Abstract

Virtual models are increasingly employed in STEM education to foster learning about spatial phenomena. However, the roles of the computer interface and students’ cognitive abilities in moderating learning and performance with virtual models are not yet well understood. In two experiments students solved spatial organic chemistry problems using a virtual model system. Two aspects of the virtual model interface were manipulated: display dimensionality (stereoscopic vs. monoscopic displays) and the location of the hand-held device used to manipulate the virtual molecules (co-located with the visual display vs. displaced). The experimental task required participants to interpret the spatial structure of organic molecules and to manipulate the models to align them with orientations and configurations depicted by diagrams in Experiment 1 and three-dimensional models in Experiment 2. Co-locating the interaction device with the virtual image led to better performance in both experiments and stereoscopic viewing led to better performance in Experiment 2. The effect of co-location on performance was moderated by spatial ability in Experiment 1, and the effect of providing stereo viewing was moderated by spatial ability in Experiment 2. The results are in line with the ability-as-compensator hypothesis: participants with lower ability uniquely benefited from the treatment, while those with higher ability were not affected by stereo or co-location. The findings suggest that increased fidelity in a virtual model system may be one way of alleviating difficulties of low-spatial participants in learning spatially demanding content in STEM domains.

Keywords: interface design, spatial ability, virtual environments, chemistry education, molecular models, stereo
1. Introduction

Imagine you are in an organic chemistry classroom looking at a diagram of a complex molecule drawn on the blackboard. Would you be able to form an accurate mental representation of the three-dimensional molecular structure that the diagram represents? How about performing mental transformations in order to infer the reactivity of the molecule? These are the type of skills that organic chemistry students must master. Now imagine again you are in the classroom using a virtual molecular model. Would your understanding of the spatial structures improve if you could interactively transform the molecule using a virtual display? Does the level of perceptual fidelity or realism of the model affect performance and learning?

These sorts of questions are becoming important as computer-based virtual models are now widely used in science, technology, engineering, and mathematics (STEM) education and research (Dede, 2009). Virtual models allow for interaction with representations of many scientific phenomena. For example, anatomists can be found dissecting virtual organs (Grantcharov, et al., 2004), geologists building virtual representations of outcrop formations (Thurmond, Drzewiecki & Xu, 2005), and biochemists viewing virtual ligand binding sites (Anderson & Weng, 1999). Current hardware and software technologies allow learners and researchers to interactively manipulate, navigate, and create digital renderings of high fidelity scientific visualizations. Three-dimensional (3D) virtual models have been shown to be effective in conveying a number of difficult spatial concepts such as the structure of organic molecules (Stull, Barrett, & Hegarty, 2013), geological formations (Piburn, et al., 2005), medical anatomy (Nicholson, Chalk, Funnell, & Daniel, 2006), and architecture and planning (Horne & Thompson, 2008). However, empirically driven guidelines for the design of educational virtual environments are lacking. Moreover, given the interest in the importance of spatial ability and participation in STEM domains (National Research Council, 2006), it is important to design virtual environments that do not place undue cognitive demands on students with poor spatial abilities.

The virtual model interface often includes the visual display and the device used to interact with the model. Interface designs often vary from one system to another, making it difficult to generalize findings (Limniou, Roberts, & Papadopoulos, 2008; Reimer & Moyer, 2005; Trindade, Fiolhais, & Almeida, 2002). For example, the computer display can vary by properties such as resolution, field of view, color reproduction, refresh rate, and whether the image is
monoscopic or stereoscopic. Virtual environments also differ in the affordances of various interaction devices such as the mouse and keyboard, joysticks, or motion-tracked direct manipulation devices and in the location of the interface relative to the display.

The purpose of this work was to investigate the effects of two aspects of virtual interface design: (1) the dimensionality of the visual display and (2) the position of a motion tracked, direct manipulation device (hereafter referred to as the interaction device) on performance in a virtual organic chemistry task. With respect to display fidelity dimensionality, we compared performance with monoscopic viewing versus stereoscopic viewing. Regarding the position of the interaction device, we compared performance when the device and virtual image were co-located versus when the device was in a different spatial location from the virtual image. We also examined the effects of spatial ability on task performance and the role it plays in moderating effects of virtual interface design on performance.

1.1. Organic Chemistry as a Test Bed

The task used in the present study is situated in the domain of organic chemistry. Understanding molecular structures is an essential skill that all organic chemists must possess in order to learn, research, and communicate about their science. Diagrams and 3D models serve as a mode of communication, and are therefore vital in understanding organic molecular structures and making advancements in the field (Kozma & Russell, 2005). Organic chemistry representations are an excellent real-world context for studying spatial reasoning and problem solving. Virtual molecular models are commonly employed in chemistry education to teach students about the 3D structure of molecules and for introducing various spatial representations (diagrams, formulas etc.) (Barnea, 2000; Limniou et al., 2008). However, little research has investigated issues of interface design for interacting with virtual models in chemistry. We chose the task of matching the orientation and conformation (i.e. local configuration) of molecular diagrams and models because this is an important and common exercise to teach the correspondence between diverse representations in chemistry (Barrett, Stull, Hsu, Hegarty, 2015; Stull, Hegarty, Dixon, & Stieff, 2012; Wu & Shah, 2004). The findings of this work inform the specific design of cognitively supportive virtual models for chemistry education and also reveal more general design implications for virtual interactions in other disciplines that rely on spatial representations.
2. Virtual Model Interface Design

2.1 Stereoscopic displays

Investigations of performance with 3D stereo displays span multiple literatures including human-computer interaction and human factors (Patterson, 2007; Lambooij, Fortuin, Heynderickx, & IJsselsteijn, 2009), vision science (Held & Banks, 2008, Shibata, Kim, Hoffman & Banks, 2011), and medicine (Hofmeister, Frank, Cuschieri, & Wade, 2001; Held & Hui, 2011). A wide variety of experimental tasks, such as navigation of virtual environments (Dahmani, Ledoux, Boyer, & Bohbot, 2012), identification of objects (Dent, Braithwaite, He, & Humphreys, 2012), and spatial manipulations such as translation (Arsenault & Ware, 2004) and pointing (Hu & Knill, 2011) have been employed in assessing the impact of providing stereo viewing over traditional 2D monoscopic displays. A recent review of the research on stereoscopic displays found that 3D stereo viewing improved performance over monoscopic viewing in 60% of the 184 experiments reviewed, 15% found mixed results, and 25% found null effects (McIntire, Havig, & Geiselman, 2014). The authors point out that stereo viewing was most beneficial for tasks involving spatial manipulations, with 67% of studies finding performance benefits. However, the types of spatial manipulations involved were not specified.

2.2 Co-location of visual and haptic workspaces

Research on performance with co-located haptic and visual workspaces, a less extensive body of literature, has produced similarly mixed findings. Ware and Rose (1999) demonstrated that co-locating the interaction device and virtual image led to 35% faster response times compared to displacing the interaction device from the virtual image for a virtual object orientation matching task. Arsenault and Ware (2004) observed disrupted performance when the task involved rotational mismatches between the interaction device and virtual image, and disruption increased with the magnitude of the mismatch. Swapp, Pawar, and Loscos (2006) found significant benefits of co-location in two 3D tasks with varying levels of difficulty. Lev, Rozengurt, Gelfeld, Tarkhnishvili, and Reiner (2010) reported co-location led to faster performance during an endoscopic surgery task. Other researchers have found no benefit of co-location in both 2D (Sprague, Po, & Booth, 2006) and 3D (Teather, Allison, and Steurzlinger, 2009; van Liere, Kok, & van Tienen, 2005) tasks. Further, combining high fidelity display and input technologies has been shown interact to greatly boost performance on specific spatial tasks (Ragan, Kopper, Schuchardt, & Bowman, 2013).
2.3 Motivation

In considering the effects of interface design on performance in virtual environment tasks, it is important to distinguish between different types of spatial manipulations, because various perceptual cues may have differential effects on components of spatial manipulations, such as translating, rotating, and scaling (Ware, 2012). The benefit of providing stereo viewing may be less pronounced for virtual object rotations than for translations in depth. For example, monocular depth cues such as relative motion, relative size, linear perspective, and shadow can provide rich spatial information depending on the characteristics of the virtual environment. Stereo viewing provides an additional depth cue of binocular disparity compared to monoscopic viewing; however, this additional cue may not always be task relevant. Ware (2012) has suggested that monocular cues alone often provide sufficient information to support successful virtual object manipulations, and that the additional information provided by stereo viewing may not always be worth the cost of implementation.

Previously studied virtual object rotation tasks have typically looked at global (i.e. rigid) manipulations of a solid virtual object using a hand-held 3D interaction device (e.g. Arsenault & Ware, 2004; Ruddle & Jones, 2001; Ware & Rose, 1999). In the present studies, inspired by the demands of learning about molecular structure and reactivity in organic chemistry, a more complex virtual object rotation task was examined. The task involved typical manipulations of molecules requiring additional twisting or “local” rotations of parts of the molecule with respect to each other, in addition to rigid global rotations of the whole molecule. The two halves of the interaction device were rotated using a twisting motion in order to configure each half of the virtual model about the long axis. The additional local rotations contributed externally valid task complexity and allowed for examination of the effects of stereo viewing and co-location in a more challenging task scenario than previously examined.

3. Theoretical Framework and Predictions

3.1 Perceptually mediated interfaces

From human-computer interaction work on direct manipulation interfaces (Huchins, Hollan, & Norman, 1985) to more recent work on tangible user interfaces (Shaer & Hornecker; 2010), a strong case has been made for the importance of embodiment and naturalism in designing digital interactions. Multimedia and educational interface researchers have employed concepts of
cognitive load theory (Sweller, 1994) to form a human-centered design framework (Chalmers, 2003; Oviatt, Coulston, Lunsford, 2004; Hollender, Hoffman & Deneke, 2010).

Virtual interactions impose varying levels of cognitive demand depending on whether the interface is more cognitively or perceptually mediated. More demanding cognitively mediated interfaces often provide conflicting sensory cues, whereas more perceptually mediated interfaces reduce extraneous cognitive load by providing analogous perceptual cues to the real-world situation. For example, using a mouse to rotate a virtual object in 3D space is one example of a cognitively mediated interaction. Physical interaction with the mouse produces translations on an x-y plane which are often mapped to rotations about the x- and y-axes of the virtual object. While most people are able to adapt to the interface, this type of mapping between different spatial transformations increases cognitive demand due to the additional mental operations required to reconcile the mismatched perceptual cues from the interaction device (object translation as input) and the visual representation (object rotation as output). On the other hand, more perceptually mediated interfaces, often characterized as high fidelity or naturalistic interfaces, often provide sensory cues that are analogous or similar to the corresponding real-world interaction. When input motion from the interaction device is perceptually matched to the outcome, for example, rotating a motion tracked stylus to rotate a virtual object, demanding mental operations can be offloaded onto more automatic perceptual systems and reduce extraneous cognitive load. In the context of ultrasound guided surgery, Klatzkly, Wu and Stetton (2008) demonstrated that an ‘in-situ’ perceptual display improved performance and aided in forming accurate spatial mental representations.

Cognitive load theory also predicts that providing perceptually mediated interfaces should foster better performance and learning over cognitively mediated interfaces, because when cognitive demand is offloaded onto automatic systems, resources can be re-allocated to improve task performance and/or learning. Within this framework, providing stereo viewing and co-locating the interaction device and virtual image can be interpreted as methods of increasing perceptual mediation of a virtual model, and consequently, we predicted better performance in our virtual molecule manipulation task over the more cognitively mediated interfaces (i.e., monoscopic viewing, displaced interaction device).
3.2 Aptitude-treatment interactions

The heterogeneity of findings regarding the effectiveness of interface design factors such as display fidelity and co-location in the virtual interface design literature may be due in part to interface design elements differentially affecting individuals at different levels of spatial ability. While it is widely accepted that individual differences in spatial ability play an important role in understanding 3D representations, the nature of aptitude-treatment-interactions (ATIs) is often disputed in the educational literature (Höffler, 2010). Two general hypotheses have been offered in literature spanning research on multimedia learning, animations, and interactivity.

On one hand, the ability-as-enhancer hypothesis predicts that high ability is required to benefit from interfaces with greater fidelity, whereas low ability individuals are hindered due to increased demands of interaction or processing of visual detail. Studies have provided evidence supporting the ability-as-enhancer hypothesis by demonstrating that high ability individuals uniquely benefit when provided with additional verbal information (Mayer & Sims, 1994), visualization interactivity (Keehner, Montello, Hegarty, & Cohen, 2004), and 3D models in virtual biology lessons (Huk, 2006).

In contrast, the ability-as-compensator hypothesis predicts that individuals with greater ability should be able to compensate for impoverished perceptual cues, and perform similarly regardless of the affordances of the interface. According to this hypothesis, individuals with lower ability should uniquely benefit from perceptually mediated interfaces because the external representation may supplant or compensate for lack of ability. Previous studies have found ATIs supporting the ability-as-compensator hypothesis by demonstrating cases where low ability individuals differentially benefited from spatial contiguity of verbal and visuospatial information (Lee, 2007) and when viewing animations rather than diagrams (Hays, 1996; Höffler & Leutner, 2011).

The dichotomy between ability-as-enhancer and ability-as-compensator as described in the paragraphs above is an oversimplification, and the moderating role of ability and task difficulty should be considered. For example, one might expect that for an easy virtual manipulation task, increased fidelity may be a compensator for those lacking the resources to perform the task with more cognitively demanding interfaces, while those with high ability may perform at near ceiling regardless of the interface. As task difficulty increases, those with high ability may be more challenged by the task and benefit from higher fidelity interfaces though offloading cognitive
processes onto more perceptual systems. In contrast, a difficult task might cognitively overload low ability individuals such that any effect of the interface would be unobservable. This interaction can also be thought of as a task demand threshold that may result in ATIs supporting the ability-as-compensator hypothesis if task demands falls below it, or offer support for the ability-as-enhancer hypothesis if task demands exceed this threshold.

3.3 Predictions

We hypothesized that any observed ATIs in the present tasks should be in line with the ability-as-compensator hypothesis. In comparison to previous ability-as-enhancer ATI findings that often measured challenging learning outcomes (see Höffler, 2010), the present experiments measured performance on a relatively easy spatial manipulation task that did not require prior knowledge or involve other cognitive faculties such as memory. Therefore we expected high spatial individuals to be able to accomplish the task regardless of the interface, however, we predicted that low spatial ability individuals, who would be more challenged by the task, would differentially benefit from increased perceptual mediation. Specifically, we expected that greater spatial working-memory resources (Shah & Miyake, 1996) and ability to comprehend spatial relations would allow high spatial ability individuals to more effectively handle the increased cognitive load of the lower fidelity interfaces. We hypothesized that the high fidelity, perceptually mediated interfaces would selectively benefit low spatial ability individuals who have fewer available spatial reasoning resources. In the context of the present task, we predicted increased perceptual mediation may circumvent limitations of low spatial ability and lead to performance more similar to those with high spatial ability.

4. Experiment 1

In Experiment 1, participants used a motion tracked 3D interaction device to manipulate and align 3D molecular models to match the global orientation and internal configuration of 2D diagrammatic representations of molecules. Participants were required to understand the 3D molecular structure represented by the model and the conventions of three different diagram formats in order to successfully match the orientation and configuration depicted in the diagrams.

4.1. Method

4.1.1. Experimental design

Experiment 1 had a two (display fidelity: stereo or mono) by two (interaction device location: co-located or displaced) between subjects design. The dependent variables were
response time and accuracy on an orientation matching task involving both global and local rotations. Response time was defined as the time taken to reach the angular error threshold (20 degrees) and accuracy was defined as the number of trials for which the angular threshold was reached. Individual differences in spatial ability, computer experience, and attitudes and subjective ratings of task demand were also measured.

4.1.2. Participants

One hundred twenty five college students (64 Female) (age: \(M = 18.8, SD = 1.8\)) from the psychology subject pool at a research university participated in the study in return for course credit. None of the participants had previously studied organic chemistry. All participants had normal or corrected-to-normal vision. Participants were randomly assigned to conditions.

4.1.3. Apparatus

A desktop virtual reality system was constructed to allow for co-located manipulation of a virtual molecular model and stereoscopic 3D viewing (Stull, Barrett, & Hegarty, 2012). As shown in Figure 1, the display was mounted horizontally on an aluminum frame 53 cm above the
desk and faced downward towards a full-silvered mirror mounted at 45° which projected the virtual image to the viewer. This configuration allowed participants to manipulate the hand-held interaction device in the same location as the perceived virtual image of the model and provided an experience similar to direct manipulation of a physical model. In the displaced condition, the hand-held interaction device was located to the left and below the image in the natural computer mouse location (15” total displacement, hands occluded). WorldViz Vizard© (Santa Barbara, CA) virtual reality software was used for stimuli presentation and data collection, and Nvidia 3D Vision Wireless Glasses © (Santa Clara, CA) provided stereo viewing. The interaction device, shown in Figure 1b, was composed of a cylinder with approximately the same overall dimensions as the virtual models, and consisted of two halves that freely rotated about the long axis of the device. One half of the device contained a 3-degree-of-freedom (3DOF) motion sensor to track yaw, pitch, and roll of the device, and controlled global rotations of the virtual models. The opposite half was attached via an optical encoder that tracked twisting rotations of the interaction device halves, and controlled local rotations that changed the configuration of the molecule itself (as was necessary on some of the experimental trials).

4.2. Materials

The study materials included a consent form, video tutorial, a set of instructions with descriptions of the task and diagrams, 24 orientation matching problems, two measures of small-scale spatial ability, a subjective task load measure, a computer experience and attitudes questionnaire, and a post-task questionnaire to collect demographic information.

4.2.1. Tutorial video

A 10-minute instructional video explained the conventions of the models and diagrams. The video explained how to find and understand important features of the model (e.g., central carbon-carbon bond), the chemical formulas for each molecular subgroup (e.g., CH₃ for a methyl group made up of a carbon atom and three hydrogen atoms), the color conventions for the different atoms (e.g., black for carbon, red for oxygen, etc.), and how to structurally align the models to match the orientation and configuration of the diagrams.
4.2.2. Diagram orientation matching trials

The diagram orientation matching task required rotation of the virtual model to match one of two target diagram types shown in Figure 2: dash-wedge notation (side-view) and Newman projection (end-view). Both diagrammatic representations are essential in chemistry education and follow different conventions for conveying 3D structure in two dimensions. In dash-wedge notation, solid wedges represent bonds coming out of the plane towards the viewer and hatched dashes represent bonds that extend beyond the plane. Newman projections represent a perspective equivalent to viewing the dash-wedge diagram from either the left or right end (i.e. 90° yaw rotation). Figure 2 shows examples of successful orientation matches for each diagram type. There were 24 orientation matching problems total, half with dash-wedge target diagrams and half with Newman target diagrams. The starting orientation of the model and interaction device was vertical such that there was a 90° angular difference between both target orientations depicted by the diagrams. Half of the trials required a local rotation around the central bond of the model (via rotation along the long axis of the interaction device) in addition to a global rotation of the whole model, and half did not. Six different molecules were used in the 24 trial problems and were systematically crossed with the target diagram type. All participants received
the trials in the same order, in which two consecutive trials never showed the same molecule. Participants were instructed to work as quickly and accurately as possible and terminated the trial by pressing a key on a response pad when they believed they had successfully matched the orientation and configuration of the depicted by the diagram.

4.2.3. Spatial ability measures

Spatial ability was measured by the Mental Rotation Test (MRT) (Vandenberg & Kuse, 1978) and the Visualization of Viewpoints (VoV) (Guay & McDaniels, 1976). The MRT consists of 20 items (two blocks of ten items; three minutes for each block) that involve comparing a 3D (standard) figure with four other figures and selecting which two of the four show identical figures to the standard figure that are rotated in space. The VOV is a three-dimensional perspective taking test which consisted of 24 items (ten minutes total) with two identical 3D figures shown from different orientations and involved inferring the viewpoint from which to view the first figure such that the perspective matches the view depicted by the second figure.

4.2.4. Subjective task load measure

Items from the NASA Task Load Index (TLX) (Hart & Staveland, 1988) were administered to assess participants’ subjective experience of the task with regard to six criteria: mental demand (How mentally demanding was the task?), physical demand (How physically demanding was the task?), temporal demand (How hurried or rushed was the pace of the task?), own performance (How successful do you feel you were at accomplishing what you were asked to do?), effort (How hard did you have to work to accomplish your level of performance?), and frustration (How insecure, discouraged, irritated, stressed and annoyed were you?). Each item as rated on a scale from 0 to 100.

4.2.5. Computer attitudes and experience measure

Computer attitudes and experience were measured using Waller’s (2000) computer use questionnaire. Ten statements (see Appendix A) were rated on a scale of 1 (completely disagree) through 7 (completely agree).

4.3. Procedure

After signing an informed consent form, participants viewed the tutorial video describing the diagram and model conventions. Next, participants read the task instructions and were given one minute to practice manipulating the virtual model. Next, they completed three practice
orientation matching problems in which they were allowed to ask questions and received feedback about their performance from the experimenter. Then they completed the 24 diagram orientation matching problems at their own pace. Finally, participants completed the NASA-TLX, MRT, VoV, computer attitudes and experience questionnaire, and the post-experiment questionnaire.

5. Results and Discussion

5.1. Data coding and screening

Response time was measured as the time taken to reach and sustain (until trial termination) an angular difference of less than 20º from the optimal orientation (global rotation) and configuration (local rotation) for each diagram matching trial. Angular difference was defined as the average angular disparity between the two sides of the model and the target orientation depicted by the diagram. Analysis of participants’ median response time (RT) was calculated based on successful trials in which the 20º angular difference threshold was reached and sustained for the remainder of the trial. Analyzing response times in this manner mitigated differences in participants’ preference to fine-tune their responses and thereby minimized outlying response times (as suggested by Ruddle & Jones, 2001).

Data from five participants were excluded from the analyses as their median RT was over 2.5 standard deviations from the group mean. The following results analyzed data from 120 participants (60 female), with an equal number of males and females in the four interface condition groups. The groups did not significantly differ on the MRT, $F(3,116) = 0.8$, ns, VoV, $F(3,116) = 1.6$, ns, computer experience, $F(3,116) = 0.1$, ns, or attitudes toward computers, $F(3,$

### Table 1: Spatial ability and computer use measures by condition in Experiment 1

<table>
<thead>
<tr>
<th>Measure</th>
<th>Mono</th>
<th>Co-located</th>
<th>Stereo</th>
<th>Co-located</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Displaced $(n=30)$</td>
<td>(n=30)</td>
<td>Displaced $(n=30)$</td>
<td>(n=30)</td>
</tr>
<tr>
<td>Mental Rotation</td>
<td>33.7, 21</td>
<td>31.4, 20.3</td>
<td>32, 20.5</td>
<td>38.7, 18.2</td>
</tr>
<tr>
<td>Visualization of Viewpoints</td>
<td>14.3, 6</td>
<td>12.2, 7.2</td>
<td>11.5, 6.4</td>
<td>14.4, 6.3</td>
</tr>
<tr>
<td>Computer Experience</td>
<td>3.5, 1</td>
<td>3.5, 1</td>
<td>3.5, 0.8</td>
<td>3.4, 0.9</td>
</tr>
<tr>
<td>Computer Attitude</td>
<td>4.4, 0.9</td>
<td>4.2, 1.2</td>
<td>4.5, 0.9</td>
<td>4.3, 0.9</td>
</tr>
</tbody>
</table>

$116) = .33$, ns, as summarized in Table 1.
5.2. Response time

We predicted that providing stereo viewing and co-locating the interaction device would lead to faster response times due to increased perceptual mediation. Median response times (RTs) for the different experimental groups are shown in Table 2 and were analyzed in a 2x2 ANOVA. A significant effect of interaction device location (co-located vs. displaced) was found on RT, $F(1, 116) = 7.27, p = .008, \eta_p^2 = .06$. Marginal means showed that participants provided with the co-located device completed trials 5s faster ($M = 18.4s, SD = 9.1$), than those with the displaced device ($M = 23.4s, SD = 10.9$). No significant effect of display fidelity (stereo vs. mono), $F(1, 116) = 0.05, ns$) or interaction between display fidelity and device location ($F(1, 116) = 1.0, ns$) was observed on RT. As predicted, increasing perceptual mediation by co-locating the interaction device and virtual image led to faster response times, with a medium effect size ($d = 0.50$). Stereo viewing had no effect on response time, which suggested that the increased perceptual mediation provided by stereo did not affect efficiency for this task.

<table>
<thead>
<tr>
<th></th>
<th>Co-located</th>
<th>Displaced</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stereo</strong></td>
<td>17.7 (9.2)</td>
<td>24.5 (11.1)</td>
</tr>
<tr>
<td><strong>Mono</strong></td>
<td>19.1 (9.2)</td>
<td>22.3 (10.1)</td>
</tr>
<tr>
<td><strong>d</strong></td>
<td>0.15</td>
<td>0.20</td>
</tr>
</tbody>
</table>

5.3. Accuracy

Due to the difficulty associated with interpreting and relating novel chemistry diagrams and models, we predicted that increased perceptual mediation should lead to a higher accuracy rate. Overall, participants reached the 20º angular difference threshold on 17.7 ($SD = 4.8$) trials out of the 24 total trials. Table 3 shows mean accuracy levels across the different interface conditions. A 2x2 ANOVA indicated a marginal effect of interaction device location, $F(1, 116) = 3.59, p = .06, \eta_p^2 = .03$. Those who received the co-located device had an average of about 2 more successful trials ($M = 18.1, SD = 4.4$) than those with the displaced device ($M = 16.4, SD = 5.1$), although this finding did not reach significance the effect size was small to medium ($d = 0.36$). No significant effect of stereo viewing was observed, $F(1, 116) = 0.03, p = .57$. A significant interaction between display fidelity and device location was found, $F(1, 116) = 5.40, p = .02, \eta_p^2 = .04$. Pairwise comparisons revealed that participants with stereo viewing were significantly more accurate with the co-located interaction device ($M = 19.3, SD = 3.8$) than those with the
displaced device \((M = 15.7, SD = 5.2)\), \(F(1, 116) = 8.89, p = .003, \eta^2_p = .07\). When provided monoscopic viewing, no significant effect of interaction device location was observed on accuracy. The prediction that increased perceptual mediation should lead to greater accuracy was partially supported in that a beneficial effect of co-locating the interface was observed for those in the stereo viewing condition with a large effect size \((d = .81)\).

5.4. The moderating role of spatial ability

We predicted that those with higher spatial abilities should have better overall performance, and that perceptual mediation (resulting from stereo viewing and co-location) would provide greater benefit for low spatial ability individuals. Scores on the MRT and VoV spatial ability measures were moderately correlated \((r = .44, p < .001)\). In order to provide a more stable measure of underlying ability, a composite spatial ability scale was created by summing the standardized MRT and VoV scores. Multiple regression analyses were conducted to determine whether the relationship between response time and interaction device location (co-located vs. displaced) was moderated by spatial ability. Interaction device location and spatial ability were entered simultaneously with display fidelity (stereo vs mono) also entered as a covariate. Results showed significant main effects of both spatial ability \((B = -6.47, SE_B = 1.52 p < .001)\) and interaction device location \((B = -4.66, SE_B = 1.73, p = .008)\) on RT. Participants with high spatial ability and those provided with the co-located interaction device had faster RTs. The interaction between spatial ability and interaction device location was also significant \((B = 5.44, SE_B = 2.09)\) and explained approximately 5% additional variance in RT \((F(1,116) = 6.77, p = .01)\), indicating the effect of interaction device location on RT was moderated by spatial ability.

<table>
<thead>
<tr>
<th></th>
<th>Co-located</th>
<th>Displaced</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Stereo</td>
<td>M (SD) = 19.3 (3.8)</td>
<td>M (SD) = 15.7 (5.2)</td>
<td>d = 0.81</td>
</tr>
<tr>
<td>Mono</td>
<td>M (SD) = 16.8 (4.7)</td>
<td>M (SD) = 17.2 (4.9)</td>
<td>d = -0.08</td>
</tr>
<tr>
<td>d</td>
<td>0.58</td>
<td>-0.30</td>
<td></td>
</tr>
</tbody>
</table>
Following a procedure suggested by Hayes (2013), conditional effects of interaction device location on RT were plotted across the observed range of spatial abilities, as shown in Figure 3a. The moderator value 0.21 defined the Johnson-Neyman significance region, suggesting a significant effect of interaction device location on performance for individuals below the 58th percentile of spatial ability level. This finding is consistent with the prediction that increased perceptual mediation would be most beneficial for those with the lowest levels of spatial ability.

We also conducted a regression analysis with display fidelity, spatial ability, and their interaction as predictors. No significant interaction with spatial ability was observed, nor did the three-way interaction between display fidelity, interaction device location, and spatial ability. Figure 3b shows response time as a function of spatial ability and display type in Experiment 1.

5.5. Subjective Task Load
If participants aware of the differences in task demands of the different interfaces, this should be reflected in their ratings of task demand. NASA-TLX ratings and the ANOVA table for these measures are shown in Table 4. On average, participants provided with the co-located interaction device reported experiencing significantly less physical demand and frustration during the task than those with the displaced interaction device. Further, participants provided with the co-located interaction device rated their own performance to be significantly better than those with the displaced interaction device. Participants provided with stereoscopic viewing reported experiencing significantly more temporal demand during the task than those with monoscopic
viewing. No effects of stereo or device location were observed on the other ratings of task demand. A significant interaction between display fidelity and interaction device location was found on frustration, \( F(1, 116) = 8.1, p = .005, \eta_p^2 = .065 \). When provided with stereo viewing, participants with the co-located interaction device reported less frustration (\( M = 19.4, SE = 4.9 \)) than those with the displaced interaction device (\( M = 46.3, SE = 4.9 \)). No other significant interactions were observed on the other five items. While significant differences in task demand ratings were only evident on a small subset of items, this analysis demonstrates that perceptual mediation impacted some aspects of subjective task load.

### Table 4: Effect of display fidelity and interaction device location on NASA-TLX ratings in Experiment 1

<table>
<thead>
<tr>
<th>Task Demand</th>
<th>Co-located M (SE)</th>
<th>Displaced M (SE)</th>
<th>ANOVA df</th>
<th>F</th>
<th>p</th>
<th>( \eta_p^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mental</td>
<td>48.4 (3.1)</td>
<td>54.9 (3.1)</td>
<td>116</td>
<td>2.1</td>
<td>0.15</td>
<td>-</td>
</tr>
<tr>
<td>Physical</td>
<td>23.7 (3.1)</td>
<td>33.7 (3.1)</td>
<td>116</td>
<td>5.4</td>
<td>.02*</td>
<td>0.044</td>
</tr>
<tr>
<td>Temporal</td>
<td>45.1 (3.0)</td>
<td>44.4 (3.0)</td>
<td>116</td>
<td>0.3</td>
<td>0.87</td>
<td>-</td>
</tr>
<tr>
<td>Performance</td>
<td>82.0 (2.2)</td>
<td>74.7 (2.2)</td>
<td>116</td>
<td>5.6</td>
<td>.02*</td>
<td>0.046</td>
</tr>
<tr>
<td>Effort</td>
<td>62.5 (3.2)</td>
<td>67.9 (3.2)</td>
<td>116</td>
<td>1.4</td>
<td>0.23</td>
<td>-</td>
</tr>
<tr>
<td>Frustration</td>
<td>23.4 (3.5)</td>
<td>36.2 (3.5)</td>
<td>116</td>
<td>6.7</td>
<td>.01*</td>
<td>0.054</td>
</tr>
<tr>
<td>Stereo</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mental</td>
<td>50.9 (3.1)</td>
<td>52.4 (3.1)</td>
<td>116</td>
<td>0.1</td>
<td>0.74</td>
<td>-</td>
</tr>
<tr>
<td>Physical</td>
<td>31.4 (3.1)</td>
<td>26.0 (3.1)</td>
<td>116</td>
<td>1.5</td>
<td>0.22</td>
<td>-</td>
</tr>
<tr>
<td>Performance</td>
<td>49.7 (3.0)</td>
<td>39.7 (3.0)</td>
<td>116</td>
<td>5.4</td>
<td>.02*</td>
<td>0.044</td>
</tr>
<tr>
<td>Effort</td>
<td>76.6 (2.2)</td>
<td>80.0 (2.2)</td>
<td>116</td>
<td>1.2</td>
<td>0.27</td>
<td>-</td>
</tr>
<tr>
<td>Frustration</td>
<td>64.2 (3.2)</td>
<td>66.2 (3.2)</td>
<td>116</td>
<td>0.2</td>
<td>0.65</td>
<td>-</td>
</tr>
</tbody>
</table>

6. Experiment 2

In order to examine the generalizability of our results, Experiment 2 required participants to manipulate the virtual molecular models to match the orientation and internal configuration of a simultaneously displayed 3D model of the same molecule. By using 3D models rather than 2D diagrams for depicting target orientations, the task in Experiment 2 focused more on the virtual manipulation itself and eliminated the potential confound of participants’ ability to interpret diagrams. Interpreting molecular diagrams has been shown to be difficult even for organic chemistry students (Stull et al., 2012; Padalkar & Hegarty, 2015), so using 3D target models allowed for a more focused examination of virtual object manipulation performance, independent of participants’ ability to interpret organic chemistry diagrams as in Experiment 1. Matching the orientation of two 3D objects is more consistent with previously studied tasks, as comparing two
3D objects is typical in the virtual object manipulation literature (Ruddle & Jones, 2001; Ware & Rose, 1999). It is important to note that due to the additional local twisting rotation required, this task is more difficult than rigid virtual object rotation tasks previously examined. We predicted that the effect of stereoscopic viewing would be apparent in Experiment 2, due to the increased demand on depth perception associated with interpreting static 3D models.

6.1. Method
6.1.1. Experimental Design

Experiment 2 had a two (display fidelity: stereo vs. mono) by two (interaction device location: co-located vs. displaced) between subjects design. Individual differences in spatial ability, computer experience and attitudes, and task demand were measured as in Experiment 1. Response time was defined as the time taken to reach the angular error threshold (20 degrees), and accuracy was defined as the number of trials in which the angular threshold was reached.

6.1.2. Participants

One hundred forty two college students (73 Female) (age: \( M = 18.8, SD = 1.6 \)) from the psychology subject pool at a research university participated in return for course credit. None of the participants had studied organic chemistry. All participants had normal or corrected-to-normal vision and were randomly assigned to condition.

6.2. Materials and Procedure

The study materials and procedure were identical to Experiment 1, except for the 3D orientation matching problems and omission of the diagram training video.

Figure 4: In Experiment 2, participants aligned the below model to match the orientation of 3D models.
1. 3D Orientation matching trials

The 3D model orientation matching task required rotation of the virtual model to match to the orientation depicted by an identical 3D model (see Figure 4). The trials were identical to Experiment 1, except the target orientations were depicted by 3D models rather than diagrams.

7. Results

7.1. Data Coding and Screening

Response time and accuracy were processed using the same method as in Experiment 1. Six participants were excluded from the analysis: two did not satisfy the 20º accuracy threshold on over 1/3 of the trials, and four had median response times that were greater than 2.5 standard deviations from the group mean. 136 participants (71 female) were included in the final analysis, with equal numbers of males and females in each condition. The groups did not differ significantly on the MRT, \( F(3, 132) = 0.5, \text{ns} \), VoV, \( F(3, 132) = 1.7, \text{ns} \), computer experience, \( F(3, 132) = 0.1, \text{ns} \), or attitudes toward computers, \( F(3, 132) = 0.7, \text{ns} \), as summarized in Table 5.

Table 5: Spatial ability and computer use measures by condition in Experiment 2

<table>
<thead>
<tr>
<th>Measure</th>
<th>Mono</th>
<th>Stereo</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Displaced ((n = 34))</td>
<td>Co-located ((n = 34))</td>
</tr>
<tr>
<td>Mental Rotation</td>
<td>32.57</td>
<td>14.76</td>
</tr>
<tr>
<td>Visualization of Viewpoints</td>
<td>8.93</td>
<td>6.83</td>
</tr>
<tr>
<td>Computer Experience</td>
<td>3.48</td>
<td>0.82</td>
</tr>
<tr>
<td>Computer Attitudes</td>
<td>3.98</td>
<td>0.9</td>
</tr>
</tbody>
</table>

7.2. Response time

We predicted that increased perceptual mediation from the co-located interaction device and stereo viewing should lead to faster response times. Response times for the different experimental groups are shown in Table 6. Median response times were analyzed in a 2x2 ANOVA. A significant effect of interaction device location was found on response time, \( F(1, 132) = 10.5, p = .002, \eta^2_p = .074 \). Marginal means showed that participants provided with the co-located device had response times that were 33% faster \((M = 18.2s, SD = 6.8)\) than those using the displaced device \((M = 23.5s, SD = 10.8)\). A significant effect of stereo was also found, \( F(1, 132) = 10.7, p = .001, \eta^2_p = .075 \). Marginal means showed that participants provided with stereo viewing median response times that were 33% faster \((M = 18.1s, SD = 8.9)\) than those provided with monoscopic viewing \((M = 23.5s, SD = 10.8)\). The interaction between display
fidelity and device location did not reach significance, $F(1, 132) = 2.1, p = .15$. The prediction that increased perceptual mediation would lead to faster response times was supported by the medium sized effects of co-location ($d = 0.59$) and stereo viewing ($d = 0.55$).

7.3. Accuracy

The difficult process of interpreting chemistry diagrams was not required for this task, and therefore high accuracy was expected. Participants reached the $20^\circ$ angular error threshold for successful trial completion on $23.6 (SD = 0.7)$ trials out of the $24$ total trials. There were no significant main effects observed on stereo or co-location, nor any significant interaction between interaction device location and display fidelity. Due to the near ceiling level of accuracy, accuracy will not be reported in subsequent analyses.

7.4. The moderating effect of spatial ability

We predicted that students with high spatial ability should have better overall performance, and that perceptual mediation would have a greater effect for low spatial ability individuals. Scores on the MRT and VoV spatial ability measures were moderately correlated ($r = .51, p < .001$). In order to provide a more stable measure of underlying ability, a composite spatial ability scale was created by summing the standardized MRT and VoV scores.

Multiple regression analyses were conducted to determine whether the relationship between RT and display fidelity (stereo vs. mono) and interaction device location (co-located vs. displaced) would be moderated by spatial ability. Display fidelity and spatial ability were entered simultaneously with interaction device location entered as a covariate. Results showed significant main effects of both spatial ability ($\beta = -5.58, SE_B = 1.15, p < .001$) and display fidelity ($\beta = -4.69, SE_B = 1.53, p = .002$) on RT. Participants with high spatial ability and those provided with the stereo display had faster RTs. The interaction between spatial ability and display fidelity ($\beta = 3.77, SE_B = 1.79$) was significant and explained approximately 2% additional variance in RT ($F(1,131) = 4.41, p = .038$), indicating that the effect of display fidelity on RT

<table>
<thead>
<tr>
<th></th>
<th>Co-located</th>
<th>Displaced</th>
<th>$d$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stereo</strong></td>
<td>M (SD)</td>
<td>M (SD)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>19.3 (3.8)</td>
<td>15.7 (5.2)</td>
<td>0.35</td>
</tr>
<tr>
<td><strong>Mono</strong></td>
<td>16.8 (4.7)</td>
<td>17.2 (4.9)</td>
<td>0.78</td>
</tr>
<tr>
<td><strong>d</strong></td>
<td>0.45</td>
<td>0.67</td>
<td></td>
</tr>
</tbody>
</table>
was moderated by spatial ability. To explore this interaction, the conditional effects of display fidelity on RT were plotted across the observed range of spatial abilities, as shown in Figure 5.

The moderator value 0.37 defined the Johnson-Neyman significance region, suggesting a significant effect of providing stereo viewing on performance for individuals below the 64th percentile of spatial ability level. In order to estimate the effect of interaction device location on performance moderated by spatial ability, the same analysis was conducted with interaction device location as the predictor variable. However, no significant interaction with spatial ability was observed, as shown in Figure 5.

7.5. Subjective Task Load Measure

We predicted that effects of perceptual mediation would be reflected in subjective measures of task demands. NASA-TLX ratings and ANOVA table are shown in Table 7. On average, participants provided with stereo viewing reported experiencing significantly less temporal demand during the task than those provided with mono viewing. No significant differences were found on the other five items. As in Experiment 1 subjective ratings of task demands provide limited partial support for the prediction.
In two experiments, effects of display fidelity and interaction device location were examined during two virtual molecule manipulation tasks. We also investigated the role of spatial ability in moderating the effects of the interface design on task performance. Experiment 1 examined performance when aligning 3D molecular models to 2D molecular diagrams, which is an important exercise in organic chemistry (Kozma & Russell, 2005). Experiment 2 examined performance in aligning two 3D molecular models. The tasks appeared to be challenging to participants, as evidenced by observed response times of about 20 seconds in Experiments 1 and 2, with accuracy rates of 75% and 98% respectively.

### 8.1. Co-location of haptic and visual workspaces

In both experiments, task performance was on average 30% faster when the haptic and visual workspaces were co-located than when they were displaced, indicating that perceptual mediation benefited performance. Subjective reports of reduced task demand associated with using the co-located interaction device corroborated its beneficial impact on performance. Similarly, Ware and Rose (1999) observed 35% faster performance with less complex rigid object rotations (global rotations only) in tasks which took only about 4s (one fifth of the time taken for our task). The result that similar large effects of co-location are observed in tasks of different

**Table 7: Effect of interaction device location and display fidelity on NASA-TLX ratings in Experiment 2**

<table>
<thead>
<tr>
<th>Task Demand</th>
<th>Co-located</th>
<th>Displaced</th>
<th>ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M (SE)</td>
<td>M (SE)</td>
<td>df</td>
</tr>
<tr>
<td>Mental</td>
<td>62.3 (2.7)</td>
<td>61.3 (2.7)</td>
<td>132</td>
</tr>
<tr>
<td>Physical</td>
<td>26.8 (2.9)</td>
<td>25.2 (2.9)</td>
<td>132</td>
</tr>
<tr>
<td>Temporal</td>
<td>44.9 (2.6)</td>
<td>46.8 (2.6)</td>
<td>132</td>
</tr>
<tr>
<td>Performance</td>
<td>81.8 (1.6)</td>
<td>79.5 (1.6)</td>
<td>132</td>
</tr>
<tr>
<td>Effort</td>
<td>70.1 (2.7)</td>
<td>70.0 (2.7)</td>
<td>132</td>
</tr>
<tr>
<td>†Frustration</td>
<td>33.7 (3.5)</td>
<td>43.3 (3.4)</td>
<td>132</td>
</tr>
</tbody>
</table>

**Stereo**

<table>
<thead>
<tr>
<th></th>
<th>M (SE)</th>
<th>M (SE)</th>
<th>df</th>
<th>F</th>
<th>p</th>
<th>ηp²</th>
</tr>
</thead>
<tbody>
<tr>
<td>†Mental</td>
<td>58.1 (2.7)</td>
<td>65.4 (2.7)</td>
<td>132</td>
<td>3.7</td>
<td>.06†</td>
<td>0.027</td>
</tr>
<tr>
<td>Physical</td>
<td>24.3 (2.9)</td>
<td>27.8 (2.9)</td>
<td>132</td>
<td>0.7</td>
<td>0.39</td>
<td>-</td>
</tr>
<tr>
<td>†Temporal</td>
<td>41.8 (2.6)</td>
<td>50.0 (2.6)</td>
<td>132</td>
<td>5.1</td>
<td>.03*</td>
<td>0.037</td>
</tr>
<tr>
<td>Performance</td>
<td>79.4 (1.6)</td>
<td>81.9 (1.6)</td>
<td>132</td>
<td>1.2</td>
<td>0.28</td>
<td>-</td>
</tr>
<tr>
<td>†Effort</td>
<td>66.5 (2.7)</td>
<td>73.5 (2.7)</td>
<td>132</td>
<td>3.5</td>
<td>.06†</td>
<td>0.026</td>
</tr>
<tr>
<td>†Frustration</td>
<td>34.0 (3.5)</td>
<td>43.0 (3.4)</td>
<td>132</td>
<td>3.4</td>
<td>.07†</td>
<td>0.025</td>
</tr>
</tbody>
</table>

**Mono**

* p < .05; † p < .10; N = 136
complexity levels suggests that the effect is practically significant and consistent across object manipulation tasks of varying difficulty. Ware (2012) offered the guideline that minimizing mismatches between visual and haptic information is especially important for virtual object rotation tasks with hand-held interaction devices. The present work supports this guideline by demonstrating the importance of providing a co-located interaction device for a more difficult rotation task than previously studied.

8.2. Stereoscopic viewing

Providing stereo viewing did not significantly decrease overall response time in Experiment 1, although participants who received stereo viewing also reported experiencing greater temporal demand during the task than those who received the monoscopic display. This finding reflects the fact that interpretation of 2D diagrams does not depend on stereo viewing and suggests that interpretation of the spatial relations of the manipulated 3D model may have been sufficiently supported by monocular depth cues such as relative motion, relative size, occlusion, perspective and shading. Any beneficial effect of stereo viewing was likely mitigated due to monocular cues affording ‘good enough’ fidelity for the diagram matching task in Experiment 1. In Experiment 2, however, providing stereo viewing led to about 30% faster performance. In this experiment, participants had to match static 3D target models, which likely placed additional demands on depth perception compared to Experiment 1. In this task scenario, additional depth cues afforded by stereo viewing contributed to task relevant perceptual fidelity compared to monocular cues alone.

The differential effects of stereo across the two experiments are consistent with previous research in which some studies have shown beneficial effects of providing stereo viewing (Wang et al., 1998; Arsenault & Ware, 2004), while some have not (Willemsen et al., 2008; Khooshabeh & Hegarty, 2010); see McIntire, Havig, and Geiselman (2014) for a recent meta-analysis. The heterogeneity of these findings can be attributed to variations in task demands and characteristics. As evidenced by the present experiments, the impact of providing stereoscopic viewing on performance depends on specific demands placed on depth perception by the task at hand. In general, beneficial effects of providing stereo viewing are likely to arise only when binocular depth cues afford additional task-relevant information beyond what is provided by the monocular cues alone. As such, designers of virtual environments should take care to examine
the demands placed on depth perception for different tasks, and whether they can be supported by monocular cues alone, so that expenses may be spared when stereo viewing is not necessary.

**8.3. Task performance comparison**

The rate of successful trial completion was 72% in Experiment 1 and 98% in Experiment 2. This difference is likely attributed to the difficulty of learning and interpreting the conventions of the Newman and Dash-wedge target diagrams as well as the process of relating multiple representations in Experiment 1 (cf., Stull et al., 2012). Nevertheless, it is interesting to note that response times were similar when matching to both 2D diagram and 3D model target orientations, at least on trials in which the accuracy threshold was met. Although students have difficulty learning and interpreting diagram conventions in organic chemistry, our research suggests that they are generally successful in matching the orientation and configuration of virtual models, suggesting that these models have excellent potential for teaching students about molecular conformations, an important topic in organic chemistry.

**8.4. Aptitude-treatment-interactions**

As predicted, individuals with higher spatial ability had better performance than those with lower spatial ability in both experiments. In Experiment 1, the predicted aptitude-treatment-interaction between spatial ability and interaction device location on performance was observed. Co-location improved performance for the majority of participants with low to average spatial ability, while no effect was found for participants with high spatial ability. It appears that high spatial ability participants were able to compensate for the increased cognitive demand associated with displaced haptic and visual information. Similarly, in Experiment 2, spatial ability moderated the effect of stereo viewing on response time. Stereo viewing led to faster performance for the majority of participants with low to average spatial ability, whereas no effect of display fidelity was found for participants with high spatial ability. Again, high spatial ability participants may have been able to handle the increased spatial processing demands associated with using the cognitively mediated monoscopic display due to greater spatial working memory and processing resources (Shah & Miyake, 1996).

This work clearly demonstrates an aptitude-treatment interaction consistent with the ability-as-compensator hypothesis. However, given that spatial ability is a relatively constant trait and the design of a virtual representation is engineered, perhaps it is useful to interpret this finding as interface fidelity compensating for low spatial ability. Interpreting ability-as-compensator
aptitude-treatment-interactions in this manner may offer important implications for educators and designers of virtual interfaces and learning environments. An implication for instructors is to examine current computer-based training technologies, as teaching with interfaces with impoverished perceptual cues may impose greater task demands on students with lower spatial ability.

In contrast to Experiment 1, co-locating the interaction device with the display improved performance across all levels of spatial ability in Experiment 2. It appears that when both the manipulated and target models were 3D models, high spatial ability participants experienced similar levels of disruption to those with low spatial ability. Further research is warranted to investigate why those with high spatial ability were unaffected by the displaced interaction device when orienting to a 2D diagram and were negatively impacted when orienting to a 3D model.

Finally, although high-fidelity interfaces were most beneficial to individuals with low spatial ability, it is important to note that a threshold likely exists where task demands may cognitively overload lower ability individuals, such that high fidelity interfaces cannot overcome the task demands. Once this difficulty threshold is crossed, it is expected that ability would play the role of an enhancer in moderating performance, causing high ability individuals to uniquely benefit. Investigating the nature of this threshold for different modalities of perceptual cues and varying levels of fidelity will be useful in clarifying the role of various spatial abilities in virtual interaction and learning environments.

8.5. Limitations and future directions

A limitation of this work is that it was focused on a relatively specific orientation matching task. Ongoing research is aimed at addressing the effects virtual models for more complex tasks, such as classifying stereochemical isomers, inferring chemical properties of the molecules and acquiring representational competence in the domain of organic chemistry (Author citation; Author citation). Another limitation of this research is sample size; a larger sample would allow for better estimates of the moderating role of spatial ability. Future work will address issues of immediate and delayed learning outcomes, academic achievement, and the generalizability of these findings to other domains such as anatomy, engineering, and design.

8.6. Practical Implications and Future Directions
Given the recent and sustained efforts to increase participation and diversity in STEM disciplines (National Research Council, 2006), the findings reported here may hold implications for educators and policy makers seeking to improve task performance and/or learning in STEM disciplines, especially for individuals with lower spatial ability. Our research demonstrates that thoughtful interface design can alleviate some of the demands faced by individuals with low spatial ability in understanding spatially rich scientific topics. While much research has focused on selection of high spatial individuals and training spatial reasoning skills in order to increase participation in STEM disciplines (Lubinski 2010; Uttal et al., 2012), perhaps gains in STEM participation and achievement may also be had by developing and improving instructional technologies that cater to those with low levels of spatial ability. Improved technology is surely not a silver bullet for increasing participation and closing the gender gap in STEM fields; however, well designed virtual interfaces may help in the education and training of more diverse student populations otherwise discouraged by the challenging spatial content of many scientific fields. Overall, this work demonstrates two clear examples of perceptually mediated interface design differentially bolstering performance of individuals with lower levels of spatial ability to that of one with higher ability. While further research should focus on a broader range of spatial tasks and learning contexts in different disciplines, circumvention of limitations imposed by spatial ability level through interface design is a promising avenue for increasing participation in the sciences.
References


Appendix A: Computer Attitudes and Experience Questionnaire

For the following items, please rate the degree to which you agree with each statement on a scale of 1 to 7 (1 = completely disagree; 7 = completely agree).

1. Computers dehumanize society by treating everyone as a number.
2. I am able to learn about computers very quickly.
3. Computers are beyond the understanding of the typical person.
4. I know how to program computers.
5. I feel at ease when I am around computers.
6. I have played a lot of computer games.
7. I feel comfortable when a conversation turns to computers.
8. Kids these days know more about computers than I do.
9. I have a lot of self-confidence when it comes to computers.
10. I think working with computers would be enjoyable and stimulating.