we did not make predictions about whether participants with access to interactivity would do better overall on this task than participants without interactive control.

### *Method*

#### *Participants*

Sixty undergraduate students were recruited from the psychology department subject pool at the University of California, Santa Barbara (UCSB), and received partial course credit for participating. Participants were randomly assigned to one of two conditions: *interactive* or *non-interactive*.

#### *Materials*

Task instructions were presented via printed materials and a custom instructional animation that explained the meaning of a cross section and the requirements of the task.

These illustrated a cross section that resulted from slicing an apple and demonstrated that a drawing of a cross section should contain the object's outer contour and any internal structures, as intersected by the cutting plane.

A novel 3-D object was created using 3-D Studio Max software. The object was egg-shaped with a transparent exterior revealing an internal network of duct-like structures (Figure 1a & b). Pictorial depth cues, such as highlights, shadows, and visual occlusion, suggested spatial depth in the figure, which was modeled in two hues to differentiate the internal and external structures.

A series of vertical and horizontal cross sections of the virtual object were produced using 3-D Studio Max software (see Figure 1g for examples). From these, 10 cross sections were selected for the experimental trials. For each trial, an 8.5 X 11 in. (21.6 X 27.9 cm) printed stimulus sheet was produced, which showed a static view of the object (identical for all trials), with a horizontal or vertical line superimposed at the location of the virtual cross section (which was different on every trial). An arrow positioned above or to the left of the figure, pointing towards the cutting line, indicated the perspective from which the participant was to imagine viewing the cross section (see Figure 1a). Answer templates showing the true spatial relations in each cross section were derived from the virtual object and used as scoring guides.

QuickTime Movie Player software was used to present two 3-D visualizations of the object on a computer screen (although standard computer monitors are capable of representing only pictorial or 2.5-D views, we adopt the convention of referring to 2.5-D computer visualizations as 3-D visualizations). The interactive visualization comprised two windows, each containing a visualization of the object. One visualization could be rotated

in depth around the horizontal axis and the other could be rotated in depth around the vertical axis, using a slider-bar controlled via a mouse. The non-interactive visualization comprised a continuously rotating visualization of the object that looped repeatedly through alternating horizontal and vertical rotations. This condition was created by splicing together one complete loop of each rotation axis from the two interactive visualizations. Thus, the object traveled through the same trajectories in the non-interactive visualization that were available in the interactive condition, but the non-interactive condition did not allow participants to control the speed or pause the rotations on a specific view. Figure 1b shows sample screen shots illustrating different on-screen views of the object.

Participants completed a standardized test of spatial ability, which was an adapted version of Guay's (1976) Visualization of Views Test (Eliot & Smith, 1983). In this paper and pencil test, a three-dimensional object is depicted in the center of a transparent cube. The same object from a different viewpoint is depicted below the cube. The task is to indicate the corner of the cube from which the new view of the object is taken (24 items, 8 minutes). A sample item is shown in Figure 2.

### *Procedure*

We tested participants individually. An experimenter explained the task verbally and with an illustrated written description, and showed the instructional animation. Participants were then introduced to the stimulus figure and told to imagine that the object had been sliced at the location of the vertical or horizontal line in the printed images of the stimulus structure. Their task was to draw the cross section at that point, as if seen from the viewing perspective specified by the arrow (see Figure 1a). For participants in the non-interactive condition, the experimenter demonstrated the looping visualization and

explained that it would run continuously during the task. For participants in the interactive condition, the experimenter demonstrated the use of the visualization slider bar. All participants were told that the dynamic visualization on the screen was designed to help with the task, and they were instructed to use it as much as they wanted, without being specifically instructed as to how they should use it.

Each participant completed a sample item prior to beginning the drawing task, and the experimenter checked this for general understanding of the task requirements, clarifying any misconceptions if necessary. Once the experimenter was satisfied that the participant understood the general procedure, the experimental trials proceeded. Participants worked through ten trials at their own pace with no time limit. During the task they were free to look at the non-interactive visualization, or rotate the interactive visualization, at will. Following the trials, the Guay spatial ability test was administered.

### *Scoring Methodology*

Examples of participants' drawings for one trial are shown in Figure 1d-f. Figure 1c shows the correct answer for this trial. Drawings were assessed for spatial understanding using four standardized criteria. They were scored on each criterion separately using a binary scheme: A score of 1 was allocated if the drawing passed, 0 if it failed.

*Number of ducts.* We scored whether the drawing of the cross section showed the correct number of ducts, which was either one, two, or three on different trials. Of the sample participant drawings in Figure 1d-f, all three passed this criterion.

*Outer shape.* We scored the outer shape of the cross section. There were two types of test items, horizontal and vertical slices, which resulted in circular and oval cross sections, respectively. Scorers measured the maximal width and height of the drawing and

divided the width by the height, to determine whether the outer shape of the drawing fell within an acceptable width to height ratio. The correct ratio for oval slices was 0.74, permissible range 0.54 to 0.94. The correct ratio for circular slices was 1.00, permissible range 0.83 to 1.20. In Figure 1, only the sample drawing in Figure 1d passed this criterion.

*Duct angles.* We measured the spatial relationships among the ducts on all cross sections containing more than one duct. Scorers measured the angles between ducts by connecting the centers of the ducts in each drawing. For cross sections containing two ducts, a line was drawn connecting the ducts, a horizontal line (aligned with the page) was drawn through the center of the lower duct, and the angle between these two lines was measured. If the difference between this angle and the correct angle did not exceed 20°, a pass was awarded. For three-duct cross sections, lines were drawn connecting the three ducts, the three resulting angles were measured, and for each the error was calculated as the difference between the correct angle and the drawn angle. These were summed, and if the total error did not exceed 60° (average of 20° at each angle), a pass was awarded. If the wrong number of ducts was drawn, a score of 0 was recorded for this measure. In Figure 1, the correct duct angles are shown by the dotted line superimposed on Figure 1 c; of the sample participant drawings shown, *d* and *f* passed this criterion, but *e* failed.

*Duct positions.* Finally, we assessed whether the ducts were located in the correct region(s) of the slice, i.e. in relation to the outer contour of the cross section. If the position of a drawn duct differed from the correct position by no more than one tenth of the drawing's width or height, a pass was awarded. All ducts in a drawing had to pass in order for a point to be assigned for a given trial. This criterion was scored independently of the *duct angles* criterion. If the wrong number of ducts was drawn, a score of 0 was recorded

for this measure. In Figure 1, the 10 x 10 grid superimposed on Figure 1 *c* shows the correct duct coordinates; of the sample participant drawings shown, *d* passed this criterion, but *e* and *f* failed.

In a sample of 180 trials, inter-rater reliability across two scorers was 97.3% for the number of ducts measure, 91.6% for the outer shape measure, 95% for the duct angles measure and 91.6% for the duct positions measure. To validate our pass-fail criteria, on a sample of 900 drawings from another experiment we conducted a measurement procedure to precisely quantify errors on the four criteria (e.g., we measured *duct angle* error as the absolute difference in degrees between the measured angle(s) in the drawing and the correct angle(s) in the stimulus). Correlations between these quantitative errors and the pass/fail scores for the four criteria were -.98, -.86, -.78, and -.80, respectively (one measure was error and the other was proportion correct, so negative relationships were expected). This level of agreement indicates that the pass/fail measures provided an assessment of performance that was comparable to using exact quantitative measurements, so the pass/fail criteria were used to score the remaining data from all experiments. Chronbach's alpha (a measure of internal consistency) for the four measures was .68 for number of ducts, .88 for outer shape, .71 for duct angles and .69 for duct position, indicating satisfactory reliability of these measures.

### *Results*

Means (and *SDs*) for the interactive and non-interactive groups, respectively, on the Guay Visualization of Views test were 11.5 (7.3) and 10.6 (7.4), which did not differ significantly,  $F(1, 58) = 0.22, ns$ . In general, individuals drew slices in which the number of ducts and the outer shape were correct, so performance for these two variables was at or

near ceiling (mean proportions correct = .89,  $SD = .14$ , for number of ducts and .79,  $SD = .28$ , for outer shape). These measures were excluded from further detailed analysis in this and subsequent experiments, although for comparison we report descriptive statistics for all four drawing measures.

Duct angles ( $M = .55$ ,  $SD = .27$ ) and duct positions ( $M = .49$ ,  $SD = .24$ ) were the most difficult of the four criteria. A correlational analysis (see Table 1) showed that duct angles and duct positions were the most strongly related of the four measures. Thus we assumed that they measured related processes and combined them into a single measure, hereafter referred to as *duct relations*, which became our primary measure of task performance. This combined measure captures participants' ability to infer the spatial relations revealed by the cutting plane, both among the ducts, and between the ducts and the whole cross section.

Figure 3 shows proportion correct in the two conditions for duct relations broken down by spatial ability, and indicates that performance was poorer in the non-interactive condition than in the interactive condition. To assess the effects of interactivity and spatial ability on performance, a univariate analysis of variance (ANOVA) was performed on duct relations. Spatial ability was entered as a fixed factor in this analysis, dichotomized into higher and lower ability groups via a median split. Performance on the duct relations criterion showed a main effect of interactivity,  $F(1, 56) = 14.51$ ,  $p < .001$ ,  $\eta_p^2 = .21$ , and a main effect of spatial ability,  $F(1, 56) = 8.42$ ,  $p = .005$ ,  $\eta_p^2 = .13$ , but no interaction between the two,  $F(1, 56) = .35$ ,  $p = .56$ ,  $\eta_p^2 = .01$ . An analysis of the relationship between spatial ability (not dichotomized) and duct relations for the two conditions showed that performance was correlated with spatial ability in both the interactive condition,  $r = .39$ ,  $p =$

.04, and the non-interactive condition,  $r = .47$ ,  $p = .007$  (correlation coefficients do not differ significantly).

### *Discussion*

Experiment 1 revealed large individual differences in performance on the cross-section drawing task. Whereas participants could generally infer the correct number of ducts and the outer shape of a cross section, they differed widely in their ability to infer the relative locations of the ducts, both within the slice and with respect to each other. Performance on this aspect of the task was correlated with spatial ability, supporting our hypothesis that success in this task depends on the ability to construct and manipulate accurate spatial representations.

Duct relations (the combined duct angles and duct positions measures) was the most challenging measure of performance and was affected by both interactivity and spatial ability. Participants who had access to interactive visualizations of the structure performed significantly better on this measure than participants who had access to a non-interactive, continuously rotating visualization of the structure. There was no interaction between spatial ability and condition, indicating that the interactive visualization was equally effective for high- and low-spatial individuals.

Why did participants in the interactive group perform better than participants in the non-interactive group? In giving participants interactive control of the visualization, we also gave them the ability to slow or pause it at specific views of the structure. By contrast, in the non-interactive condition, the visualization continuously rotated, so that non-interactive participants did not see the same information that interactive participants saw, confounding interactivity with visual information. Therefore, we cannot determine on the

basis of Experiment 1 whether the interactive group did better because of interactivity per se, or because they had access to more useful views of the object.

### Experiment 2

The aim of Experiment 2 was to disambiguate the role of two factors in reasoning with external visualizations—interactivity and access to informative views. As in Experiment 1, we compared two groups, one of which had access to interactive control of the visualization while the other did not. Unlike Experiment 1, however, the visual information on screen was identical in the two conditions. We digitally recorded the interactions made by each of the interactive participants and later played them back to non-interactive participants in a yoked-pairs design. This ensured that the available visual information about the object's structure in the two conditions was identical. If the advantage in Experiment 1 was due to interactive control, the interactive group should score higher on this task than the non-interactive group. If the advantage was due only to the quality or quantity of visual information available, rather than interactive control per se, the two groups should not differ.

These digital recordings of participants' interactions also allowed us to explore questions relating to interactive behavior. In Experiment 2, we examined whether there are substantial individual differences in how people use interactive visualizations, and whether different ways of manipulating the visualization are related to success on the task and to spatial ability. In particular, we were able to examine whether the most successful participants preferentially rotated the display to show what we assumed would be the most informative view for our task, namely the view from the perspective indicated by the arrow in the printed stimulus. We called this the *arrow view*. We predicted that success on this

task would correlate with the extent to which participants rotated the display in order to access the *arrow view*.

Finally, we also provided participants in the interactive condition of Experiment 2 with a more intuitive interface for interacting with the computer visualization which allowed more interactive control. The interface was a motion sensor encased in an egg-shaped object, which could be rotated around any axis in 3-D space and which produced corresponding real-time rotations of the computer visualization (see Figure 4).

### *Method*

#### *Participants*

Sixty undergraduate students were recruited from the psychology department subject pool at UCSB and received partial course credit for participating. None had participated in any other experiments in this series. Participants were randomly assigned to one of two conditions: *interactive* or *non-interactive*. Each participant in the interactive condition was paired with a participant in the non-interactive condition using a yoked-pairs experimental design. Participants were unaware of the yoked design.

#### *Materials*

Task instructions were presented via the same printed materials and instructional animation as in Experiment 1, and the same novel 3-D object was used (see Figure 1a & b). This was presented in the form of static 2-D printed images and a dynamic 3-D computer visualization. As in Experiment 1, printed stimulus sheets showed a static view of the object with a horizontal or vertical line superimposed at the location of the virtual cross section and an arrow indicating the imagined viewing perspective. The 10 cross sections were similar to the trials used in Experiment 1 but were not all identical: in Experiment 2,

the arrow indicating the imagined viewpoint could be positioned above, below, to the left, or to the right of the object in the printed stimulus.

The structure represented in the 3-D computer visualization was identical to the stimulus object used in Experiment 1, but there was just one visualization, which could be rotated in any direction (i.e., rotations were not confined to the vertical and horizontal axes). The control interface also differed. The visualization was presented using a custom program written in *Python* script run on *Vizard* software. In the interactive condition, participants manipulated the visualization using a 3 degrees-of-freedom motion sensor, the *InterSense Intertiacube2*, mounted inside an egg-shaped casing that resembled the egg-shaped object in the display (see Figure 4). The participant held the interface in one hand and manipulated it naturalistically, i.e., there were no system-specific constraints on how the device could be handled or rotated, so that participants could manipulate it in the same manner as any other object of that approximate size, shape, and weight. As the interface was rotated, the object on the screen rotated identically in real time; translations were not replicated, so that the visualization did not drift from the center of the screen. In the non-interactive condition, the program was used to play back the previously recorded movements of a yoked interactive participant, and the interface was kept out of sight. As in Experiment 1, participants completed the adapted version of Guay's (1976) Visualization of Views Test (Eliot & Smith, 1983).

### *Procedure*

Participants were tested individually. They received the same task instructions as in Experiment 1. Prior to the interactive condition, the experimenter demonstrated the use of the hand-held interface, and the participant was given a short period to practice

manipulating it. Prior to the non-interactive condition, the experimenter explained that the object would move on the screen during the task; these movements were demonstrated using a pre-recorded sample. Participants in both conditions were told that the dynamic visualization on the screen was designed to help them gain more information about the structure of the object, and were instructed to use it as much as they wished during the task, but were not given instructions regarding its use.

Each participant completed a sample item prior to beginning the drawing task. The experimenter checked the sample for general understanding of the task and clarified any misconceptions. Participants worked through ten trials at their own pace with no time limit. During the task they were free to look at (non-interactive condition) or rotate (interactive condition) the visualization at will. At the start of each trial, the experimenter pressed the left mouse button, producing a tone, which was the auditory signal to begin the trial. In the interactive condition, this initiated a digital recording of any manipulations of the visualization produced by the participant. In the non-interactive condition, this initiated a playback of the previous interactive participant's manipulations for that trial. When the participant finished drawing, they indicated this verbally to the experimenter, who then pressed the right mouse button. This produced a different tone indicating the end of the trial, and stopped the recording process or the playback, respectively<sup>3</sup>. Non-interactive participants could ask to watch the playback again if they wished to; the number of times this occurred was not recorded, but the majority of participants did not ask to watch the playbacks again. The mouse clicks at the start and end of each trial also controlled an automatic timing mechanism to record approximate time on task for each trial. Prior to beginning each interactive trial, the interface was returned to its "home" position on a

stand, which returned the on-screen visualization to its default start position. Following the drawing trials, the Guay test was administered.

### *Results*

#### *Drawing Accuracy and Psychometric Data*

The scoring criteria were the same as in Experiment 1, yielding accuracy measures for number of ducts, outer shape, duct angles, and duct positions. Mean proportions correct (and *SDs*) for the four scoring criteria were as follows: number of ducts = .91 (.10); outer shape = .76 (.28); duct angles = .50 (.18); duct positions = .42 (.18), showing levels of performance similar to those in Experiment 1. Relationships among the four scoring criteria were also similar to those found in Experiment 1 (see Table 1). As in Experiment 1, duct angles and positions were combined into an aggregate *duct relations* measure and number of ducts and outer shape were not analyzed further. Means (and *SDs*) for the interactive and non-interactive groups on the Guay test were 11.2 (7.2) and 11.7 (7.4) respectively; the groups did not differ on this measure,  $F(1, 47) = .06, p = .81$ .

#### *Effects of Interactivity and Spatial Ability on Drawing Performance*

Means (and *SDs*) for the interactive and non-interactive groups on the duct relations measure were very similar, .47 (.15) and .46 (.18), respectively. A 2 X 2 univariate ANOVA was performed in which the factors were interactivity and spatial ability, which was entered as a fixed factor, dichotomized into higher and lower ability groups via a median split. Consistent with Experiment 1, there was a significant main effect of spatial ability on performance,  $F(1, 45) = 13.93, p = .001, \eta_p^2 = .24$ , but in contrast to Experiment 1 there was no significant effect of interactivity,  $F(1, 45) = .02, p = .88, \eta_p^2 = .00$ . There was no significant interaction between the two factors either,  $F(1, 45) = .07, p = .79, \eta_p^2 = .00$ .

We correlated spatial ability (not dichotomized) and performance, for the two conditions separately. Performance was correlated with spatial ability in both the interactive condition,  $r = .53$ ,  $p = .008$ , and the non-interactive condition,  $r = .49$ ,  $p = .013$ .

### *Patterns of Interactivity*

We expected that the most informative view of the object for the purpose of our task would be the view from the perspective of the arrow in the stimulus—the *arrow view*. An informal examination of patterns of interactivity indicated that a common strategy was to rotate the computer visualization to this view. We specified a window  $20^\circ$  to either side of the arrow view on each trial, and derived three measures of interactivity, which we used to examine how the interactive participants manipulated the visualization: (1) amount of time spent on the arrow view, (2) time taken to reach the arrow view (which was missing data for participants who did not reach the arrow view on any trials), and (3) time spent on views other than the arrow view. These values were averaged across 8 trials; we excluded two trials in which the arrow view was the same as the start view for the visualization, because in these instances the arrow view was on screen from the start of the trial. Mean overall time per trial for these eight trials, which includes both time spent interacting with the visualization and time spent drawing the cross section, ranged from 23.4 to 133.1 seconds ( $M = 60.1$ ,  $SD = 26.3$ ) and was not normally distributed<sup>4</sup>,  $skewness = 1.08$ ,  $SE = .43$ . Mean time spent on the arrow view ranged from 0.00 to 35.80 seconds ( $M = 10.42$ ,  $SD = 10.91$ ; one statistical outlier of 57.95 seconds was excluded), and was not normally distributed,  $skewness = .89$ ,  $SE = .43$ . Mean time taken to reach the arrow view ranged from 8.24 to 48.43 seconds ( $M = 25.23$ ,  $SD = 12.42$ , normally distributed). Mean time spent on

views other than the arrow view ranged from 18.2 to 86.3 seconds ( $M = 47.2$ ,  $SD = 17.1$ , normally distributed).

In order to compute the *number of trials* on which the arrow view was accessed, we defined a categorical variable with a binary scheme of 0 or 1. We set a minimum threshold for this variable because random rotations of the interface could inadvertently cause the visualization to pass through the arrow view at some point during a trial. An interactive participant was deemed to have deliberately accessed the arrow view on a given trial if the visualization was maintained within the  $20^\circ$  window for a total time of two seconds or longer over the course of that trial. Based on this definition, individuals varied widely in the proportion of trials on which they deliberately accessed the arrow view ( $M = .40$ ,  $SD = .32$ ). As Figure 5 shows, participants fell largely into two groups: those who selected the arrow view on most trials, and those who selected the arrow view on few or no trials. The bimodal distribution suggests that this measure should be treated as a dichotomous variable. We therefore divided participants using a median split on this measure. The mean proportion of trials in which the two groups accessed the arrow view was  $.67$  ( $SD = .11$ ) and  $.10$  ( $SD = .14$ ). We called these two groups the *high access* and *low access* groups, respectively.

#### *Relationship Between Interactivity and Task Performance*

An independent samples t-test on the duct relations scores showed that participants in the *high access* group ( $M = .54$ ,  $SD = .13$ ) significantly outperformed participants in the *low access* group ( $M = .38$ ,  $SD = .12$ ),  $t(28) = 3.67$ ,  $p = .001$ ,  $CI_{.95} = .07$  to  $.26$ . Because a parametric correlation analysis was not appropriate we applied a Spearman rank-order correlation, which confirmed that proportion of trials on which the arrow view was accessed was correlated with performance,  $Rho = .55$ ,  $p = .002$ .

Performance was also significantly correlated with interactivity measures expressed as time, namely time spent on the arrow view,  $Rho = .47$ ,  $p = .01$ , time spent on other views,  $Rho = .43$ ,  $p = .02$ , and total time on task,  $Rho = .57$ ,  $p = .001$ . The correlation with time taken to *reach* the arrow view,  $Rho = .35$ ,  $p = .09$ , did not reach statistical significance.

Total time on task covaried with time spent on the arrow view ( $Rho = .64$ ,  $p < .001$ ), time spent on other views ( $Rho = .90$ ,  $p < .001$ ), and time to reach the arrow view ( $Rho = .79$ ,  $p < .001$ ). This suggests that all of these measures might be driven by a common factor, such as motivation. If poor motivation was the key factor causing some individuals not to access the arrow view, then presumably these individuals would also spend less time exploring the structure in general. We therefore compared time spent on views *other than* the arrow view for the high and low access groups. An independent samples t-test showed that these groups did not differ significantly in terms of time spent on other views,  $t(27) = .84$ ,  $p = .54$ . Mean time per trial in seconds spent on other views was 49.1 ( $SD = 17.8$ , range 18.2 to 86.3) for the high access group, and 45.1 ( $SD = 16.6$ , range 23.3 to 80.5) for the low access group. Thus, individuals who accessed the arrow view less often did not show an associated disinclination to explore other, non task-relevant views of the structure (this was further supported by the fact that time spent on the arrow view did not correlate significantly with time spent on other views,  $Rho = .30$ ,  $p = .11$ , *ns*). High access and low access groups differed *only* in time spent on the arrow view itself, and this alone caused the difference in total time on task. There is no evidence that a third factor, such as motivation, underlies both total time on task and time spent on the arrow view.

Spatial ability did not correlate with any of the interactivity measures,  $Rho = .06$  to  $.31$ ,  $p = .13$  to  $.77$ , except for a moderate correlation with time spent on views other than

the arrow view,  $Rho = .42$ ,  $p = .04$ . In a regression analysis, spatial ability and the proportion of trials on which the arrow view was accessed jointly accounted for 42% of the variance in performance ( $R^2_{adj} = .42$ ;  $R = .69$ ). The overall relationship was significant,  $F(2, 20) = 8.25$ ,  $p = .003$ , and performance was reliably predicted by both spatial ability,  $t(23) = 2.36$ ,  $p = .03$ , and proportion of trials on which the arrow view was accessed,  $t(29) = 2.62$ ,  $p = .02$ . None of the other interactivity measures accounted for significant additional variance.

#### *Effects of Interactivity Patterns on Performance of Non-interactive Participants*

We also examined whether the patterns of interactivity predicted performance of the non-interactive participants who passively viewed the interactions. If the key predictor of success was the visual information available, rather than control of the visualization, the manner in which the structure was manipulated should predict performance even in the non-interactive condition.

There was a significant correlation between performance on duct relations and the time taken to reach the arrow view, that is, the later the visualization reached the arrow view, the better the non-interactive participants performed,  $Rho = .58$ ,  $p = .003$ . Non-interactive participants' performance was also predicted by the amount of time the visualization showed the arrow view,  $Rho = .35$ ,  $p = .06$  (marginal), amount of time the visualization showed other views,  $Rho = .43$ ,  $p = .02$ , and total time on task,  $Rho = .48$ ,  $p = .007$ . The performance of non-interactive participants who viewed the interactions of *high access* interactive partners ( $M = .49$ ,  $SD = .21$ ) did not differ significantly from those who viewed the interactions of *low access* partners ( $M = .41$ ,  $SD = .15$ ),  $t = 1.13$ ,  $p = .27$ .

The non-interactive participants' spatial ability and the measures of interactivity were entered into a hierarchical regression analysis. Spatial ability and time taken for the visualization to reach the arrow view together produced a significant model,  $F(2, 18) = 6.79, p = .007$ , jointly accounting for 39% of the variance in performance of the non-interactive participants ( $R^2_{adj} = .39; R = .68$ ), and time to reach the arrow view was a significant predictor within the model,  $t(23) = 2.37, p = .03$ . The addition of time spent on the arrow view and time spent on other views did not significantly improve the predictive power of the model ( $R^2$  change = .000 and .017, respectively).

*Post Hoc Comparison of Interactive Conditions from Experiments 1 and 2*

The interactive conditions of Experiments 1 and 2 differed in both level of interactivity and type of interface. Five of the 10 trials in these two experiments were common to both experiments. Examining only the data from these five trials, we conducted a *post hoc* comparison of interactive participants in Experiments 1 and 2. A 2 X 2 univariate ANOVA with the factors of experiment (Experiments 1 & 2) and spatial ability (higher and lower ability groups defined by median split) showed a main effect of experiment,  $F(1, 49) = 8.11, p = .006, \eta_p^2 = .14$ , and a main effect of spatial ability,  $F(1, 49) = 5.28, p = .03, \eta_p^2 = .10$ , but no interaction between the two,  $F(1, 49) = 2.40, p = .13, \eta_p^2 = .05$ . This indicated that the interactive participants in Experiment 1 ( $M = .63, SD = .18$ ) outperformed the interactive participants in Experiment 2 ( $M = .47, SD = .15$ ). The two groups did not differ in spatial ability ( $t[51] = .15, p = .88$ ), precluding this as a possible explanation; mean (and *SD*) spatial ability scores for interactive participants in Experiments 1 and 2, respectively, were 11.5 (7.3) and 11.2 (7.2).

*Discussion*

In Experiment 2, the visual information available to participants was the same in the interactive and non-interactive conditions, and interactive and non-interactive participants performed equally well. This contrasts with Experiment 1, where the visual information was not the same for the two conditions, and interactive participants performed better than non-interactive participants. These results suggest that interactivity per se is not the critical factor in the performance of our cross-section task. Rather, the quality of the visual information available predicts success on this task, regardless of whether participants have control over it. Apparently, interactive participants in Experiment 1 performed better than non-interactive participants because interactivity allowed them to access key information for completing the task, not because they had active control.

There were large individual differences in use of the interactive controls in Experiment 2. Thus, not all individuals offloaded the *cognitive* process of mentally imagining the stimulus from the perspective of the arrow onto the *perceptual-motor* process of rotating the external stimulus to the arrow view. These results are not consistent with theories stating that people always minimize reliance on internal computations and offload these onto perceptual motor processes (e.g. Kirsh & Maglio, 1994). Instead, our results suggest that individuals differ in their relative reliance on internal visualization processes versus manipulation of external visualizations.

The patterns of unconstrained interactions observed in Experiment 2 provide insights into the nature of the key information for completing the task. We predicted that the critical view for the task is the view from the perspective indicated by the arrow in the stimuli. We found that the performance of interactive participants was significantly related to how much they rotated the visualization to this view during the task. That is, interactive

participants who accessed this view more often and spent more time on this view performed better. Interestingly, interactive participants' spatial ability did not relate to the degree to which they selected the arrow view, so that patterns of interactivity and spatial ability made independent contributions to performance.

Consistent with the relevance of the arrow view is the finding that seeing this view is also important for non-interactive participants, even though they do not actively select it. For non-interactive participants, the timing of the movements is critical. If the arrow view is reached too early in the trial, the non-interactive participants cannot use it effectively. But the later in the trial the arrow view is reached, the better non-interactive participants perform, suggesting that they can benefit from this view when they are ready to make use of it. One possible reason for this finding is that trials in which the object took longer to reach the arrow view involved slower transitions towards the target view. These more gradual changes may be easier to understand for individuals who do not have active control. By contrast, the person planning and executing the transitions has meta-knowledge of what is coming next and how each movement relates to subsequent goals, so that the speed of the navigation does not impact his or her ability to understand the changes occurring on-screen.

The result that interactive participants in Experiment 1 outperformed interactive participants in Experiment 2 is also consistent with the notion that the arrow view provides the most relevant information for this task. Note that participants in Experiment 1 used a less direct interface that constrains rotations to two axes, whereas participants in Experiment 2 used a more direct interface that can be manipulated freely and naturalistically, as with any real object. This finding is contrary to theoretical claims that

more naturalistic interfaces should generally trump less direct interfaces, because of their lower cognitive demands (Hutchins, Hollan, & Norman, 1985; Shneiderman, 1983), and consistent with suggestions that constrained interactivity might be more effective (and less cognitively demanding) for some tasks and users (Krygier et al., 1997; Betrancourt, 2005).

The constrained vertical and horizontal rotations that participants could make in Experiment 1 allowed the object to be rotated only around the most relevant axes for the task (the vertical and horizontal axes). The unconstrained rotations of Experiment 2, in contrast, allowed participants to rotate the object around irrelevant axes that did not provide helpful views of the object.

If our interpretation is correct, then it is access to informative views of the structure, not active control, per se, that predicts performance. It should therefore be possible to optimize performance by restricting visual information to views that are particularly informative for a given task. We examine this possibility in Experiment 3 by introducing a condition in which some participants are presented with “optimal” movements of the visualization, based on the most effective interactions that we observed in Experiment 2.

### Experiment 3

In Experiment 3, we introduced a non-interactive condition in which the movements presented to participants mimicked the most effective manipulation strategies observed in Experiment 2. In this condition, the visualization rotates into the arrow view, pauses, and then intermittently rocks back and forth to either side of this view. Participants in this *arrow views* condition had no control over the visualization; they were exposed to the “optimal” manipulation strategy on every trial.

In a similar study on 3-D object recognition, James, Humphrey, and Goodale (2001) identified that the critical information for learning the structure of novel 3-D objects comes from “plan” views, such as the front and side of the object, as these provide the most information about 3-D structure with the smallest amount of movement. In a later experiment, they restricted participants’ interactions with computer visualizations of objects to either these plan views or to less informative intermediate views, and found that the former led to significantly better learning of the structure (James et al., 2002). In Experiment 3 we used the same methodology, taking into account that the perspective of the arrow, not the structure of the 3-D object, determines the optimal view for our task. In Experiment 2, we found that successful interactive participants commonly rotated the visualization into the arrow view and maintained it there for a period of time, during which they repeatedly tilted the object back and forth a small amount. These movements are consistent with manipulation patterns observed by James et al., who noted that while participants dwelled primarily on “plan” views, they also repeatedly moved the object back and forth around these views. These intentional *wobbles* typically had a range of less than 45° around a central point ( $\pm 22.5^\circ$ ). They presumably provide more 3-D spatial information than a static view alone because of the depth cues provided by motion. We based the movements shown in the non-interactive arrow views condition on these types of manipulations.

We compared performance in this non-interactive condition to performance in an interactive condition. As in Experiment 2, we classified participants in the interactive condition into those who accessed the arrow view on most trials and those who accessed the arrow view on few trials. We predicted that participants in the non-interactive arrow

views condition would perform similarly to interactive participants who made effective use of the visualizations, but they should perform better than interactive participants who did not make effective use of the visualization. As in our earlier experiments, we also examined the role of spatial ability in the two conditions.

### *Method*

#### *Participants*

Sixty undergraduate students were recruited from the psychology department subject pool at UCSB, and received partial course credit for participating. None had participated in any other experiments in this series. Participants were randomly assigned to one of two conditions: *interactive* or *arrow views*.

#### *Materials*

Task instructions were presented via the same printed materials and custom instructional animation as in Experiments 1 and 2. The same novel 3-D object was used (Figure 1a & b), and the experimental trials were the same 10 cross sections as used in Experiment 2.

Two visualizations were created using a custom program written in *Python* script run on *Vizard* software. The visualizations started all trials at the same orientation as the printed stimulus. The interactive condition was identical to Experiment 2. In the arrow views condition, when the movement was activated the visualization first rotated into the arrow view. It paused on the arrow view for 5 seconds, then it rocked once back and forth 35° to each side of that view (following the same axis as the initial movement), and then it returned to the arrow view where it paused again. This pause-rock-pause movement was repeated continuously until the visualization was restarted or until the participant had

finished that trial. As in Experiments 1 and 2, participants completed the adapted version of Guay's (1976) Visualization of Views Test (Eliot & Smith, 1983).

### *Procedure*

Instructions were the same as in Experiments 1 and 2. The procedure for the interactive condition was identical to the interactive condition in Experiment 2. In the arrow views condition, the experimenter first explained that the object would move on the screen during the task; this was demonstrated using a pre-recorded sample. At the start of each trial, the experimenter first ensured the participant was looking at the computer screen (drawing his or her attention to the screen if necessary). The movement was then initiated with a computer key; the object rotated into the arrow view and repeated the pause-rock-pause motion two or three times. After 15 to 20 seconds the movement was halted and the visualization was returned to its starting position. The movement was initiated from the beginning a second time during each trial. Because the experimenter waited until the participant's full attention was on the screen, the precise timing of the two movements varied slightly from trial to trial. The visualization continued to repeat the pause-rock-pause motion until the participant finished the trial. The participant could also ask to watch the entire movement again from the beginning at any time during the trial; the number of times participants asked to see the movement again was not recorded, but the majority of participants did not request additional viewings. In both conditions, the visualization was reset to the default start position at the beginning of every trial (the same orientation as in the printed stimuli). Participants worked through ten trials at their own pace. Following the drawing trials, the Guay test was administered.

### *Results*

### *Drawing Accuracy and Psychometric Data*

The drawings were assessed for accuracy using the same scoring scheme as in Experiments 1 and 2. Mean proportions correct (and *SDs*) for the four scoring criteria were as follows: number of ducts = .90 (.14); outer shape = .82 (.24); duct angles = .62 (.27); duct position = .51 (.25), showing similar levels of performance to Experiments 1 and 2. Similar relationships among the four scoring criteria were found (see Table 1). As in Experiments 1 and 2, duct angles and positions were amalgamated into *duct relations*, and this was the primary dependent variable in subsequent analyses. Means (and *SDs*) for the interactive and arrow views conditions respectively on the Guay test were 11.1 (7.8) and 12.1 (6.7), which did not differ significantly,  $t(57) = -.54, p = .59, ns$ .

### *Comparison of Interactive and Arrow Views Conditions*

We compared the performance of participants in the non-interactive arrow views condition (who were exposed to the arrow view on every trial) with participants in the interactive condition (who could select the arrow view or not). A 2 X 2 univariate ANOVA was performed, with type of computer visualization (interactive, arrow views) and spatial ability (dichotomized via a median split) as factors. There was a marginally significant main effect of condition, with participants in the non-interactive arrow views condition ( $M = .61, SD = .23$ ) scoring higher overall than participants in the interactive condition ( $M = .52, SD = .25$ ),  $F(1, 55) = 4.01, p = .05, \eta_p^2 = .07$ . Consistent with Experiments 1 and 2, there was a significant main effect of spatial ability,  $F(1, 55) = 29.08, p < .001, \eta_p^2 = .35$ , and no interaction between spatial ability and condition,  $F(1, 55) = .01, p = .94, \eta_p^2 = .00$ . The correlations between task performance and spatial ability in the interactive and arrow

views conditions were  $r = .57$ ,  $p = .001$  and  $r = .79$ ,  $p < .001$  respectively (these coefficients are not significantly different).

To shed further light on these results, we used the same binary categorical variable as in Experiment 2 to compute the number of trials in which interactive participants deliberately accessed the arrow view, in order to identify *high access* and *low access* groups. The arrow view was accessed on a mean of .42 of all trials ( $SD = .30$ ; range = 0.0 to 1.0), and a Spearman rank-order correlation confirmed that proportion of trials on which the arrow view was accessed correlated with performance,  $Rho = .50$ ,  $p = .005$ . A median split divided those who selected the arrow view on more trials ( $>.4$  of the trials) from those who selected the arrow view on fewer trials ( $<.4$  of the trials), with participants who fell at the median excluded. This produced *high access* and *low access* interactive groups ( $n = 11$  in both cases), who accessed the arrow view on an average of .72 ( $SD = .19$ ) and .13 ( $SD = .14$ ) of trials, respectively. Figure 6 shows the performance of these two groups alongside participants in the non-interactive arrow views condition. The non-interactive arrow views participants performed significantly better than the low access interactive group,  $t(39) = -2.80$ ,  $p = .008$ ,  $CI_{.95} = -.38$  to  $-.06$ . By contrast, the performance of non-interactive arrow-views participants was nearly identical to the high access interactive participants and these two groups did not significantly differ,  $t(39) = .09$ ,  $p = .93$ . In neither case did the groups differ in spatial ability,  $p = .38$  to  $.45$ , precluding this as a possible explanation for any difference.

#### *Interactivity Data*

Using the same criteria as in Experiment 2, we analyzed patterns of interactivity among participants in the interactive condition. Mean time per trial was 65.0 seconds ( $SD =$

40.8, range = 12.7 to 182.6, one outlier removed). A Spearman rank-order correlation showed that mean time on task was correlated with performance,  $Rho = .48, p = .008$ . Performance was also correlated with time spent on the arrow view,  $Rho = .51, p = .005$ , time taken to reach the arrow view,  $Rho = .59, p = .001$ , and time spent on other views (one outlier removed),  $Rho = .47, p = .013$ . Spatial ability did not correlate with any of the interactivity measures,  $Rho = -.01$  to  $.14, p = .47$  to  $.95$ .

In a regression analysis, spatial ability and proportion of trials on which the arrow view was accessed jointly accounted for 48% of the variance in performance of interactive participants ( $R^2_{adj} = .48; R = .72$ ). The overall relationship was significant,  $F(2, 27) = 13.36, p < .001$ , and performance was reliably predicted by both spatial ability,  $t(28) = 3.65, p = .001$ , and proportion of trials on which the arrow view was accessed,  $t(28) = 3.22, p = .004$ .

### *Discussion*

In Experiment 3, non-interactive participants who were presented with the arrow view on every trial performed virtually identically to interactive participants who accessed the arrow view on most trials, and actually performed better than interactive participants who did not access the arrow view on many trials. This finding supports our hypothesis that seeing the arrow view rather than interactivity per se is important for performance on this task. Like the two previous experiments, Experiment 3 indicates that spatial ability is an important predictor of success on this task, but is not related to how effectively the interactive visualization is used.

### *General Discussion*

Over the course of three experiments, we found that providing participants with active control of a computer visualization does not necessarily enhance performance on a

spatial reasoning task, whereas seeing the most task-relevant views of the structure does, regardless of whether these views are obtained actively or passively. When participants are given interactive control, there are large differences in how effectively they use the interactivity and these differences predict performance on the task. Spatial ability also makes an independent contribution to performance on the spatial reasoning task, but does not predict patterns of interactive behavior.

In Experiment 1, participants who were allowed to control an external visualization performed better overall than participants who saw a continuously rotating version of that visualization. However, Experiment 2 showed that there was no difference in the performance of interactive and non-interactive participants if the movements of the visualization were held constant in a yoked design. A more important predictor of performance was the quality of visual information that participants received, regardless of whether they controlled the visualization. For participants who had control of the visualization, success was predicted by whether they accessed the same view of the structure as indicated by the arrow in the stimuli. For participants who did not have control, the timing of when they saw this arrow view was important, with better performance when the visualization took longer to reach the arrow view. In Experiment 3, a non-interactive condition was created to mimic the “ideal” movements of the visualization, based on the most successful patterns of interactivity recorded in Experiment 2. Participants in this condition had no control over the visualization but were exposed to the arrow view on every trial. These non-interactive participants performed better than interactive participants who had the means to control the visualization but did not use it in a strategically useful way, and just as well as interactive participants who did use the visualization in a

strategically useful way. These findings confirm our hypothesis that seeing the critical view of the object's structure is more important than interactivity per se. That is, merely having active control over the visualization does not guarantee success on the task, whereas seeing the critical view does.

To better understand these results, it is important to separate two aspects of use of the external visualization: 1) getting to the arrow view, and 2) benefiting from seeing this view. The first refers to how the visualization is manipulated, specifically the cognitive, perceptual and motor processes involved in rotating the visualization to the arrow view. The second refers to how the information provided by this view benefits the task of inferring and drawing a cross section. We discuss each of these in turn.

*Effective Manipulation of the Visualization – Getting to the Arrow View.* According to our task analysis, the process of rotating the visualization to the perspective of the arrow view is the one step in the task that can be externalized. We formulated two alternative hypotheses regarding how participants would behave when given the interactive visualization. One hypothesis, related to the minimum memory hypothesis (cf. Kirsh & Maglio, 1994; Ballard, Hayhoe, Pook & Rao, 1997), predicted that all interactive participants would offload this step onto a perceptual-motor action (i.e., select the arrow view). The second hypothesis, related to the soft constraints hypothesis (Gray & Fu, 2004; Gray et al., 2006) allowed for the possibility that different individuals would rely differentially on internal visualization to *imagine* the arrow view versus use of the external visualizations to *see* this view. The results are clearly more consistent with the second hypothesis. Some individuals did reliably rotate the visualization into the perspective of the arrow in the stimulus, matching what we would expect from individuals who are

intelligently offloading internal processes onto the external world. However, many individuals did not use the visualization in this way.

Why did some interactive participants fail to rotate the visualization to the arrow view? In terms of the minimum memory hypothesis, this is always the optimal strategy. However, use of this strategy requires recognizing how the external visualization can be used to achieve the task goals. One possibility is that some participants did not have, and did not acquire, a metacognitive understanding of how to use the visualization in a helpful way for this task. This is supported by the result that many participants manipulated the external visualization, but did not rotate it to the arrow view. If participants do not access the arrow view, they cannot experience its benefits. Reliance on perceptual-motor processes rather than internal cognition was proposed by Kirsh & Maglio (1994) to explain skilled performance. Our research suggests that with a novel task, even if reliance on the external visualization is the optimal strategy, not all individuals will discover this strategy.

An alternative explanation for why people did not access the arrow view can be formulated within the soft constraints hypothesis (Gray & Fu, 2004; Gray et al., 2006). According to this hypothesis, external resources are not privileged, and the individual chooses on a momentary basis whether to rely on internal or external resources. In this framework, a user performing our task must decide whether to imagine the arrow view using internal resources (e.g., spatial working memory) or rotate the external model to see the arrow view. Whereas the internal computations are clearly demanding, in our task, manipulating the external visualization to the arrow view is also non-trivial. For example, the user has to mentally compute the axis of rotation, generate a motor program to execute that rotation and use feedback from the visual display on the monitor or from the

musculature to decide when the model has been rotated by 90 degrees from its starting position. Thus, even manipulation of the external visualization depends on internal cognitive resources. It is possible then that when users weigh the cognitive costs of imagining the arrow view versus manipulating the external visualization to see the arrow view, there is not a clear winner.

It was somewhat surprising that effective use of the external visualization was unrelated to spatial ability in Experiments 2 and 3. In theory, we might expect high-spatial individuals to be better able to discover the strategy of rotating the visualization to the arrow view, because this strategy appears to depend on an understanding of spatial relations. One possibility is that this strategy requires a metacognitive understanding of the task that is not systematically related to spatial ability, and that this metacognitive understanding was lacking in some individuals in both the high-spatial and low-spatial groups. Alternatively, a soft constraints account would argue that decisions about whether to use internal or external resources depend on the relative effort of each (Gray et al., 2006). Both internal visualization and use of external visualizations are likely to be very effortful for low-spatial individuals and less effortful for high-spatial individuals, but the *differential effort* involved in internal versus external computations might not vary as a function of spatial ability. This might explain why some participants in both the high-spatial and low-spatial groups chose to access the arrow view while others did not, presumably choosing to rely on internal computations instead. Finally, it is possible that the soft-constraints trade-off applies only to the higher-spatial individuals, for whom this choice was apparent and thus easily made, whereas among lower-spatial individuals a failure to access the arrow view may have been due to metacognitive limitations. From the

present data we cannot establish which account provides the best explanation, but this is an important issue for future research.

*Benefiting From the Information Provided by the Visualization.* A striking result of our experiments is that viewing rotations of the external representation into the viewpoint of the arrow helped performance in general. Importantly, this was true regardless of whether the user actively produced that rotation or passively viewed it. Even non-interactive participants, who did not physically execute the perceptual-motor actions to get to this view, benefited from the critical information.

It is somewhat surprising that actively selecting the arrow view does not enhance performance relative to passively viewing it, given that interactive and non-interactive conditions differ not just in agency, but also in control over the timing of the manipulations. Interactive control allows the user to rotate the external visualization in his or her own time, such that the actions performed on the external visualization are tightly coupled with his or her internal cognitive processes (cf. Hollan et al., 2000). One of our results suggested that the timing of viewing the task-critical information may be somewhat important (although it did not lead to an overall significant difference between interactive and non-interactive participants). Notably, for the yoked non-interactive participants in Experiment 2, the arrow view was more beneficial if it was reached later in the trial. It is possible that on these trials, the rotation of the visualization to the arrow view was more gradual. Unlike individuals who are executing the rotations, passive observers do not have meta-knowledge of the user's goals and planning processes, and therefore cannot predict what will happen next. Thus they may need more time to make sense of what is seen on-screen and to update their internal representations to match the unpredictable changes occurring in the external

information. This finding may have implications for the design of non-interactive visualizations generally.

Spatial ability did not affect whether participants accessed the arrow view, but did it affect how much they benefited from this view once it was accessed? It is important to remember that seeing the visualization from the perspective of the arrow is just one step in the task, and participants still have to imagine the visualization being sectioned, etc. once they have accessed this view. In a post-hoc analysis of the combined data from the interactivity conditions of Experiments 2 and 3 (which were identical in terms of experimental procedure), we examined whether task performance of high-spatial and low-spatial individuals was differently affected by amount of access to the critical view (groups were formed using a median split;  $N = 26$  in both). Among higher spatial participants, time spent on the arrow view was strongly correlated with task performance,  $r_{ho} = .84, p < .001$ , whereas among lower spatial participants time spent on the arrow view was not significantly related to performance,  $r_{ho} = .28, p = .08, ns$ . These correlation coefficients differ significantly,  $z_{rho} = 3.08, p < .01$  (Feiller, Hartley, & Pearson, 1957).

Earlier, we raised three possible ways in which external visualizations might affect individuals with different spatial abilities (Hegarty, 2004). They may compensate for the poorer internal visualization skills of low-spatial individuals, their use may depend on the stronger internal visualization skills of high spatial individuals, or they may augment performance for individuals of all levels of spatial ability. In general, the effects of spatial ability were independent of the experimental manipulations in our experiments. Thus there is no evidence that an external visualization acts as a cognitive “prosthetic” such that it can compensate for low internal visualization abilities. In fact, our post-hoc analysis of the

combined interactive conditions of Experiments 2 and 3 indicates that high spatial individuals in the interactive groups benefited more from seeing the arrow view compared to low spatial individuals in the same group. Moreover, the very high correlation of spatial ability with performance in the arrow views condition of Experiment 3 (.79) suggests that when the critical view was available on every trial, as in this condition, spatial ability was especially important to performance. Thus, if anything, being able to benefit from the information provided by external visualizations may depend on high spatial ability, rather than external visualizations compensating for low spatial ability.

*Implications.* Our results may help to explain why there have been inconsistent findings in previous studies on interactivity, with some showing that interactivity can help, others showing that it can hurt, and still others showing no effects. Our experiments indicate that if the most appropriate manipulations of an interactive visualization are not transparent, as in our task, then more interactivity may not help because people may not use it effectively or because the costs of using the external visualization are not significantly lower than the costs of performing internal computations. Furthermore, there are individual differences between people in how effectively they use interactive visualizations, so the inconsistent results might reflect differences among participants in different studies.

How might we enhance performance with this type of visualization? One possible method of improving performance might be to redesign the visualization to make its task-relevant affordances more intuitively obvious. Our comparison of Experiments 1 and 2 is relevant to this question. Interactive participants in Experiment 1 performed better than their interactive counterparts in Experiment 2. This difference may be due to the more constrained interface in Experiment 1, which effectively restricted possible rotations to the

two axes that contained the arrow view for horizontal and vertical cross sections. It is likely that this “hard constraint” (Gray et al., 2006) changed the metacognitive demands of the task, allowing individuals to more easily discover the critical view, and thus reduced the cognitive costs of using the external visualization. Consistent with this interpretation, in a related study of the cross section task in which participants used the same interface as in Experiment 1 (Cohen & Hegarty, 2007), the majority of participants accessed the key view on at least three quarters of the trials, a rate of access that is substantially higher than we observed in the present study. Thus our results are consistent with the somewhat counterintuitive theoretical claim (Betrancourt, 2005; Krygier et al., 1997) that limited interactivity may sometimes be more effective than full interactivity. This point is reminiscent of research by Smallman and St. John (2005) regarding misplaced faith in 3-D displays, in which they show that performance with 2-D displays is superior for many tasks.

Another possible method of improving performance with interactive visualizations is to instruct people on how to use them effectively. It seems plausible that one could teach the types of metacognitive understanding necessary to use external visualizations effectively. In the interactive conditions of our experiments, participants had no chance to discover the benefits of seeing the arrow view if they did not rotate the visualization to that view. One interesting question is whether merely exposing students to the task relevant information (as in the non interactive condition of Experiment 3) would be sufficient for them to seek it out later when given an interactive visualization, or whether students need to be explicitly taught strategies for accessing and using the critical information.

Although our research suggests that a non-interactive visualization showing the task-critical information can be superior to an interactive visualization, it is important to

acknowledge that this may be true only if the visualization is to be used only for a specific task (as in these experiments) or a small number of specific tasks. Of course, an important benefit of interactivity is that it allows users the freedom to manipulate visualizations differently to solve different problems. The relative advantages of interactive versus non-interactive visualizations probably depend on the range of tasks for which these visualizations will be used, and more interactive visualizations may be more effective when they must be used to solve a range of problems. Nevertheless, in these situations it is important to keep in mind, as we have learned, that people will not always spontaneously discover the affordances of interactive visualizations and may need to be explicitly taught how to use them.

*Conclusion.* In conclusion, this research has shown that, contrary to many people's intuitions, providing interactive control of computer visualizations does not necessarily enhance performance in spatial reasoning tasks. There are individual differences in how effectively people use these visualizations and the extent to which they offload cognition onto these external aids, rather than performing internal computations. As a result, different individuals receive different visual information from external visualizations, and in the end what matters is whether they access the key task-relevant information, not whether they had active control. As we continue to design interactive visualizations for cognitive performance, it is important to recognize that more interactivity may not always be better, manipulating external visualizations may not always be less effortful than relying on internal visualization, and people may not always discover the affordances of external visualizations without explicit instruction.

## References

- Attree, E. A., Brooks, B. M., Rose, F. D., Andrews, T. K., Leadbetter, A. G., and Clifford, B.R. (1996). Memory processes and virtual environments: I can't remember what was there but I can remember how I got there. *Proceedings of the 1<sup>st</sup> European Conference on Disability, Virtual Realty & Associated Technologies* (pp.117-121).
- Ballard, D. H., Hayhoe, M. M., Pook, P. K., and Rao, R. P. N. (1997). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences*, 20, 723-742.
- Betrancourt, M. (2005). The animation and interactivity principles in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 287-296). New York, NY: Cambridge University Press.
- Card, S. K., MacKinlay, J. D., & Shneiderman, B. (1999). *Readings in information visualization: Using vision to think*. San Francisco: Morgan Kaufmann.
- Christou, C. G., & Bühlhoff, H. H. (1999). View dependence in scene recognition after active learning. *Memory & Cognition*, 27, 996-1007.
- Cohen, C. A., & Hegarty, M. (2007). Individual differences in use of an external visualization to perform an internal visualization task. *Applied Cognitive Psychology*, 21, 1-11.
- Eliot, J., & Smith, I. M. (1983). *An international directory of spatial tests*. Windsor, Berks: Nfer-Nelson.
- Feiller, E. C., Hartley, H. O., and Pearson, E. S. (1957). Test for rank correlation coefficients. I. *Biometrika*, 44(3/4), 470-481.

- Feldman, A., & Acredolo, L. P. (1979). The effect of active versus passive exploration on memory for spatial location in children. *Child Development, 50*, 698-704.
- Foreman, N., Sandamas, G., and Newson, D. (2004). Distance underestimation in virtual space is sensitive to gender but not activity-passivity or mode of interaction. *CyberPsychology & Behavior, 7*(4), 451-457.
- Gordin, D. N., & Pea, R. D. (1995). Prospects for scientific visualization as an educational technology. *The Journal of the Learning Sciences, 4*, 249-279.
- Gray, W. D., & Fu, W.-T. (2004). Soft constraints in interactive behavior: The case of ignoring perfect knowledge in-the-world for imperfect knowledge in-the-head. *Cognitive Science, 28*(3), 359-382.
- Gray, W. D., Sims, C. R., Fu, W.-T., & Schoelles, M. J. (2006). The soft constraints hypothesis: A rational analysis approach to resource allocation for interactive behavior. *Psychological Review, 113*(3) 461-482.
- Harman, K. L., Humphrey, G. K., & Goodale, M. A. (1999). Active manual control of object views facilitates visual recognition. *Current Biology, 9*, 1315-1318.
- Hegarty, M. (2004). Commentary: Dynamic visualizations and learning: Getting to the difficult questions. *Learning and Instruction, 14*, 343-351.
- Hegarty, M. & Waller, D. (2005). Individual differences in spatial abilities. In P. Shah & A. Miyake (Eds.). *The Cambridge Handbook of Visuospatial Thinking*. Cambridge University Press (pp. 121 – 169).
- Hollan, J., Hutchins, E., & Kirsh, D. (2000). Distributed cognition: Toward a new foundation for human-computer interaction research. *ACM Transactions on*

*Computer-Human Interaction*, 7(2), 174-196.

Hutchins, E. (1995). How a cockpit remembers its speed. *Cognitive Science*, 19, 265-288.

Hutchins, E. L., Hollan, J. D., & Norman, D. A. (1985). Direct manipulation interfaces.

*Human-Computer Interaction*, 1, 311-338.

James, K. H., Humphrey, G. K., & Goodale, M. A. (2001). Manipulating and recognizing virtual objects: Where the action is. *Canadian Journal of Experimental Psychology*, 55, 111-120.

James, K. H., Humphrey, G. K., Vilis, T., Corrie, B., Baddour, R., & Goodale, M. A.

(2002). "Active" and "passive" learning of three-dimensional object structure within an immersive virtual reality environment. *Behavior Research Methods, Instruments, & Computers*, 34, 383-390.

Kali, Y., & Orion, N. (1996). Spatial abilities of high school students in the perception of geologic structures. *Journal of Research in Science Teaching*, 33, 369-391.

Kirsh, D. (1997). Interactivity and multimedia interfaces. *Instructional Science*, 25, 79-96.

Kirsh, D. (2004). Metacognition, distributed cognition and visual design. In P. Gardinfors & P. Johansson (Eds). *Cognition, Education, and Communication Technology*. Lawrence Erlbaum.

Kirsh, D. & Maglio, P. (1994). On distinguishing epistemic from pragmatic action.

*Cognitive Science*, 18(4), 513-549.

Krygier, J., Reeves, C., Cupp, J., & DiBiase, D. (1997). *Multimedia in geographic*

*education: Design, implementation, and evaluation*. Retrieved March 16, 2006,

[http://horizon.unc.edu/projects/monograph/CD/Science\\_Mathematics/Krygier.asp](http://horizon.unc.edu/projects/monograph/CD/Science_Mathematics/Krygier.asp)

- Lord, T. R. (1985). Enhancing the visuo-spatial aptitude of students. *Journal of Research in Science Teaching*, 22, 395-405.
- Lowe, R. K. (1999). Extracting information from an animation during complex visual learning. *European Journal of Psychology of Education*. 14, 225-244.
- Lowe, R. K. (2004). Interrogation of a dynamic visualization during learning. *Learning and Instruction*, 14, 257-274.
- Marchak, F. M., & Marchak, L. C. (1991). Interactive versus passive dynamics and the exploratory analysis of multivariate data. *Behavior Research Methods, Instruments, & Computers*, 23, 296-300.
- Marchak, F. M., & Zulager, D. D. (1992). The effectiveness of dynamic graphics in revealing structure in multivariate data. *Behavior Research Methods, Instruments, & Computers*, 24, 253-257.
- Melanson, B., Kelso, J., & Bowman, D. (2002). *Effects of active exploration and passive observation on spatial learning in a CAVE*. Retrieved October 2005, from <http://eprints.cs.vt.edu/archive/00000602/>
- Peruch, P., Vercher, J.L., & Gauthier, G. M. (1995). Acquisition of spatial knowledge through visual exploration of simulated environments. *Ecological Psychology*, 7(1), 1-20.
- Philbeck, J. W., Klatzky, R. L., Behrmann, M., Loomis, J. M., & Goodridge, J. (2001). Active control of locomotion facilitates nonvisual navigation. *Journal of Experimental Psychology: Human Perception and Performance*, 27, 141-153.

- Richardson, B. L., Wuillemin, D. B., & MacKintosh, G. J. (1981). Can passive touch be better than active touch? A comparison of active and passive tactile maze learning. *British Journal of Psychology*, *72*, 353-362.
- Rieber, L. P., Tzeng, S-C., & Tribble, K. (2004). Discovery learning, representation, and explanation within a computer-based simulation: Finding the right mix. *Learning and Instruction*, *14*, 307-323.
- Schwan, S. & Riempp, R. (2004). The cognitive benefits of interactive videos: learning to tie nautical knots. *Learning and Instruction*, *14*, 293-305.
- Schwartz, D. L. & Holton, D. L. (2000). Tool use and the effect of action on the imagination. *Journal of Experimental Psychology: Learning Memory and Cognition*, *26*, 1655-1665.
- Shepard, R. N. & Cooper, L. A. (1982). *Mental images and their transformations*. Cambridge, MA: MIT Press.
- Shneiderman, B. (1983). Direct manipulation: A step beyond programming languages. *Computer*, *16*(8), 57-69.
- Smallman & St. John (2005) Naïve realism: Misplaced faith in realistic displays. *Ergonomics in Design: The Quarterly Journal of Human Factors Applications*, *13*(3), 6-13.
- Wang, R. X.-F., & Simons, D. J. (1999). Active and passive scene recognition across views. *Cognition*, *70*, 191-210.
- Wexler, M., Kosslyn, S. M. & Berthoz, A. (1998). Motor processes in mental rotation. *Cognition*, *68*, 77-94.

- Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin and Review*, 9, 625-636.
- Wilson, P. N. (1999). Active exploration of a virtual environment does not promote orientation or memory for objects. *Environment and Behavior*, 31, 752-763.
- Wilson, P. N., Foreman, N., Gillett, R., & Stanton, D. (1997). Active versus passive processing of spatial information in a computer simulated environment. *Ecological Psychology*, 9, 207-222.
- Wilson, P. N. & Peruch, P. (2002). The influence of interactivity and attention on spatial learning in a desk-top virtual environment. *Current Psychology of Cognition*, 21(6), 601-633.
- Wohlschlager, A., & Wohlschlager, A. (1998). Mental and manual rotation. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 397-412.
- Zhang, J. & Norman, D.A. (1994). Representations in distributed cognitive tasks. *Cognitive Science*, 18, 87-122.

## Author Note

Madeleine Keehner, Mary Hegarty, Cheryl Cohen, and Peter Kooshabeh, Department of Psychology, University of California, Santa Barbara; Daniel R. Montello, Department of Geography, University of California, Santa Barbara. Madeleine Keehner is now at the School of Psychology, University of Dundee.

This research was supported by National Science Foundation grant 0313237. We thank Jerome Tietz for his programming expertise and Meg Vallas for assistance with data coding. We are also grateful to Wayne Gray, Dan Schwartz, and three anonymous reviewers for their very insightful and expert reviews, which led to significant improvements especially to the theoretical aspects of this paper. Portions of this research were presented at the 2005 meeting of the Psychonomic Society in Toronto, Canada, the 2004 meeting of the Cognitive Science Society in Chicago, IL, the Multi-Disciplinary Workshop “Teaching Geoscience with Visualizations: Using Images, Animations, and Models Effectively” at Carleton College in Northfield, MN (2004), and at the Fall Meeting of the American Geophysical Union in San Francisco, CA (2004).

Correspondence concerning this article should be addressed to Dr Madeleine Keehner, School of Psychology, University of Dundee, Dundee, DD1 4HN, Scotland, UK. Email: M.M.Keehner@Dundee.ac.uk.

## Footnotes

<sup>1</sup> The steps are not necessarily carried out in this order. Note that although we describe the task in terms of mental imagery processes, we do not claim that it is solved exclusively by analog transformations of holistic images (cf., Shepard & Cooper, 1982) and the experiments reported here are not designed to directly address the nature of the internal representations and processes involved.

<sup>2</sup> Computer visualizations can also use stereopsis to convey depth, but we did not, because we wished to generalize to situations in which people learn about 3-D structures in anatomy, geology, and data analysis, in which stereo viewing is not typically available.

<sup>3</sup> This end tone would sometimes sound while a non-interactive participant was still working, depending on the duration of the playback and the time they took to complete the task.

<sup>4</sup> In all analyses in this paper that used parametric tests, the data were normally distributed. Non-parametric tests were used, as noted, when the distribution was not normal.

Table 1. *Correlations Between Drawing Measures and Spatial Ability in All Experiments*

	Number of Ducts	Outer Shape	Duct Angles	Duct Positions
Experiment 1 (N = 60)				
Spatial ability	.30*	.38**	.38**	.42**
Number of ducts		.20	.58**	.50**
Outer shape			.36**	.41**
Duct angles				.84**
Experiment 2 (N = 60)				
Spatial ability	.33*	.08	.41**	.51**
Number of ducts		.24	.48**	.27*
Outer shape			.24	.24
Duct angles				.68**
Experiment 3 (N = 60)				
Spatial ability	.48**	.17	.65**	.63**
Number of ducts		-.02	.68**	.53**
Outer shape			.31*	.35**
Duct angles				.82**

\*p&lt;.05. \*\*\*p&lt;.01.

## Figure Captions

*Figure 1.* *a)* Printed stimulus of 3-D object showing the cutting plane and viewing-direction arrow for one trial. *b)* Illustrations of the on-screen object in various orientations. The arrow view for the trial shown is the fourth image down. *c)* Correct cross section for the trial shown. The superimposed dotted lines show the correct angles among the ducts and the superimposed 10 x 10 grid shows the correct positions of the ducts within the cross-section. A drawing was correct on the *Duct Angles* criterion if the angles among the drawn ducts differed from the correct angles by not more than 20 degrees on average per angle (up to 60 degrees total error was therefore acceptable for three-duct cross-sections). A drawing was correct on the *Duct Positions* criterion if the locations of the drawn ducts differed from the correct locations by not more than one grid square, or one tenth of the cross-section's overall area, in any direction. *d – f)* Sample drawings by participants for the trial shown. The drawing in example *d* passed both duct angles and duct positions; the drawing in example *e* failed both duct angles and duct positions; the drawing in example *f* passed duct angles but failed duct positions. All three drawings passed the number of ducts criterion but only drawing *d* passed the outer shape criterion. *g)* A range of sample cross-sections used in the studies. The cross-sections were derived by slicing the virtual 3-D object horizontally or vertically at different points, and rendering the resulting cross-section as if viewed from above, below, left, or right of the object.

*Figure 2.* Sample item from the adapted Guay's Visualization of Views Test. The correct answer is the lower right corner.

*Figure 3.* Proportion correct on the combined duct relations measure in the two conditions of Experiment 1 (interactive and non-interactive), by spatial ability (high-low median split). Error bars represent +/- 1 standard error of the mean.

*Figure 4.* Experimental set up in the interactive conditions of Experiments 2 and 3. The egg-shaped control device contained a 3 degrees-of-freedom motion tracker, so that rotations made with this device were replicated by the on-screen object in real time.

*Figure 5.* Distribution of trials in which interactive participants in Experiment 2 accessed the arrow view. The vertical axis shows frequency (number of participants). The horizontal axis shows proportion of trials on which the arrow view was deemed to have been purposely accessed, i.e., the visualization remained within the critical window (arrow view angle +/- 20°) for at least 2 seconds in total.

*Figure 6.* Proportion correct on the combined duct relations measure in the two conditions of Experiment 3 (interactive, arrow views). Interactive participants are separated by amount of access to arrow view, i.e., individuals who accessed the arrow view more than the median versus those who accessed it less than the median (participants who accessed it the median number of times were not included). Error bars represent +/- 1 standard error of the mean.

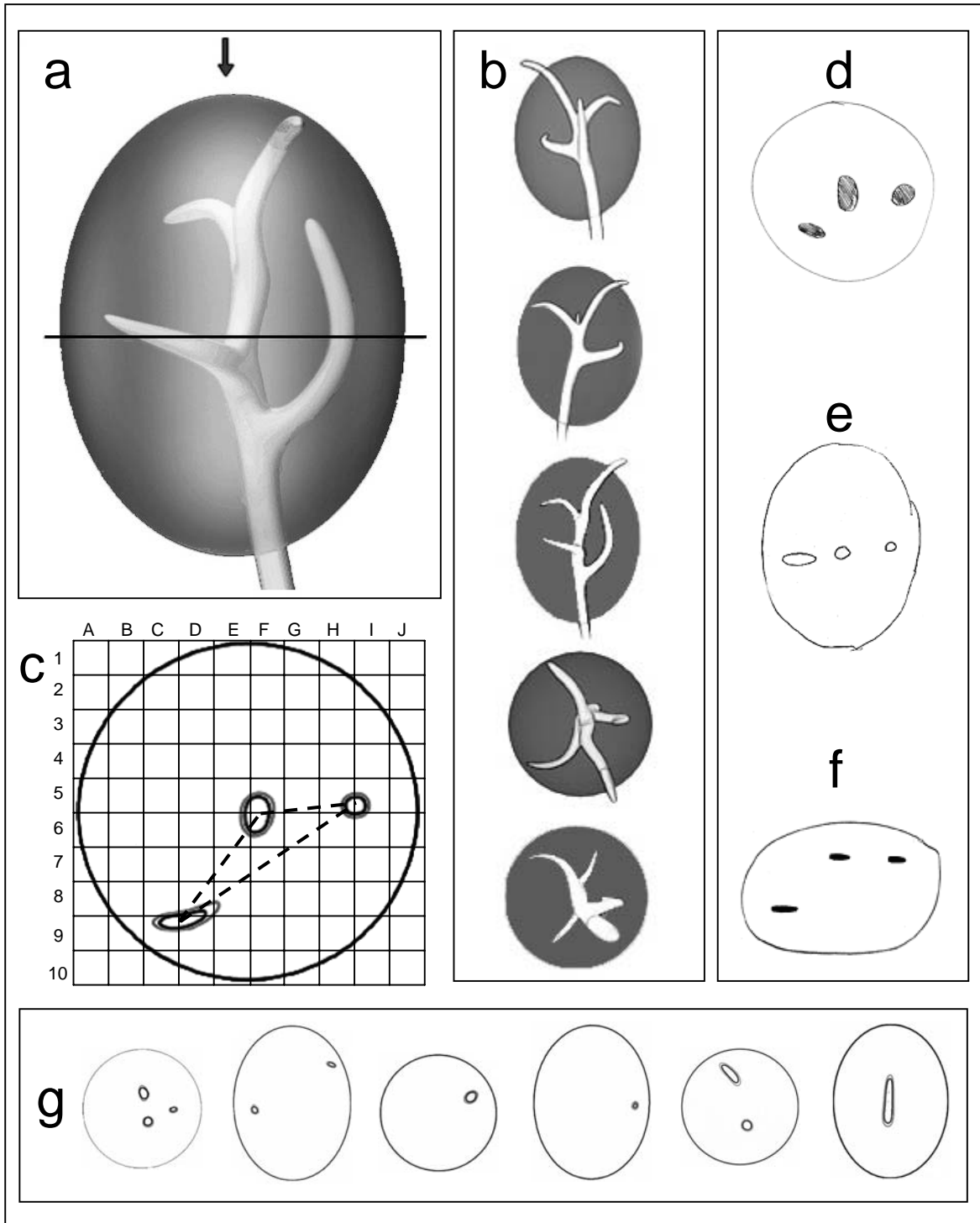


Figure 1

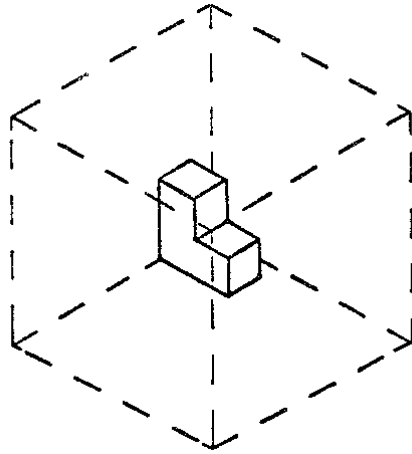


Figure 2

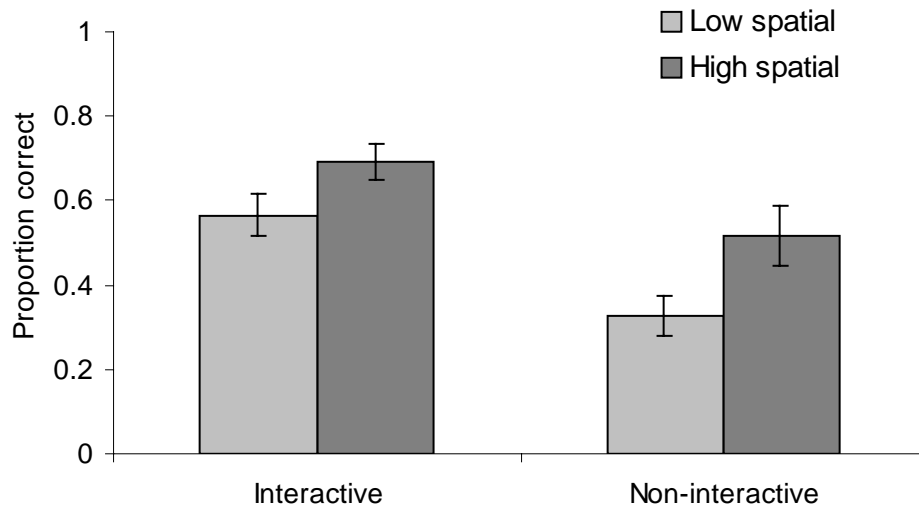


Figure 3



Figure 4

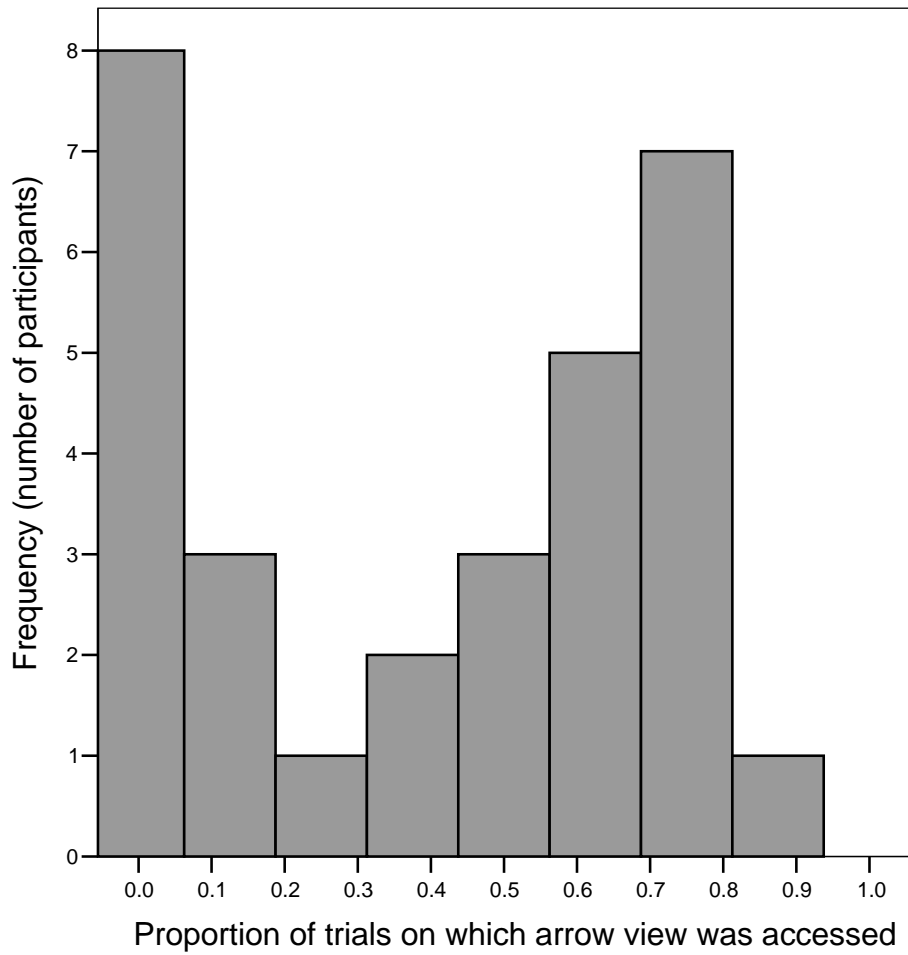


Figure 5

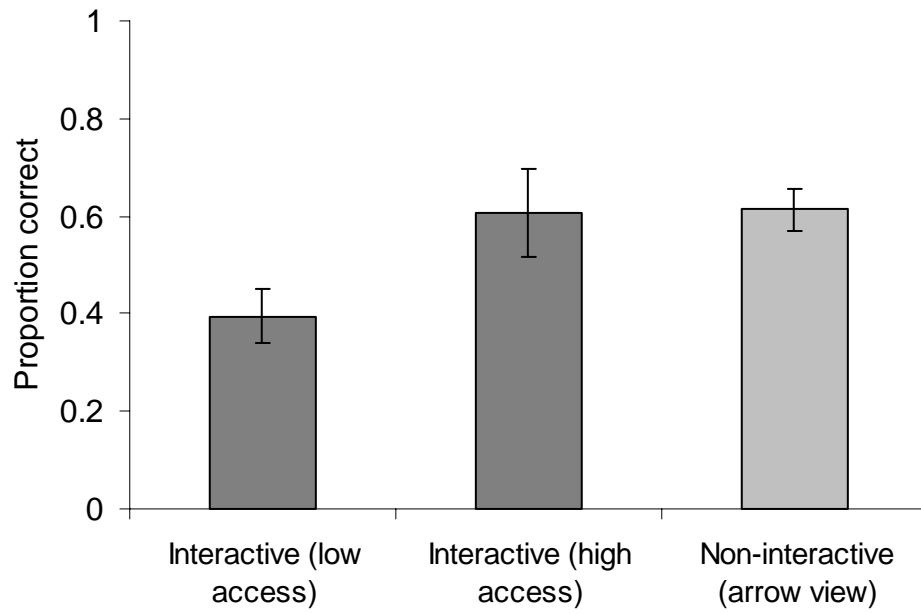


Figure 6