Identifying Representational Competence With Multi-Representational Displays

Mike Stieff

University of Illinois, Chicago

Mary Hegarty

University of California, Santa Barbara

Ghislain Deslongchamps

University of New Brunswick

Increasingly, multi-representational educational technologies are being deployed in science class-rooms to support science learning and the development of representational competence. Several studies have indicated that students experience significant challenges working with these multi-representational displays and prefer to use only one representation while problem solving. Here, we examine the use of one such display, a multi-representational molecular mechanics animation, by organic chemistry undergraduates in a problem-solving interview. Using both protocol analysis and eye fixation data, our analysis indicates that students rely mainly on two visual–spatial representations in the display and do not make use of two accompanying mathematical representations. Moreover, we explore how eye fixation data complement verbal protocols by providing information about how students allocate their attention to different locations of a multi-representational display with and without concurrent verbal utterances. Our analysis indicates that verbal protocols and eye movement data are highly correlated, suggesting that eye fixations and verbalizations reflect similar cognitive processes in such studies.

As part of their normal course of instruction in science, students must develop a set of skills for constructing, interpreting, transforming, and coordinating domain-specific external represen-tations for learning and problem solving. Collectively, this constellation of skills in mathemat-ics and science has been referred to as *representational competence* (Kozma & Russell, 1997; Kozma, Chin, Russell, & Marx, 2000; Wu, Krajcik, & Soloway, 2001; Nathan, Stephens, Masarik, Alibali, & Koedinger, 2002). Achieving representational competence in mathematics or science is non-trivial as scientists have developed a host of diagrams, graphs, models, and pictures to rep-resent the objects of study in their respective domains. Although these representations have been

Correspondence should be addressed to Mike Stieff, Department of Chemistry, Learning Sciences Research Institute, University of Illinois–Chicago, 845 W. Taylor St., Room 4376, Chicago, IL 60607. E-mail: mstieff@uic.edu

developed over hundreds of years, students are asked to apprehend them in a few short months during instruction. The sheer number of representations available and their unique applications for problem solving pose a unique challenge to both the beginning and advanced science student.

Nowhere is this challenge more evident than in the study of chemistry. From the first day of instruction, students enrolled in chemistry are introduced to multiple representational systems that will serve as a common language for students and teachers. For example, students must learn symbolic notations for indicating the identity of atoms, the internal structure of molecules, and the properties that emerge from both molecular and atomic interactions. Although the number of representations learned in early instruction is limited, as students proceed to advanced study, the variety and complexity of relevant representational systems grow. Thus, by the time students reach the second year of chemistry instruction at university, they may be assessed on their skill at interpreting and constructing no less than 15 unique representations for one molecule. Moreover, students must learn the limited applicability of each representation for solving unique problems in the domain.

The challenge of developing representational competence in chemistry results in part from the imperceptibility of the phenomena of interest in the domain, namely atoms and molecules. Chemists' inability to perceive atoms and molecules directly has led them to develop the wide range of external representations seen today to facilitate scientific practice. Indeed, major discoveries and advances in chemistry historically have followed the invention of new representational systems to characterize the phenomena under study (Hoffmann, 1997; Hoffmann & Laszlo, 1991). In the classroom, students are asked not only to learn the formalisms and application of each representational system, they are asked to coordinate multiple representations to explain how imperceptible molecular objects and interactions result in observable physical and chemical properties in the laboratory. Moreover, students must not only master representational systems unique to chemistry, they must also concurrently apprehend mathematical representations of quantitative data (e.g., graphs and equations) that result from laboratory experiments. Understandably, students face significant challenges achieving representational competence in chemistry (Kozma & Russell, 1997; Kozma et al., 2000; Wu et al., 2001).

To address this difficulty, instructional designers are now investigating new methods for supporting students' development of representational competence in chemistry by providing scaffolds for interpreting and coordinating multiple representations. In chemistry (and several other science disciplines) students now routinely learn and solve problems while engaging with dynamic, multi-representational visualizations of previously imperceptible objects and phenomena. Chief among these are innovative educational technologies that include animations, simulations, and virtual laboratories (e.g., Russell, Kozma, Jones, Wykof, Marx, & Davis, 1997; Stieff, 2005; Stieff & Wilensky, 2003; Wu et al. 2001). These environments attempt to make explicit the information embedded in external representations with interactive visual displays. Many of these tools include multiple representations that are dynamically linked to help students both perceive the relationship between the representing and represented world and connect various external representations together.

The development and increased use of these tools not only offer new avenues to improve student achievement in chemistry, they also provide opportunities to enrich our understanding of how students interpret and employ external representations for problem solving in authentic settings. Evaluation studies both in and out of the classroom have begun to indicate that student achievement and understanding of domain concepts improves after using these

multi-representational displays (Wu et al., 2001; Stieff & McCombs, 2006). At a more fine-grained level, analyses of student artifacts in the classroom reveal that students who learn chemistry while using multi-representational displays show major improvements in their ability both to generate accurate chemical representations and to use them appropriately for describing imperceptible interactions (Stieff & McCombs, 2006). Indeed, Linn, Lee, Tinker, Husic, and Chiu (2006) have reported that students who learn chemistry using novel technologies in the classroom display more integrated knowledge of both scientific representations and domain knowledge.

Although these current technologies allow for the creation of complex interactive visualizations of scientific concepts, and there is much excitement over their classroom application, there have been few studies of how students actually use these new technologies for learning and problem solving. Extant studies have mainly employed interactivity logs and verbal protocols to examine how students process the information presented in multi-representational displays. For example, Kozma and Russell (1997) and Wu et al. (2001) have used verbal protocols to reveal that novice students often lack the necessary knowledge to coordinate the use of multiple representations to problem solve effectively in multi-representational displays. These verbal protocol studies have indicated that one's ability to interact with and learn from multi-representational displays may be limited both by content knowledge and by representational competence in the domain.

However, other researchers have pointed out that studies of representational competence have problematically assumed a deficit model in their approach (e.g., diSessa, 2004; diSessa & Sherin, 2000). That is, these prior studies have focused on the mistakes that problem solvers make when using multiple representations without acknowledging the nascent skills that problem solvers have to select, interpret, and modify representations within and across academic domains. For example, Stieff and McCombs (2006) found that students who learn chemistry with multi-representational displays produce rich pictorial representations of atoms and molecules, although they are unable to provide accurate verbal descriptions of their pictures. Laboratory studies may underestimate problem solvers' inability to coordinate multiple representations effectively because they study situations that are disconnected from authentic contexts and disciplinary knowledge. Likewise, the types of data collected in previous studies may also lead to an underestimation of students' representational competence.

In particular, most studies that have focused on the use of multi-representational displays in authentic contexts have been limited to verbal protocols that produce limited qualitative descriptions of representational use. To be sure, these methods of investigating multi-representational displays have yielded useful data; however, there are lingering concerns about the validity and reliability of verbal self-reports of problem solving. There remains doubt about whether a student's reference to a particular representation reliably indicates that the representation was useful for problem solving; more problematic is that verbal protocols often produce incomplete data, such as when students solve problems in multi-representational displays without referring to any particular representation. Furthermore, verbal protocols only reflect those representations that students are conscious of using; however, it is possible that students attend to representations that they do not mention in their protocols. This may be particularly true of the visual–spatial representations seen in educational technologies designed for teaching chemistry, which are likely to be associated with non-verbal thinking processes that are not reflected in the verbal report. For instance, studies of student gestures while solving spatial thinking problems often reveal thought processes that were not evident from utterances (Alibali, 2005; Hegarty, Mayer, Kriz & Keehner,

2005; Emmorey & Casey, 2001) suggesting that verbal protocols may provide a less complete record of spatial thinking.

Given these limitations, outstanding questions remain regarding the role of students' representational competence in learning and problem solving with multi-representational displays. For example, contrary to previous findings, students may indeed coordinate multiple representations in multi-representational displays although they produce verbal utterances that suggest they use only one representation. Likewise, the extent to which students make use of any one representation cannot be reliably known given that the frequency of mention of a specific representation may not correlate strongly with the students' perceived relevance of that representation for problem solving. Thus, new complementary methods for characterizing students' representation use with multi-representational displays are needed to provide richer models that describe the role of new technologies in developing representational competence.

One such approach is the use of eye fixation data to characterize problem solving. Interpretation of eye fixation data is based on the "eye-mind" assumption, which states that "the eye fixates the referent of the symbol currently being processed if the referent is in view" (Just & Carpenter, 1976, p. 441). While recent research has suggested some qualifications of this assumption (Irwin, 2004), researchers still agree that when people are free to move their eyes anywhere in a visual scene or display, the location of their current gaze is a good indication of where they are attending (Rayner, 1998; Henderson & Ferreira, 2004). For example, in previous studies of graph comprehension (Carpenter & Shah, 1998; Peebles & Cheng, 2003), comprehension of text and diagrams (Hegarty & Just, 1993; Graesser, Lu, Olde, Cooper-Pye, & Whitten, 2005), mental animation of static diagrams (Hegarty, 1992), and troubleshooting (Graesser et al., 2005; van Gog, Paas, van Merriënboer, 2005), eye fixation data have specified the precise amount of time that students attend to a particular representation or parts of a representation in a display. Using such data, researchers have inferred how students judge the relative priority of particular representations and how they coordinate representations by transitioning their gaze between areas of the display.

However, despite evidence that eye fixations are related to attention (Just & Carpenter, 1976; Rayner, 1998; Henderson & Ferreira, 2004), like protocol measures, eye fixation measures also suffer concerns about their validity. Like verbal protocols, eye fixation data can be incomplete, due to loss of calibration. More problematically, they offer incomplete information about the cognitive processes students employ. For example, some students may fixate for longer durations on representations they find confusing rather than those they find particularly useful for problem solving. Specifically, eye fixations only tell us *where* people attend in a display, but not necessarily *why* they choose that representation or *how* they use that representation.

Consistent with prior calls for examining the resources students use for problem solving with multiple representations (diSessa, 2004; diSessa & Sherin, 2000), we argue that students actively and effectively coordinate multiple representations when using novel educational technologies. Further, we suggest that prior research claiming that students' representational competence limits their ability to learn with these technologies is itself limited by the exclusive use of one method. Specifically, we propose that the limitations of both verbal protocols and eye tracking measures used in earlier studies can be partially overcome through the complementary application of both methods. Such complementary approaches have been examined elsewhere (e.g., Cook, Wiebe, & Carter, 2008; Patrick, Carter, & Wiebe, 2005; von Gog et al., 2005) and are beginning to offer a more holistic account of student representation use. However, eye fixations and verbal protocols
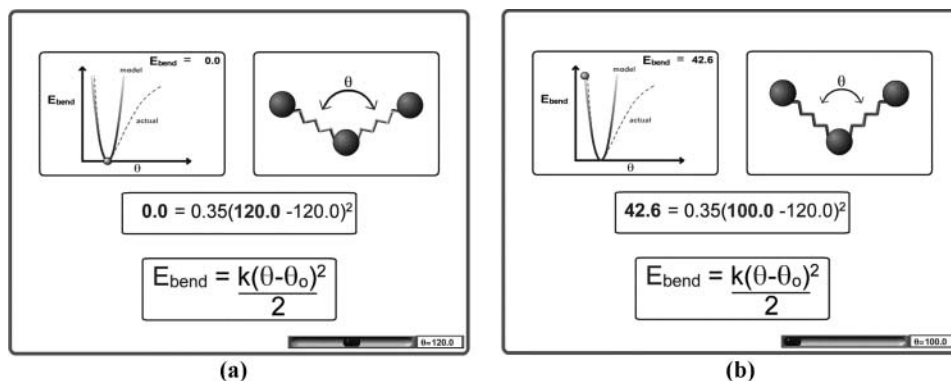
FIGURE 1   Examples of two frames from the Angle Bend Visualization. a) frame depicting lowest energy bond angle (note molecular model, low ball position on graph, and zero value on numerical solution); b) frame depicting compressed bond angle (note molecular model, high ball position on graph, and high energy value on numerical solution).

have been compared only qualitatively in previous studies. Here we compare these methodologies quantitatively.

The main goal of the present study is to evaluate students' ability to coordinate representations using multi-representational displays for teaching chemistry. We present an analysis of student question answering using a novel multi-representational dynamic display containing four representations: a molecular model, a general equation, a numerical equation, and a graph (see Figure 1). The model emphasizes the qualitative internal spatial relationships in the molecule. The general equation and numerical equation emphasize the quantitative relationship between specific spatial conformations and corresponding quantities of energy. The graph displays the corresponding function of the numerical equation by emphasizing visually the relationship between the spatial conformations and corresponding quantities of energy. Although related, the form and formalisms of such representations render them non-equivalent to novices (Ainsworth, 2006; Larkin & Simon, 1987).

While viewing the displays, students were asked questions that could be answered by viewing the model alone, questions that required coordination of the model and graph, and questions that required coordination of all four representations. Given the format of the interview, and previous research on representational competence, we can consider two possible approaches students might take to responding to each type of question. First, students might select only one representation, which may or may not provide sufficient information for answering any question, due to a lack of knowledge about the particular affordances of specific representations (cf., Kozma & Russell, 1997). An alternative possibility is that students will adjust which representations they consult as a function of the question they are asked, and answer all of the questions correctly. Of course, these possibilities represent extremes of poor and excellent representational competence, and actual student performance may lie somewhere in between.

Furthermore, we predict that in scenarios where students focus on only one representation, they should be drawn to the most realistic, dynamic, and salient representation in the display; that is, they should focus on the molecular model, which is relatively more rich in color and intensity, as opposed to the mathematical or symbolic representations (the graphs and equations). We base

this prediction on research demonstrating that people prefer realistic displays of data, even when they are better served by more abstract displays (Smallman & St. John, 2005), and research indicating that novices are particularly drawn to the most dynamic and salient aspects of displays (Lowe, 2004). Cognitive models of visual attention and perception also propose that attention is drawn to salient regions of displays (Itti & Koch, 2000; Parkhurst, Law, & Niebur, 2002; Treisman & Gelade, 1980; Wolfe, 1998). In cases where students focus on only the molecular models, they should perform poorly on questions that require the information in the molecular models to be integrated with the graph and equations. In contrast, if students are able to use the different representations effectively, they should consult the graph and equations (and not just the model) when the question requires information from these representations, and they should be able to answer all of the questions accurately.

A secondary goal of this study is to examine the complementary contributions of eye fixation data and verbal protocols quantitatively. We argue that the co-application of these methods can offer evidence that corroborates the claims that result from each method as well as provide redundant sources of information that permit inferences about student reasoning in cases where the use of only one method would otherwise prove unreliable. To that end, we first examine how the independent application of each of these methods can offer partial insight into how students select unique representations for problem solving. Second, we examine the consistency between the conclusions reached from the results produced by each method by examining the correlations between measures of representation use derived from two independent analyses. Specifically, we hypothesize that data produced from the two sources will be highly correlated because these approaches both offer insight into the same cognitive processes. Finally, we compute measures of the frequency with which each approach provides information that is not given by the other, to offer a richer account of student problem solving with multi-representational displays.

METHOD

Participants

Ten participants were recruited on a volunteer basis for the study. All students had completed a minimum of one year in university instruction and at least one year of instruction in university general chemistry. Each student was enrolled in the second quarter of a three-quarter course in organic chemistry and had recently received prior instruction in the chemistry topic discussed in the interview and eye fixation experiment. We selected this novice population of students for two reasons. First, the students had received the minimum amount of instruction necessary to comprehend the interactive visualizations used in the experiment; second, they represented the population of students for whom our interactive visualizations are designed. Participants each received USD$10.00 after completing the study.

Materials and Instruments

Interactive Visualizations.   Three interactive Flash animations were adapted from a series of interactive courseware for teaching computer-based molecular modeling to undergraduate

chemistry students. The animations depicted the components of a typical molecular mechanics equation, which is used for modeling chemical structures and their corresponding energies via a simple ball-and-spring approach. Three common energy terms from the molecular mechanics equation were presented as three separate multimodal animations: a "stretch animation" to model changes in energy due to bonds between two atoms, an "angle bend animation" to model changes in energy due to changes in angles between sets of three contiguous atoms, and a "dihedral animation" to model changes in energy due to rotations about bonds.

Within each animation, four different representations of the molecular mechanics component were shown: a ball-and-stick model of the molecular system, the general mathematical model used to model its energy, a specific numerical example of the equation, and a graph of the equation with a ball to depict the position on the graphical function corresponding to the specific numerical example; in this article, the four regions will be referred to as *model*, *general equation*, *numerical equation*, and *graph*, respectively. Figure 1 illustrates selected frames of the Angle Bend animation. The multimodal representations were synchronized via an interactive mouse-driven slider, allowing the user to change the molecular model and to observe the resulting change in energy as it appeared on the graph and on the numerical example. The Flash animations were then converted into a series of static images (17 frames per animation) and input into the eye tracking system. Subjects could interact with the animation constructs via keyboard arrow keys to advance or reverse the static images.

Apparatus. Eye movements were monitored using an SMI EyeLink I head mounted eye tracking system, which was spatially accurate to within $0.5°$ and had with a sampling rate of 250 Hz. An eye movement was classified as a saccade if acceleration exceeded $9,500°/sec^2$ and velocity exceeded $30°/sec$. Fixations were defined as time between saccades (see Salvucci & Goldberg, 2000). Participants viewed images presented on a computer screen while resting their chins on a chin rest, set 30 inches from the screen. The computer screen measured 15 inches horizontally, by 11.5 inches vertically. The monitor's screen resolution was set to 800 by 600 screen pixels, with a refresh rate of 75 Hertz.

Verbal Protocols. Participants were asked to view the animations and respond verbally to four questions presented orally by an interviewer. Each problem, listed in the Appendix, was adapted from typical organic chemistry assessment problems regarding molecular mechanics and validated by three organic chemistry instructors to ensure that participants had received sufficient instruction to attempt the problems. Participants were instructed to think aloud as they solved each problem, in accordance with Ericcson and Simon (1980).

The first question regarding each animation asked the student to explain verbally what the animation was representing, and was included primarily for the purpose of familiarizing the student with the animation. The next two problems for each animation asked the students specifically to consider the relationship between the internal spatial relationships of the depicted molecule and the conformational energy: Students were asked either to explain how the energy of the molecule changed after a specific spatial transformation or how a change in energy affected the spatial configuration of the molecule. The second question could be answered using information presented in the model and *either* the graph or numerical equation. The third question about each molecule

could be answered qualitatively from the model and the graph; alternatively, students could provide quantitative values in their response by integrating the model and the numerical equation. Finally, the fourth question about each molecule asked students to consider the specific properties of a molecule similar to the model presented in the animation and determine how the energy would change following a specific spatial transformation. To answer this question, participants had to integrate the information in the model with *either* the graph or the numerical equation.

Procedure

Participants took part in the experiment one at a time. After they arrived in the laboratory the eye tracker was calibrated. Then they performed a task that involved making judgments about diagrams of organic chemistry molecules; this task was unrelated to the specific molecular structures, representations, and relevant content knowledge to the tasks discussed in this paper and will not be analyzed here. Next they were shown the three interactive animations in turn and were allowed to interact with these visualizations by pressing the forward and backward arrow keys on the keyboard to move through the frames at their own pace. While they were viewing each animation, an experimenter read them the four specific questions about that visualization and the participant answered orally. When students ambiguously referred to "the model" or "the picture," the interviewer asked for a specific clarification of the term. In 8% of cases, the interviewer did not ask one of the four problems directly because a participant had answered it spontaneously when responding to an earlier problem. Student's eye fixations were tracked and digitally recorded as they viewed the visualizations and answered the questions. In addition, all verbalizations were recorded on digital video camera, which also provided a recording of student interactions with each animation on the computer display.

RESULTS

Eye fixation data and verbal protocol data were sequestered and analyzed by two independent researchers to increase the objectivity of each analysis. To analyze the eye fixations, we defined rectangular areas of interest on the displays corresponding to the graph, the ball-and-stick model, the general equation, the specific-values equation, and the slider bar. Note that while objects moved and changed within these areas, the different representations (graph, model, equation, etc.) themselves did not move, so the same areas of interest were valid for all frames of the animation. Figure 2 shows an example of one frame of the Angle Bend animation with the regions of interest (indicated by thin rectangles) and the eye fixations of one participant (indicated by points) while viewing this animation. A computer software program was used to compute the number of fixations and the total fixation duration (sum of the durations of all fixations) on each of these areas of interest for each participant on each trial of the experiment. Thus the coding of the eye fixations was completely objective. Measures of the number of fixations in an area were highly correlated with measures of total fixation duration (median correlation = .97) and showed the same patterns, so we present only the data for total fixation duration here. Finally, we examined transitions (saccades) between the areas, that is, the number of times participants made an eye movement from each area to each of the other areas.

The videotaped verbalizations of each participant were transcribed for analysis. One of the videotapes (Subject 3) contained no audio data for analysis, thus nine interviews were transcribed.
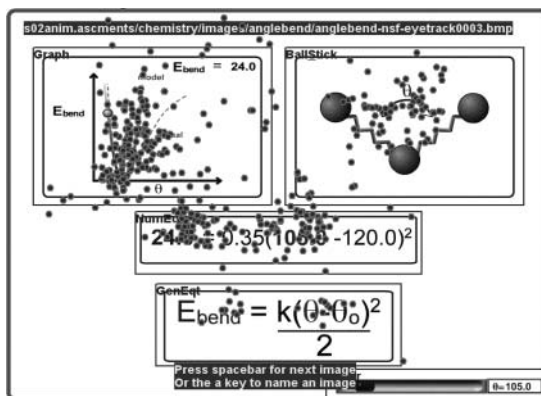
FIGURE 2 Example of a frame from the Angle Bend Visualization showing the areas of interest used in the eye fixation analysis and the eye fixations of one participant while viewing this visualization.

After transcription, each interview was reviewed to identify participants' specific utterances regarding the representations contained in the multi-representational display. Each response was coded by a single reviewer for (a) verbal references to each specific representation in the display (e.g., "I can see in the molecular model") and (b) verbal references to information embedded in a representation (e.g., "The curve is hyperbolic," indicating a reference to the graph). Further, the rater coded responses to indicate whether one representation or multiple representations were referenced. Because participants frequently referred to "the equation," it was not possible to distinguish references to the general equation versus the numerical equation; thus, only three representation codes were applied to the verbal protocol data (i.e., graph, model, equation). Likewise, codes for references to multiple representations did not attempt to capture precisely which representations were used due to occasional indirect references to multiple representations (e.g., "it is comparing the theoretical versus the specific [reference to graph or equation] as ethane rotates around its axis [reference to the model]"). Example excerpts and their codes are detailed in Table 1. The transcripts were also reviewed for accuracy. Each response was given a binary score (correct or incorrect) according to a rubric established by an organic chemistry instructor. A subset of the verbal data (10 items, 12%) was analyzed by a second reviewer to establish the reliability of the coding procedure (% agreement $= 85\%$; $\kappa = 0.71$).

Question Accuracy

We scored the accuracy of student answers to 9 of the 12 questions using a binary rubric provided by the chemistry instructor (because the first question regarding each animation simply asked the students to explain what they were viewing, it was not scored for accuracy). Thus students could receive a maximum score of 9 and a maximum score of 3 on each of the 3 types of questions. Overall, students were reasonably accurate in their responses to the nine questions ($M = 6.17$, $SD = 2.12$). Accuracy of specific questions across all three animations was non-uniform. The average number of correct answers to Question #2, which students could answer using any single

TABLE 1
Illustrative Examples of Transcript Codes

Subject	Question	Verbal Utterance	Code
S1	Stretch #1	Because it's moving equidistant from the zero point, just in the other direction. And that model line will intersect at the same rate. So, the energy should be the same. No, actually ... I'm not really sure, because the actual model doesn't fit that. On the actual line it would say, it would be higher.	Graph/Single
S4	Dihederal #2	Right here ... *she moves the animation forward several frames.* ... The staggered one at 60 degrees.	Model/Single
S6	Stretch #4	By the equations. Since the radii distance is squared, it doesn't matter if it increases or decreases.	Equation/Single
S9	Dihedral #1	I'm trying to picture the positions. If you have 240, should be pretty much be the exact opposite of 60, on a unit circle or whatever [ref. to model]. It would be—although, I guess if you have different molecules. ... It would be about the same 'cause if you think of it as a sine curve [ref to graph], like we have here, you go through another cycle and you get to 240, it's going to be the lowest energy, which is like our 0 degrees.	Model/Multiple

representation, was quite high ($M = 2.67$ out of 3, $SD = .71$). For Question #3, which students could answer by using either the model or the graphical representation, the average number of correct answers was somewhat lower ($M = 2.3$, $SD = .72$). Finally, for Question #4, which required the coordination of two or more representations, student accuracy was the lowest ($M = 1.2$, $SD = .97$). A one-way analysis of variance showed that the accuracy rate differed between the questions, $F(2,26) = 7.04$, $p < .05$, and a post-hoc analysis using the Scheffé test indicated that the average number of correct responses to Question #4 was significantly lower than the average number of correct responses collapsed across Questions #2 and #3, $F(1, 17) = 12.69$, $p < .05$. Thus, students were more accurate on problems that could be answered from the model alone than on problems that required integrating the model with one of the symbolic representations.

Viewing of Interactive Animations

Students spent up to 3 minutes viewing and answering questions about each animation. Their total fixation duration was 179.64 sec ($SD = 64.1$) on the Stretch animation, 109.83 sec ($SD = 36.14$) on the Angle Bend animation, and 121.19 sec ($SD = 49.5$) on the Dihedral animation. Mean total fixation durations for each of the ten participants on the different regions of interest (the graph, the model, the general equation, the numerical equation, and the slider bar) are presented in Table 2. As this table shows, participants spent approximately one third (34%) of their time (47.2 sec, $SD = 17.0$) viewing the graph and another third (32%) (43.2 sec, $SD = 20.7$) viewing the molecular model, while spending less time on the equations (11% or 17.0 sec, $SD = 16.2$ on the specific values equation and 4% or 6.1 sec, $SD = 6.6$ on the general equation). Percentage of time spent viewing the different representations was compared to chance, which was defined as the percentage of the screen area taken up by each representation. The percentages of time

TABLE 2
Time (in Seconds) Spent Viewing Each of the Representations, the Slider Bar, and Other Regions of the
Screen During the Interviews

Participant	General Equation (%)	Numerical Equation (%)	Graph (%)	Model (%)	Slider Bar (%)	Total[a]
			Total Fixation Duration (sec)			
1	5.1 (3)	9.3 (6)	54.1 (34)	45.38 (31)	.69 (0)	151.28
2	2.4 (2)	17.0 (14)	56.9 (46)	19.18 (15)	.11 (0)	135.35
3	1.9 (2)	6.8 (7)	35.4 (39)	25.49 (29)	.86 (1)	90.87
4	1.0 (1)	7.4 (10)	12.1 (17)	27.64 (49)	.70 (1)	63.65
5	1.5 (1)	11.0 (7)	36.1 (29)	50.69 (40)	.69 (0)	133.64
6	18.4 (8)	44.6 (18)	61.6 (28)	82.52 (40)	.55 (0)	219.55
7	3.0 (3)	8.5 (8)	35.2 (41)	33.19 (42)	.15 (0)	84.61
8	9.2 (6)	13.3 (8)	66.4 (44)	38.30 (25)	1.59 (1)	153.40
9	1.4 (1)	3.3 (2)	52.6 (32)	73.80 (50)	.21 (0)	155.98
10	17.0 (9)	49.0 (26)	61.5 (32)	35.43 (23)	2.49 (1)	180.55
Mean	6.08 (.04)	17.01 (11)	47.18 (34)	43.16 (.35)	.80 (1)	136.89

Note. [a]Times spent viewing the individual representations do equate to the total time, because this time also reflects time spent viewing other areas of the display.

spent viewing the graph and the model were both greater than chance, which was 18% for both of these representations, (one-sample $t(9) = 6.03$, $p < .001$ for the graph; $t(9) = 4.49$, $p = .003$ for the model). Proportion of time spent viewing the numerical equation was not different to chance (8%) and people spent less time viewing the general equation than would be predicted by chance (9%, one-sample $t(9) = -5.65$, $p < .001$).

There were large individual differences in allocation of attention, for example participant 10 spent as much as 35% of her time (66 sec) viewing the equations. Participants who spent more time on the graph tended also to spend more time on the general equation ($r = .64$, $p < .05$) and time spent viewing the general equation was highly correlated with time spent viewing the numerical (specific value) equation ($r = .93$, $p < .01$).

Figure 3 illustrates the transition data, showing the mean number of transitions (saccades from each fixation to the next) that occurred within and between the areas of interest corresponding to the graph, model, numerical equation, and general equation (averaged across both participants and interactive animations). The size of the arrows between each two representations is roughly proportional to the average number of transitions made. Although most transitions were within representations,[1] participants also made several transitions between the four representations

---

[1]The transition data were analyzed by comparing the observed transitions to those that would be expected by chance, taking into account the total number of fixations in each region of interest (using the method described by Bakeman & Gottman, 1997). This analysis revealed that transitions within representations and transitions between the two equations were more frequent than chance ($z > 3.5$, $p < .001$ in all cases) whereas all other transitions between representations were less frequent than chance ($z < -3.5$, $p < .001$ in all cases).
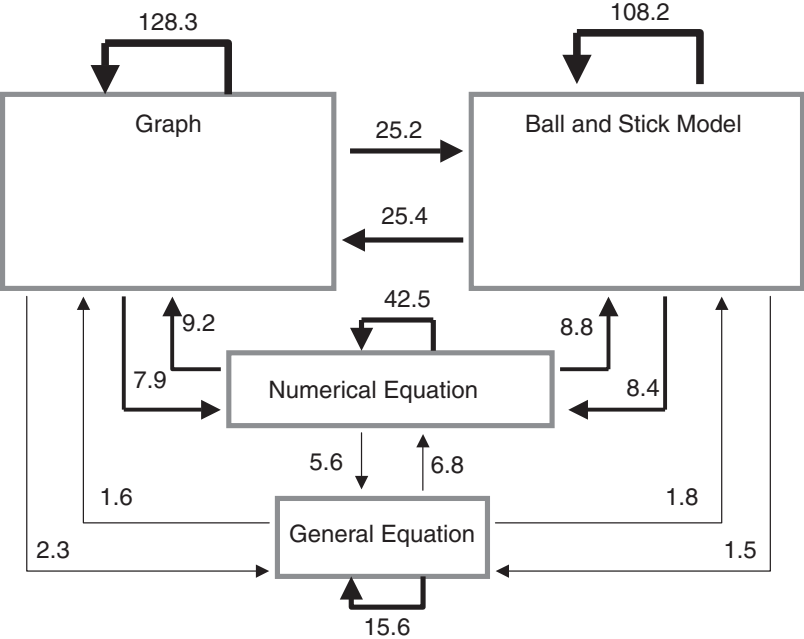
FIGURE 3 Schematic representation of the interactive animations showing the mean number of transitions (averaged across students and interactive animations) between the four areas of interest representing the graph, model, numerical equation, and general equation. The size of the arrow is roughly proportional to the number of transitions made.

per interactive animation, suggesting they were clearly attempting to relate and integrate the information in the four different representations.

In summary, the eye fixation data indicated that participants did not just look at the model, as might be expected if they had poor representational competence. Instead, they looked at the graph just as much as the model, although they looked less at the equations. All participants viewed the model, while there were individual differences in viewing the more symbolic representations, especially the equation. The transition data indicated that participants were clearly attempting to integrate the information in the different representations, rather than relying exclusively on the most realistic representation (i.e., the model).

Verbal Protocols

Analysis of the verbal protocols also yielded information regarding the relative frequency with which participants (a) used a specific representation by referring directly to one representation or information included in only one representation, (b) used multiple representations by referring to information presented in two different representations, and (c) qualitatively integrated information across multiple representations. The raw frequency data for references to each representation from the multi-representational display are listed in Table 3. Although each participant varied in the

TABLE 3
Frequency of Verbal References to Specific Representations in the Interview

| Participant | Verbal References | | | |
	Equation (%)	Graph (%)	Model (%)	Multiple (%)
1	1 (9.1)	6 (54.5)	4 (36.4)	5 (41.7)
2	1 (10)	5 (50)	4 (40)	5 (41.7)
3	—	—	—	
4	0 (0)	1 (16.7)	5 (83.3)	3 (25)
5	0 (0)	6 (54.5)	5 (45.5)	5 (41.7)
6	5 (26.3)	6 (31.6)	8 (42.1)	8 (66.7)
7	1 (12.5)	3 (37.5)	4 (50)	2 (16.7)
8	1 (10)	4 (40)	5 (50)	7 (58.3)
9	1 (9.0)	3 (27.3)	7 (63.6)	5 (41.7)
10	6 (46.2)	5 (38.5)	2 (15.4)	6 (50)
Total	16 (16.1)	39 (39.4)	44 (44.4)	46 (42.6)

number of references made, most participants made some verbal reference to each of the displayed representations during the interview. The participants referred primarily to either the graph or the model. As indicated in Table 3, the majority of references (44.4%) to a specific representation across the interviews were to the model and the fewest references (16.1%) were to the equation, as predicted. The participants also made frequent references (39.4%) to the graph during the interview. Again, individual differences in references to each representation were evident in the protocols. Three students (1, 2, and 5) referred more frequently to the graph when answering questions, while the remaining six students referred more frequently to the model. There was no significant relationship between student choice of representation and accuracy.

Additional examination of participants' references to multiple representations throughout the interview indicated that each student attempted to integrate information presented in more than one representation on multiple occasions. As indicated in Table 3, participants referred to information in multiple representations directly on 42.6% of questions. Individual differences in multiple representation use were evident and dramatic. Frequent references to multiple representations were seen among some students whereas others rarely referenced more than a single representation. For example, one student referred to multiple representations on 8 of the 12 questions (66.7%) whereas another mentioned more than one representation on only 2 of the questions (16.7%). Here, we discuss the frequency of representation coordination and offer illustrative cases of student responses for Questions 2, 3, 4; Question 1 was not analyzed in detail due to the open ended nature of the question.

Participant responses referenced multiple representations on 16 of the 27 (59%) responses to Question #2, which asked to identify the most stable geometry of a molecule. Here, participants appeared to use the model representation to identify the geometry that resulted when either the graph or the equation displayed the relative energy minimum. S8's response to Stretching Question #2 illustrates how the students coordinated multiple representations to generate a solution. By advancing the representation forward and backward, she is able to deduce the specific radius between the nuclei in its most stable state. In the transcript, the student appears to use the model representation to interpret the variables in the equation. First, she uses the model representation

to identify the meaning of the variable r, "Okay, so r is the distance between these atoms." At this point, she is able to manipulate the animation to display the stable configuration and report a specific value for r that is only displayed in the numerical equation. When questioned on how she arrived at her answer, she again references the atoms in the model representation as well as the energy values reported in the equation.

> **S8:** Oh! Okay . . . *she keys the animation forward, then backward several times* . . . Okay, so r is the distance between these atoms. And, as it [r] increases there's a point at when . . . let's see . . . *she manipulates the animation until the E_{str} is equal to 0* . . . The bond is more stable when the radius, or the distance, is 1.5.
>
> **I:** How do you know that?
>
> **S8:** It seems the energy required to keep them in that position is zero . . . *she moves the animation backward and forward* . . . As they get closer together, more energy is required to keep them together, and as they go apart more energy is required to stay at that distance.

In response to Question #3, students referenced multiple representations in 15 responses (57%),[2] a proportion similar to that observed for Question #2. In contrast to Question #2, in the case of Question #3 the participants appeared to use the equation to provide a specific value in their responses. In the example below, S2 displays the coordination of two representations to determine the minimum dihedral angle in response to Bending Question #3. As noted in the transcript, after moving the animation to a position close to the minimum, she definitively states that the dihedral angle is 12. Her response suggests that she used only the numerical equation representation to respond since the specific value of the dihedral angle was only displayed in the equation. Yet, when asked a follow up question to clarify how she had determined her answer, she reveals that also used the model representation by noting, "they [the atoms] are overlapping as it shows in the model." With this last reference, S2 demonstrated that she coordinated information in the numerical equation and the molecular model to decide upon the dihedral angle presumably by moving the atoms close together in the molecular model and then stating the value reported in the numerical equation.

> **I:** At what dihedral angle are the atoms closest to each other?
>
> **S2:** *She moves the animation to about 4 slides from the beginning* . . . 12.
>
> **I:** Ok, 12. How do you know?
>
> **S2:** Because they [the atoms] are overlapping as it shows in the model.

The least number of references to multiple representations occurred in response to Question #4. Note that this question required the integration of the model with either the graph or the equation. Only 5 tasks (21.7%)[3] included references to multiple representations. Most protocols included only a brief reference to the graph to infer qualitatively the relative energy. In the few protocols that included clear references to multiple representations, the participants appeared to coordinate the model representation with the graph. For example, when asked to explain his

---

[2]26 responses were coded for Question 3. One response was not coded due to the lack of information provided by the participant S10 when responding to this question in regard to the Angle Bend animation.

[3]23 responses were coded for Question 4. On four occasions, participants produced ambiguous answers that could not be coded definitively.

answer, S9 indicated that he was first trying to determine the spatial configuration in the model by manipulating the animation and stating that he was, "trying to picture their positions." Following this, he refers directly to the graph representation noting, "if you think of it as a sine curve, like we have here," one could determine the relative energy.

I: If I tell you that the energy is at a minimum in ethane with a dihedral angle of 60 degree, what would you expect the energy to be at a dihedral angle of 240 degrees? Higher, lower, or about the same?

S9:   240? Um . . . 240 . . . um . . . Wait—am I supposed to say lower or higher?

I: Uh-huh. Higher, lower, or about the same?

S9:   Can I move this? . . . *He advances the animation forward, then backward a few slides until the dihedral angle is 60 degrees.* I think . . . about the same.

I: Why would you say that?

S9:   *He advances the animation forward two slides.* I'm kind of thinking as . . . I'm trying to picture their positions. But, if you have 240, should be pretty much the exact opposite of 60, on a unit circle or whatever. It would be—although, I guess if you have different molecules—Oh! It would be about the same. Cause if you think of it as a sine curve, like we have here, you go through another cycle and you get to 240, it's going to be the lowest energy, which is like our 0 degrees.

Analysis of the verbal protocols yielded clear evidence that students recruited and integrated information from multiple representations to respond to roughly 40% of the responses in the interview. However, it also appeared that students relied primarily on one representation despite coordinating across representations to support their reasoning when asked to explain how they had arrived at an answer. Interestingly, students referred to multiple representations, either directly or indirectly, on Questions 2 and 3 more frequently than on Question 4. This is notable since Question 4 included specific quantities in the problem statement for the students to consider and suggests that direct integration of the numerical equation and the model is required to generate a solution. Students' failure to integrate the equations with the model representations on this task may explain the low accuracy on this question. Regardless of the apparent need for appealing to one or more representations, the participants displayed a willingness for using multiple representations to respond to each question in addition to the clear aptitude for obtaining information from the array of representations to formulate a solution. In sum, the analysis of the verbal protocols provided rich examples of students' use of multiple representation as well as some indication of the relative frequency of representation (specific and multiple) use; however, the verbal protocol analysis did not provide a clear explanation for students' apparent failure to coordinate multiple representations on tasks that seemingly mandated coordination.

Quantitative Relations Between Eye-Fixation Measures and Verbal Protocols

Table 4 reports the correlations between the total fixation duration on the graph, ball-and-stick model, and the combined equations and the number of times each of these was mentioned in the verbal protocols. Note that the amount of attention to the different representations, as measured by total fixation duration, was highly correlated with how often they were mentioned in the verbal protocols (correlations ranged from .63 to .96) indicating that students' eye fixations and

TABLE 4
Correlations Between Total Fixation Duration and Number of Mentions of the
Different Representations in the Protocols

Number of Mentions	Total Fixation Duration		
	Graph	Model	Equations
Graph	.63 ($p = .07$)	.30 n.s.	.47 n.s.
Model	.03 n.s.	.80 ($p < .01$)	−.06 n.s.
Equations	.55 n.s.	−.31 n.s.	.97 ($p < .001$)

verbalizations reflected similar cognitive processes. Moreover, the high degree of correlation between the verbal reports and eye fixation data indicate that both methods can serve as valid measures of students reasoning in multi-representational displays.

To further explore the synergistic power of eye fixation data and verbal protocols, we examined the specific student responses to Question 4 for each animation.[4] As with the overall analysis, the verbal protocols and eye fixation data were analyzed independently and compared. Analysis of the verbal protocols included determining the accuracy of participant responses and the primary representation that the participants referred to when generating an answer. Precise time stamps for when the experimenter began asking the student Question 4 and when the student finished answering this question were also coded from the videotapes, and a focused eye fixation analysis was conducted to indicate amount of time spent viewing each of the representations during that specific time. This analysis allowed us to determine independently the primary representation used by the student from the eye fixation data. This was defined as the representation viewed for the longest time during this specific time frame.

As shown in Table 5, analysis of the verbal protocols provided a definitive claim on the primary representation used by participants for 15 (50%) problem solving trials. The remaining 15 protocols did not yield enough data to make firm claims about which representation students were referencing during their problem solving. In 12 of the remaining cases, participants referred to information that was present in more than one representation, were not clear on how they came to their answers, reported that they guessed, or in three cases the verbal protocol was not recorded due to equipment error. These responses were coded as unknown.

Comparatively, there was a high level of agreement between the conclusions from the independent coding based on the verbal protocols and eye fixation data. As Table 5 shows, the results disagreed on only one trial from the 14 for which we had codes based on both data sources (the Angle Bend animation for Participant 9). In that case, the participant reported using the graph to answer the question and in fact spent more time viewing the model on this trial (15.4 sec) but also spent a sizeable amount of time viewing the graph (10.8 sec). Perhaps more interesting, is that in cases where the verbal protocols failed to produce data that could be analyzed, the eye fixation data offered an alternative method for determining the representation that was used by the student. Likewise, in one case (the Dihedral Animation for Participant 8), the eye tracking

---

[4]We concentrated on this question for the focused analysis, because it apparently depended most on the integration of information from the different representations. In addition, a focused analysis was not relevant for Question 1 as it was a very general question, and some students spontaneously answered Questions 2 and 3 when asked the first question.

TABLE 5
Primary Representation Used to Answer Question 4 Based on Independent Coding of the Verbal Protocols and Eye Fixations

	Stretch Animation		Dihedral Animation		Angle Bend Animation	
Participant	Verbal Protocol	Eye Fixations	Verbal Protocol	Eye Fixations	Verbal Protocol	Eye Fixations
1	Graph	Graph	Model	Model	Graph	Graph
2	Graph	Graph	Graph	Graph	Graph	Graph
4	Unknown	Equation	Unknown	Model	Unknown	Equation
5	Unknown	Equation	Unknown	Model	Graph	Graph
6	Equation	Equation	Model	Model	Equation	Equation
7	Unknown	Graph	Unknown	Model	Unknown	Graph
8	Model	Model	Model	Unknown	Graph	Graph
9[a]	Unknown	Graph	Model	Model	Graph	Model
10	Unknown	Equation	Unknown	Model	Unknown	Equation

Note. [a]One disagreement between eye fixation and verbal protocol analysis occurred for the Angle Bend animation of Participant 9.

apparatus failed to provide reliable data for analysis (because of loss of calibration) and the verbal protocols were able to make a definitive claim.

In a final analysis, we examined the average amount of time spent viewing each of the representations (graph, model, or the two equations combined) for trials on which the verbal protocols indicated that the primary representation used was the graph (8 trials), the model (4 trials), and the equations (2 trials). These data are graphed in Figure 4. They provide strong evidence for the consistency of the verbal protocol data and eye fixation data, as it is clear that
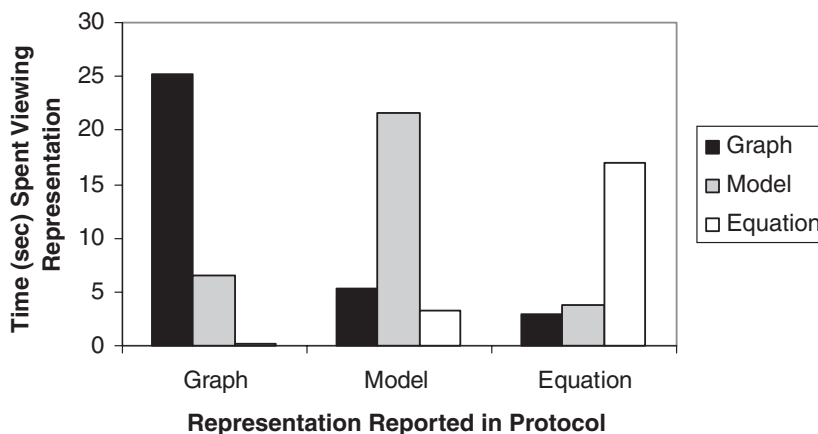

FIGURE 4 Graph showing the time spent viewing each representation (the graph, the model, or combined time spent viewing the two equations) for trials in which the verbal protocols indicated that participants were consulting each of these representations.

in each case, participants spent most time viewing the representation identified as primary based on the verbal protocols. But they also show that in each case, the representation reported in the protocol was not the only one viewed. For example, on trials in which participants reported using the model to derive their answer, they spent an average of 21.7 sec viewing the model, but also spent 5.3 sec on average viewing the graph and 3.2 sec on average viewing the equations. These data add to our understanding of why students appeared not to integrate representations in answering Question 4 if we rely on the verbal protocols as a source of data. Specifically, they suggest that the verbal protocols seemingly underestimate the extent to which students attend to multiple representations. They demonstrate that in addition to validating verbal protocol data, eye fixation data can augment these data by indicating not just the *major* representation used to answer a question, but also the other representations that were consulted but not mentioned in the verbal protocols.

DISCUSSION AND CONCLUSIONS

In summary, the present study indicates that students are capable of coordinating diverse representations in multi-representational displays to answer questions in advanced chemistry. In addition, our analysis offers quantitative data to support the complementary power of verbal protocols and eye fixation data that in turn supports the validity of each for measuring students' representation use. Students did not rely primarily on the visual representation of the molecules that were rendered with richer colors and included more prominent animations of multiple atoms. Instead they gave equal weight to the model and graph and made frequent transitions between these two representations. On the other hand, the participants made very few references to the equations. As a result, participants performed well on questions that could be answered from the model and graph, and performed poorly on questions that required the integration of information from the model, graph, and equation. The results from the verbal protocols and from the eye fixation data were highly correlated. This correlation suggests that both methods reflect the underlying cognitive processes students use when answering questions with multi-representational displays: namely each method offers data that speak to attention and information gathering strategies used.

Our results offer insights into students' developing representation use at the undergraduate level when using advanced educational technologies, that is, the development of *representational competence* (diSessa, 2004; Kozma et al., 2000; Wu et al., 2001). Although prior work has suggested that students may lack the ability and knowledge to coordinate multiple representations in such displays, we did not find this to be the case. Instead, the results from the eye fixation data reported in Figure 3 indicate that students made transitions between the model and graphical representations when solving each problem, suggesting that they were attempting to relate these two representations to generate an answer to each problem. Moreover, the data reported in Table 5 confirm that students do not treat the representations as equivalent. Instead they make explicit choices about the affordances of particular representations for solving unique tasks. As illustrated in the table, the students spent relatively more time viewing the molecular model to reason about the dihedral animation whereas the graph was used more extensively to reason about the bending and stretching animations. This finding is of note as the dihedral animation uniquely attempts to illustrate the internal three-dimensional spatial relationships in a complex molecule whereas the stretching and bending animations display only two-dimensional relationships. Here, the eye fixation data suggest that students preferentially make use of the model for reasoning about the

three-dimensional spatial relationships involved in molecular mechanics. This finding represents a potential level of sophisticated representational competence due to the fact that the majority of students appeared to refer explicitly to the most appropriate representation for reasoning about each animation.

While this behavior indicates good representational competence, it was also clear that there was room for improvement in students' use of representations, in that they paid much less attention to the equations and performed poorly on questions that required information from the equations to be integrated with information in the model and graph. During debriefing, three of the ten students explicitly stated that they did not need to use the equation to answer the questions and one student avoided the representation after claiming, "I am bad at math." While this study offers only anecdotal evidence for why students do not use the equations, this is an important topic to address in future studies and points. Our findings are consistent with previous research that has illustrated that although designers may perceive equivalent information in multiple representations, students may not (Ainsworth, 1999; Ainsworth, Bibby, & Wood, 2002), and that users perceive information displayed in specific representations as useful for some tasks, but not others (Ainsworth & Peevers, 2003; Gilmore & Green, 1984). In general, our results indicate that students at this level are somewhat in between the extremes of poor and excellent representational competence outlined earlier. Although they have good understanding of graphical representations, their understanding of more abstract representations (equations) needs to be developed further.

In addition to providing new information regarding representational competence, this paper makes an important methodological contribution by demonstrating that verbal protocols and eye fixations can provide complementary insights into students' problem solving with multiple representations. First, independent analyses of these two data sources leads to highly consistent characterizations of the major representations used by students to solve different chemistry problems. Thus each source of data serves to validate conclusions from the other. Second, because this consistency has been established, we can rely on the record from the eye fixation data when there are missing data from the verbal protocols and vice versa. Third, each source of data provides unique information that is not provided by the other. For example, the eye fixation data indicated that in addition to the major representation used to solve each problem that was mentioned in the verbal protocols, participants also attended to the other representations to some extent. Thus, eye fixations can provide more fine-grained information about the allocation of visual attention while students solve problems using multimedia displays. Similarly, the verbal protocol data provided information about *how* people used the representations to which they attended, such as what information they extracted from these representations and whether they used this information accurately to answer the questions.

While the verbal protocols suggest that students may ultimately choose one representation to support their arguments regarding changes in energy, the eye fixation data indicate that the students interpret and coordinate the information displayed in more than one chemical representation. For example, the protocol analysis indicated that students primarily used the graph to respond to Question 4 without referring to other representations in the display. This result is surprising given that Question 4 seemingly required students to use information from multiple representations to respond. Analysis of the eye fixation data revealed that the students did indeed attend to multiple representations in the display even when they referenced only one representation in their verbal response (although their poor performance on this question suggests that they may not have always been successful in their integration attempts). In this way, we find the complementary

results of the two methods particularly informative for identifying students' coordination of multiple representations as the isolated use of verbal protocols in the study would have indicated that students use only one representation and invited the conclusion that students lack even basic levels of representational competence.

We recognize that the small sample size in this study may lessen the external validity of our findings. However, small sample sizes are not unusual in studies that provide detailed analyses of either verbal protocols or eye fixations and several of our results were highly significant despite our small sample size, which speaks to the strength of the relationships observed here. The high degree of correlation also suggests that concurrent application of both methods may be particularly useful when access to large populations of students is limited.

It is possible that the size of the representations and their positions on the page also influenced the amount of time spent viewing the representations. Thus, students spent more time viewing the graph and model, which were larger and presented at the top of the display. Likewise, students may have found it easier to relate representations in close proximity on the display (e.g., model and graph, specific and general equation). Our analyses indicated that the frequency of fixating the different representations was not merely a function of their size. Future research should evaluate the extent to which the page layout might influence attention to different representations by manipulating the location of each representation on the display. However it is notable that we found differences in relative attention to the graph and model as a function of the content of the animation, even though these representations were about the same size and the position of these representations on the display was kept constant across the three animations.

The observations reported here suggest that students have available certain resources that have been presumed missing. Namely, our results show that second-year undergraduate students who receive a normal course of instruction in introductory organic chemistry are able to make appropriate claims about molecular mechanics when using multi-representational displays. Of course, we are not able to discern whether the level of competence we observed among students was due to their prior training, the affordances of the technology, or an interaction between the two. Moreover, the extent to which students fully integrated the information across representations was unclear given the ambiguity in students' responses. At minimum our findings illustrate that students are attending to information in multiple representations, and in many cases coordinating this information. Perhaps more important, our findings suggest that students at this level tend not to make use of mathematical representations of phenomena when visual displays are present. Future studies that examine how students respond to questions that explicitly require the use of information in the mathematical representations and questions that mandate integration of information from each representation are necessary to characterize fully student use of multi-representational display such as the animations examined here.

Finally, the results of this work indicate that multi-representational displays can be highly effective in the science classroom, but such displays must be situated in activities that both support students' utility and coordination of multiple representations and clearly define the utility of specific representations for unique problem solving and learning tasks. Our analysis shows that individual differences in representational competence can be significant among students, and the simple incorporation of multi-representational display into science instruction is not likely to help students develop skills related to representational competence (cf., Cox & Brna, 1995; Kozma & Russell, 1997). Novel curriculum materials that include multi-representational displays should offer explicit information about the affordances of specific representations for solving specific tasks.

Likewise, instructional activities using such displays should include learning objectives regarding the relationships between specific representations used. Our recommendations are congruent with those of design frameworks that recommend careful scaffolding of multiple representations during instruction (e.g., Ainsworth, 2006; Brünken, Plass, & Leutner, 2003; Quintana et al., 2004). Ultimately, our data support the perceived potential of multi-representational displays in science instruction and suggest that future investigations of their affordances and constraints may best be explored through the complementary application of verbal protocol and eye movement measures.

ACKNOWLEDGMENTS

REFERENCES

Ainsworth, S. E. (1999). The functions of multiple representations. *Computers & Education*, *33*, 131–152.

Ainsworth, S. (2006). DeFT: A conceptual framework for learning with multiple representations. *Learning and Instruction*, *16*, 183–198.

Ainsworth, S. E., Bibby, P., & Wood, D. (2002). Examining the effects of multiple representational systems in learning primary mathematics. *Journal of the Learning Sciences*, *11*, 25–61.

Ainsworth, S. E., & Peevers, G. J. (2003). The interaction between informational and computational properties of external representations on problem-solving and learning. In R. Altmann & D. Kirsch (Eds.), *Proceedings of 25th Annual Conference of the Cognitive Science Society*. Boston, MA: Cognitive Science Society.

Alibali, M. W. (2005). Gesture in spatial cognition: Expressing, communicating, and thinking about spatial information. *Spatial Cognition and Computation*, *5*, 307–331.

Bakeman, R., & Gottman, J. M. (1997). *Observing interaction: An introduction to sequential analysis*. New York: Cambridge University Press.

Brünken, R., Plass, J., & Leutner, D. (2003). How instruction guides attention in multimedia learning. In H. Niegemann, R. Brünken, & D. Leutner (Eds.), *Instructional design for multimedia learning* (pp. 113–126). Münster: Waxmann.

Carpenter, P. A., & Shah, P. (1998). A model of the perceptual and conceptual processes in graph comprehension. *Journal of Experimental Psychology: Applied*, *4*, 75–100.

Cook, M., Wiebe, E. N., & Carter, G. (2008). The influence of prior knowledge on viewing and interpreting graphics with macroscopic and molecular representations. *Science Education*, *92*, 848–867.

Cox, R., & Brna, P. (1995). Supporting the use of external representations in problem solving: The need for flexible learning environments. *Journal of Artificial Intelligence in Education*, *6*(2/3), 239–302.

diSessa, A. A. (2004). Metarepresentation: Native competence and targets for instruction. *Cognition and Instruction*, *22*, 293–331.

diSessa, A. A., & Sherin, B. L. (2000). Meta-representation: An introduction. *Journal of Mathematical Behavior*, *19*, 385–398.

Emmorey, K., & Casey, S. (2001). Gesture, thought and spatial language. *Gesture*, *1*, 35–50.

Ericcson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, *87*, 215–251.

Gilmore, D. J., & Green, T. R. G. (1984). Comprehension and recall of miniature programs. *International Journal of Man-Machine Studies*, *21*, 31–48.

Graesser, A. C., Lu, S., Olde, B. A., Cooper-Pye, E., & Whitten, S. (2005). Question asking and eye tracking during cognitive disequilibrium: Comprehending illustrated texts on devices when the devices break down. *Memory & Cognition*, *33*, 1235–1247.

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

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

Hegarty, M., Mayer, S., Kriz, S., & Keehner, M. (2005). The role of gestures in mental animation. *Spatial Cognition and Computation*, *5*, 333–356.

Henderson, J. M., & Ferreira, F. (2004). Scene perception for psycholinguists. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision and action: Eye movements and the visual world* (pp. 1–58). New York: Psychology Press.

Hoffmann, R. (1997). *The same and not the same*. New York: Columbia University Press.

Hoffmann, R., & Laszlo, P. (1991). Representation in chemistry. *Angewandte Chemie International Edition*, *30*, 1–16.

Irwin, D. (2004). Fixation location and fixation duration as indices of cognitive processing. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision and action: Eye movements and the visual world* (pp. 105–134). New York: Psychology Press.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, *40*, 1489–1506.

Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology*, *8*, 441–480.

Kozma, R. B., Chin, E., Russell, J., & Marx, N. (2000). The role of representations and tools in the chemistry laboratory and their implications for chemistry learning. *Journal of the Learning Sciences*, *9*, 105–144.

Kozma, R., & Russell, J. (1997). Multimedia and understanding: Expert and novice responses to different representations of chemical phenomena. *Journal of Research in Science Teaching*, *34*, 949–968.

Larkin, J., & Simon, H. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, *11*, 65–99.

Linn, M. C., Lee, H.-S., Tinker, R., Husic, F., & Chiu, J. L. (2006). Teaching and assessing knowledge integration in science. *Science*, *313*, 1049–1050.

Lowe, R. (2004). Interrogation of a dynamic visualization during learning. *Learning and Instruction*, *14*, 257–257.

Nathan, M. J., Stephens, A. C., Masarik, D. K., Alibali, M. W., & Koedinger, K. R. (2002, October). Representational fluency in middle school: A classroom based study. In D. Mewborn, P. Szatajn, D. White, H. Wiegel, R. Bryant, & K. Nooney (Eds.), *Proceedings of the Twenty-fourth Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Columbus, OH: ERIC Clearinghouse for Science, Mathematics, and Environmental Education.

Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research*, *42*, 107–123.

Patrick, M. D., Carter, G., & Wiebe, E. N. (2005). Visual representations of DNA replication: Middle grades students' perceptions and interpretations. *Journal of Science Education & Technology*, *14*, 1573–1839.

Peebles, D., & Cheng, P. C-H. (2003). Modeling the effect of task and graphical representation on response latency in a graph-reading task. *Human Factors*, *45*, 28–46.

Quintana, C., Reiser, B. J., Davis, E. A., Krajcik, J., Fretz, E., Duncan, R., Kyza, E., Edelson, D., & Soloway, E. (2004). A scaffolding design framework for software to support science inquiry. *Journal of the Learning Sciences*, *13*(3), 337–386.

Rayner, K. (1998). Eye movements in reading and information processing. *Psychological Bulletin*, *124*, 372–422.

Russell, J., Kozma, R., Jones, T., Wykof, J., Marx, N., & Davis, J. (1997). Use of simultaneous-synchronized macroscopic, microscopic, and symbolic representations to enhance the teaching and learning of chemical concepts. *Journal of Chemical Education*, *74*, 330–334.

Salvucci, D. D., & Goldberg, J. H. (2000). Identifying fixations and saccades in eye-tracking protocols. In A. Duchowski (Ed.), *Proceedings of the Eye Tracking Research and Applications Symposium* (pp. 71–78). New York: ACM Press.

Smallman, H. S., and St. John, M. (2005). Naïve realism: Misplaced faith in the utility of realistic displays. *Ergonomics in Design*, *13*, 6–13.

Stieff, M. (2005). Connected chemistry—A novel modeling environment for the chemistry classroom. *Journal of Chemical Education*, *82*, 489–493.

Stieff, M., & McCombs, M. (2006). Increasing representational fluency with visualization tools. In S. Barab, K. E. Hay, & D. T. Hickey (Eds.), *Proceedings of the Seventh International Conference of the Learning Sciences (ICLS)* (Vol. 1, pp. 730–736). Mahwah, NJ: Erlbaum.

Stieff, M., & Wilensky, U. (2003). Connected chemistry—Incorporating interactive simulations into the chemistry classroom. *Journal of Science Education and Technology*, *12*, 285–302.

Treisman, A. M., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, *12*, 97–136.

van Gog, T., Paas, F., & Van Merriënboer, J. J. G. (2005). Uncovering expertise-related differences in troubleshooting performance: Combining eye movement and concurrent verbal protocol data. *Applied Cognitive Psychology*, *19*, 205–221.

Wolfe, J. M. (1998). Visual search. In H. Pashler (Ed.), *Attention* (pp. 13–74). London: University College London Press.

Wu, H.-K., Krajcik, J. S., & Soloway, E. (2001). Promoting conceptual understanding of chemical representations: Students' use of a visualization tool in the classroom. *Journal of Research in Science Teaching*, *38*, 821–842.

APPENDIX

List of Questions Asked While Participants Were Viewing Each of the Interactive Visualizations

Stretch animation:

1. Tell me what this is showing. (pre eye tracker)
2. What is the most stable geometry?
3. As the distance between the two atoms increases/decreases what happens to the energy?
4. The normal O–O bond distance for HO–OH is 1.3Å; at a bond distance of 1.5Å its energy increases by 6 kcal/mol. According to the model, what would you expect the energy to be if the bond distance is 1.1Å?
   a) higher   b) same   c) lower

Dihedral animation:

1. Tell me what this is showing. (pre eye tracker)
2. What is the most stable conformation?
3. At what dihedral angle are the sets of atoms closest to each other?
4. In ethane (H_3C–CH_3), the energy is at a low at a dihedral angle of $60°$. What would you expect the energy to be at a dihedral angle of $240°$?
   a) higher   b) same   c) lower

Angle Bend animations:

1. Tell me what this is showing. (pre eye tracker)
2. What is the most stable geometry?
3. As "θ" decreases, what happens to the distance between atoms 1 and 3?
4. The normal bond angle for H_2O is $100°$; at $110°$ its energy increases by 10 kcal/mol. According to the model, what would you expect the energy to be if the angle is $90°$?
   a) higher   b) same   c) lower