TAPPING INTO AUGMENTED VIRTUALITY LABORATORIES: INVESTIGATING THE IMPACT OF USER INTERFACE ON STUDENT LEARNING IN SECONDARY SCIENCE

A Dissertation Presented to

The Faculty of the Curry School of Education

University of Virginia

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

by

Crystal J. DeJaegher

May 2015

ABSTRACT

Physical laboratories have been incorporated in K-12 science classrooms for the better part of the last two centuries, but research demonstrates that students need help developing and retaining deep understanding of observable phenomena during physical labs. Virtual laboratories help students interact with and manipulate unobservable levels of phenomena in ways that physical laboratories do not, yet students can fail to connect these experiences to the real world. Leveraging affordances of physical and virtual manipulatives in a mixed-reality environment may help students develop deep understanding of science by facilitating connections between observable and unobservable levels of phenomena.

This dissertation uses an explanatory-sequential mixed-methods approach to explore the effect of augmented virtual laboratories on high school students' conceptual understanding of gas properties. Building upon embodied cognition and knowledge integration perspectives, this dissertation investigates the effects of a specific augmented virtual technology, the Frame, on students' ability to make connections between observable gas properties and molecular-level behavior through a comparison study involving high school chemistry classes using the Frame to classes using purely virtual labs. The dissertation examines differences in student performance on conceptual assessments and uses video and interview data to explore students' interactions with the technologies. Research findings indicate that the Frame labs were just as effective as purely virtual labs for students' development of molecular-macroscopic connections regarding gas behaviors. Observations reveal that students working with the Frame tended to use non-prescribed, innovative investigative activities while students in the purely virtual condition focused on task completion. Results point to the need for the development of more sensitive assessments for scientific practices and suggest opportunities for further investigation.

TABLE OF CONTENTS

TABLE OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Lackluster student performance in science at the national level has prompted a push for more research into science education (National Research Council [NRC], 2010). Eighth grade students scored worse than 6 economically competitive countries in Europe and Asia and were outperformed in the physical sciences by students in 16 other education systems (Provasnik et al., 2012). Data from the 2009 Program for International Student Assessment (PISA) indicate that United States' students' average scientific literacy performance was exceeded by students from 18 other countries (Fleischman, Hopstock, Pelczar, & Shelley, 2010).

Results from the 2011 National Assessment for Educational Progress (NAEP) demonstrate that only 31% of 8^{th} grade public school students in the nation performed at or above the *Proficient* level on measures of scientific literacy (National Center for Education Statistics [NCES], 2012). In the context of the physical sciences, *proficiency* in the physical sciences involves students identifying chemical compounds, understanding chemical reactions, describing changes between kinetic and potential energy, and basic aspects of force and motion (NCES, 2012, p. 14). Among the scientific knowledge required for students to demonstrate an *Advanced* level of performance are the skills of predicting, observing, and explaining phenomena at multiple levels; all of which are hallmarks of authentic scientific practices.

These stagnant figures represent a less than progressive outlook for our nation's science education and, inherently, the United States' ability to produce citizens prepared compete and innovate in the global economy. As a result, politicians, educators and researchers have called for science education reform in the Next Generation Science Standards (NGSS; NGSS Lead States, 2013). The NGSS create challenging guidelines for instruction to integrate core ideas, scientific practices, and crosscutting concepts. The NGSS place particular importance on their definition of scientific practices, which includes: asking questions and defining problems; developing and using models; planning and carrying out investigations; analyzing and interpreting data; using mathematics and computational thinking; constructing explanations and designing solutions; engaging in argument from evidence; and obtaining, evaluating and communicating information.

With the aim of engaging students in these kinds of authentic scientific practices, science classes incorporate physical laboratories across grade levels. Physical laboratories provide students with direct, hands-on access to scientific phenomena. Physical lab experiences enable students to act as scientists, to investigate the world around them, find patterns and develop explanations of natural phenomena. In addition to acculturating students in authentic scientific practice, physical labs also give students the opportunity to inspect, touch, and manipulate physical objects and phenomena, which in itself can lead to increased engagement (Feisel & Rosa, 2005). Research also demonstrates instruction with physical or tangible objects can benefit learning (Gire, Carmichael, Chini, Rouinfar, & Rebello, 2010; Zacharia & Olympiou, 2011).

However, science laboratories are not always effective in building students' conceptual understanding (Finkelstein et al., 2005). The National Research Council has identified laboratory experiences as one facet of science education that needs improvement (NRC, 2005). In reality, students often follow rigid cookbook-style directions and leave with little understanding of the content or the practices of science

(Charen, 1970; Cracolice & Monteyne, 2004; NRC, 2005). For example, students engaging in a lab about the Gas Laws can successfully complete the lab by following directions to take measurements of pressure at different volumes to "discover" that when volume decreases, pressure increases. Although students may write down the right numbers and correctly identify a relationship, this does not mean that students have developed explanations to why this is the case or have developed skills that would enable them to investigate these kinds of questions independently.

Additionally, research documents the varied ideas that students have after typical instruction with physical laboratories (e.g., Levy, Novak, & Wilensky, 2006). In particular, students have difficulty connecting molecular-level properties to macroscopic phenomena (Ardac & Akaygun, 2004; Ben-Zvi, Eylon, & Silberstein, 1986) and similarly have difficulty developing explanations that connect macroscopic or observable phenomena with molecular-level behaviors. Sophisticated understanding of many science domains such as chemistry relies on students making normative connections among these levels (Gabel, 1999; Johnstone, 1993). For example, Ben-Zvi et al. (1986) found that students studying atomic structure typically ascribe macroscopic properties to the molecular level. In other words, students have trouble visualizing matter as a collection of many components that make up a substance, often focusing on an individual component, such as an atom or a molecule, to represent the whole substance (i.e., students refer to the states of matter in terms of a singular "gaseous atom" or a "solid atom", p. 65).

Educational technologies, such as simulations and visualizations, can provide students with direct access to scientific phenomena that is otherwise not readily available or easily accessible (Honey & Hilton, 2011). Simulations refer to computational models of systems or processes (de Jong & van Joolingen, 1998) where users can investigate novel situations and relationships between dynamic, real life variables and effectively test hypotheses (Lee, Plass, & Homer, 2006). Simulations help students develop knowledge and skills to comprehend authentic scientific problems (Barab & Dede, 2007) and can provide a foundation from which students develop informed scientific explanations. For example, a PhET simulation of gas molecules (http://phet.colorado.edu) enables students to investigate gas properties by manipulating variables (e.g., volume, temperature, and gravity) and see what happens to the molecules in a closed container.

As is the case with PhET, many simulations incorporate dynamic visualizations. Dynamic visualizations represent phenomena that are typically too small or large to be directly observed, such as cell division or the orbits of planets (Linn, Chang, Chiu, Zhang, and McElhaney, 2010). For example, the PhET gas properties simulation uses dynamic visualizations of molecules moving around a closed container instead of only numerical outputs. Dynamic visualizations can be valuable tools to motivate student learning (Corliss & Spitulnik, 2008) and have been successfully employed in science education to provide students with access to otherwise unseen phenomena (Honey & Hilton, 2011).

Many computer-based or technology-enhanced instructional units use simulations as part of virtual labs. I define virtual laboratories as computer-based scientific investigations that encourage direct manipulation of phenomena through simulations and visualizations. For example, virtual laboratories using Molecular Workbench can help students investigate, visualize, and understand chemical reactions (Xie & Tinker, 2006) through dynamic visualizations of diatomic molecule dissociations (e.g., the breakdown of H_2) and collisions between free radicals and diatomic molecules leading to the formation of a new molecule (e.g., chemical reactions leading to the formation of HCl). Such virtual laboratories provide access to otherwise unobservable phenomena using a safe, cost-effective medium of delivery.

Virtual laboratories using visualizations have proven effective for student learning in science (Bell & Trundle, 2008; Honey & Hilton, 2011; Klahr, Triona, & Williams, 2007; Zacharia & Olympiou, 2011), and they are especially helpful to students when they are asked to make connections between macroscopic and microscopic levels of phenomena (Kozma & Russell, 1997; Levy & Wilensky, 2009; Wu, Krajcik, & Soloway, 2001). For example, Connected Chemistry (Levy & Wilensky, 2009) anchored a visualization-based gas laws curriculum in the macroscopic world and used instruction to help students (n=904) connect virtual NetLogo models to the particulate level. Students working with visualizations in activities that drilled down into the particulate-level behavior, improved their ability to bridge molecular and observable levels. Exposure to virtual scientific phenomena can lead to increased student performance in science when key elements of scientific practice, such as developing models and constructing scientific explanations, are emphasized (Bell & Trundle, 2008; Levy & Wilensky, 2009; Blikstein, Fuhrmann, Greene, & Salehi, 2012).

Although several empirical works have shown the benefits of using visualization-based instruction in science, visualizations can also lead students to engage in less meaningful, off-task behaviors such as clicking without thinking (Pillay, 2010). Additionally, students find it difficult to translate between types of visual representations (Ainsworth & van Labeke, 2004) especially in science, when conceptualizing complex phenomena across levels (i.e., microscopic, submicroscopic, macroscopic, and symbolic; Wu & Shah, 2004). Visualizations of complex phenomena are often simplified for instructional purposes (Ainsworth, 2006) and without adequate support, students may not understand the representation (van der Meijj & de Jong, 2006). Students working with visualizations often believe that they understand more than they actually do (Linn et al., 2012) and fail to make connections between real-world experiences and the visualization (Chiu, 2010). Similarly, students can become focused on superficial elements of visualizations (Lowe, 2004), such as the color selection for representing molecules or lines representing bonds, instead of developing a deeper conceptual understanding.

Because research documents benefits and drawbacks to both physical and virtual labs in classrooms, several researchers have attempted to leverage the affordances of physical and virtual labs in various combinations. Some researchers have investigated providing students with sequential combinations of physical and virtual labs (e.g., students engage in a physical lab, then a virtual lab, or vice versa). Research demonstrates that some sequential implementations of virtual and physical labs can positively affect student learning (Gire et al., 2010; Zacharia, Olympiou, & Papaevripidou, 2008). For instance, undergraduate physics students using both real and virtual labs in succession had improved conceptual understanding compared to the real lab control condition (Zacharia, 2007). Some studies find benefit using virtual materials before physical materials (Zacharia & Olympiou, 2011), whereas others find no difference in the order (Smith & Puntambekar, 2010), but overall benefit of the combination of physical and virtual (de Jong, Linn, & Zacharia, 2013).

Researchers have also combined virtual and physical labs by side-by-side *bifocal modeling.* In a side-by-side or bifocal approach, students construct virtual models at the same time as seeing a physical model instead of having a physical and then a virtual experience (Blikstein & Wilensky, 2007). For example, using NetLogo and a GoGo Board (Sipitakiat, Blikstein, & Cavallo, 2004), an interface that translates analog to digital representations, students were tasked with building models of scientific phenomena (e.g., acid-base reactions). Students successfully investigated their hypotheses and engaged in problem-solving tasks with their physical models and computer-generated algorithms on the visual display. By having the virtual and physical lab available at the same time, students can compare their virtual model to the real phenomena develop understanding, similar to practicing scientists. Studies suggest that this kind of bifocal approach can be successful with science students (Blikstein, Fuhrmann, Greene, & Salehi, 2012).

Emerging mixed-reality (MR) technologies render it possible to leverage the affordances of both virtual and physical labs into a singular experience. MR technologies capitalize on the affordances of both physical and virtual experiences (Abelson et al., 2008) and have shown promise in science education (Johnson-Glenberg, Koziupa, Birchfield, & Li, 2011; Novellis & Moher, 2011). The capability of these types of MR technologies to enhance educational experiences may be best understood in considering the combination of virtual and physical elements in terms of a reality-virtuality continuum (Milgram & Kishino, 1994; Figure 1). This continuum categorizes different types of MR environments spanning a two-dimensional plane from reality to complete virtuality, including augmented reality (AR) and augmented virtuality (AV). The difference between AR and AV is whether the target phenomena are real or virtual, and thus whether reality or virtuality is enhanced (Liu, Cheok, Mei-Ling, & Theng, 2007).

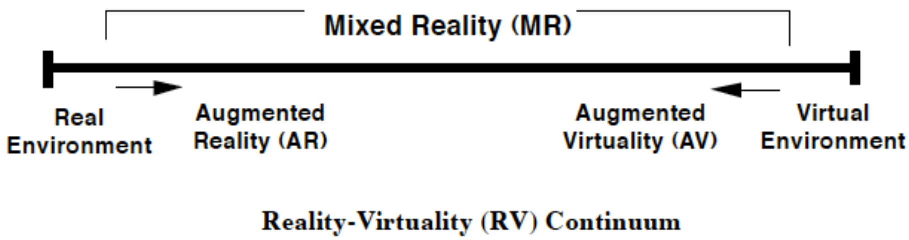


Figure 1. Milgram and Kishino's (1994) Virtuality Continuum Model

AR technologies used in education typically make use of virtual overlays in a real world environment. For instance, students can use cell phones or hand-held tablets to overlay information on real-life objects. Implementations of augmented reality technologies in education have mostly been limited to university-based research and isolated clinical studies due to the constraints of head-mounted or eyeglass displays (e.g., Kaufmann & Dünser, 2007), the application of AR in classrooms is continuing to evolve (e.g., Yuen, Yaoyuneyong, & Johnson, 2011). Additionally, most implementations enhance reality with macroscopic or symbolic notations. At best, augmented reality would enable students to zoom into an object to see the underlying process, but students would not be able to directly interact with the visualization and see how it relates to the real-world object.

Augmented virtuality (AV) approaches may benefit students learning science. Such approaches combine affordances from physical and virtual labs into a singular experience. In particular, augmented virtual technologies use *virtual* phenomena that provide learners with direct experience with unobservable levels, such as molecular behavior through simulations, drawing upon affordances of virtual labs. Additionally, the physical controls give students real objects to manipulate, drawing upon affordances of physical labs. The simultaneous connection between the real-world controls and the molecular visualization can help students develop connections between molecular and macroscopic levels. As AV technologies are relatively new, there is a dearth of empirical studies that examine augmented virtual technologies in science classrooms (Chao et al., 2014; DeJaegher, 2014; DeJaegher, Chiu, & Chao, 2014).

Purpose of the Study

Students often struggle to make connections between observable levels of scientific phenomena and the underlying molecular mechanisms that affect what can be directly observed (Finklestein et al., 2005). This lack of connection exists even in the presence of laboratory activities that are designed to connect scientific ideas being studied in a textbook to authentic science practice. The structure and design of laboratory experiences in secondary school may impact the connections that students can reasonably be expected to make. Several alternative approaches to traditional physical labs have been implemented in science classrooms with mixed results.

This dissertation investigates how augmented virtual technologies might be used in authentic classrooms to improve students' understanding of complex phenomena. Specifically, this study compared learning outcomes between students using augmented virtual approaches to purely virtual technologies in high school chemistry classes.
Past studies with the Frame have demonstrated learning outcomes (DeJaegher, Chiu, & Chao, 2014) and added benefit over traditional physical lab instruction (Chao et al., 2014). Additionally, I explore the kinds of interactions that students have with the different

technologies. Analyzing the way students interact with the interface may provide meaningful insights as to how they use augmented virtual approaches to help construct understandings of complex science phenomena. In particular, this dissertation addresses the following research questions:

1. What differences, if any, are there between students using the Frame and students using virtual labs on students' conceptual understanding of gas properties?
2. What differences, if any, are there between students using the Frame and students using virtual labs on conceptual retention over time?
3. What characterizes students' interactions using the Frame compared to students' interactions when using the visualization only and how might students' interactions explain learning outcomes?

Hypotheses and Expected Outcomes

The hypotheses corresponding to the research questions are grouped in terms of conceptual understanding (i.e., research questions 1 and 2) and interactions (i.e., research question 3). My general hypothesis is that students using the Frame will perform better than the visualization only group and exhibit more on-task behaviors. For the analysis of interactions, the triangulation of quantitative data (i.e., KI scores), video observations, and the semi-structured interview following the laboratory informed the interpretation of the results.

Conceptual Understanding

Increased conceptual understanding in Frame condition. Embodied cognition posits that learning is embedded in physical actions (Barsalou, 2008). Existing research involving tangible user interfaces (TUI) and haptic feedback suggests that students learn better when they are physically connected to their learning (Han & Black, 2011). The Frame is an example of a TUI, and it provides an experience that is likely to result in increased students' conceptual understanding (in terms of KI score gain) when compared to students using the graphical user interface (GUI), also known as the visualization only condition. This may be related to physical interactions facilitated by the interface. The TUI encourages students to interact with the computer in a novel way via external inputs. The visualization asks no more of students than to click a mouse—in this way, the interaction with the visualization represents a ubiquitous application of technology that is typically used in homes and classrooms. It is possible that by prompting students to physically engage with the Frame apparatus, their tactile experiences and knowledge may be coded simultaneously, facilitating greater recall for students using the Frame than those assigned to the visualization only condition. This is one expected outcome in response to the first research question of this study. Regarding the second research question, this hypothesis posits that the students in the Frame condition would have greater retention over time than students in the visualization only condition.

No difference in conceptual understanding. While the existing research using augmented virtuality applications in authentic classroom environments is scant, several studies involving physical and virtual manipulatives indicate that there is no difference in student learning outcomes between these types of conditions. It may be the case that no significant difference will be discovered; that is, students using an interface with physical controls will demonstrate a similar level of conceptual understanding as the students using an interface with graphics manipulated using a mouse. One reason for this could be the treatment duration; a 90-minute block of time may not be enough for a student to become familiar and to meaningfully engage with a novel technological application in a secondary science laboratory setting. This is another expected outcome in response to the first research question of this study. Regarding the second research question, this hypothesis posits that the students' retention in both conditions would not exhibit any statistically significant differences when compared.

Interactions

GUI interactions show more off-task behaviors. It is possible that students may show more off-task behaviors with the GUI. Off task behaviors can be defined as behaviors in which the students engage that do not pertain directly to the use of the GUI, the curriculum packet, or the laboratory. Student interactions are expected to be more off-task (e.g. students may manipulate the visualizations color scheme, drag objects, or click buttons randomly) in the GUI environment. This may be because of the ubiquitous nature of the setup—it looks just like any other computer, and students may perceive the activities to be monotonous. Student-student interactions are also expected to be more off-task (e.g. students socializing) for similar reasons. As a result, it may be the case that student-teacher interactions reflect the redirection of behavior instead of casual troubleshooting or asking deep questions to encourage student thinking about the topic of gas laws. This hypothesis is one expected outcome in response to the third research question that seeks to identify the behaviors exhibited by students in each group.

TUI interactions show students more on-task, inquisitive behaviors. Both empirical and qualitative studies involving hands-on learning opportunities often yield positive results for student affect, creativity (Mostmans, Vleugels, and Bannier, 2011), and constructing meaning (Varelas, Pieper, Arsenault, Pappas, & Keblawe-Shamah, 2014) Considering the posited framework of embodied cognition, I expect that students will be more engaged with the Frame in the TUI condition. I believe that student-technology interactions may be reflective of play and exploration in an attempt to understand the relationships of the properties of gases, and that most of these interactions will be on-task and related to the instructional activities. Students in the TUI condition may have a similar level of off task behaviors, but I think it's more likely that the conversations taking place will relate to the task. Finally, I think student-teacher interaction will be minimal, and the student-teacher interaction that is present will represent troubleshooting or content facilitation. This hypothesis represents one expected outcome in response to the third research question that seeks to identify the behaviors exhibited by students in each group.

Significance of the Study

This study aims to contribute to an emerging body of literature involving mixed-reality environments and student learning outcomes (Johnson-Glenberg, Birchfield, Tolentino, and Koziupa, 2014; Lindgren & Moshell, 2011; Novellis & Moher, 2011; Price & Falcão, 2011; Price, Sheridan, Falcão, & Roussos, 2008). Currently, there is a dearth of empirical studies looking at mixed-reality technologies in authentic classroom settings (Lindgren & Johnson-Glenberg, 2013; Lindgren & Moshell, 2011). Existing empirical work tends to focus on augmented reality or virtual reality, which involves supplementing the existing reality or environment by integrating virtual interactions (e.g., Arvanitis et al., 2009; Dunleavy, Dede, & Mitchell, 2009; Ibáñez, Di Serio, Villarán, & Delgado Kloos, 2014; Squire & Jan, 2007; Squire & Klopfer, 2007). In contrast, few research studies have explored augmented virtual technologies on science learning (Chao et al., 2013; Chao et al., 2014; DeJaegher, Chiu, & Chao, 2014; Lindgren & Moshell, 2011).

This comparison study was designed to isolate the component of physical augmentation via comparison to a virtual lab using a standard, graphical user interface (GUI) with the same visualization to determine whether there are any differences in students' performance. The explanatory-sequential mixed-methods approach aims to provide insight into not only what but also how students learn with these technologies. This work has the potential to influence future implementations of augmented virtual technologies in classrooms by documenting what works and what may not work in authentic settings. The knowledge gained from this study will likely have benefits extending beyond science education into other academic disciplines that look to use augmented virtual technologies for instruction.

Delimitations and Limitations

This study investigated the effect of augmenting virtual labs on student learning of gas laws phenomena in a mixed-reality environment. It is not clear to what extent any findings from this study can be directly translated to other subject areas or instantiations of mixed-reality environments. While it may be possible to apply findings to comparable mixed-reality environments and to other science classrooms, this study is highly situated and context dependent. Detected effects may be present based on the population makeup or ability grouping of students, which may not be representative of a larger population demographic.

Students in this implementation worked in groups of 2-4 students. Because working in groups is naturally a social and collaborative effort, it may be that the group dynamics influenced students' understanding more than the technology itself, for better or for worse. The effect of students' collaboration will not fully be investigated in this study; the nature of students' collaboration in this environment may be an avenue for future work.

The usability of each type of user interface may influence the learning that takes or fails to take place. The user interface is an important component of this technology, and as such, could be instrumental in determining how users receive and interpret information from the connected experiences. The interface may also play a role in how users transition between and interpret the mixed-realities. Part of this investigation is concerned with the role that user interface may play in terms of how students develop conceptual understanding of complex science phenomena and what interactions students have with the technology. The second part of this dissertation will provide information about how users apply information they receive in both conditions. As this is an observational aspect, what is known is only what the student says or does, and this is not necessarily a direct reflection of their thinking or their learning.

CHAPTER 2

REVIEW OF LITERATURE

To understand the current state of affairs in science laboratory education, I review existing literature on the impact of physical and virtual laboratory experiences on students' understanding of complex phenomena. This is followed by mixed-reality implementations inside and outside of science education, in order to position a specific type of mixed-reality (i.e., augmented virtuality) as a viable option for student learning in education. Literature focused on the particulate nature of matter (PNM) is interspersed throughout this portion of the review, as this is the content area of focus for this study. To understand the potential impact of the user interfaces involved in this study regarding users' experience and interactions, I review literature from the field of Human-Computer Interaction (HCI) and Computer Supported Collaborative Learning (CSCL). Finally, this chapter concludes with an overview of the theoretical frameworks that guide the study by offering perspectives on how students learn through the manipulation of physical and virtual tools followed by the introduction of key terms and definitions necessary prior to the methods chapter.

Background

The Particulate Nature of Matter

This study focuses on student learning relative to the particulate nature of matter (PNM), which is central to the molecular study of gas properties. The PNM is considered to be foundational knowledge for the development of complex scientific understanding (Özmen, 2013). The PNM is essential to build normative understandings in chemistry

(Gabel et al., 1997; Nakleh, 1992; Snir et al., 2003) as well as in physical, life, and earth sciences (Benson et al., 1993; Bouwma-Gearhart et al., 2009; Lee et al., 1993; Noh & Scharmann, 1997). Because of the importance of the PNM, it is a subject of emphasis in secondary school curricula (e.g., NGSS, 2013). In spite of the PNM being an essential building block across science topics, research demonstrates that students at all ages, even chemistry graduate students and teachers, have a variety of ideas about the PNM, gases, and gas laws (Bodner, 1991; Krajcik, 1995; Lin, Cheng, & Lawrenz, 2000; Nakhleh, 1992). Each of the following sections relates existing research on pedagogical approaches the PNM where applicable.

Physical Laboratories in Science

For two centuries, laboratory experiences have been an established component of American K-12 science education at all levels. Laboratory experiences in science classrooms give students an opportunity to interact directly with scientific phenomena and engage with systematic, scientific practices (NRC, 2005) and can help students grasp difficult scientific concepts (Tobin, 1990). Laboratory experiences are generally most helpful for student learning when embedded within science instruction (Sadler & Tai, 2001), however, extensive evaluation of lab experiences in American schools suggests that laboratory experiences are unsatisfactory for most students (NRC, 2005).

Typically, physical laboratory activities include concrete objects that assist in evoking students' experiential knowledge and to engage students with scientific phenomena. Physical laboratories allow students to experiment with real materials that facilitate direct interaction with scientific phenomena (Gire et al., 2010) and concrete objects can also lead students to become more engaged in science by way of tactile experience and physical examination (Feisel & Rosa, 2005).

Although physical laboratories have been widely used for decades, existing research points to the persistence of student misunderstandings and alternative conceptions, especially regarding scientific phenomena with underlying, invisible mechanisms, such as the particulate nature of matter, gas laws, kinetic molecular theory (e.g., Novick & Nussbaum, 1981). For example, Levy, Novak, and Wilensky (2006) discovered students' ideas about particulate motion to be inconsistent with scientifically normative explanations. Many students believed that particle collision results in increased particle speed or that an increased number of particles in a sealed container decrease the speed at which particles move. Additionally, several other research studies indicate that students have trouble distinguishing between macroscopic and molecular levels of phenomena (Ardac & Arkaygun, 2004; Ben-Zvi, Eylon, & Silberstein 1986) and confuse characteristics of each level (e.g., Wilensky & Resnick, 1999). Students' alternative ideas are not at all surprising given that physical laboratories generally do not include representations of scientific phenomena at the atomic level. The atomic level is not readily and visibly accessible for observation. Students can generally articulate an understanding of macroscopic phenomena, yet they often fail to make macroscopic-molecular connections (Gabel, 1999; Johnstone, 1993). For example, students conducting a lab about the gas laws may not make connections to the underlying explanation of the kinetic molecular theory (Liu, 2006). Even when subsequent instruction targets molecular-level explanations, students have difficulty making connections among scientific phenomena studied in a physical lab setting, classroom instruction, and real-world experiences (Finkelstein et al., 2005).

Visualizations in Science

Similar to physical laboratories, visualizations can provide students with direct experiences with virtual scientific phenomena. Linn et al. (2010) define visualizations as "computer-based animations of scientific phenomena" (p. 235). Research demonstrates that such technology-enhanced approaches can help students understand chemical phenomena on a molecular level (e.g., Ardac & Akaygun, 2004; Jones, Jordan & Stillings, 2005; Kelly & Jones, 2007; Madden, Jones & Rahm, 2011; Sanger et al., 2000; Stieff & Wilensky, 2003). Visualizations provide pictures of phenomena that are typically unseen (Winn et al., 2006) while generally providing multiple representations. Visualizations have been used extensively in science education, as visualizations allow students to investigate and manipulate unobservable levels (Honey & Hilton, 2011). Visualizations are particularly useful in helping students comprehend molecular concepts, including the particulate nature of matter and gas laws (Chiu & Linn, 2014; Kozma & Russell, 1997, Levy & Wilensky, 2009; Wu, Krajcik & Soloway, 2001). For example, Ardac and Akaygun (2004) compared a multimedia approach with traditional classroom instruction (i.e., lecture and discussion) with middle school students (n=49) across two weeks of class time (~450 total minutes) to see if a multimedia unit emphasizing visualizations of molecular representations would improve students' understanding across levels of phenomena. The technology used in this study enabled students to navigate between different representational levels. For instance, students could see chemical reactions displayed macroscopically (e.g., $AgNO_3$ and NaCl mixing in a flask) and microscopically (e.g., 3-D molecular structures illustrating the mix of chemicals). Student performance on pre-/post- test assessments and interviews demonstrated that the multimedia approach helped students understand molecular phenomena across molecular, macroscopic, and symbolic levels. Similarly, Wu, Krajcik, and Soloway (2001) successfully implemented a software visualization tool, eChem, which lets users construct molecular models and view a variety of representations at the same time. With 71 eleventh grade students, Wu et al. (2001) investigated whether the software helped students translate representations and what particular features of the software were instrumental in this process. Students' scores improved 2.25 standard deviations on conceptual assessments from the pretest to the posttest.

Visualizations can also help students make visual, conceptual, and referential connections to molecular phenomena through the use of linked representations (Kozma, 2000) or leveraging explicit guidance during instruction (Chiu & Linn, 2014; Zhang & Linn, 2013). For example, Connected Chemistry (Levy & Wilensky, 2009) encourages high school students to make connections between molecular and observable levels by grounding investigations in the real world with instruction involving atomic-level NetLogo models. Using a pre/post design, Levy and Wilensky (2009) found a marked improvement in students' understanding of the microscopic level of particulate phenomena. Additionally, students made increased connections between the submicroscopic and macroscopic levels. Instructional guidance that shepherds students to make these types of connections from molecular-level visualizations to observable phenomena can improve conceptual understanding (Chiu & Linn, 2014).

Visualizations in science address another gap exposed in students' understanding through physical laboratories. Where students often fail to make connections to the real-world following physical laboratories (see Finkelstein et al., 2005), visualizations can help students make connections between scientific phenomena and their own experiential understandings. Using two versions of the same simulation in a semester long computer-enhanced curriculum, Clark and Jorde (2004) found that 8^{th} grade students (n=120) using a simulation of energy transfer incorporating a representation of a human hand performed better on posttests than their peers who used the simulation version without the explicit macroscopic-level component of the human hand. Additionally, the students assigned to the condition with the representation of the hand also showed statistically significant improvement in written explanations of thermal concepts covered in the unit. These findings suggest that students may be able to refine their understandings of complex phenomena when visualizations are connected explicitly to experiential ideas.

Although visualizations have demonstrated benefit for scientific learning (Honey & Hilton, 2011; Bell & Trundle, 2008; Dori & Belcher, 2005; Jaakola, Nurmi, & Veermans, 2011; Korakakis, Boudouvis, Palyvos, & Pavlatou, 2012; Lee, Linn, Varma, & Liu, 2010; Zacharia & Anderson, 2003; Zhang & Linn, 2011), other studies point out several drawbacks to visualization-based instruction. Dynamic visualizations can foster *deceptive clarity* and provide students with a false sense of understanding (Linn et al., 2010). For instance, students using a molecular simulation may be able to recall that there were several yellow spheres bouncing around on the screen, but they may understand little beyond what is represented by surface features. Simulations may decontextualize the phenomena being studied, contributing to students' difficulties with conceptual learning (Hofstein & Lunetta, 2004). Visualizations may also distract students, promoting passive learning or off task behaviors such as clicking without thinking (Pillay, 2010), and students sometimes focus on superficial elements of visualizations instead of learning outcomes (Lowe, 2004).

Additionally, some research suggests that there may be no difference between physical and virtual labs (Triona & Klahr, 2003; Klahr, Triona, & Williams, 2007; Zacharia & Constantinou, 2008). For example, Klahr, Triona, and Williams (2007) tasked middle school students (n=56) with designing a mousetrap car and determining the distance it would travel. Klahr et al. compared students' knowledge gains between the physical manipulatives only condition and the sequential ordering of physical and virtual manipulatives. Using a pre/post design, specifically examining the change in number of correct answers from the pre to the posttest, results indicate that while students in the virtual condition built their cars faster (as expected), there was no difference in student learning between physical and virtual conditions.

Combining Physical and Virtual Science Laboratories

Because both virtual and physical labs have some benefit for science learners, researchers have investigated how physical and virtual laboratory experiences can be used sequentially to assist students in understanding scientific phenomena (Gire et al., 2010; Olympiou & Zacharia, 2012; Zacharia & Anderson, 2003; Zacharia, 2007; Zacharia, Olympiou, & Papaevripidou, 2008b). For example, Zacharia et al. (2008b) investigated undergraduate students' (n=62) conceptual understanding of heat and temperature in an introductory-level physics course. The control group engaged in a physical lab and the experimental group used a physical lab followed by a virtual lab.

Qualitative analysis of the explanations provided by students in open-response questions across tests indicated that students with the combination of virtual and physical materials outperformed students in the control condition. Zacharia et al. (2008b) suggest that this difference was due in part to the speed at which the virtual components could be manipulated.

Combining physical and virtual approaches has also been beneficial for students studying gas laws. For example, Liu (2006) had high school students sequentially engage in both computer-based molecular simulations and a hands-on lab to learn about gas laws. Results demonstrated that the combination of a physical and virtual lab was better than either approach alone to develop conceptual understanding of gas laws.

Across several different content areas and classroom contexts, it seems that the combination of virtual and physical experiences has merit for improving student learning (Gire et al., 2010; Price & Rogers, 2004; Zacharia, 2007). The combination of physical and virtual experiences affords greater opportunities for students to learn and make connections than any singular approach (de Jong, Linn, & Zacharia, 2013). However, research findings concerning the effects of sequential order is mixed (Smith & Puntambekar, 2010). For example, using the ComPASS curriculum content of simple machines (Puntambekar, Stylianou, & Goldstein, 2007), Smith and Puntambekar (2010) investigated students' learning about pulleys in a physical to virtual (PV) sequential condition (n=45) compared to a virtual to physical (VP) sequential condition (n=17). Overall results revealed that students learned more in the PV condition, suggesting that virtual experiments are more helpful for learning when following the physical lab. Further analysis of the pre/mid/posttest questions showed that the PV condition was more beneficial for students learning specific concepts (i.e., making comparisons between fixed and moveable pulleys, with regard to force and mechanical advantage), while other questions comparing fixed pulleys in different configurations showed that students' gained from the virtual component, in the PV sequence. Students performed equally as well in all other questions. The conclusion of this study is that the order of implementation and the learning gains are dependent upon the ways in which cognition is grounded in perceptual experiences; whether learners are able to ground their knowledge better in physical or virtual conditions varies across contexts.

A related approach to sequential implementation is Blikstein's (2007) *bifocal modeling,* which encourages student participation in authentic scientific practices using side-by-side physical and virtual models. Blikstein, Fuhrmann, Greene, and Salehi (2012) had 9th-11th grade students construct physical models that they used to compare and contrast with virtual models to refine their understanding of complex science concepts. Preliminary conclusions, gathered from classroom observations during interventions, point to the specific benefit of having students compare physical and virtual models to build understanding. The side-by-side approach may be particularly beneficial to help students understand and engage in authentic scientific practices, such as modeling phenomena (Blikstein & Wilensky, 2007).

Mixed Reality Technologies

Mixed-reality (MR) technologies can take advantage of the combination of both real and virtual environments through a seamless integration of virtual and physical interaction (i.e., physical and digital objects exist simultaneously and can be used together). MR technologies capitalize on the affordances of both physical and virtual experiences (Abelson et al., 2008) and have shown promise in science education (Johnson-Glenberg, Koziupa, Birchfield, & Li, 2011; Novellis & Moher, 2011). MR approaches include augmented reality (AR) and augmented virtuality (AV) technologies.

Augmented Reality. AR allows the real world to be the main driver for interaction with virtual objects. Existing research has investigated AR by way of user experience (Kaufmann & Dünser, 2007; Martìn-Gutièrrez, Contero, & Alcañiz, 2010; Olsson, Kärkkäinen, Lagerstam,Ventä-Olkkonen, 2012; Theng, Mei-Lei, Liu, & Cheok, 2007), embodied learning (Johnson-Glenberg, Birchfield, Tolentino, & Koziupa, 2014), and the affordances of AR technologies for science learning (Cheng & Tsai, 2013). In the field of education, AR dominates the types of mixed-reality applications for learning.

One example of AR in education is *Mad City Mystery*, a game leveraging mobile technologies (i.e., iPods, iPads, mobile phones) that allow students to investigate a murder mystery at Lake Mendota, an urban watershed region in Madison, Wisconsin (Squire & Jan, 2007). Students working in teams are assigned a role and tasked with physically observing environmental phenomena on-site and conducting virtual interviews with key stakeholders. Through the use of mobile devices, students collect data to solve the mystery while developing scientific argumentation skills and a conceptual understanding of geochemical water cycles. The physical location of this game is the primary driver of the interaction, which is augmented by virtual data collection components.

AR technologies typical use virtual overlays in a real-environment; the overlays can be superimposed using a head mounted display (i.e., a camera mounted on a device that is worn by the user; i.e., image-based AR) or implemented using a mobile GPScapable device (i.e., location-based AR). Researchers have identified several affordances of AR that may benefit student learning, including real-time contextualized augmentation (Squire & Klopfer, 2007), virtual object interaction (Kerawalla, Luckin, Seljieflot, & Woolard, 2006), and potential to scaffold spatial recognition and visualization of abstract concepts (Arvanitis et al., 2007).

Notable drawbacks to using AR technologies in a classroom include cost and school-based logistical issues, participant sickness (Kaufmann & Dünser, 2007), computing processing delays between real-world information and virtual overlays, and a lack of a dynamic feedback system, potentially leading to a less authentic experience for learners (Rosenbaum, Klopfer, & Perry, 2007). While researchers seem to agree that AR is a promising technology for teaching and learning, to fully leverage the affordances of the AR, a better understanding of instructional design and the pedagogical strategies necessary to implement these technologies (Dunleavy, Dede, & Mitchell, 2009).

Augmented Virtuality. AV differs from AR in that it uses a virtual-world as the main driver of interactions with physical objects. For example, *SMALLab* (Johnson-Glenberg et al., 2011) encourages students to engage with multiple modalities in an open space with infrared cameras that are used to detect students' movements. Students can interact with virtual molecules by moving their hands and see the result of their actions in a simulation projected on the ground. Similarly, AquaRoom (Novellis & Moher, 2011) is a spatial simulation of water flow that helps students understand topographical concepts and directional flow in the context of hydrological practices (i.e., using dye tags and examining water samples from aquifers).

Scant research has used AV technologies in authentic classroom contexts (e.g., Johnson-Glenberg et al., 2014; Tolentino et al., 2009; Yannier, Koedinger, & Hudson, 2013). However, studies that have been implemented in real classrooms demonstrate potential to help students learn STEM content. Johnson-Glenberg et al. (2014) specifically examined the potential of AV technologies through implementations of *SMALLab*, using data from two separate high school science classroom implementations. The first implementation contained 9th and 11th grade chemistry students (n=51) learning about chemical titration using a virtual tracking wand to move molecules in space to interact with a virtual flask. The second involved biology students (n=56) learning about disease transmission through the manipulation of virtual avatars to understand the spread of infection. Both implementations spanned 3 days, with comparison classes (i.e., lecture and hands-on lab sessions). Results found that students learned more in the SMALLlab classes than did their peers with regular classroom instruction across groups participating in both studies. The SMALLab condition registered moderate to large effect sizes (i.e., 0.53 to 1.93) as an indication of relatively large learning gains, while the classroom condition's effect sizes were substantially smaller (i.e., 0.09 to 0.37).

Aside from this data regarding mixed-reality environments (largely AR focused) in classrooms, there is a definite gap in the literature regarding AV technologies in classrooms. It may be that AV technologies have potential address several of the identified shortcomings of AR. Research findings in science education, regarding physical and virtual experiences and student learning, make it apt to suggest that combining these experiences in a connected, simultaneous manner may go a step further to help develop student understandings of complex science topics, including making particularly molecular-macroscopic connections across multi-level phenomena.

Interactions

Mixed-reality approaches to science learning point to the need to clearly articulate why and how a designer augments the virtual or physical environment. These augmentations change how learners interact with the technology. In this section I define and review literature about learner interactions with technology.

The term *interaction* is widely applied as a theoretical construct to make sense of exchanges between two or more people, objects, and events. Literature concerning interactions in educational research is generally constrained within web-based online environments or distance education and tends to focus on social exchanges (Kim & Hannafin, 2011). Classifying interactions in this domain has commonly been in accordance with Moore's (1989) typology: learner-learner, learner-teacher, and learner-self. Learner-interface interactions were addressed several years later (Hillman, Willis, and Gunawardena, 1994), and this consideration also rests primarily in distance education. Outside of distance education, there is a lack of empirical work examining students' interactions with technology in smaller scale implementations, such as the intervention in the proposed study.

In spite of the lack of explicit empirical work on interaction, Wagner (2006) has identified two commonly held assumptions about learning and interaction: (1) the quality of a learning experience is are generally perceived to be proportional with the level of interaction, and (2) maximizing benefits of interaction is necessary if technology-mediated learning initiatives are to impact teaching practices. Based on these observations, the relationship between technology, learning, and interaction is that an increased number of interactions leads to positive learning experiences and learning experiences leveraging technology must be interactive to have a lasting pedagogical impact.

Literature in Computer-Supported Collaborative Learning (CSCL) has traditionally focused on social interactions that take place between people using technology (Kreijns, Kirschner, & Jochems, 2003). Rather than assuming that interaction spontaneously happens as a part of a learning environment, CSCL studies place interaction as an intrinsic component of any instructional design (Northrup, 2001; Sims, 1997).

In the field of educational research, the term *interactivity* has also gained steam; however, it is not clearly defined and not distinctly separated from *interaction* in the literature. What is clear is that the definitions for both terms vary across content domains and lack consistency; several metaphorical interpretations of *interaction* have been identified (Wagner, 2005). For the purposes of this study, the term *interaction* will adopt one of two metaphorical interpretations, depending upon contextual references: interaction as transaction or interaction as experience.

In the transactional sense, interactions in this study will be used to describe a reciprocal exchange of mutual influence that requires a minimum of two objects and/or actions. These descriptions will primarily be based in the context of student-student or student-technology transactions in accordance with Moore's typology (1989).

Interfaces

The way that a user interacts with a technology is through an interface. This section discusses interfaces drawing on literature from computer science and human-computer interaction.

A computer interface, or user interface (UI), generally consists of hardware, software, and information inputs and outputs. Inputs allow users to interact with or manipulate variables within the system while outputs produce the results of user manipulation. Existing UIs are based on metaphors (Erickson, 1990). Interfaces operate on common metaphors in an effort to anchor new concepts, ideas, and actions, into a user's existing knowledge base. Several researchers oppose the use of such metaphors because they can restrict innovation of functionality (Gentner and Nielsen, 1996), literal translation from physical to virtual can lead to poor designs (Mullet & Sano, 1995), and violate basic design principles (Nelson, 1990). These metaphors are largely familiar to adults and occasionally problematic for novices—for example, the virtual desktop metaphor display on most computer operating systems make use of file folders for which a novice may not have a representational analog due to a lack of exposure or inexperience. Schneider (1996) discovered that novices using interfaces incorporating unfamiliar metaphors could use them successfully, if they are direct and predictable.

The proliferation of innovative technologies has paved the way to rethinking interface interaction paradigms. These paradigms represent designers' conceptualization of end-user interaction with the technology and ultimately the design for such interactions. Interface interactions translate to user experiences, and unique interface components combine to create unique user experiences (Hinckley & Wigdor, 2012) and as such, each interface has affordances and constraints.

Interactive technologies should be designed for learner's experiences to help them develop knowledge about a given domain (Hourcade, 2007). Applications that involve students interacting with real-time data may help learners confront their existing beliefs about phenomena and consolidate their experiential and conceptual understandings. Additionally, innovative interfaces can capitalize on different data sources and multimodal representations that have potential to increase opportunities and access for students to develop understandings of complex science concepts. Interfaces may be used to optimally constrain student thinking, which may encourage students' persistence. This study builds upon literature from two different types of user interfaces, Graphical User Interfaces (GUI) and Tangible User Interfaces (TUI).

Graphical User Interface (GUI). The graphical user interface is a standard paradigm for user interaction. GUIs typically require a device, such as a mouse, which acts as a remote control to select, or manipulate variables on the display. Figure 2 shows

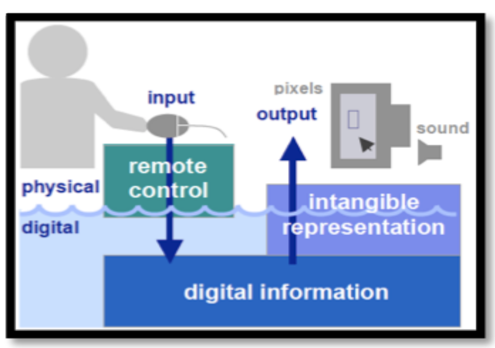


Figure 2. Graphical User Interface Interaction Map (Ishii & Ulmer, 1997)

the standard user-input machine-output relationship in the graphical user interface configuration across physical and digital boundaries. The end-user manipulates a remote control, for example, a computer mouse, in the physical world which, when used (e.g., clicking on a desktop icon), is translated to digital information (e.g., program execution resulting from a mouse-click) in the form of an intangible representation (e.g., program startup screen) which is displayed on the computer screen in pixels, perhaps with accompanying sound.

The use of the GUI has become somewhat ubiquitous, and users generally appear to rely on automaticity for navigation; that is, everyday experiences with this type of interface do not require end-users to think deeply about how these interactions may map onto existing frameworks for making sense of the world. Once the basic features of the interface are mastered, the lack of connection between the interface and a user's everyday experiences encourages rote action over meaningful learning. For example, Oviatt, Arthur, & Cohen's (2006) study investigating the role of user interface and students' mathematics learning outcomes discovered that student performance in the GUI (i.e., Graphic Tablet Interface) group decreased by 50.3% compared to students in the pencil-and-paper problem solving group. This may be a result of the activity not mapping onto students' existing knowledge and more familiar, traditional representations of mathematical calculations.

Tangible User Interface (TUI). Nearly two decades ago, this UI entered the world of Human-Computer Interaction (HCI), and in recent years, this interface has emerged a novel application in the field of education and learning. Ishii & Ulmer (1997) drew a comparison between bit-mapped applications of information in the GUI to explain TUIs as *tangible bits*. *Tangible bits* are a manifestation of digital information that acts as both a physical representation and a digital control (see Ishii, 2003). This coupling of seamless representation and control descends from the ubiquitous computing paradigm (Weiser, 1991) where technology is integrated into the physical world, facilitating seamless interaction, while the computer "fades into the background" (p. 94). Transparent interaction can immerse students in the learning environment to provide direct experiences with phenomena without requiring students to have knowledge of abstract symbols with which to make sense of the phenomena (Dede, 1995).

TUIs may be especially beneficial to learning, given the unique interplay of augmented objects, environment, and embodied controls (Antle & Wise, 2013). TUIs can have multiple inputs, which are believed increase student productivity and satisfaction (Inkpen, 1997) TUIs can facilitate users' direct-manipulation of input variables (Figure 3), which may or may not include manipulatives (Price et al., 2012).

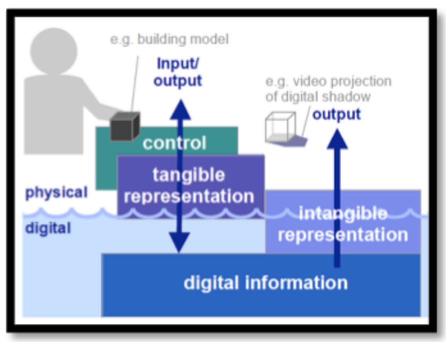


Figure 3. Tangible User Interface Interaction Map (Ishii & Ulmer, 1997)

TUIs can increase students' spatial problem solving and strategy development (Antle, Droumeva, & Ha, 2009), user exploration (Pouw, van Gog, & Paas, 2014), accessibility (Marshall, 2007), hands-on engagement (Price, Sheridan, Falcão, & Roussos, 2008; Zuckerman, Arida, & Resnick, 2005) and bolster students' task efficiency and technological usability in accordance with Norman's (1988, 1993) *principles of recognition and affordance* where users can immediately recognize functionality, especially when the control's design corresponds to its function. Designing action-representation links that meaningfully connect user actions to feedback received are especially important if students are to engage in "conceptual reflection" (Price et al., 2012, p. 151).

A paucity of empirical work exists in education comparing effects of interfaces on learning. Existing empirical work suggests the singular input feature of the GUI has several disadvantages for learning, including the lack of invitation for users to engage, collaborate, and actively explore instructional content (Horn, Solovey, Crouser, & Jacob, 2009). However, tangible interfaces (i.e., touchscreens) implemented in classrooms may not necessarily provide a benefit to children who have already learned to use a mouse (Romeo, Edwards, McNamara, Walker, & Ziguras, 2003).

Antle, Droumeva, and Ha (2009) examined interactional patterns of children (n=132) using tangible user interfaces and graphical user interfaces during the completion of a spatial task (i.e., assembling a jigsaw puzzle). Using a semi-experimental design, Antle et al. analyzed task completion times of paired student groups and compared them across three conditions (i.e., traditional hands-on, tangible user interface, and mouse-click GUI interactions) to completely assemble a puzzle. Children in the GUI condition took longer to solve the puzzle, had fewer puzzle completions in their first attempt, and engaged more in trial-and-error approaches when compared to students in other conditions. The success of a user's interaction with the GUI interface appears dependent upon prior experiences; these experiences may limit the users' ability to gain explicit conceptual understanding through this interface.

While existing research appears to favor TUIs for learning, Marshall, Cheng, and Luckin (2010) discourage making too many assumptions. Investigations of TUIs for learning often overlook the role of physicality in learning; Marshall et al. (2010) refined their investigation to determine the aspects of physicality—control and manipulation—to examine the benefits of using physical versus graphical materials. In this study, adult subjects (n=34; median age = 20 years) participated in pre/post design experiment to complete a discovery learning balance beam task (i.e., balancing weights on a beam). Findings indicate that there were no effects due either the learning medium or the level of user agency; neither had a significant impact on participants' understanding.

Haptics. The term, haptics, describes "manual interactions with environments" and can reflect user exploration to seek out information or user manipulation that changes the environment (Srinivasan & Basdogan, 1997, p. 393). Haptics can refer to a range of interactions; the most relevant to this study is haptic feedback. Haptics rely on touch technology for the provision of feedback, which directly relates to the TUI examined in this study.

Haptic feedback leverages interface design and humans' sense of touch in providing end users with information in the form of tactile feedback. The addition of haptic feedback to the action-representation pairing may be more effective than non-haptic simulations as perceptual anchors. For example, Han and Black (2011) compared the effects force-kinesthetic, kinesthetic, and non-haptic experiences on 200 6th grade students' learning of simple machine movements using simulations of mechanical gears in two class sessions over a two-week timeframe. Three simulation conditions in this study were force and kinesthetic (FK), kinesthetic (K), and non-haptic (NH). The FK simulation provided visual, auditory, and force feedback through a joystick. The K simulation is the same as the FK simulation, except that it provided non-force feedback, only kinesthetic movement to accompany visual and auditory feedback. Finally, the non-haptic version provided students with visual and auditory feedback only. Students used these simulations to learn about simple machines (i.e., window winders and salad spinners) and mechanical gears. Posttest performance revealed that the FK and K simulations were more effective than the non-haptic version for understanding of simple machines. Results suggest that the kinesthetic simulation was not sufficient for students to generate multimodal representations that can benefit future learning. Force feedback was necessary in students' creation of multimodal representations to understand targeted concepts.

Likewise, Bivall, Ainsworth, and Tibell (2011) found that adding haptic feedback to the protein-ligand recognition process improved students' conceptual understanding of molecular interactions. Schönborn, Bivall, and Tibell (2011) examined students' interactions with a visuohaptic model simulating protein-ligand docking. The visuohaptic simulation used in this study involved students in the process of biomolecular binding; students manipulate a ligand molecule using a haptic stylus device and space mouse to control a visually displayed protein molecule. Using a pre/posttest experimental design, students were randomly assigned to two conditions (i.e., haptic and nonhaptic) and tasked with using the simulation finding an appropriate binding site for the ligand on the protein molecule. Schönborn et al.'s (2011) work examined students' interactions with the technology and the relationship of those interactions to Bivall et al.'s (2011) documentation of conceptual gains. Schönborn et al.'s (2011) findings suggest haptic feedback was particularly beneficial because it provided "explicit perceptual experiences" (p. 2103) which can draw students' attention to important elements of the model (i.e., protein binding sites).

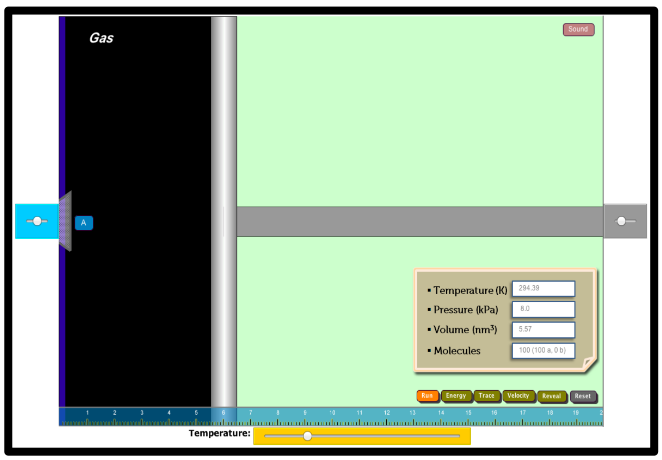


Figure 4. The GUI visualization

The present study seeks to add to existing findings through a comparison study using the Frame (TUI) and a virtual-only condition (GUI; Figure 4) in an authentic classroom environment. The Frame uses atypical tangible items (i.e. a syringe, a thermistor, and a spring; see Table 1) requiring direct physical manipulation to control features of a dynamic molecular visualization.

The virtual only condition uses a similar dynamic molecular visualization with a traditional mouse input. This study examines and compares chemistry students' learning outcomes using in each of these conditions using a mixed-reality lab activity investigating gas laws. Interpreting the types of interactions afforded and constrained by these conditions in an authentic classroom environment may guide future implementations of MR technologies.

Table 1
Frame sensor and input correspondence

Sensor Type	External Input
Pressure	Syringe
Force	Spring piston controller
Temperature	Wire thermistor

Theoretical Framework

I have adopted a constructivist, knowledge-integration perspective to understand the benefits of mixed-reality approaches for student learning. This section identifies and describes the learning frameworks employed in this study, which draw from both embodied and knowledge integration approaches to learning.

Embodied Cognition

I believe that learning is not based strictly on a cognitivist perspective; knowledge is connected to the sensory and motor experiences present during knowledge acquisition (i.e., *situational activities*; Anderson, 2003), and is not limited to prior cognitivist brain-centric conceptions of knowledge. Knowledge construction is driven by an individual's grounded experiences and perceptual understandings.

Embodied cognition posits that we formulate conceptual understandings and make inferences based on the inputs and actions in our perceptual and motor systems (Lakoff & Johnson 1980; 1999) where mental schemata are developed based on the interactions between an individual and objects in their environment (Barsalou, 2008;

Wilson, 2002). Barsalou (2008) indicates perceptual anchors or representations in multiple modalities, (i.e., situated activities, mental images, and physical states) can help learners perceptually “ground” their understandings in real world experiences. For example, a pilot operating a flight simulator emulates actions that are required to fly a plane. These actions are procedures that are physically enacted through the manipulation of simulation controls. Embodied cognition adopts the perspective that these physically enacted operations are encoded in the tactile experiences. This encoding can create, add, or reinforce existing mental schemata corresponding to the given action. The idea that physical actions contribute to an individual’s understanding of a concept, idea, or operation, is the crux of this theoretical foundation.

Some researchers believe that mixed-reality experiences can provide unique affordances of embodiment based on the types of interactions and feedback such systems can facilitate (Birchfield et al., 2008; Price & Rogers, 2004). Students who learn using their physical senses (e.g., feeling seeing, and hearing) can improve their conceptual understanding (e.g., Abrahamson et al., 2012; Anastopoulou, Sharples, & Baber, 2011; Tolentino et al., 2009). For instance, Anastopoulou et al. (2011) examined university students’ (n=18) learning and ability to connect multiple representations of motion using a motion-tracking device and LabView software that offered graphical depictions of distance-time and velocity-time relationships in an effort to understand how multimodal learning may lead to the development and implementation of effective computer-aided learning. Students participated individually in a singular 50-minute clinical study session where they completed a worksheet that served as their pretest and posttest; they were asked to draw and interpret motion graphs in addition to describing how they might move their hand to generate a particular type of graph. Participants were placed either into active or passive roles as doers or watchers, when instruction was provided for the worksheet. The researchers found that the physical manipulation in a multimodal environment allowed learners to make connections among different representations of motion (i.e., graphical, linguistic, kinesthetic). Combining sensory modalities and virtual representations allowed students to transition between the symbolic and real world, positively affected their understanding of motion.

Instruction using embodied interactions may help students recognize and resolve alternative ideas about concepts. For example, Howison, Trninic, Reinholz, and Abrahamson's (2011) application of the *Mathematical Imagery Trainer* (MIT) focused heavily on embodied learning in a lab setting. The MIT relied on a computer-based hand tracking system (i.e., Wii remote) to ground mathematics concepts in dynamic imagery to provide students with rich, grounded experiences that they could use to construct conceptual understanding of proportionality. Students were provided with instructions to complete a series of embodied tasks that involved changing the colors on the corresponding visual display (e.g., "make the screen green" or "find green somewhere else") using arm movement. The visual display showcases colors from red to green that change as the students hand heights reach the selected ratio being tested (e.g., "raising the hand trackers to 3" and 12" will turn the screen red or yellow for a 1:2 ratio," p. 1993) and changed to display a grid as the activities progressed. The study included 22 students in 4^{th}-6^{th} grades who were guided through a series of activities using a Wii remote to represent the designated ratio using the visual display. Results suggest students were successful in developing strategies using embodied interaction to solve proportion problems with some degree of mathematical reasoning. In this case, the students' interactions with and manipulations of the remote were more connected, indicating that embodied interactions may help students' recognize and revise alternative ideas.

Existing mixed-reality experiences designed for student engagement and learning suggest that student learning is improved by incorporating embodied interactions into design (Lindgren & Moshell, 2011; Pillat, Nagendran, & Lindgren, 2012). Although the degree to which concepts are embodied, under what circumstances, and how seems to vary between individuals (Hauk & Tschentscher, 2013). Incorporating mixed-reality experiences in the classroom that create opportunities for novel interactions may leverage embodied interactions afforded by different interfaces to positively affect students' learning gains.

Knowledge Integration

This study leverages the knowledge integration (KI) framework to structure activities within the mixed-reality environment. The KI perspective is based on decades of classroom research that incorporates computing technology in science classrooms (Linn & Eylon, 2011). KI encourages students to elicit experiential knowledge to provide a foundation for students to make connections between ideas and develop deeper conceptual understandings (Clark, 2006; Linn & Eylon, 2006). The following section provides a description of knowledge integration and its application as a tool to make sense of students' conceptual understandings of the scientific phenomena being investigated in this study.

I believe students come to class with existing knowledge and intuitive understandings from everyday experiences (Brown, Bransford, & Cocking, 2000).

These everyday experiences provide a framework through which students interpret new information and construct knowledge, including what they acquire through classroom instruction. In a classroom setting that facilitates knowledge integration, students experiential ideas are elicited, new scientific ideas are added through instruction, and students distinguish between ideas, providing a framework where students give consideration to all ideas and sort them out to refine their conceptual understandings (Linn & Eylon, 2006).

Imagine a scenario where a teacher presents a student with a metal spoon and a wooden spoon at room temperature and asks which item is colder. These items combined with this prompt may elicit students' experiential knowledge. A student who has touched both items independently may perceive the temperature as different, based on the sensation that they feel. This rationalization of equating sensation with temperature, although common, is not scientifically normative. Providing a student with additional instruction about the concept of heat transfer can add ideas to their understanding of the relationship between heat and temperature. Following this instruction, this student is given a task that causes them to reflect and apply their understanding; ideally, this is where they can sort all of these ideas, distinguish among them, and construct a more cohesive and complex scientific understanding based on information they have acquired. One of the many challenges in teaching and learning, especially in science, is that students come to class with a variety of experiential understandings and preconceived ideas that may need reconceptualization (Klahr, Zimmerman, & Jirout, 2011) to align with scientifically normative understandings. Knowledge integration encourages students to revisit, reflect, and revise their understandings in light of new information.

In spite of receiving information that contradicts existing understandings, we know alternative conceptions are persistent; students can be provided time and again with normative knowledge, but direct transfer of that knowledge rarely occurs to change students' ideas. It is possible that activating students' prior science knowledge may provide an experiential foundation, a framework from which students can construct more normative, enduring complex understandings of scientific phenomena. Experiential knowledge is critical for successful knowledge integration. Students' experiential and scientific knowledge are called to the fore where cognitive conflicts are addressed and links between ideas are weakened or strengthened to evolve into more cohesive understanding (Linn & Eylon, 2011). Research using physical and virtual materials with the KI perspective has demonstrated that students develop increased connections between ideas, indicating the development of a more cohesive understanding (Chiu & Linn, 2014; McElhaney & Linn, 2011; Zhang & Linn, 2011).

Technology offers additional opportunities for students to interact with a learning domain through the use of sound, visualizations, physical manipulatives, and novel interfaces. Students encouraged to use technology often engage in exploration, test hypotheses, and revisit content as needed (Brown et al., 2000). These learning processes invite students to evaluate their existing knowledge and construct new knowledge based on their experiences with technology. It is possible that technology can encourage students to engage these learning processes in ways that didactic instruction does not. Emergent mediums like mixed-reality, and augmented virtuality in particular, may be more effective in engaging students in knowledge integration through the embodied nature of the physical controls connected to the scientifically normative simulations. The embodied nature of the physical controls may elicit students' prior knowledge and experiences in different ways than a traditional mouse or GUI interface. Tactile interaction afforded by the Frame may elicit students experiential ideas along with ideas related to students' actions (i.e., pushing the spring to control the piston, using a hot jar to affect the temperature). The physical controls can foreground or highlight experiential ideas that students may hold about a phenomena such as pressure or temperature. When students use the physical controls to directly interact with the scientific simulations, they are presented with the normative molecular view. Thus, students can test their existing experiential ideas about pressure or temperature against what happens in the simulation at a molecular level. Having both experiential and normative ideas presented at the same time may encourage students to recognize and resolve conflicts through knowledge integration.

Frame Pilot Studies

Initial research studies using the Frame in authentic classroom contexts indicate that augmented virtual technologies have promise in secondary science education. One pilot study explored if the Frame can help students make macro-molecular level connections concerning the properties of gas and help students refine alternative ideas about gas properties (DeJaegher, Chiu, and Chao, 2014). Two physical science teachers and their students (n=45) in four middle school classes participated in the study. Students were guided through an investigation of gas properties with the Frame over the course of two 90-minute class sessions. The curriculum was developed by researchers and teachers and followed a scaffolded knowledge integration approach (Linn, Davis & Eylon, 2004). Pretest and posttests were coded using a knowledge integration rubric (Liu, Lee, Hofstetter, & Linn, 2008). Students' responses were also categorized in terms of normative, alternative, partial, and vague ideas to get at whether and how students refined their initial explanations.

Students improved their understanding of scientific concepts covered on the assessment from pre to posttest. Further analysis demonstrated that students progressed in refining alternative ideas in molecular-level explanations of temperature and volume relationships and incorporated more normative and partial ideas on the posttest. The results from the exploratory pilot study suggested that students could use the Frame to help them develop explanations of complex gas phenomena.

The Frame was also implemented in two high school chemistry classes with 10^{th} and 11^{th} grade students (n=30) using a pre/post comparison design to (1) determine if students working with the Frame could improve their understanding of gas laws and (2) compare students working with the Frame to students receiving traditional instruction with physical labs and discussion (Chao et al., 2014). Two classes were randomly assigned to treatment and control; both classes had the same instructor, who has over 10 years of chemistry teaching experience. Both groups spent a total of two 90-minute class sessions completing either the Frame activity or the traditional instruction.

Students in the Frame condition worked in groups of 2-3 students to complete the guided lab activity, receiving requested support from teachers and researchers during the intervention. The control group did not use the Frame; instead they conducted a traditional physical lab covering the gas laws, as well as lecture and discussion activities, where students constructed explanations of everyday phenomena focused on the same topics as the treatment group. The control group used a closed syringe and a pressure sensor to experiment with the gas volume-pressure relationship; molecular explanations were not emphasized during the physical lab activity. Results demonstrate that students in the treatment condition improved their performance on seven of the nine questions with medium to large effect sizes from pre to posttest, indicating that instruction with the Frame helped students understand gas laws. A comparison between treatment and control group performance on the posttest indicated that while both groups performed equally on most items, the Frame group outperformed the control group questions addressing an understanding of pressure on the molecular level. Results suggest that the Frame may help students develop more nuanced explanations of complex molecular-level phenomena, in this case, a deeper understanding of gas pressure on a molecular level.

Although these results provide initial evidence about possible benefits of the Frame to physical labs, the differences in responses between the treatment and control were mainly the addition of molecular ideas with the Frame. Students may get the same benefit from using visualizations of molecular behaviors as with the Frame, as “seeing” and being able to interact with the molecular level may be the main advantage. Thus, this dissertation compares the Frame (a TUI) to just the simulation (GUI). Comparing the Frame to the GUI version of the simulation will help isolate the effects of physical augmentation. This study seeks to capture not only what students learn from each condition but if there are any differences in the way students interact between conditions. Results from this study will contribute to the existing body of knowledge regarding the design and implementation of mixed-reality technologies in authentic classroom environments as well as advance our understanding of a need (if any) for physical augmentation.

Summary of Existing Research

Recent research in science education focuses on the relative benefits of physical or virtual materials for students to learn science. Several researchers have investigated the benefits and drawbacks to physical labs versus virtual labs, either by comparing separate conditions, presenting representations simultaneously, or presenting physical and virtual representations of phenomena side by side. Some studies report no difference between virtual and physical materials (Klahr, Triona, & Williams, 2007) while others indicate that using both provides benefits to student learning (Olympiou & Zacharia, 2011).

Mixed-reality technologies attempt to leverage affordances of both physical and virtual materials, representing a range of technologies that are just now beginning to enter the classroom. Unlike solely virtual or solely physical labs, mixed-reality experiences provide opportunities for students to engage in both physical and virtual manipulation simultaneously. Existing implementations of science laboratories are just beginning to examine students' learning outcomes resulting from integrated experiences.

Prior investigations explored the use of a specific kind of augmented virtual technology, the Frame, in authentic classroom settings. Results demonstrate that students working with the Frame improved their conceptual understanding of gas properties (DeJaegher, 2014; DeJaegher, Chiu, & Chao, 2014), and outperformed students exposed to more traditional instruction on some complex science topics (Chao et al., 2014). This work highlights potential benefits in connecting the physical and virtual experiences. However, further investigation of interface mediums may help us pinpoint what it is about the combined experience that helps students learn. Specifically, this dissertation will compare students using the Frame to students using the same simulation without the physical augmentation to tease apart benefits on student learning and interaction, if any, that are added through a mixed-reality approach.

Limitations of Existing Research

The literature has two major gaps addressed by this study: (1) the lack of empirical literature reporting on investigations combining physical and virtual labs into a singular, connected experience (2) a majority of research being done is in a clinical or university setting—few studies have been conducted at the secondary level. Not only is there a dearth of research examining mixed-reality technologies in science education, few studies examining mixed-reality technologies have taken place in authentic classroom settings. If we are looking to implement this type of technology in schools, it is important to understand not only how the target population makes use of this tool, but also to understand how the constraints of the classroom environment may have on such an implementation. Additionally, research on the effects of user interfaces on student learning in science is more anecdotal than empirical. Most studies compare GUIs (visualizations) to traditional instruction; this dissertation will help the field understand the particular affordances of tangible user interfaces for learning about molecular phenomena.

Key Terms and Definitions

Frame Apparatus

The Frame apparatus is a mixed-reality technology classified as augmented virtuality that blends physical and virtual experiences permitting the simultaneous functioning of hands-on manipulation and dynamic visualization. The Frame apparatus is a comprised of two main parts: (1) a Lenovo ThinkPad tablet computer and (2) a boxed enclosure upon which the computer rests. The box contains one force sensor, one temperature sensor, and one pressure sensor. An external input attaches to and corresponds with each sensor (Table 1) and extends beyond the box to the outside. The boxed enclosure hides all internal components from view and fits this model of computer so that the edges of the computer are flush.

Tangible User Interface (TUI)

A tangible user interface refers to a human-machine interface that uses physical controls for end-user interactions. This basis of this interface is natural or intuitive human action. The type of TUI employed in this study allows the end-users to physically interact with the Frame, using the aforementioned external outputs, to engage with the dynamic visualization being displayed. This version will be used in the treatment as the Frame; the Frame and TUI will be referred to interchangeably.

Graphical User Interface (GUI)

A graphical user interface refers to a human-machine interface where end-user interaction is based on visual components. The type of GUI used in this study belongs to the WIMP (i.e., window, icon, menu, pointing device) paradigm where a mouse is used to interact with the visual components of the dynamic visualization. In this implementation of the Frame, the control group will use a GUI in the form of a scientific visualization that is the same as the visualization in the Frame.

Interaction

Interaction is defined in the literature review to be a distinct yet simple blend of computer-supported collaborative learning and human-computer interaction terminology. Interaction involves the ways in which students engage either with the Frame or the visualization technologies, their peers, and classroom instructors. In this study, the specific types of interactions have been defined based on observation. A complete framework for classifying interactions by theme is detailed in the data analysis segment of Chapter 3.

Conceptual Understanding

In the context of this study, conceptual understanding is represented in terms of ideas and connections among ideas according to the Knowledge Integration (KI) framework. The KI framework categorizes scientifically normative ideas on a 6-point interval scale from 0-5. The students' elicited ideas were classified as either normative or alternative, and links between idea types helped determine if students have developed more cohesive understandings of the particulate behaviors and properties of gas molecules. The difference in KI score across assessments designed to measure conceptual understanding of kinetic molecular theory served as a basis for examining differences between the TUI and GUI treatments in this study.

CHAPTER 3

METHOD

My dissertation investigated the research questions in Table 2 using an explanatory-sequential mixed-methods approach that relies on both quantitative and qualitative methods. The quantitative analyses involved descriptive analyses and a multilevel linear growth model designed to account for data nesting. The resultant outcomes addressed the first two research questions informed the selection of video cases to be qualitatively analyzed (Creswell & Clark, 2007; Hesse-Biber, 2014; Ivankova, Creswell, & Stick, 2006). Qualitative data were analyzed using analytic induction to categorize observational data. Multiple iterations of analysis led to the emergence of themes arising from established categories. These themes were used to present a more comprehensive understanding of students' interactions between conditions.

The primary purpose of the quantitative aspect of this study was to identify differences in learning outcomes, should they exist, and the qualitative aspect focused on generating and interpreting interaction/response themes to help researchers in the field develop a broader understanding of students' mixed-reality experiences. It was expected that students in the TUI condition would perform similarly or better than the students in the GUI group regarding conceptual understanding. Additionally, it was believed that students in the TUI group would be more focused and on task than students in the GUI group.

My investigation occurred in authentic classrooms situated in three public high schools in Central Virginia; these data were gathered in eighteen honors/standard chemistry courses taught by a four teachers. This chapter details the context and participants involved in the study, investigative methods, student performance assessment instruments, types of data being collected, and the methods of analysis used for the collected data. Schools and teachers are identified by pseudonyms throughout this report to protect the identities of the participants.

Table 2
Research questions addressed in this study

Research Questions
1. What differences, if any, are there between students using the Frame and students using virtual labs on students' conceptual understanding of gas properties?
2. What differences, if any, are there between students using the Frame and students using virtual labs on conceptual retention over time?
3. What characterizes students' interactions using the Frame compared to students' interactions when using the visualization only? How might students' interactions relate to learning outcomes?

Context and Participants

This study was conducted across eighteen honors and standard high school chemistry classes in three high schools in Central Virginia during the 2013-2014 academic year. The honors course classification is an intermediate level between standard and accelerated placement, indicating that typical students' overall academic performance is above average. Students in honors courses at these particular schools are students with a GPA of 3.0 or better and are on a college preparatory academic track. Standard courses are similar to honors courses in terms of content, but the pacing is different and advanced enrichment activities are generally supported at the honors-level. These schools operated on modified block scheduling with Chemistry courses meeting 3 times a week; one meeting lasts 45 minutes and the other two meetings last 90 minutes. Science credits in chemistry are required to graduate from high school in Virginia, however; the level of placement is student elected or teacher recommended based on prior academic performance.

Students

Sampling Method. An explanatory-sequential mixed-methods balanced design with non-random convenience sampling (Creswell, 2011, p. 145) was used in this study. The sampling method was based on the research team's existing relationship with the local chemistry teachers who agreed to participate in this project.

Study Sample. Data from 307 students were analyzed for this study. This sample included students from both standard level and honors classes; in total, there were 12 regular classes and 6 classes housed in the Mathematics, Engineering, and Science Academy (MESA). Regular courses include students enrolled in the required science courses for the standard terminal high school credential (i.e., diploma). Of these courses 10 (7 regular; 3 MESA) were randomly assigned to the NUI condition and the remaining 8 (5 regular; 3 MESA) were randomly assigned to the GUI condition.

Students in MESA courses were drawn from a cohort intent on pursuing an engineering career trajectory. These courses were designed to incorporate traditional curriculum with increased opportunities for students to develop authentic problem-solving and collaborative skills through the integration of mathematics and science disciplines. In some ways these courses represent a more advanced track, but students in these classes possessed a variety of academic strengths and interests. Within both types of classes, regular and MESA, students represented a variety of backgrounds and levels of academic achievement.

Teacher and Courses. All teachers participating in this study have advanced degrees in chemistry (i.e., M.S.), and one holds a Ph.D. in Microbiology with over 35 years total experience between them in teaching high school chemistry (Table 3).

Table 3
Teacher experience and participation breakdown

Teacher	School	Semester	Years of K-12 Experience	Highest Level of Education Completed	Number of Courses Included in this Study
A	Madison	Fall 2013	5 years	M.S. Chemistry; M.A. Teaching	2
B	Jefferson	Fall 2013/Spring 2014	10 years	M.S. Chemistry	4
C	Washington	Fall 2013	15 years	M.S. Curriculum & Instruction	6
D	Jefferson	Spring 2014	9 years	Ph.D. Microbiology	6

The teachers were part of the development and implementation of prior versions of the Frame, including participation in multiple summer and mid-year professional development workshops. Prior to the semester of implementation, all teachers participated in a daylong professional development workshop where they were exposed to the technology, the dynamic visualization, the proposed curriculum packet, and the proposed assessments. This professional development opportunity familiarized teachers with the technologies, whereby the teachers provided useful feedback to the research team regarding usability and implementation in the context of their students and courses, respectively, as well as revisions to the accompanying curriculum packet.

District Demographics. The student sample is taken from school district population that contains approximately 13,108 students in total. The demographic of the district is majority male (51.1%) and white (69.6%), with 11.4% of the population classified as black, 9.0% as Hispanic, 4.5% Asian, and 5.0% other. Eight percent of the district's population consists of ESL students and 25.4% of the students enrolled in this district are eligible for free or reduced price meals.

Across the district there are approximately 1,200 teachers; 99.5% are "highly qualified" according to No Child Left Behind. Over half of these teachers (58%) hold advanced degrees. The teaching population of this school district averages 15 years of teaching experience. The population, demographic information, and number of teachers participating at each school that participated in this study are detailed in Table 4.

Table 4
Participating school demographics

School[a]	No. of Teachers	Total Population	Gender	Majority	Minority	FRL[b]
Washington	1	1000	52% F	White (82%)	16%	<10%
Jefferson	2	1791	50.8% M	White (69.6%)	22.8%	21.3%
Madison	1	1101	52.5% M	White (73.2%)	22.9%	27.5%

[a] Schools identified by pseudonyms
[b] Proportion of students eligible for free and reduced lunch based on United States' federal criteria

Curriculum

Curriculum Development. The accompanying curriculum was developed in collaboration with several educational researchers, scientists, and classroom teachers to target common misunderstandings in physical science. The two curricula are supplemented in Appendices A and B.

Curriculum Content. Misunderstandings specific to the nature of gases and temperature were identified in research and in practice (e.g., AAAS Project 2061; Clough & Driver, 1985). The curriculum also draws upon existing successful approaches to visualization-based chemistry laboratories (e.g., Chang, Quintana, and Krajcik, 2009; Levy & Wilensky, 2009; Olympiou & Zacharia, 2012). The curriculum guides students through a series of five activities: (1) Behaviors of gas molecules (2) Temperature (3) Pressure (4) Relationship between Temperature and Volume (5) Relationship between Pressure and Volume.

In each activity, students apply predict-observe-explain strategies (White and Gunstone, 1992) organized cyclically throughout the lab. For example, after students have the tutorial activities in the lab, they proceed to learning more about temperature, volume, and pressure by first considering temperature. Students are asked to make a prediction about molecular distribution and movement in a hot room versus a cold room. Students draw representations of both instances to indicate what they think will happen in each scenario. Following the prediction phase, students are directed to use the simulation to observe what actually happens to the molecules when the temperature is manipulated. As a follow up students are asked to compare their observations with their predictions to explain and support their final conclusions. This pattern is applied to several investigations during the laboratory.

The activities use a guided inquiry framework where investigations start with simpler phenomena and gradually become more complex (see Hmelo-Silver, 2006;

Hmelo & Gudzial, 1996; Minstrell & Kraus, 2005; Trundle, Atwood, Christopher, & Sackes, 2010) as well as a scaffolded knowledge integration perspective to help students make connections among their ideas (Linn, Eylon & Davis, 2004). The curriculum has been tested and refined with prior studies using the Frame with secondary science students (see Chao et al., 2014; DeJaegher, Chiu, & Chao, 2014).

Instruments

Pretest/Posttest/Delayed Posttest

Assessments (Appendix D) administered prior to the implementation consisted of 11 questions (a mixture of selected and open-ended response items) adapted from assessments designed by the American Association for the Advancement of Science (AAAS). Assessment items were chosen to capture common alternative conceptions in science (e.g., molecules clump together in cold temperatures) and concepts targeted in this laboratory, namely, the particulate behavior of gases relating to temperature, pressure, and volume. The items have been piloted, refined, and validated through other classroom implementations with similar populations over the course of the previous year (see Chao et al., 2013; Chao et al., 2014; DeJaegher, Chiu, & Chao, 2014). Pilot results found no ceiling effect with secondary level honors/standard courses. Based on piloting these test items, revisions included eliminating redundancy in response stems and rewording existing questions to incorporate appropriate terminology for the age group. Feedback from the participating high school teachers who have been involved with this project guided many of those revisions. Test-retest reliability was calculated for past implementations, and Cronbach’s alpha was greater than 0.60 for all assessment items.

Video Data

Video data were collected from the selected student groups to capture students' interactions with the technologies in an authentic classroom setting. Video data was captured using cameras placed next to the group. This data documented what students said and did with the technologies. I worked with the teachers to identify two participant groups from each class to be video recorded; the identification of groups was based on video consent as well as the teacher's selection of students based on cooperative compatibility. Following the implementation I selected and analyzed video cases based on identified trends from pretest, posttest, and delayed posttest assessments.

Semi-structured Interviews

Participants in video groups were invited to participate in semi-structured interviews. These interviews were conducted with members of each respective student workgroups simultaneously. The purpose of these interviews was to elicit student experiences with and perceptions of the technologies. Prior to the interviews, I reviewed group video data and selected segments student interaction for targeted questioning. During each interview I shared the selected segment with group members and asked them to explain of their interaction with the technology. The remainder of the interview focused on specific features of the technology that students perceived as contributing to their learning, as well as eliciting their perceptions of the laboratory activity, in general.

The students included in the video observations were selected to participate in the semi-structured interviews; students were grouped in the interviews, reflecting the group arrangement during the classroom laboratory implementation. Interview data focused on students' responses to the interview protocol (Appendix C). As suggested by Bailey (2007), some questions were altered, added, or skipped entirely, adjusting for students' interview responses to either (1) probe students about something which had been mentioned or (2) to avoid redundant questioning after students provided relevant information in a prior response. Categories were generated from the video data and the interview data independently and later combined to determine themes (see Table 14).

Conditions and Procedures

Treatment

Students assigned to the treatment group used the Frame (TUI). Students used an accompanying paper-based curriculum to help them progress through the inquiry-based lab. The curriculum guided students to interact with the external inputs built into the Frame (i.e., the syringe, the spring piston, and the temperature sensor).

Control

Students in the control condition used the GUI version of the Frame, which includes the same basic visualization as the Frame but uses graphical inputs and a mouse (Figure 2). Students used the same curriculum as the treatment group with small wording modifications tailored to the GUI. Procedural wording modifications directed students to click appropriate places on the screen instead of using augmented physical controls. For example, when students were asked to "push on the spring" in the hands-on treatment, they were asked to "click on a slider bar" to adjust pressure in the visualization only model.

Procedure

Approximately 2 weeks prior to running the study, I met each teacher in person to confirm dates and times, disseminate consent forms, and review a scripted implementation protocol in conjunction with an overview of the laboratory activity. The teacher supplied a bell schedule and class rosters, whereby I randomly assigned classes to each condition. Approximately 1-3 days before each implementation, teachers administered the pretest (electronic or pencil-and-paper), and the teacher established working groups to be recorded, based on completed consent forms and their understanding of student relationships. Digital video, screencasts, and post-implementation interview data were collected from students selected to participate in video recording.

The day of the implementation, the teachers conducted a brief whole class review including: (a) definitions and properties of gases (i.e., temperature, volume, pressure, and number of molecules) and (b) goals of the laboratory which are to understand gas laws using molecular explanations. The teacher introduced me, and I oriented students with the project components, the curriculum packet and the technology (either the Frame or the GUI). The curriculum packet overview covered basic instructions for students, such as writing their names on their packets. I performed a whole class demonstration of the controls and their respective functions; for the Frame this meant indicating what the external inputs did, and for the GUI condition, this meant indicating what slider bars on the dynamic visualization did when they are clicked on using the mouse. Students assembled into their teacher-assigned lab groups and began working on the activities. I visited each student group selected for video recording and gave a brief explanation of their participation, the collection of digital video data, and answered any questions prior to recording. As the students worked through the laboratory activities, both the teacher and a member of the research team remained in the classroom, answering students' questions, troubleshooting technological difficulties, monitoring student progress, and observing the implementation. Upon completion of the laboratory activities, students submitted completed curriculum packets to a member of the research team. Each implementation took 1 block (~90 minutes) of instructional time.

At the conclusion of the implementation, the researcher conducted semi-structured interviews with those students who participated in digital video recorded groups for both conditions to learn more about their experiences. Each interview followed a protocol (Appendix C) designed to elicit students' thoughts about the laboratory and lasted approximately 20 minutes. Of particular interest to this study was the portion of the interview where students were asked to talk about their experience with the technology and to review researcher-selected video segments and comment on their specific interactions with the technology.

Approximately 1-2 days following the implementation, students in all participating classes took the posttest as administered by their classroom teacher (electronically or paper-and-pencil). One month after the implementation, all participants completed a delayed posttest, also administered by the classroom teacher. The posttest and delayed posttest were identical to the pretest measure. After all assessment and video data was recorded, data from students missing general consent forms were securely destroyed and eliminated from the dataset. Student data included in the analysis of this study are representative only of participants who completed the pretest, posttest, and delayed posttest assessments. I transcribed and prepared all test, video, curriculum packet responses, and interview data in preparation for analysis.

Data Analysis

Quantitative Data Analysis

This study used a quantitative approach to understanding students' conceptual understanding and performance on the pretest, posttest, and delayed posttest assessments. The quantitative analysis of conceptual understanding was based on the knowledge integration assessment framework (Liu et al, 2008), which determined the complexity of students' understanding through scientific explanations. Students' responses were assigned a score based on a knowledge integration scale, and these scores were evaluated using a series of multilevel models applied to longitudinal data, from here on referred to as multilevel modeling (MLM).

Conceptual Understanding. The first research question targeted differences in students' conceptual understanding comparing the Frame and control (GUI) conditions. Assessments were coded using a knowledge integration rubric (Table 5; Liu et al., 2008). I used and refined codebooks from prior analyses of the assessments (DeJaegher et al., 2014). Part of the coding involved categorizing ideas as irrelevant, alternative, partial, and normative. Irrelevant ideas are those that do not answer the question, such as, "I love Buffy the Vampire Slayer". Alternative ideas are those considered to be scientifically non-normative, such as "molecules shrink when it's cold." Partial ideas are not entirely non-normative, and may be scientifically correct but not central to the targeted phenomenon. For example, when asked how gas molecules cause pressure, a student may state, "gas molecules are constantly moving". While this is scientifically normative, this idea is partially linked to the target concept; it does not fully express the relationship of molecular collisions to gas pressure. Normative ideas are scientifically correct ideas.

Two independent raters coded 20% of the total data and established the Cohen's Kappa interrater reliability statistic ($\kappa > 0.6$). I coded the remaining data and assigned KI scores to students' responses. Pretest, posttest, and delayed posttest KI composite scores were calculated for each student across each of the six assessment items.

Table 5
Adapted knowledge integration rubric (from Liu, Lee, Hofstetter, & Linn, 2008)

KI Level	Score	Response Characteristics	Examples Using Student Work
Complex-link	5	Students elicit and connect three or more normative and relevant scientific ideas	When it was day and it was hot the molecules moved faster hitting the walls inflating the mattress a little. During the night the slowed down deflating it a little.
Full-link	4	Students elicit and connect two normative and relevant scientific ideas	The particles cooled & slowed down creating less force, which created less pressure deflating the mattress.
Partial-link	3	Students elicit normative and relevant scientific ideas	When the mattress was heated it had more kinetic energy witch mad it hard and tall. Then when it lost kinetic energy it deflated.
No-link	2	Students elicit non-normative ideas or make invalid connections between non-normative ideas or between normative and non-normative ideas	There were enough molecules during the day so it was inflated then at night they bunched together and it deflated.
Irrelevant	1	Students do not elicit scientific ideas	I love Buffy the Vampire Slayer.
No response	0	No response is provided	No response or "I don't know"

Multilevel Modeling (MLGM). In addition to these procedures, a series of linear regression models (MLGM) were used to estimate the effects of students' KI scores by condition, while considering variances between subjects and within subjects over time in each condition. This approach was selected over a univariate General Linear Model (GLM) because the structure of the GLM does not assume independence while adjusting for the covariance structure of the data. Therefore, the effects of students in the same class, as well as performance growth over time, as measured by students' KI composite scores on pretest, posttest, and delayed posttest assessments, formed the focus of this analysis. The purpose of employing MLGM in this context were twofold; (1) to account for the nested data structure in this study and (2) to elicit any patterns in student conceptual growth over time, evaluating performance across conditions. Several multilevel linear models were examined using the data from the study to capture variance, within class effects, between class effects, and change over time. A longitudinal approach was used to determine students' conceptual understanding in terms of KI score growth over time, specifically across three measured time points, in each condition; this approach suggests that the data is not strictly linear comparing student scores across time. For this reason the trajectory of scores was divided at the posttest to compare growth rates between the pretest and posttest assessments and between the posttest and delayed posttest. Thus, multilevel models were constructed to fit the data.

The Application of MLGM. Based on the nature of the data, nested configurations were determined based on the classification of condition (i.e., GUI and TUI) as a fixed effect. Other variables were categorized as random effects (Raudenbush, 1993; Raudenbush & Bryk, 2002) to establish multilevel growth models for analysis. The KI scores from each time point pretest (TIME=1), posttest (TIME=2) and delayed posttest (TIME=3) were first grand mean centered and compared across conditions as well as over time. Grand-mean centering (GMC) is a common practice in multilevel modeling (Raudenbush & Bryk, 2002) because it gives the intercept parameters a more useful interpretation. Grand-mean centering does not affect the relationship between predictors and may reduce multicollinearity (Raudenbush & Bryk, 2002). Grand-mean centering is generally regarded as important because higher-level models are explained by lower level coefficients.

A three-level MLGM was used to describe the achievement growth (as measured by KI score) while incorporating the effects of condition, gender, and class. The level-1 (L1) variable consists of time, as this represents measurement data. Level-2 (L2) variables include all student level variables, *KIGMC,* the grand mean centered KI scores, and *GENDER,* a dummy coded variable representing students as either male (0) or female (1). Level-3 (L3) estimates differences between classes of students using the variable, *CONDITION,* which was randomly assigned at the class level, to explore whether the TUI (1) or GUI (0) condition affected students' learning outcomes (*KIGMC)*. A total of four models were constructed using the *lme4* library package in R during the exploratory data analysis phase. To make fair comparisons of classes, students' composite KI scores were grand mean centered for each assessment. Effects of student level variables are treated as fixed across classes; KI scores and growth over time vary within and among classes. L1 accounts for time, L2 accounts for between and within-student variance, respectively, and L3 accounts for variance between classes.

L1 and L2 Variables.

Knowledge-Integration Score. The KI scores calculated from the aforementioned procedures involving the KI rubric and rater coding were used to represent outcome variables in the L1 model. Each student included in the analysis has a total of three KI scores, one for each assessment. These scores are represented in the dataset as *KIGMC*.

KIGMC scores were represented over time. The time points are represented by the variable, *TIME*, which represents each assessment; any variation in assessment administration between classes was due to variables beyond my control (i.e. classroom pacing, snow days, assemblies, etc.) and are reflective of an authentic classroom environment. The administration time between the laboratory activity and the posttest and the time between the posttest and the delayed posttest remained comparable across groups, therefore, the variable, *TIME*, was coded to represent each of the time points instead of the time as a measurement of days passed between measures. The trajectory of these scores over time is examined through the multilevel linear growth model to address the first two research questions.

Gender. This variable doesn't specifically address the current research questions; however, it is being included to examine whether gender has had any significant effect on students' outcomes. This is especially relevant given several extant studies highlighting the existence of a *gender gap* between males and females relative to science learning and technology-based learning (e.g., Imhof, Vollmeyer, & Beierlein, 2007; Sanchez & Wiley, 2010)

L3 Variables.

Condition. There is one L3 variable under examination in the modeling procedures for this study. The variable, *CONDITION* was randomly assigned by class a priori. *CONDITION* was dummy coded, where 0 denotes the reference (GUI) group and 1 is the comparison (TUI).

Multilevel Latent Growth Model: Theory and Construction

This comparison using a hierarchical growth model, or multilevel latent growth model, (MLGM), examines group-level variables (L3) to explain variance in parameters at the individual level (L2) across variation in L1. Gender and condition are fixed and used in this analysis to investigate potential influence on learning outcomes.

Unconditional Model. The MLGM analysis began with an Unconditional Model. This model does not include predictors applied at any level of the model. The general equations for this model are as follows:

Level 1 (within-student): $Y_{ij} = \pi_{0i} + \pi_{1i}TIME_j + \varepsilon_{ij}$

where Y_{ij} is the KI score for a given student and *TIME* is the predictor; below is the equation breakdown for each measured time point ($where\ t_i\ denotes\ time\ nested\ within\ individuals$):

$$Y_{i1} = \pi_{0i} + \pi_{1i}t_1 + \varepsilon_{i1}$$
$$Y_{i2} = \pi_{0i} + \pi_{1i}t_2 + \varepsilon_{i2}$$
$$Y_{i3} = \pi_{0i} + \pi_{1i}t_3 + \varepsilon_{i3}$$

Level 2 is represented by using level 1 intercepts (π_{0i}) and slopes ($\pi_{1i}t_1$) as outcomes:

$$\pi_{0i} = \beta_{00j} + r_{0j}$$

$$\pi_{1i} = \beta_{10} + r_{1j}$$

Level 3 is represented by using level 2 intercepts (β) as outcomes:

$$\beta_{00j} = \gamma_{000} + u_{00j}$$

$$\beta_{10j} = \gamma_{010} + u_{01j}$$

Unconditional Growth Model (m_0). The unconditional latent growth model time as a predictor variable at Level 1 in order to estimate the random effects for the intercept and the slope at Level 2; it is essentially a one-way ANOVA procedure with random effects. The purpose of this initial model is to investigate the random effects associated with this data; these random effects include variances of students nested in classes. There are two levels in this model, represented by the following equations:

Level 1: $KIGMC_{ti} = \pi_{0i} + \pi_{1i}TIME_j + \ e_{ti}$, where ~ N(0, σ_ε^2) represents within-student variance

Level 2: $\pi_{0i} = \ \beta_{00} + \ r_{0i}$ where ~ N(0, σ_0^2) representing between-student variance

$\pi_{1i} = \ \beta_{10} + \ r_{1i}$

The composite unconditional means model is as follows:

$$KIGMC_{ij} = \ \beta_{00} + \beta_{10}(TIME_j) + \ e_{ij} + r_{0i} + r_{1i}$$

where Y_{ij} denotes the response variable, KI score for the i^{th} student at the j^{th} assessment occasion (i.e. pretest=1, posttest=2, and delayed posttest=3); β_{00} represents the slope of the fixed effect, β_{10} represents individual student-specific means (intercept) across assessments, e_{ij} indicates within-student (i) variance across assessment time points (j), and the $r_{01} and\ r_{1i}$ terms account for random effects present at Level 2.

The purpose of this model is to test the relationship between the outcome variable (i.e., KIGMC) and time. By using a hierarchical linear growth model to estimate student performance by condition in the context of class and other students within the same class, it can be determined to what extent these variables may have affected students' performance in terms of conceptual understanding as represented by KI score. This initial model represents an estimate of students' test score growth over time; it is important because it indicates that students have varying scores at the pretest level and that individual students show progress at different rates over time. The unconditional growth model was used to compute the intraclass correlation coefficient (ICC) was calculated to determine the variation not accounted for and to determine if a hierarchical linear model is appropriate for this dataset.

The estimated level 1 variance (σ_0^2) is 5.178 and the estimated level 2 variance (σ_ε^2) is 1.239. The ICC indicates that 81%, over half, of the variation in KI scores is due to differences among classes.

$$\rho = \frac{\sigma_0^2}{\sigma_0^2+\sigma_\varepsilon^2} = \hat{\rho} = \frac{5.178}{5.178+1.239} = 0.81$$

The initial model assumes a linear relationship between scores at all assessment time points; because this model has three data points, it is not reasonable to model a quadratic term, as a sufficient number of parameters to estimate a reliable quadratic term is not present in the data.

The ICC indicates that it is necessary to account for the within-class variance in the model. A progressive series of multilevel models were constructed to account for condition, gender, as well as the interaction effects, condition*time and gender*time, to isolate effects of condition and gender, respectively. The specifications of these exploratory models are outlined in Table 6.

Table 6

Model Construction and R Syntax

Model #	R Syntax
M_0 unconditional means model	m0<-lmer(KIGMC ~ 1 + (1|CLASSID/STUDENTID), data=mydata, REML=FALSE)
M_1 unconditioal growth model	m1<-lmer(KIGMC ~ 1 + TIME + (1|CLASSID/STUDENTID), data=mydata, REML=FALSE)
M_2 gender as a level 1 predictor	m2<-lmer(KIGMC ~ 1 + TIME + GENDER + TIME*GENDER + (1|CLASSID/STUDENTID), data=mydata, REML=FALSE)
M_3 condition as a level 1 predictor	m3 <-lmer(KIGMC ~ 1 + TIME + CONDITION + TIME*CONDITION + (1|CLASSID/STUDENTID), data=mydata, REML=FALSE)
M_4 full model	m4 <-lmer(KIGMC ~ 1 + TIME + GENDER + CONDITION + TIME*CONDITION + TIME*GENDER +

including all predictors	(1|CLASSID/STUDENTID), data =mydata, REML=FALSE)

Model Building Description. Each model was evaluated following each subsequent addition of terms and compared using analysis of variance (ANOVA) using R. The second model in the series incorporates gender and the interaction of time*gender to account for any additional effects that may not be accounted for in the evaluation of condition. The third model accounts for the condition only terms, including condition*time interaction, as this is the most pertinent in addressing the research questions in this study. An analysis of the third model revealed no significant difference when analyzing the predictors associated with condition alone. The fourth and final model retains the terms of time, gender, condition, students nested within class, and the gender*time and condition*time interactions. Two different linear regression lines were developed to fit the data and to represent the change in KI score from pretest to posttest and from posttest to delayed posttest for each condition. The model outputs and comparison details are provided in Chapter 4.

Limitations of Growth Models

This application of a multilevel latent growth model allows for the analysis of associations between condition (GUI or TUI) within-student and between-student variance in each participating class block, and change in conceptual understanding over time based on the repeated assessments. Any existing correlation between levels cannot be reasonably interpreted as a causal relationship. Variability in human performance is too complex to be understood completely by the limited variables analyzed in this model. The degree of association may be established using multilevel modeling; however, this model alone is not sufficient to establish causal links. It is important to note that existing, yet unobserved factors or predictors not considered in this model may also bear responsibility for the resultant outcomes.

Qualitative Data Analysis

My role in the qualitative data analysis was one of non-participant observer during the video data collection, and I was the acting interviewer for all student groups participating in this study. As a researcher I tried to remain conscious of my perspective as a former classroom teacher and allow students to proceed through the laboratory without obstruction and without unnecessary assistance. In cases where students were having technical difficulties or did not understand something in the packet, I assisted them, though only to the extent that I did not provide students with any answers and deferred all content queries to the classroom teacher. During the laboratory most students remained standing at the laboratory stations in the classroom, and gathered around the computer to work on the activity.

The goal of examining specific cases was to better understand the types of interactions that students have with mixed-reality technologies in an authentic classroom setting. As a former classroom teacher, I am particularly interested in knowing how students may use technology to develop conceptual understanding of complex content. Qualitative analysis examined students' behaviors during the gas laws laboratory in each condition, TUI and GUI. Also included in the analysis were the follow up semi-structured interviews with student groups, which provided additional insight into students' interactions with the technology and with their peers while using technology.

Video Selection. The third research question targets student interactions with the technologies. To answer this question, I analyzed video recordings using a categorical approach to thematic analysis (Ryan & Bernard, 2003). A categorical approach to analyzing qualitative data generally involves generating themes, or categories (Rossman & Rallis, 2012, p. 269). In this particular study, I systematically employed analytic induction, with a focus on analyzing specific events or behaviors related to a phenomenon with the goal of creating some generalized principles, or themes, in light of the data (see Znaniecki, 1934). My review focused on students' behavioral and social interactions using the technologies throughout the duration of this laboratory and relating those interactions to students' test performance. While this cannot establish a causal relationship, the analysis of this data may provide useful insights for the implementation of innovative and integrative technologies in classroom environments.

The initial dataset included 21 video cases. Seven cases were removed due to a lack in identifiable membership (i.e., file labeling error in one class); packets or assessments in the dataset were not present or could not be matched with any reasonable level of certainty to participants of record in the video. This could be due to students not submitting packets or submitting packets without names, student names not corresponding to teacher-provided roster lists, or extreme legibility issues during transcription. One class set of video data from Madison High School was eliminated from consideration in the subset due to a labeling error; two separate groups of students were inadvertently assigned the same numerical identifier. The remaining cases were compared using average KI score gains from the pretest to the posttest and from the pretest to the delayed posttest. In reviewing these gains, four video groups were selected to represent each high performing and low performing ranges in each condition.

The purpose of reviewing video data was to describe students' interactions in context of the existing quantitative analysis to better understand how students work with emergent technologies. For this analysis I observed students' interactions with the technology by iteratively screening of all relevant video data. The initial review was focused on making sense of what transpired during the laboratory. The second review was where initial observations were transcribed and recorded as text and analytical memos were generated for each video recording (Appendix E). These written observations were reviewed and augmented by an additional review of the video data. Final video observations were collected based on the analytical memo data and organized in concept map form in order to make sense of any existing patterns in group similarities and differences (Appendix F) through the application of informal pawing, where certain types of ideas were color coded on the branches of the concept map. The concept map focused on student-technology relationships as well as social and behavioral outcomes influencing the laboratory group dynamic. Behaviors categorized as similarities and differences subsequently evolved into themes. I systematically compared and coded videos between conditions using the generated these interaction themes to triangulate the findings from the quantitative analysis and to develop a more detailed, balanced picture of what happened when students used the different interfaces for this laboratory activity.

The interview data collected following the laboratory activity underwent iterative review in a similar process as the video data. This data was systematically reviewed three times, with an initial sense making review followed by another review where I took notes, collected quotations, and began to organize concepts present in the data. All relevant student quotations, including portions of the interview where students were asked to review a portion of the in-class video recording and explicitly prompted to explain their interactions with the technology in the context of the lab, were transcribed and the *pawing* method of organization was used to organize the information for analysis (Ryan & Bernard, 2003). Formalized pawing typically involves a cut and sort approach to data analysis. For this study the quotations collected from the interview data were cut out and categorized according to questions asked in the interview protocol. Following that, I reviewed the data several times and developed categories. These content of these categories was reviewed in conjunction with the categories established from the video observations.

Themes

The focus of digital video and interview analysis was narrowed to 8 different student groups in both conditions (4 GUI, 4 TUI) to construct an understanding of learner-technology interactions that takes place in each condition. These 8 cases were selected to represent 2 high performing and 2 low performing groups in each respective condition; the rationale was that using students of varying performance levels may provide a more accurate snapshot of what type of learning does or does not occur when students use these types of technology to answer the third research question. Information gleaned from thematic analysis reflects the data triangulation of quantitative and qualitative data as described in greater detail in the following chapter.

CHAPTER 4

RESULTS AND CONCLUSIONS

The results and conclusions of this study have been founded on both quantitative and qualitative analysis. The quantitative analysis included (1) descriptive analysis and (2) inferential analysis. The descriptive analysis included a description and demographics of the study sample. Inferential analysis used Multilevel Modeling (MLM). The linear mixed growth model was constructed to account for performance in the context of nested data. For example, the model accounted for differences between students within the same class and the times when assessments were administered. The variance attributed to teacher and to time within each class was modeled and serves as evidence for any potential contribution to observed effects. The qualitative analysis examined video and interview data from the implementation of Mixed-Reality Laboratories in an authentic secondary classroom setting. This analysis was based on the explanatory-sequential mixed methods approach whereby qualitative data is used to provide a broader context to supplement or refute what is known from quantitative data analysis.

Descriptive Analysis

This section reviews the sample used for this study. The descriptive analysis provides an overview of the student population and the variables examined through the presentation of descriptive statistics. The purpose of the descriptive analysis is to create a more focused picture of the participants involved in this study. As such, the descriptive analyses will not be used to draw conclusions about the data; inferential statistics used for that purpose are to be addressed later in this chapter.

Sample

The study sample consisted of 307 students. Participants missing any one of the assessments, along with those who completed less than 50% of the items on any of the three assessment time points were excluded from this analysis (n=101; see Table 7) After eliminating data not subject to analysis, a total of 206 students remained and were included in the final analysis for this study. Classes were randomly assigned to condition; students across eleven classes (n=106) participated in the treatment group (Frame) while seven classes (n=100) were assigned to the control group (visualization-only). The treatment and control groups completed the laboratory activity working in independent groups of 2-4 students, with a maximum of 9 groups per class. The number of maximum groups was determined by Frame apparatus availability.

Table 7
Determining data for study inclusion

Total No. of Students in Classrooms	307
< 50% Assessment Completion	5
Missing Pretest	26
Missing Posttest	48
Missing Delayed Posttest	22
Total No. of Students (included in final dataset)	206

Demographics

The distribution of students by gender and condition can be found in Table 4. Of the 206 students participating in this study, 43.2% (n=89) were males while 56.7% (n=117) were females. In the overall dataset, a larger proportion of females is present in this sample. Reviewing the gender breakdown by condition, 52.9% (n=47) of all males are represented in the GUI condition while 47.1% (n=42) of all males are represented in the TUI condition. 45.3% (n=53) of all females are represented in the GUI condition while 54.7% (n=64) of females are represented in the TUI condition. Examination of the proportions of male to female representation in each condition reveals that 47% of participants in the GUI condition are male, and 53% of participants in the GUI condition are female. In the TUI condition, there are 40% males and 60% females.

Data related to students' race/ethnicity was not collected individually. Therefore, it may be helpful to refer the demographics of each school (see Table 3) and the following proportions of students in this sample from each respective site: 38.8% (n=80) were from Washington High School, 47.1% (n=97) were from Jefferson High School, and 14.1% (n=29) were from Madison High School.

Table 8
Study sample gender distribution by condition

	Overall Sample	Proportion of Total Gender in Each Condition		Proportion of Each Gender in Condition Total	
Gender	Proportion	GUI	TUI	GUI	TUI
Male	43.2% (n=89)	52.9% (n=47)	47.1% (n=42)	47%	40%
Female	56.7% (n=117)	45.3% (n=53)	54.7% (n=64)	53%	60%

Inferential Quantitative Analysis

Missing Data

Listwise deletion was employed to eliminate students with missing or incomplete tests. It should be noted that while the listwise deletion approach reduces statistical power due to the elimination of participants, it allows for reasonable comparisons across analyses. The missing data (n=101) includes 5 students who had incomplete assessments and 28 students did not complete a pretest; therefore, the final analysis of missing data included students who completed the pretest (n=68) and may have been missing either a posttest (n=47) or delayed posttest (n=21). To test whether significant performance differences existed between groups, an independent sample t-test was conducted to compare KI pretest composite scores of included students and excluded students across all six assessment questions analyzed in this study. Levene's test for equality of variances was not found to be violated for the present analysis, $F(1, 272) = 1.753$, $p = .187$). There was no significant difference in the composite KI scores for included participants (M=2.40, $SD = 0.841$) and participants excluded from the study due to missing data (M=2.463, $SD = 0.671$); $t(272) = -0.595$, $p = 0.552$. These results suggest that data included in the final analysis is not significantly different from data excluded from this study.

Multilevel Latent Growth Model

Multilevel modeling was used in this analysis to address aspects of the first two research questions. The first two research questions ask whether students' conceptual understanding as measured by KI improves from pretest to posttest and in the delayed posttest for retention. A hierarchical linear growth model was employed to estimate the relationship of variables to student performance in each respective condition by partitioning the variance within and between students in a nested in a particular class. In the analysis process, four distinct models were evaluated and fitted to the data (Table 9).

Table 9
Models and parameter descriptions

Model	Description
M_1	Unconditional Growth Model (base model without predictors + time)
M_2	Gender + Gender*Time interaction terms added to the growth model
M_3	Condition + Condition*Time interaction terms added to the growth model
M_4	Full model containing all predictors at all levels (combines M_2+M_3)

Model Comparison. Nested models were compared to using analysis of variance (ANOVA) in R (i.e., M_1-M_2, M_2-M_4). The comparison of M_1-M_2 indicates that the residual sum of squares was significantly lowered with the addition of gender and gender* time into the model (χ^2=8.396; *p=0.015*).

To further assess the goodness-of-fit for these models, the Log Likelihood Ratio (LR) where LR = 2 * (LLmodel1 - (-LLmodel2) where model1 is the first model being compared and model2 represents the second model being compared. This statistic was calculated using the Log Likelihood (LL) statistics calculated in R as a part of the multilevel modeling package lme4. The log likelihood ratio calculations are as follows:

M_1-M_2: 2*(-659.68)-(-655.48)=2*-4.2=-8.40 (p-value < 0.01)

M_2-M_4:2*(-655.48)-(-654.99)=2*-0.49=-0.98 (p value > 0.05)

The corresponding p-values were derived from a chi-square distribution, as the outcome from the LR calculations is a chi-square distribution with 2 degrees of freedom.

Table 10
Multilevel Model Analysis

Fixed effects	**M_1**		**M_2**		**M_3**		**M_4**	
	β (SE)	***t***	**β (SE)**	**t**	**β (SE)**	***t***	**β (SE)**	***t***
Intercept	-0.004 (0.112)	-0.039	-0.061 (0.138)	-0.447	-0.085 (0.440)	0.530	0.031 (0.178)	0.177
Time	0.214 (0.029)	7.492	0.286 (0.043)	6.622	0.223 (0.041)	5.435	0.290 (0.051)	5.694
Gender			0.104 (0.140)	0.742			0.105 (0.140)	-0.786
Time*Gender			-0.127 (0.057)	-2.214			-0.126 (0.057)	-2.198
Condition					-0.165 (0.219)	-0.751	-0.173 (0.220)	-0.786
Time*condition					-0.017 (-0.057)	-0.299	-0.008 (0.057)	-0.136
Random effects	***SD***	***Variance***	***SD***	***Variance***	***SD***	***Variance***	***SD***	***Variance***
STUDENTID:CLASSID	0.440	0.193	0.435	0.189	0.440	0.194	0.434	0.189
CLASSID	0.368	0.135	0.370	0.137	0.353	0.125	0.357	0.127
Goodness of Fit								
AIC	1329.4		1325.0		1332.2		1328.0	
BIC	1351.5		1355.9		1363.2		1367.8	
Log Likelihood	-659.7		-655.5		-659.1		-655.0	
Deviance	1319.4		1311.0		1318.2		1310.0	

LR statistics as calculated suggest that the largest model, M_4,has best fit. Based on this data, I have concluded at Model M_4, the full model is the best fitting model to represent any change in relationship between condition and pretest to posttest growth as well as between condition and posttest to delayed posttest growth.

KI Score Trajectories

An examination of average KI score trajectories across conditions indicates students in the GUI group generally began with a higher conceptual understanding than that of the TUI group (see Figure 5). The GUI group hit a plateau and exhibited a slight downward turn from the posttest to the delayed posttest. The TUI group started at a lower average KI level than the GUI, demonstrated a minor average gain than the GUI from pretest to posttest, and performance remained steady, without much change from the posttest to the delayed posttest.

Table 11

KI composite score means by condition

Condition	N	Assessment Time	Mean KI Score	SD
		Pretest	2.51	0.938
GUI	100	Posttest	2.96	0.749
		Delayed posttest	2.98	0.954
		Pretest	2.28	0.725
TUI	106	Posttest	2.99	0.679
		Delayed posttest	2.71	0.706

Condition	**N**	**Assessment Time**	**Mean KI Score**	**SD**
		Pretest	2.51	0.938
GUI	100	Posttest	2.96	0.749
		Delayed posttest	2.98	0.954
		Pretest	2.28	0.725
TUI	106	Posttest	2.99	0.679
		Delayed posttest	2.71	0.706

Overall, KI trajectories are a surface reflection of conceptual changes or understandings as they have been presented in individual student responses. A lack of large significant shift in the TUI and GUI conditions were observed in a majority of students' responses regarding the particulate nature of matter. As research indicates, alternative conceptions related to the particulate nature of matter are persistent and difficult to change. The results of this study iterate similar findings. For example, one TUI student's pretest response to "…how do gas molecules cause pressure?" states that, "gas molecules have pressure because they are being pressed together." This is fairly typical across the data set; students often do not drill down to the molecular level and often try to make sense of what they see relative to a macroscopic view. In this case the context of the question asks about people sitting on an air mattress, and like many other students in this study, this response focuses on the macroscopic element of something being pushed onto something else, indicating that this push creates pressure. In the posttest, this student extended their pretest response by saying, "the molecules come closer together, creating more pressure." This shows a nuanced shift in understanding, or at least in communicating the idea, yet it still falls into the alternative conception of how gas molecules cause pressure. This student's final response in the delayed posttest is more normatively developed in stating that, "they are in constant motion and there is limited volume."

As this response indicates, the shift from non-normative to scientifically normative ideas can be a challenging, slow process. Though several students did make great strides in improving their KI scores, for example, one GUI student submitted the following response for the same question asking how gas molecules cause pressure:

> They are constantly rebounding off each other and their surroundings, in this case the air mattress. They get sent back outwards because of this rebounding, and force the mattress outwards. I think.

This particular response received a 2. A score of 2 indicates that a student attempted to answer the questions, but they may have presented ideas that were alternative or not directly connected to the targeted conceptual answer. The reason that this response received a 2 was because of the mention of molecule-molecule collisions; while these collisions do occur, these collisions do not contribute directly to pressure. This student moved to a score of 3 in the posttest on the same question with this response:

> They cause pressure because they are constantly hitting the sides of the area they are contained within, in this case the mattress. this bouncing around causes some molecules to head outwards, which causes the container to have a force almost pushing out, which is the pressure within the mattress. Even if there is a force on the mattress, the molecules are still bouncing around and pushing back on the innards of the mattress, creating pressure.

The score went up to a 3 in this case, because the student has clearly omitted the earlier reference to molecule-molecule collisions, thus making it more aligned with a scientifically normative explanation of how gas molecules cause pressure.

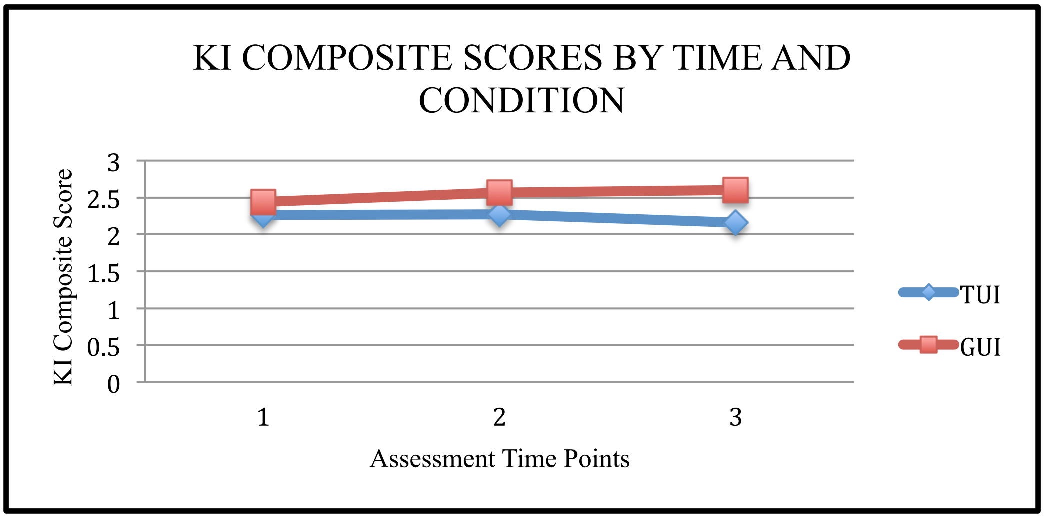


Figure 5. KI Composite Scores by Time and Condition

Interpretation

The first research question investigated whether the effect of condition (TUI or GUI) was significant as evidenced by students' KI score growth from pretest to posttest. My initial hypothesis, when beginning this study, was that the technological condition emphasizing direct hands-on manipulation corresponding to the visualization would be as good as or more beneficial for student learning than the GUI condition. The graph comparing growth between conditions (see Figure 5) indicates that there is comparable growth across each condition, and model comparisons revealed that there were no statistically significant differences in the study sample between students' performance in the TUI condition when compared to student performance in the GUI condition from pretest to posttest assessments (Appendix F).

The second research question examines whether there were any differences between groups from the posttest to delayed posttest time points. There was approximately one month between the posttest and delayed posttest assessments. Ideally, it would be expected that student scores increase from pretest to posttest and then those acquired gains would plateau through the delayed posttest time point and perhaps taper off. As with the first question, quantitative analysis reveals that there is no statistically significant difference between groups in this study sample, in terms of KI score, from the posttest to the delayed posttest. Students' KI scores plateaued in both conditions from posttest to delayed posttest and remained largely the same. Visually, a greater increase is observed for the GUI condition compared to the TUI condition, however, this apparent difference between conditions was not found to be statistically significant in this study sample.

For both research questions, the model investigated the interaction between gender*time as well as that of condition*time. Gender was added to the analysis to see if there were any statistically significant differences on assessment performance by gender. In this particular study population, there was no detectable difference between males and females.

Qualitative Analysis

The emphasis of the qualitative analysis was to describe how students interacted with the technologies in both groups and explore any key aspects or behaviors that may have differed across groups. The quantitative data informed the selection of two high performing and two low performing groups in each condition; overall KI gain scores from pretest to posttest and from pretest to delayed posttest were compared across groups.

Groups classified as high performing had positive average gains in KI score from pretest to posttest and retained an overall positive average gain remaining in the delayed posttest. Groups classified as low performing had marginal or no gain from pretest to posttest in addition to a lower KI score in the delayed posttest. The video cases for the final analyses were purposefully selected to represent a range of prior knowledge; 2 high performing and 2 low performing student groups were selected from each condition for a total of eight video groups (Table 12). The video case selection captured a range of student performance in the sample, and iterative review of video revealed commonalities and differences that may influence students' performance and conceptual understanding of the content.

Table 12
Video groups included in final analysis

Group Identifier	Condition	Performance Category[1]	Notes
O37	GUI	HP	Positive average gains from pretest to posttest; overall positive average gain remaining in delayed posttest
P38	GUI	HP	Positive average gains from pretest to posttest; overall positive average gain remaining in delayed posttest
G82	GUI	LP	Marginal gain from pretest to posttest; loss on delayed posttest
O77	GUI	LP	No gain from pretest to posttest; loss on delayed posttest
O19	TUI	HP	Positive average gains from pretest to posttest; overall positive average gain remaining in delayed posttest
G62	TUI	HP	Positive average gain from pretest to posttest; loss equivalent to half of that gain on the delayed posttest
P28	TUI	LP	Loss from pretest to posttest; increased loss in the delayed posttest
G57	TUI	LP	No gain from pretest to posttest; increased loss in the delayed posttest

[1]HP denotes "high-performing" groups, LP "low-performing"

Thematic analysis. Eight student groups were selected from the overall sample for analysis based on performance across pretest, posttest, and delayed posttest measures. Two high performing groups and two low performing groups were selected from each condition. Across student groups there were total of 23 student participants. Of the 23 in video data collection, 11 are female and 12 are male. To explain the qualitative findings, these themes were grouped by interaction type: student-student and student-technology (Moore, 1989). A systematic review led to the reorganization and consolidation of categories and the further development of themes (Table 13).

Overview of Student-Student Interaction Themes

Themes categorized as student-student interaction themes represent the interactions that students had with one another. These themes were derived primarily from the video recorded laboratory, as the interview focused largely on students' experiences with the technology. Group dynamics and social communication emerged pertaining to the relationships of grouped students when interacting with the technology. Each of these two themes addressed aspects of students' ability to cooperate with others to complete an assigned task.

As with any group arrangement, previously existing relationships or familiarity has some bearing on the outcomes of students working in a group setting. In this laboratory there were many cases where students self-selected their groups. In the case of several video recorded groups, some students were simply arranged in a group due to their status as a participant with video consent on file. Teachers determined group assignments based on the receipt of signed consent. While it is not possible to ascertain which students' prior relationships directly affected their performance in the technology-based laboratory, it is an important consideration when making sense of the social aspect of these data.

Group Dynamics. Group dynamics appeared to play a role in the degree of participation of members of particular groups in this study sample. Group dynamics of all of the video groups can be characterized as either alpha-dominant or mutually cooperative. Groups that exhibited alpha-dominance had one key member who acted as leader and authority over the laboratory procedures; this person became the primary controller of the technology. This was most notable in the low performing GUI groups, where a male student typically overran the ideas of others to retain control of the laboratory. As previously mentioned, gender was not a variable of focus in the initial line of research questions; however, analyzing behaviors and interactions in terms of gender provided interesting insight into group dynamic and observed laboratory performance.

Table 13
The derivation of analytic themes

Initial Categories	Consolidated Categories	Emergent Themes	
		Student-Technology	Student-Student
• Technology Preference • Likes • Dislikes • Social Interactions • Hands-on vs. Visualization • Play • Confusion about Controls • Preplanning & Innovation • Intergroup Communication • Experiential Ranking • Easiest Feature • Pressure	• Preplanning and Innovation • Social Interaction • Usability • Learning with Technology • Perception vs. Performance	• Investigative Procedures • Limitations of Technology • Perception vs. Performance	• Group Dynamic • Social Communication

Alpha-dominant groups. One such GUI group, identified as O77, included one male and two female group members. One female, Jessica, read the packet aloud to the other members of the group. The boy, James, talks loudly and asserts confidence in his knowledge, referring explicitly to Newton's Third Law of Motion several times as an explanation for molecular motion occurring within the visualization. When examining the *trace* function of the simulation, he stated, "an object in motion will stay in motion until hit hits something else". James also linked phenomena observed in different states of matter to the gas laws content (i.e., solid and liquid), using his prior knowledge to dominate the group opinion as seen in the following excerpt:

> James: When the temperature increases, volume will increase
> Jessica: Really?
> James: They'll move around a lot faster
> Jessica: (looks at simulation) I suppose that's true
> Heather: You suppose?
> James: Well, gases move around faster in a solid and [a] solid's colder so…
> Jessica: I don't think that's always true though
> James: Ice to water vapor
> Jessica: Oh…yeah
> James: Which is colder?
> Jessica: Okay! (loudly)
> James: Ice is cold, and it's dense
> Jessica: Okaaaaaay!!! (louder still; defensive tone)
> James: And it's like this [makes a block shape with his hand] and water vapor is like this [gestures using a hand waving motion]
> Jessica: ((sigh))
> Heather: So we can agree that when temperature increases volume will increase as well.
> James: That's not a properly stated hypothesis--
> Heather: But *they* have it this way…
> James: It's supposed to be *'if'* or *'may'*

This conversation evolved into the following exchange between group members as they were constructing hypotheses about the behavior of gas molecules as temperature increases:

James: The piston is going to go from here to here [gestures with hands indicating that the piston will move to the right and the chamber will be larger].
Jessica: If he says [it], it's right.

This deference to the dominant individual was common across GUI groups categorized as alpha dominant. In group O37, a low performing GUI group, the membership consisted of 2 males and 1 female. This group lacked direct peer-to-peer communication; they rarely spoke to each other throughout the duration of the laboratory. These three students were arranged in front of a laptop computer running the virtual simulation. One student was situated to the right of the machine, one in front of the machine, and one to the left. The student in the middle was directly facing the computer and served as the primary driver for the simulation activities. The other students did not touch or interact with the simulation at all, although the students in this group sometimes referred to it using hand gestures to assist them in the construction of verbal explanations, as evidenced by the following exchange about molecular motion:

Jane: Are they all moving the same speed?
John: Well, [pointing at screen] the red ones are moving a little faster.

On other occasions, such as in the exploration of partial pressure and volume involving the relationship of different types of molecules, hand gestures were used without direct reference to the simulation apparently for emphasis or illustration:

John: Oh, the volume stays the same…the volume's constant so I think the bigger they are the higher the pressure…so as the molecular mass increases, the pressure will also increase because they are bigger, they repel more [pushes his fists, knuckle on knuckle against each other to show the molecular collision]. They're bigger so they take up more space.
Jane: Sounds reasonable.

The members of O37 spent a lot of time writing in their packets independently, as opposed to interacting with the simulation. When they do discuss what is going on with the laboratory, the student driving the simulation dominates the conversation and compared to other group members expresses more of his understandings verbally. These understandings go unchallenged by the other group members, and the explanations provided in the laboratory packets are largely dictated by the thoughts of this one individual. Discussions within the group are limited to phrases like "I got this" or "okay".

The dominant student in this group resolves one discrepancy that came up in the lab regarding group members' graphs. Students were having difficulty figuring out how to plot data for a laboratory investigation that was designed to illustrate the volume-pressure relationship. Instead of engaging in a debate or a discussion whereby thoughts and ideas are brought to the fore, these differences were aligned by the dominant member using a fact-checking approach. For instance, he would say, "I got this number for pressure" while pointing at the numerical data in his table as a reference to his fellow group members. The members of the group acquiesced, and even though the simulation was right in front of them, they did not go back and rerun it to double check the data that had been previously recorded in their packets.

This observation of alpha dominance is also consistent with the high performing GUI group, G82. This group contained only two members, 1 male and 1 female. The male in the group assumed control and was in charge of running the simulation for the duration of the laboratory. In this case, the male student appears to do a majority of the work. He talks aloud as he is completing the packet, and the female transcribes the answers he provides without question and without attempting to engage with the simulation directly to discover or confirm answers for herself. In at least two instances, the boy did ask the girl for an opinion, and she offered little in return. It was not immediately clear if this was due to her lack of understanding or to other variables not considered in this investigation. In any case, this group was male dominant and the female accepts his assumptions and assertions as correct; she doesn't appear to make her own hypotheses and they do not confer as a group to discuss the hypotheses. She never challenges what he says and goes so far as to copy his drawings to represent her observations of the molecular phenomena.

In these three examples, it seems that the alpha dominance quality is exhibited in varying degrees. In the low performing GUI groups, one male asserts authority over what the group does. Whether this is based on direct knowledge of the content is not immediately clear, but being that this degree of dominance was observed specifically in low performing groups, one might argue that this is not the case. In the final example, this alpha dominance seems to be due to the default of another member. That is, a perceived lack of interest, attention, and understanding required the male in G82 to engage with the laboratories largely independently.

This alpha dominance quality seemed to differ from leadership, as exhibited in mutually cooperative groups, based on students' relationships or the lack of established working relationships which appeared to substantially influence the group dynamic. When group members were asked whether they had worked together previously, they noted that they were familiar with the people with whom they were working. However, they did not necessarily indicate that they had jointly participated in laboratory activities in their classes prior to this implementation.

To summarize, alpha dominance was observed in two GUI low performing groups and one GUI group classified as high performing. This behavior was typically characterized by rigid group roles, mechanistic fact-checking behavior, and ideas going unchallenged. For example, Group O37, a high performing GUI group with 2 males and 1 female assumed rigid roles for the laboratory (e.g., one person is the recorder, another conducts investigations, etc.). In these roles students relied on mechanistic fact-checking behavior primarily focused on laboratory packet completion as opposed to understanding. For example, one male and one female in this group often worked together toward completion by asking each other questions like, “what did you put down for this question?” The response from the student being queried more often than not went unchallenged and the group proceeded to the next investigation while appearing to bypass broader conceptual thinking about the topics in question.

Mutually cooperative. Mutually cooperative groups rotated roles through investigational tasks and cultivated interdependent relationships among members. All TUI groups, high performing and low performing, (i.e., O19, G62, G57, and P28) and one high performing GUI group (i.e., P38) involved in this analysis were classified as mutually cooperative. For example, in the high performing TUI group identified as G62, there 2 males and 1 female who took turns using the various components of the simulation and conducting investigations. The female in the group began by narrating activities for the rest of the group; she later traded places with the males in the group to try the spring and to help advance the laboratory. Each person in this group experimented with the TUI inputs and no one was limited to a prescribed role.

Mutually cooperative groups made regular progress throughout the laboratory activity; however, in the observed groups, those categorized as mutually cooperative were more likely to engage in play than alpha dominant groups. While it is important to note that only 2 mutually cooperative groups did engaged in play, play was not an element observed in groups classified as alpha dominant. Out of the mutually cooperative groups where play was observed, students who had good rapport with one another seemed to tinker and create games to play, specifically using the TUI interface.

Group G62, a high performing TUI group, created a game resembling tug-o-war using the external inputs to represent opposing sides. The piston/spring mechanism represented humans and the syringe, which was used to increase the number of molecules, was used to represent the molecules in this game, even though it was human controlled. The syringe was controlled by one member of the group and used to increase the number of molecules in the chamber while another member of the group controlled the piston and attempted to reduce the volume in the chamber. The students referred to this game as "humans versus molecules" where each student was involved; two were interacting directly with the simulation while the remaining group member took on the role of cheerleader. The apparent objective of this game was to see who would win the invented competition. This game was developed during a brief break between activities when the female in the group went to ask her instructor a question. It began with the male group members inventing and playing the game, and this group resumed the laboratory activities upon the girl's return. However, it should be noted that once this group had completed the packet, they returned to this game and all group members were involved in play at the conclusion of the laboratory.

Another TUI group, P28, classified as low performing, consisted of 3 males and 1 female and engaged in play with the external inputs driving the simulation. In this case it was also primarily the males who took to playing by adding and subtracting different molecules using the syringe and the button on the screen that allowed students to switch molecule types. A and B molecule types in this simulation represent lighter and heavier molecules, respectively.

In playing with the simulation, students discovered that B molecules move much more slowly and begin thinking about what might happen with a change in temperature. One student expressed eagerness and enthusiasm for understanding how these molecules might react when temperature increases, "imagine *that* with heat!" They were curious to know about this specific interaction, though these students were not explicitly asked to investigate partial pressure (i.e., the activity that asks students to use two different molecule types) in this class period. Interestingly, at least one student in this group did not appear to link the physical and virtual worlds successfully, which may have contributed to the low performance as measured by the assessments. Following group play at the conclusion of the laboratory, one male student poses the question, "are there actual gas molecules in there?" which indicates an uncertainty about the mechanism involved in representing this phenomena on screen.

Social Communication. The repetition of ideas related to a group's collective social behaviors (e.g., intergroup communication, social interaction, etc.) led to the development of an overarching theme, social communication, to encapsulate observations of student communication within their workgroups. Within social communication observations fell into two general categories: beneficial social communication and limited social communication. Beneficial social communication generally involved students trying to reach a consensus about a given answer or idea presented by the laboratory simulation or packet. Additionally, there were groups whose social interactions and attempts at social communicate appear limited within this context.

Beneficial social communication. Beneficial social communication was characterized by laboratory discussion involving 2 or more members. In groups where this was observed, students within an assigned group posed questions and discussed findings in the laboratory to come to a consensus on the development or retrospective reflection of the hypotheses being tested. Group O19 primarily used laboratory discussion to reach agreement on hypotheses, specifically when deciding whether molecular pressure would increase with temperature.

Likewise, the mutually cooperative group, GUI P38 (2 males, 2 females), engaged in discussion using the simulation as evidentiary support for claims and occasionally asked for outside help. An initial discussion by this group involved comparing what they observed in the simulation to initial hypotheses. Two students in this group proposed that molecules bunch up together in the center when temperature decreases.

> Jacob: They don't look like they bunch together.
> Monty: The question's not about the chamber.
> Jacob: Yes, it is.
> April: They don't fill the chamber as before.
> Jacob: Why not?
> April: Because the chamber's bigger.
> Monty: That's irrelevant. Look at how much more spread out they are (*after the piston was adjusted to make the chamber larger)*

The discussion about this interpretation of the visualization is resolved largely because students were able to observe the behavior of the simulation to use them as evidence for their claim that the hypothesis they made required revision. Dynamic visualizations bridged a gap in student-student communication of idea. This type of discussion is contrary to what was observed in many other groups; this will be detailed in terms of limited social interaction.

Limited social interaction. Limited social interaction was characterized by task-based roles and individual work dominating group effort. Task-based roles are considered to be unchanging and inflexible roles assumed by a student member in a group. The task-based nature of these roles is guided by the prescribed laboratory activities, and each of the roles is heavily procedural and scripted based on the laboratory. Alpha dominant GUI groups tended to focus on task delegation and individual work rather than engaging in content discussions to compare hypotheses and share ideas amongst the group. For example, in the GUI group O77, the task delegation was divided up into three: (a) 1 member read directives (b) 1 member was the simulation driver and (c) 1 member was a passive participant. Similarly, GUI group O37 was driven by a leader dictating how things were, which appeared to reinforce the lack of communication between members of the group. The leader went largely unchallenged and the extent of the conversation was limited to rudimentary fact checking and copying work from other group members, as was the case in both GUI 037 and G62.

The input from students, aside from the alpha dominant group member, was generally limited to one or two word answers, such as “yeah”, “not sure”, “okay”, and “yep”. When there were discrepancies between the graphs of each group member, there was not a discussion concerning the cause of the difference. The conversation was limited to, “I got 6.5 for pressure” and the other student says, “okay” without inquiring about how that reading was derived. Each member in this group completed explanations and drawings independently, though students in this group sought confirmation simply by looking at, and occasionally copying, the work of fellow group members, as was the predominant method for packet completion in G62.

Overview of Student-Technology Interactions

Non-prescribed investigative procedures. Students in the TUI condition tended to engage in task relevant, non-prescribed investigative activities that were not included in the lab guide. For instance, when investigating the relationship between temperature and pressure, students were asked to take multiple temperature readings, two of which using jars containing hot and cold water and to record data from the investigation in a corresponding data table. The laboratory suggested taking other readings at room temperature, placing fingers on the thermistor, etc. Group O19, as a part of their preplanning for this investigation as previously mentioned, talked through an actionable plan that was used to collect multiple readings:

Linda: (finishes writing previous answer) Okay

John: Reset the frame and record the volume at five different temperatures...So I think the sensor's over here (looking around the frame; settling on syringe side where the thermistor is located)—I don't see anything else that looks like a sensor
Linda: Yeah.
John: So I'm thinking we get the cold water, and just put it up against it
Linda: Well, we need five so we're going to need another two.
John: I'll go get the cold...We only have two jars
Linda: Do we want to start it at room temperature?
John: Yeah. So for number 5, we're going to do room temperature?
Linda: Room temperature, hot water, cold water, human fingers, and?

In this dialogue, students are trying to come up with ideas for things to measure. Unlike TUI groups, GUI groups were not observed having similar conversations with each other; they were not observed preplanning prior to actually doing the laboratory activity, and rarely did they seek outside help. TUI groups were more likely to put a plan in place and then act on it, as opposed to just going through the motions of the laboratory. In this particular set of activities, it was not uncommon for TUI groups to think a few steps ahead and determine what they would use for measurement. This particular group was approached by a member of the research team asking if they needed assistance; Linda and John were both looking around the room and appeared to be lost:

Researcher: Questions?
Linda: We're trying to think of things to measure temperature,
John: Five different
Linda: So we've got room temperature, hot water, cold water, and human fingers.
Researcher: Mixing hot water and cold water?
Linda: Oh, we could do that!

While engaging in preplanning, students were brainstorming about how they could collect five different measurements. Each group was provided with two jars; each respective jar was to be used one time in the data collection process. However, after consulting with a member of the research team who inquired if the group had any questions because they were both looking around the room during the laboratory, this group determined that they could get multiple readings using the jars by altering their contents. Students in this group took turns using the jars and proceeded to use one of the jars for the hot water trial by emptying the contents in exchange for warm water from the faucet.

Likewise, in the same laboratory investigation group G62, consisting of 2 males and 1 female, discovered another way to gather similar data. When conducting the temperature and volume investigation, students in this group combined the contents of both jars for an additional measurement:

Gina: So what do we do with the last one [measurement]? We've already done room temperature, our fingers, and cold and hot water, so?
Brett: I don't know

Brian: Can we use another part of our bodies?
Brett: Can mix the hot and cold water?
Gina: Sure! (girl mixes the water)
Brett: I wanted to see a reaction! ((laugh))
Gina: So now we have icy hot water! (swirls the water around inside the jar)
Brett: It's so cold—the top is cold and the bottom is warmer
Gina: That's weird!

After the students in this group agree on a plan of action, Gina mixes the water from two jars into one. They appear to be surprised by the energy transfer from the water to the material of the jar, and they proceed as normal throughout the remainder of this activity. It should be noted that this group was a different class entirely than the previous lab group.

The TUI condition provided opportunities for interaction outside of the GUI. These opportunities seemed to facilitate preplanning and innovation within TUI groups; this affordance of the TUI condition also allowed students to try to enable and discover new methods of interaction. Group G57, for instance, tried to pair the jars in the same laboratory activity with the Frame apparatus in novel ways. The female student in this group was attempting to put the hot jar on top of the touchscreen interface to register temperature change within the simulation. While this action is certainly attributable to confusion about the sensors, which will be discussed later in this chapter, it also seems to indicate that students are exploring unconventional, non-prescribed ways of interacting with the device as a method of problem solving.

Technology limitations. In the course of this study, several limitations of both technologies were apparent in video observations as well as identified by students in interviews. These limitations have been categorized as usability issues, representational incongruence, and students' perception versus performance measures.

Usability. To some extent the usability of the technology was an issue for all of the groups participating in this study. A common usability theme across all groups, TUI and GUI, high performing and low performing, involved confusion about controls. This section highlights some of the usability issues that had arisen during the implementation of the GUI and TUI laboratory activities, beginning with a breakdown of usability issues by condition and concluding with a segment highlighting usability issues that were present regardless of condition.

GUI condition usability. Students in the GUI condition were accustomed to this kind of interface as computers in homes and schools are fairly ubiquitous. However, there were a few simulation controls that were difficult for students to manipulate, specifically the temperature slider, which included a small indicator button that increased navigational difficulty with the mouse. A male student from the high performing GUI group, P38, indicated that, "you couldn't really like slide it [temperature slider] to where you wanted…it jumped to like 6 or 8 different spots along the side [when clicked]". Likewise, students from O37 agreed this feature was difficult to use because "the sliders were small" and students from G82 said that they "were having a hard time finding out where we could increase or decrease the temperature" due to the size of the slider. The ability to manipulate the slider by clicking on a small button was problematic for many students, at least initially. This feature seemed to cultivate frustration, but fortunately, it was not an insurmountable obstacle; students were able to continue with the lab and complete it regardless.

The add/remove molecule feature in the GUI was also difficult for students because it was a checkbox that made it difficult for students to attain measurement precision when trying to add a certain number of molecules to the simulation. In the laboratory packet, there was an activity that asked students to add 100 more molecules to the simulated chamber. In the GUI version of this laboratory, the check boxes were frustrating and proved to be a time consuming aspect of the lab for a few groups. A female from the low performing GUI group O77 indicated that the group spent “5 to 15 minutes” tinkering with this to match the request, “click to 99, click to add more, now we got 110” and eventually the group gave up “because it was 24 or 26” molecules added to the system instead of preferred lower increments. The female student went on to elaborate that “[she] got really frustrated; I walked away and took off my sweatshirt.” She mentioned in the follow up interview that she didn’t think adding a physical component to this (i.e., a syringe to add molecules) would be “any more precise than if you could just type in a number.” A male student from group G82 echoed this sentiment. He noted that “whenever you say remove and click that again [to stop removing the molecules], it still removed them…for some reason, ours kept going.” In his group’s follow-up interview, he also suggested the approach of being able to “type in the amount” to avoid the hassle.

Another issue limited to the GUI condition involved the force slider, which was intended to reflect the relationship of atmospheric pressure on the internal pressure of the chamber. The laboratory investigation of volume and pressure asked students to hold the volume constant and review the pressure, but there was no way, using the spring, that students could keep the volume completely constant without fluctuation. While this is not completely representative of the molecular mechanisms underlying this concept, it proved difficult to get students to understand this in using the technology without a piston lock in the activity as it was presented to students. A student from a P38, a high performing GUI group, expressed that she “didn’t really get what it was used for” and that this was the most difficult thing about using the simulation because she wasn’t “sure if the molecules were doing something to the piston or if it was the slider” affecting the piston. Another classmate from G82, a low performing GUI group, went a step further to admit that “the biggest thing [he] didn’t like was how you can’t lock the piston.” Additionally, the pressure reading was not an accurate reflection of pressure; sometimes it read as a negative integer, which was difficult for the students to comprehend. This was a simulation glitch that appeared in a few different groups throughout the implementation. This was resolved by resetting the simulation.

TUI condition usability. Students, especially those in the TUI condition, had to become accustomed to the operations of the novel external input controls. Once these were fully explained to the student, through the introductory walkthrough, student exploration, and addressing questions, this was less of an issue as students progressed through the laboratory packet. As one student explained,

> this isn’t a piece of technology that we’ve seen before like it’s not every day that you come into contact with the box part of like the parts with the syringe built in…it added to kind of like exploring it more like because we weren’t exactly familiar with the technology.

The lack of immediate familiarity was evident when students, such as those in the TUI group G57 (1 male, 1 female), were confused as to the appropriate use of the sensors. After students in this group filled their jars with hot and cold water, respectively, they had difficulty finding the temperature sensor. At two different points, the female student was observed attempting to place the jar on top of the touch screen:

> Jeremy: I think we kind of struggled in the beginning of like how to do it. We really didn't think that you just touch it together [the jar to the box]. We were kind of confused like how to use that but we figured it out.
>
> Janice: It was weird…well, like it's just normally you put something in the water then that tells you the temperature; you don't put it next to it.

The male in this group also exhibited a similar confusion about the temperature sensor when he held a piece of PVC plastic tubing in lieu of the temperature sensor's thermistor. Although both students expressed that the opportunity for hands-on interaction was positive for them. When asked what they would have thought if hand controls were removed and replaced with buttons on a screen, they agreed that the hands on made more sense:

> Jeremy: I would have been much more confused, I think, because when you make the connection between the two, like you see this as a spring [on screen], you just know to push the spring [the input] in and it will move the spring [on screen]…when it's actually interactive like this, it just makes more sense.
>
> Janice: And I'm better when you can like touch something and not press a button, it just makes more sense to me…it seems like you're actually doing something real and you're not just like on a computer.

One notable usability issue specific to the TUI was representational incongruence. Representational incongruence occurs when the physical representation of the action or concept does not correspond intuitively to the action or concept. For example, a syringe is used to add/remove molecules from the simulation. However, when presented with a problem that requested students to add or remove a specific number of molecules from the system, students of group O19 had difficulty understanding the correspondence of the syringe to the number of molecules added or removed. The video observation and the follow-up interview revealed that students believed the numeric labeling on the barrel of the syringe in some way corresponded to the number of molecules being added as the plunger was compressed.

Similarly, this group of students initially thought that the volume of the chamber was equal to the volume represented within the syringe when they started to pull or push on the plunger. This confusion often ended with the plunger being extracted completely from the barrel of the syringe, in which case students resumed the laboratory after the syringe components were reconnected and the simulation was reset.

These difficulties were also expressed in the TUI group, G62, in follow-up interviews:

> Jane: I didn't realize that it wasn't a correlation between like how much you pulled out the plunger and how many atoms of gas came out as well, so that was something [I found to be] a little confusing.

The novelty of the controls in the TUI condition elicited positive and negative feedback. The novelty appears to increase student engagement in the laboratory to some extent, based on observation, yet the novelty also increases the likelihood that confusion will arise, even after explicit instructions have been issued. Based on student feedback, this issue requires balance to be thoroughly considered prior to implementing similar technologies in an authentic classroom environment.

Usability issues across conditions. Across conditions, there were some common complaints about the general operation of the simulation. The primary complaint was the repetition of the activities. In both conditions, students were required to reset the simulation in between investigations, so as to start with a fresh set of variables and to ensure that one investigation would not unintentionally impact the ones that followed. Students from TUI and GUI groups became frustrated with the repetition of having to hit the run and reveal button every time a new investigation began. A student from low performing TUI group G57 remarked, "we got confused with the buttons…we kept having to reset it after everything we did, that's one of our problems," in the follow-up interview. Additionally, there was minimal confusion about the play button also functioning as the pause button even though the button label did not change to reflect the change in function. Specifically, the play arrow did not turn to two parallel vertical lines (i.e. the symbol for pause) when the simulation was running. The source of confusion in this case, could easily be addressed through a change in labeling.

Students' perception versus performance. Students' positive perception of the added value of the laboratory in the TUI condition differs markedly from the performance measured by the posttest and delayed posttest measures. For instance, students from group P28 (3 males, 1 female) thought highly of the TUI laboratory activity and perceived that they had learned from doing:

> Gary: It actually like showed if you added temperature how it was cold, the particles was slow, if it was fast—if it was hot then it would speed up. And like if you pushed the spring in, it would make them bounce more and more…so it actually showed everything.
> Mike: I would say that it definitely helped because when I'm taking notes, I feel I'm more just writing down and not paying attention to what I write. And then when I do this, I kinda learn it myself and it teaches me and I see it all happening so it's easy for me to relate that to the actual notes

> Geena: Yeah, this was certainly more memorable than seeing something up on a screen…I'll definitely, like, when I think about like trying to remember it on a test or something like this, this is more useful in helping to remember that, like actually seeing what's going on than seeing words up on a page explaining it to me.
> Garrett: I don't know about everyone else, but I'm like a more visual person—so when we're writing notes about gas laws and stuff now I can actually think about what we're doing with this and I think about like pressure change, volume change, heat, whatever—it gave us a visual, now it all makes sense

In reviewing the assessment performance from students in this group, the test scores continue to drop following the pretest through the delayed posttest, offering no significant indication of an increase in conceptual understanding or in retention over time.

Likewise, in the low performing TUI group G57, students ranked the technology very high—8 or 9 out of 10. One male student remarked, "I'd probably give it a 9 because I'm a very visual learner and I love hands on activities cause this actually helped me learn a lot." He later elaborated on the intuitive nature of this technology:

> …you see this as a spring (pointing to screen), you know just to push the spring in (pointing to physical input) and it will move the spring (pointing to screen)…when it's actually interactive like this, it just makes more sense.

One female in the same group indicated that she is "better when you can touch something and not press a button" and another female indicated that the technology was helpful:

> I thought it was helpful that you could see what was going on with all the molecules, like how they would change speed and direction after like coming into contact with something, so it was helpful for me. I'm also a visual, hands-on learner. It was good.

However, when reviewing the scoring trajectories for this group, the positive perception of these technologies exceeded students' performance in terms of conceptual understanding, as measured by KI. Students from other groups echoed that they favored hands on engagement, but some also indicated that while it was more fun, they thought they could learn whether the information was presented as it was in the GUI or with a hands-on component.

High performing groups tended to be more critical of the technological features or lack of features (e.g., buttons, slider size, etc.) than students in low performing groups. Overall, the low performing groups tended to express critical thoughts about technology, in general, more frequently than high performing groups. For example, a male student from the high performing GUI group O37 remarked,

> [the] simulation was kind of bland...it was hard to find like the things to adjust it...so like the un and the stop buttons were like really small on like the bottom of the screen and like the sliders were small...it wasn't visually stimulating.

While another male student in P38, a high performing GUI group, pointed out that, "the mechanics were very hard to work with in the simulation," referring to the slider on the end of the piston. This student also found value in having both types of experiences—visual and hands on—when asked which he thought he would like better, he responded:

> I like both…because on a computer there is [sic] aspects besides being able to see particles which are more helpful…you can change color coding to see temperature, energy, all that stuff. You have little displays that you can interact with but you have much more, a greater degree of control [in hands on].

A female student in the high performing TUI group G62 really liked the technology:

> I liked that it kind of gave us a tangible way to actually investigate the gas laws…I like labs that are more hands on better, but obviously you can't do that in this case to investigate all of these different properties at once. I liked the way that that [the Frame] showed it to us all in one.

Although she indicated that she had a positive perception of the device, she thought that, "it would be nice if there was a running average of everything [in the display]."

Low performing students in either condition appeared to put greater emphasis on technical issues or on voicing a general aversion to technology in the follow-up feedback. For example, a female student in a low performing GUI group (O77) that experienced some technical issues explained, "I want to make a mistake because I made a mistake, not because the computer made a mistake." She also admitted in the post-laboratory interview that she was not impressed with computers and if given a choice, she did not want to use them.

Like students in low performing GUI groups, some students in the low performing TUI group, G57, seemed to prefer more hands-on activities, as one female explained:

> I'm better when you can like touch something and not press a button. It just makes more sense to me. It seems like you're actually doing something real and you're not just like on a computer.

Students' prior experiences with technology may influence their thoughts about each successive piece of technology that they encounter, and this represents a possible avenue for consideration in future work.

Summary of Findings

This chapter features descriptive statistical analysis and inferential analysis, including multilevel modeling. Prior to addressing the research questions, basic descriptive statistics were used to evaluate the study sample participants' demographics relative to the demographics of the school district from which the sample was taken. Students missing test data were incorporated in this study to present more wholistic results.

In the study sample, analysis showed students in each condition (i.e., the TUI or GUI) performed comparably on pretest, posttest, and delayed posttest measures. The first research question examined if students' conceptual understanding differed between the TUI and GUI conditions from pretest to posttest while the second question was concerned with differences in retention over time between conditions. In order to get a more accurate result, teacher influence and variation of students within each class were evaluated using a three-level multilevel model. Using condition as the dependent variable, the model indicated that there is no significant effect of variance on students' KI gain scores from pretest to posttest and no significant effect on retention, as measured by KI score across assessment items from pretest to posttest. To summarize, there was no significant difference detected in this sample regarding student performance between TUI and GUI conditions from the pretest to the posttest and there were no significant differences detected in this sample regarding retention between the groups.

Qualitative analysis was conducted to uncover emergent themes in TUI and GUI conditions not revealed in the quantitative analysis. While TUI and GUI groups performed comparably on the assessments, there were notable distinctions present in the qualitative data. Themes emerged from video and interview data showing that student groups in the TUI condition tended to be mutually cooperative, engage in beneficial social communication, and conduct non-prescribed activities or play outside of the prescribed lab guide. Additionally, students in the low performing TUI groups were more likely to express positive perceptions of the technology and their learning, overestimating their actual performance on posttest and delayed posttest measures over time. Student groups in the GUI conditions tended to have one person lead the activity, limited social communication, and focused on finishing up the lab packet as prescribed. Both groups had usability issues, mostly involving controls and physical inputs; these issues differed depending on the user interface.

Student groups were categorized as either high performing or low performing. Generally, high performing groups improved from pretest to posttest and sustained improvement in the delayed posttest. High performing groups, specifically those in the study sample's TUI condition, were characterized by preplanning and innovative investigations; these groups were more likely to explore the features of the technology. High performing GUI groups in this study sample did not appear to have any defining characteristics aside from assessment scores.

Low performing groups in this study either did not have gains from pretest to posttest or gained from pretest to posttest and did not sustain these gains over time. All low performing groups engaged in play more often than high performing groups, regardless of condition. Aside from usability issues that most groups had throughout the implementation, there were no distinctive features observed in low performing TUI groups in this sample. Low performing GUI groups were all alpha dominant, where one group member took control of the investigation.

CHAPTER 5

DISCUSSION AND IMPLICATIONS

The overall focus of this research study was to examine and compare students' conceptual understanding in a sensor-augmented virtual laboratory with a virtual only condition to determine if any significant learning advantages were measured in either condition. More specifically, the research questions sought to investigate (1) any differences in conceptual understanding of gas properties between students using the Frame (i.e., sensor-augmented virtual laboratory) and students using only virtual labs and (2) whether these differences, if any, are evident in terms of conceptual retention over time, and finally, (3) to identify what characterizes students' interactions with both types of technologies and to compare these interactions qualitatively to supplement the quantitative data.

One outcome of this study, in terms of conceptual understanding as measured by knowledge integration, is that there was no difference in students' performance across conditions in this study sample. This is an important finding because it highlights that different types of technologies have the potential for success in a classroom environment. Neither technology was better for the improvement of conceptual understanding than the other, indicating that hands-on and the visualization-only condition each have the potential for assisting students in learning about gas laws. Perhaps more notable is that the lack of difference in this sample does not necessarily indicate that there was no effect. Equally as important is the idea that neither condition was found to be detrimental to student learning

Results are consistent with existing research comparing physical and virtual approaches in science education (e.g., Chiu & Linn, 2014; Kozma & Russell, 1997; Levy & Wilensky, 2009, Wu, Krajcik, & Soloway, 2001). Specifically, these findings align with research involving physical and virtual components in laboratory settings (e.g., Triona & Klahr, 2003; Klahr, Triona, & Williams, 2007; Zacharia & Constantinou, 2008). Klahr, Triona, and Williams (2007) found no difference between 7th and 8th students learning with physical manipulatives and those in a virtual condition with respect to engineering design. Students in Klahr et al.'s (2007) study built and tested mousetrap cars to see whose car could travel the farthest distance. The study involved the type of instruction and categorized students across four different conditions, largely determined by whether students had physical or virtual learning opportunities. Using several different measures, there was no significant difference in students' performance across conditions in that study sample.

Likewise, Zacharia and Constantinou's (2007) work examined whether physical and virtual manipulatives were more influential in developing undergraduate students' (n=68; 15 male, 53 female) conceptual understanding of heat and temperature. Heat and temperature, like the scientific content involved in this study, the properties of gases, elicit common alternative conceptions that do not align with the scientifically normative understanding of associated phenomena. Zacharia and Constantinou concluded that each method of experimentation, whether virtual or physical, was equally as successful in developing students' conceptual understanding. The current study corresponds with the idea that students can learn scientific concepts equally as well, regardless of the medium used for scientific investigation. However, the current study extends these findings through the comparison of a connected laboratory, where physical components are integrated with a virtual simulation (i.e., NUI) and compared to a visualization only condition (i.e., GUI).

A more recent study by Chung, Cheng, Lai, and Tsai (2014) compared high school students' performance on a simulation-based laboratory (SBL) versus performance on a microcomputer-based laboratory (MBL). Unlike prior studies that compare purely physical to purely virtual conditions, Chung et al.'s approach is very similar to the one taken in the current study. Students in the MBL condition manipulate physical components and also use handheld computers. Students in the SBL condition are limited to scientific investigations using virtual simulation. Like the current study, Chung et al. investigates secondary students' conceptual understanding of gases (i.e., Boyle's Law). The researchers concluded that there is no significant difference in terms of students' conceptual understanding. Similarly, the researchers discovered that students in the MBL might be more engaged in inquiry practices than those in the SBL condition. The findings of this study support Chung et al.'s conclusions about the outcomes of connecting physical and virtual manipulatives.

When using only the visualization, students' construction of conceptual understandings connecting molecular and macroscopic levels is just as good as the augmented approach. Each of the students' explanations was evaluated using the same KI rubric, the same raters, and identical codebooks corresponding to each question to maintain consistency across conditions. That students can construct such explanations speaks to the learning benefit of visualizations (e.g., Honey & Hilton, 2011; Bell & Trundle, 2008; Dori & Belcher, 2005; Jaakola, Nurmi, & Veermans, 2011; Korakakis, Boudouvis, Palyvos, & Pavlatou, 2012; Lee, Linn, Varma, & Liu, 2010; Zacharia & Anderson, 2003; Zhang & Linn, 2011). Findings suggest that visualizations may be the main component to help students make connections to observable levels, and underscore the idea that carefully crafted instructional guidance with visualizations assists students in making connections between visualizations and observable scientific phenomena (Chiu & Linn, 2014). Tangible augmentation may not be needed to help students achieve this learning objective. As the tangible augmentation may require more time and cost to set up, this study provides guidance for instructors trying to find effective and efficient instructional methods to help chemistry students.

Unlike existing mixed-reality technology research (e.g., Johnson-Glenberg et al., 2009, 2011; Lindgren & Moshell, 2011; Novellis & Moher, 2011), this study involves an implementation of an augmented virtual approach that takes place in authentic high school classrooms. Several researchers have examined various aspects of mixed-reality, primarily through instantiations of augmented reality. These studies are primarily proof-of-concept with small numbers of students participating in design-based research (e.g., Billinghurst & Duenser, 2012, Di Serio, Ibanez, & Kloos, 2013). Very few studies investigate augmented virtual approaches and even fewer compare augmented virtual approaches to other technologies in real classrooms. For instance, recent mixed-reality technology research focuses on the applications of immersive virtual worlds to students' learning of science concepts where physical, embodied interaction is scaffolded and studied (i.e., Johnson-Glenberg et al., 2009; Lindgren & Moshell, 2011). These immersive virtual worlds are difficult to scale and to facilitate within most traditional schools. I argue that, while beneficial, this type of implementation seems unlikely to become reality in public schools lacking in funding or direct university partnerships. This study takes place in standard classrooms, specifically in a standard science laboratory setting, using equipment most secondary science teachers can access with relative ease (e.g., scientific probeware, input technologies, laptops), which makes these findings especially relevant to current pedagogical practices and existing limitations in public school science classrooms.

Additionally, there were no statistically significant differences in this study sample's performance by gender across conditions. Simulations have been shown to increase content knowledge for low performing female students (Sadler, Romine, Stuart, & Merle-Johnson, 2013); however, students regardless of gender and performance classification performed comparably in each condition investigated in this study.

One drawback to these findings involves the type of assessment used. That there are no detectable differences between conditions in this study sample could indicate that the sensitivity of the instrument is less than ideal or that the sample size needs to be larger to differentiate signal from noise. Better assessment measures may get at the root of the effect in ways that the existing assessment measure may not. What we know from this particular study, is that it is not that augmenting the virtual lab experience does not help students' conceptual understanding; instead, students are learning just as much from the augmented version with the Frame as from the visualization. Students are still able to develop explanations and make key connections between the molecular and macroscopic levels that undergird the properties of gases.

However, qualitative data suggest that there may be differences on other important aspects of science such as planning and conducting investigations. These and other scientific practices are emphasized as one of three key components to the Next Generation Science Standards. As practices are considered to be the combination of knowledge and skill (NGSS Lead States, 2013 p. 30), this study concentrated on connecting levels of phenomena throughout the course of a scientific investigation, as that is what the technology seemed to afford. The results point to a need for developing more specific assessments aligned to both the knowledge and the skill comprising a particular scientific practice. For example, assessments that incorporate the process of investigation, not simply just the output of students' thinking may have led to a better understanding of the role that technology may play as students fully engage in that process. The data from this study indicates that the development of this type of assessment may help us drill down to the influential aspects of the technologies employed.

In response to the third research question, groups that were high performing tended to improve from pretest to posttest with retention of conceptual understanding (i.e., students retained the established molecular-macroscopic connections over time) whereas low performing students either did not establish those connections in the assessments or if they did, those connections were not retained. There were no observed distinctive characteristics of low performing TUI groups and high performing GUI groups in this sample, aside from assessment scores.

Low performing groups in this study sample tended to engage in play more often than high performing groups. This could be due perhaps (1) procrastination or (2) exploration. Students who typically struggle or those who are classified as low performing may exhibit less interest in such work and it is possible that the students engaged in play because they lacked interest in completing the assigned task. It is also possible that students engaged in play because they were legitimately interested in the capabilities of the technology and wanted to test it because they had the opportunity to do so. It does not appear that play had a negative effect on performance, though future research that is more focused on this aspect of technological influence is necessary.

Students in the low-performing GUI groups tended to be alpha dominant where a male took charge of the laboratory and maintained that control, often unchallenged. This behavior typically represented a lopsided group effort where students either appeared to withdraw from learning or were less interested in what was going on. While this type of behavior cannot be directly attributed to the technology due to other variables not considered in this study, it is interesting to note that this categorization only took place in the GUI groups. It is possible that the lack of opportunities for students to directly interact with the technology resulted in this dynamic. For example, the Frame had three additional ways to interact with the simulation as well as the touch screen interface. This affords more students to have direct interaction with the simulation, instead of one person driving the simulation, which is what typically happened in the GUI condition where there was one input device (i.e., the mouse). Providing students with additional points of interaction with a technology may add to the cohesion of the group dynamic and minimize the alpha dominant relationship that developed in all but one GUI group.

Likewise, TUI groups were classified as mutually cooperative, because they took turns engaging with the simulation and playing around with the external inputs (i.e., syringe, spring, and temperature sensor). These behaviors may have been present had the students been assigned to the other condition, simply based on personality and perhaps prior relationships. Students in mutually cooperative groups interacted more with each other, which facilitated positive social communication. These students were involved in content-related discussions and idea sharing with fellow members. This differed from alpha dominant groups where one member's expressed thoughts dominated the group's collective answers.

In general, group dynamics seemed to influence the types of social communication that were observed. The alpha dominant groups tended to lack effective social communication, such as idea sharing, as they approached the lab investigations. These groups were more reliant on fact checking and dominated by the unchallenged opinion of one member. The opposite was observed in a majority of the mutually cooperative groups. Groups who worked well together were more likely to share ideas and communicate using the visualization as an evidentiary tool to support argumentation and ideas. The mutually cooperative groups were also more likely than alpha dominant groups to deviate from the prescribed laboratory activity in ways that appeared to enrich the experience of engaging in scientific practices, as were observed in student-technology interactions. Allowing for students to have different experiences and roles throughout the duration of the laboratory may have contributed to their ability to engage with the content material and to work in conjunction with other people in their group.

High performing groups exhibited preplanning and innovation in investigative activities. This was specific to the TUI condition. Groups in the GUI condition, both low and high performing, were not observed as employing investigative methods other than those prescribed in the packet. It appears that the increased opportunity for physical manipulation in the TUI condition may have invited the innovative behaviors consisting of non-prescriptive methods of investigation, such as combining water from two jars to take an additional measurement for the temperature and volume investigation. Students in the TUI groups seemed more curious and more open to learning, as opposed to the focus on completion. While it is possible that students in the TUI condition were caught up in the novelty of the technology and it is difficult to attribute these observed behaviors solely to the assigned condition in either case, it seems likely that the TUI afforded students a laboratory experience that by the nature of its construction and presentation allowed students more freedom of choice than a visualization driven by mouse-clicking alone.

All groups in this study sample, regardless of performance or assigned condition, experienced some usability issues. These usability issues were the result of (1) technological novelty, (2) interface design, and (3) system malfunction. The technological novelty, particularly in the TUI condition, created confusion for many students. Even after a thorough explanation was provided, it seems plausible that this explanation included too much information for students to remember. The interaction with the TUI, while intuitive in many ways, was also unfamiliar, at least initially, for students, and this could have led to some delays in completing the laboratory.

The interface design posed other challenges for students in both groups. This was particularly true in the GUI groups where the size of slider bars and the functionality of the add molecules feature were the biggest issues. Some of the TUI groups struggled with the correspondence between the physical input and the dynamic visualization. This representational incongruence made it difficult for students to carry out the lab efficiently and to understand the actual connection between the physical input and the phenomena being observed. Finally, system malfunctions were few and far between, but invariably they occurred in several of the implementations, where the simulations had to be reset due to an odd molecule escaping from the chamber or due to the display of negative or zero pressure.

Usability issues that required the simulation to be reset or that affected data collection may have translated into some students' negative perceptions of technology and the effect on their learning. On the other hand, the novelty may have contributed to students' positive perceptions in at least one TUI group, which reported having a great time with the technology. This group indicated that the engagement with the external inputs really added to their experience and understanding; however, on this self-reported increase in understanding was not evident based on assessment scores over time. In this case, the students actually performed worse than they indicated. This could be a potential drawback for new technologies implemented in the classroom. Students may actually like it and want to affect how laboratories (or other activities) are taught in the classroom to the extent that they are willing to report learning benefits when they weren't actually measured.

Implications

This study has the potential to influence the research trajectory in science education and technology. Innovative technologies are changing the way we think about classroom learning. Currently, the most similar technology-related research in science education focuses on augmented reality and is typically not situated in an authentic classroom environment.

The findings of this study disrupt and confirm some conclusions from prior research studies. As mentioned in the literature review, there are mixed results from the implementations of physical and virtual laboratories in science education. As Olympiou and Zacharia (2012) note, the sequencing of physical and virtual experiences in science can be more meaningful for learning concepts than either in isolation. Blikstein and Wilensky (2009) noted that connecting these different mediums may improve learning. Along those lines, this particular study investigated whether the direct, integrated combination of physical and virtual components was beneficial for students to improve conceptual understandings of gas properties. While there was no statistically significant difference between students in this study sample based on condition, the observed behaviors are worth a closer look.

The relationship of bodily action to the cognitive acquisition of knowledge is an important aspect of embodied cognition (Barsalou, 2008), and the representations offered in this comparative need further investigation to tease out what aspects of the physical components may be related to a change in student behaviors and interactions. This appears to be the case of other content areas like aeronautical engineering, cognitive acquisition of required skills and concepts seems to be done largely through hands-on simulation and preparation. This underscores two ideas: (a) visualizations are almost essential to achieving complex understandings and (b) in some cases, physical experiences form the foundation for content learning. It is possible that an integration of physical and virtual components could help students make sense of and understand phenomena they cannot see in settings that they have not yet experienced. The relationship between physical action and cognition may have a substantial, yet wholly untapped influence on students' conceptual understanding in science, and this should definitely be a consideration for future research.

This study has broader implications in classroom practice, regarding teaching and learning in high schools, as the results indicate that students are capable of learning equally well regardless of the type of technology that is being used as a medium of investigation.

Limitations and Future Research

There are several limitations to this study worthy of consideration in light of the findings. First, the study sample is limited to a demographic that does not necessarily represent the United States high school population at large. Though it is not reasonable to make generalizations about technology in science education and apply that to a broader population using only this sample, it should be noted that the analyzed sample had complete data with regard to all assessments without differing from the excluded population.

Additionally, it should be noted that the findings in this study inadequately address what it is about the physical interaction that accounts for the observed differences in students' behaviors in the TUI groups. It seems likely that another study focusing on the physical components would be helpful in narrowing down the aspects of students' interactions and performance that can be tied to the physical manipulation of external inputs. Following that, another comparative study may be needed to compare certain pieces of the physical interaction to analogous interactions in a virtual only condition.

Regarding interfaces this study assumes that the graphical user interface is a common resource for students. Throughout this study I encountered several students with an aversion to technology and it turns out that some 'digital natives' may lack a basic understanding of computing technology. Although this specific population seems well equipped to manage the tasks within the laboratory using the technologies, it is quite possible that the GUI is not as ubiquitous in students' everyday lives as assumed.

Another possibility is that the specific kinds of interfaces presented in the GUI condition were not as intuitive as students' everyday experiences with technologies. Specific kinds of controls (e.g. sliders) were chosen to mirror the interface of the Frame in a GUI format, but may not have been the easiest to manipulate. Other kinds of controls for the GUI condition may lead to different student perceptions and uses.

The measures used to collect data in this case were pretests, posttests, and delayed posttests. Because they were identical and not widely spaced, it is possible that students could have become familiar enough with the content to influence a response. Furthermore, the instrument was designed to measure conceptual understanding of gas properties based on common misconceptions in this content area. It is quite possible that some other learning took place that the assessments did not capture. The qualitative findings demonstrate that students may have engaged in different practices during the labs, which again points to the need for future studies to use assessments that capture other kinds of learning such as understanding of scientific processes or practices.

In addition, students with missing data (n=68) were incorporated into the analysis. Missing data included students who completed 50% or fewer questions on an individual assessment; this data also included students who may have been absent for any subsequent assessments beyond the pretest. Students missing responses were assigned a KI score of 0, which may have skewed the results by drowning out any detectable signal. This data is important to include because it reflects the reality of working in an authentic environment. However, it is possible that careful scheduling (i.e., not near holidays, breaks, or weekends) in some cases may help decrease student attrition.

The qualitative aspect of this study hinges on self-report, specifically, the follow up interview. Due to a power dynamic between the researcher (i.e., an adult) and a student in a school setting, it is possible that students were not completely honest because they felt an obligation to tell an adult what they thought should be said. While I do not believe this happened all that much, it is possible that the position of the researcher/interviewer influenced what students said in the follow up interview.

Finally, it came to my attention that one teacher's students were privy to information that let them know that there were two different conditions. In the follow up interview with one of the GUI groups, one female expressed that she wished they had gotten the other version, noting that the teacher told her there were two. The extent that this teacher discussed the different treatments with the participating classes is unknown, and it is not clear what effects this had, if any, on the outcomes. Future studies should be more explicit to participating teachers on how a research study is conducted without compromising the fidelity of the experiment.

More information about students' experiences with technology, the relationships that they have with assigned workgroup peers may further inform the observed group dynamics and social communication behaviors which could shed light on how students respond to novel or standard types of technology. Additionally, further examination of the relationship of teachers and student peers on student performance is necessary. Using teachers and student peers to predict student performance on assessment measures by condition is certainly more influential if a link can be established to determine the effect of teachers and student peers on student performance. However, establishing this broader connection is outside of the scope of this study and may be considered as a future step in this line of research.

REFERENCES

AAAS Project 2061 (n.d.) [Pilot and field test data collected between 2006 and 2010]. Unpublished raw data.

Abrahamson, D., Gutiérrez, J. F., Charoenying, T., Negrete, A. G., & Bumbacher, E. (2012). Fostering hooks and shifts: tutorial tactics for guided mathematical discovery. Technology, Knowledge, and learning, 17(1-2), 61-86. DOI: 10.1007/s10758-012-9192-7.

Abelson, H., Dirks, L., Johnson, R., Koedinger, K. R., Linn, M. C., Lynch, C. A., ... & Szalay, A. (2008). Fostering learning in the networked world: The cyberlearning opportunity and challenge, a 21st century agenda for the National Science Foundation. *NSF Report*.

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, *16*(3), 183-198.

Ainsworth, S., & VanLabeke, N. (2004). Multiple forms of dynamic representation. *Learning and Instruction*, *14*(3), 241-255.

Anastopoulou, S., Sharples, M., & Baber, C. (2011). An evaluation of multimodal interactions with technology while learning science concepts. *British Journal of Educational Technology*, *42*(2), 266-290.

Anderson, M. L. (2003). Embodied cognition: A field guide. *Artificial Intelligence*, *149*(1), 91-130.

Antle, A. N., & Wise, A. F. (2013). Getting down to details: Using theories of cognition and learning to inform Tangible User Interface design. Interacting with Computers, 25(1), 1–20. doi:10.1093/iwc/iws007

Antle, A. N., Droumeva, M., & Ha, D. (2009, June). Hands on what?: comparing children's mouse-based and tangible-based interaction. In *Proceedings of the 8th International Conference on Interaction Design and Children* (pp. 80-88). ACM.

Ardac, D., & Akaygun, S. (2004). Effectiveness of multimedia‐based instruction that emphasizes molecular representations on students' understanding of chemical change. *Journal of Research in Science Teaching*, *41*(4), 317-337.

Arvanitis, T. N., Petrou, A., Knight, J. F., Savas, S., Sotiriou, S., Gargalakos, M., & Gialouri, E. (2009). Human factors and qualitative pedagogical evaluation of a mobile augmented reality system for science education used by learners with physical disabilities. *Personal and ubiquitous computing*, *13*(3), 243-250.

Bailey, C. A. (2007). *A guide to qualitative field research*. Pine Forge Press.

Barab, S., & Dede, C. (2007). Games and immersive participatory simulations for science education: an emerging type of curricula. *Journal of Science Education and Technology*, *16*(1), 1-3.

Barsalou, L. W. (2008). Grounded cognition. Annual Review of Psychology, 59, 617–45. doi:10.1146/annurev.psych.59.103006.093639

Bell, R. L., & Trundle, K. C. (2008). The use of a computer simulation to promote scientific conceptions of moon phases. *Journal of Research in Science Teaching*, *45*(3), 346-372.

Ben-Zvi, R., Eylon, B-S, & Silberstein, J. (1986). Is an atom of copper malleable? *Journal of Chemical Education,* 63, 64-66.

Billinghurst, M., Kato, H., & Poupyrev, I. (2001). The MagicBook: a transitional AR interface. *Computers & Graphics*, *25*(5), 745-753.

Birchfield, D., & Megowan-Romanowicz, C. (2009). Earth science learning in SMALLab: A design experiment for mixed reality. *International Journal of Computer-Supported Collaborative Learning*, *4*(4), 403-421.

Birchfield, D., Thornburg, H., Megowan-Romanowicz, M. C., Hatton, S., Mechtley, B., Dolgov, I., & Burleson, W. (2008). Embodiment, multimodality, and composition: convergent themes across HCI and education for mixed-reality learning environments. *Advances in Human-Computer Interaction*, *2008*.

Bivall, P., Ainsworth, S., & Tibell, L. A. E. (2011). Do haptic representations help complex molecular learning? Science Education, 95(4), 700–719. doi:10.1002/sce.20439

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. *Psychology and the real world: Essays illustrating fundamental contributions to society*, 56-64.

Blikstein, P., Fuhrmann, T., Greene, D., & Salehi, S. (2012). Bifocal modeling: mixing real and virtual labs for advanced science learning. In *Proceedings of the 11th International Conference on Interaction Design and Children* (pp. 296-299). ACM.

Blikstein, P., & Wilensky, U. (2007). Bifocal modeling: a framework for combining computer modeling, robotics and real-world sensing. *Annual meeting of the American Research Education Association*. Chicago, IL.

Bransford, J., Brown, A. L., & Cocking, R. R. (2000). *How people learn*. Washington, DC: National Academy Press.

Chang, H. Y., Quintana, C., & Krajcik, J. S. (2010). The impact of designing and evaluating molecular animations on how well middle school students understand the particulate nature of matter. *Science Education*, *94*(1), 73-94.

Chao, J., Chiu, J.L., Hazzard, E., & Xie, C. (2013). Proceedings from AERA '13: Effects of Mixed Reality Labs on Students' Conceptual Understanding, Cognitive Load, and Epistemology of Science. San Francisco, CA.

Chao, J., Chiu, J.L., Pan, E., DeJaegher, C.J., Hazzard, E., & Xie, C. (2014). Proceedings from AERA '14: The Effects of Mixed-Reality Laboratories on High School Students' Conceptual Understanding of Gas Laws. Philadelphia, PA.

Charen, G. (1970). Do laboratory methods stimulate critical thinking?. *Science Education*, *54*(3), 267-271.

Cheng, K. H., & Tsai, C. C. (2013). Affordances of augmented reality in science learning: suggestions for future research. *Journal of Science Education and Technology*, *22*(4), 449-462.

Chung, S., Chang, W. H., Lai, C. H., & Tsai, C. Y. (2014). A Comparison of Students' Approaches to Inquiry, Conceptual Learning, and Attitudes in Simulation-Based and Microcomputer-Based Laboratories. *Science Education*,*98*(5), 905-935.

Chiu, J. L., & Linn, M. C. (2014). Supporting Knowledge Integration in Chemistry with a Visualization-Enhanced Inquiry Unit. *Journal of Science Education and Technology*, *23*(1), 37-58.

Clark, D. B., & Jorde, D. (2004). Helping students revise disruptive experientially-supported ideas about thermodynamics: Computer visualizations and tactile models. *Journal of Research in Science Teaching, 41*(1), 1–23.

Clark, D. B. (2006). Longitudinal conceptual change in students' understanding of thermal equilibrium: An examination of the process of conceptual restructuring.

Cognition and Instruction, 24(4), 467–563.

Corbin, J., & Strauss, A. (Eds.). (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Sage.

Corliss, S. B., & Spitulnik, M. W. (2008, June). Student and teacher regulation of learning in technology-enhanced science instruction. In *proceedings of the 8th international conference on International conference for the learning sciences-Volume 1* (pp. 167-174). International Society of the Learning Sciences.

Cracolice, M.S., & Monteyne, K. (2004). "What's wrong with cookbooks? A reply to Ault." *Journal of Chemical Education,* 81(11), 1559.

Creswell, J. W. (2011). Educational research: Planning, conducting, and evaluating quantitative and qualitative research. Boston, Mass: Allyn & Bacon.

de Jong, T., Linn, M. C., & Zacharia, Z. C. (2013). Physical and Virtual Laboratories in Science and Engineering Education. *Science*, *340*(6130), 305-308.

Dede, C. (1995). The evolution of constructivist learning environments: Immersion in distributed, virtual worlds. *Educational technology*, *35*(5), 46-52.

DeJaegher, C. J. (2014). Proceedings from AERA '14: Effects of Mixed Reality Labs on Middle School Students' Conceptual Understanding. Philadelphia, PA.

DeJaegher, C.J., Chiu, J.L., &Chao, J. (2014) The effects of augmented virtual science laboratories on middle school students' conceptual understanding of gases. Manuscript in preparation.

De Jong, T., & Van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of educational research*, *68*(2), 179-201.

Dori, J.Y., & Belcher, J. (2005). How does technology-enabled active learning affect undergraduate students' understanding of electromagnetism concepts? *Journal of the Learning Sciences*, 14, 243-279.

Dunleavy, M., Dede, C., & Mitchell, R. (2009). Affordances and limitations of immersive participatory augmented reality simulations for teaching and learning. *Journal of Science Education and Technology*, *18*(1), 7-22.

Erikson, R.D. (1990). "Working with interface metaphors." In Laurel, B. (Ed.), *The Art of Human-Computer Interface Design* (pp. 65-73). Reading, MA: Addison-Wesley.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175-191.

Feisel, L.D., & Rosa, A.J. (2005). The role of the laboratory in undergraduate engineering education. *Journal of Engineering Education*, 94, 121-130.

Finkelstein, N., Adams, W., Keller, C., Kohl, P., Perkins, K., Podolefsky, N., Reid, S., et al. (2005). When learning about the real world is better done virtually: A study of substituting computer simulations for laboratory equipment. *Physical Review Special Topics - Physics Education Research*, *1*(1), 1–8.

Fleishman, H.L., Hopstock, P.J., Pelczar, M.P., and Shelley, B.E. (2010). *Highlights from PISA 2009: Performance of U.S. 15-year-old students in reading, mathematics, and science literacy in an international context* (NCES 2011-004). U.S. Department of Education, National Center for Education Statistics. Washington, D.C.: U.S. Government Printing Office.

Gabel, D. (1999). Improving teaching and learning through chemistry education research: A look to the future. *Journal of Chemical Education*, 76 (4), 548.

Gentner, D., & Nielsen, J. (1996). The anti-mac interface. *Communications of the ACM*, *39*(8), 70-82.

Gire, E., Carmichael, A., Chini, J.J., Rouinfar, A., Rebello, S., Smith, G., et al. (2010) The effects of physical and virtual manipulatives on students' conceptual learning about pulleys. In K. Geomez, L. Lyons, & J. Radinsky (Eds.) Learning in the disciplines: Proceedings of the 9th international conference onf the learning sciences (ICLS 2010) (Vol. 1, pp. 937-944). Chicago International Scoeity of the Learning Sciences.

Gredler, M. E. (2004). Games and simulations and their relationships to learning. *Handbook of research on educational communications and technology*, *2*, 571-581.

Han, I. (2013). Embodiment: A New Perspective for Evaluating Physicality in Learning. *Journal of Educational Computing Research*, *49*(1), 41-59.

Han, I., & Black, J. B. (2011). Incorporating haptic feedback in simulation for learning physics. Computers & Education, 57(4), 2281–2290. doi:10.1016/j.compedu.2011.06.012

Hauk, O., & Tschentscher, N. (2013). The body of evidence: what can neuroscience tell us about embodied semantics?. *Frontiers in Psychology*, *4*.

Hillman, D.C., Willis, D.J., & Gunawardena, C.N. (1994). Learner-interface interaction in distance education: An extension of contemporary models and strategies for practitioners. *The American Jounral of Distance Education,* 8(2), 30-42.

Hinckley, K. & Wigdor, D. (2012). Chapter 6: Input technologies and techniques. In J.A. Jacko (Ed.), *The human-computer interaction handbook: Fundamentals, evolving* *technologies, and emerging applications* (3rd Ed., pp. 95-132). New York: CRC Press.

Hmelo-Silver, C. E. (2006). Design principles for scaffolding technology-based inquiry. *Collaborative learning, reasoning, and technology*, 147-170.

Hmelo, C. E., & Guzdial, M. (1996, July). Of black and glass boxes: Scaffolding for doing and learning. In *Proceedings of the 1996 international conference on Learning sciences* (pp. 128-134). International Society of the Learning Sciences.

Höffler, T. N., & Leutner, D. (2007). Instructional animation versus static pictures: A meta-analysis. *Learning and instruction*, *17*(6), 722-738.

Hofstein, A., & Lunetta, V. N. (2004). The laboratory in science education: Foundations for the twenty-first century. *Science education*, *88*(1), 28-54.

Honey, M. A., & Hilton, M. L. (Eds.). (2011). *Learning science through computer games and simulations.* Washington, DC: National Academy Press.

Horn, M. S., Solovey, E. T., Crouser, R. J., & Jacob, R. J. (2009, April). Comparing the use of tangible and graphical programming languages for informal science education. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 975-984). ACM.

Hourcade, J. P. (2008). Interaction design and children. *Foundations and Trends in Human-Computer Interaction*, *1*(4), 277-392.

Howison, M., Trninic, D., Reinholz, D., & Abrahamson, D. (2011, May). The Mathematical Imagery Trainer: from embodied interaction to conceptual learning. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1989-1998). ACM.

Ibáñez, M. B., Di Serio, Á., Villarán, D., & Delgado Kloos, C. (2014). Experimenting with electromagnetism using augmented reality: Impact on flow student experience and educational effectiveness. *Computers & Education*, *71*, 1-13.

Inkpen, K. (1997, June). Three important research agendas for educational multimedia: Learning, children, and gender. In *AACE World Conference on Educational Multimedia and Hypermedia* (Vol. 97, pp. 521-526).

Ishii, H. U. (2008, February). Tangible bits: beyond pixels. In *Proceedings of the 2nd international conference on Tangible and embedded interaction* (pp. xv-xxv). ACM.

Ishii, H. U., & Ulmer, B. B. (1997). Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms. In *Proceedings of CHI 97 Conference on Human Factors in Computing systems* (pp. 22-27).

Ivankova, N. V., Creswell, J. W., & Stick, S. L. (2006). Using mixed-methods sequential explanatory design: From theory to practice. *Field Methods*, *18*(1), 3-20.

Jaakkola, T., Nurmi, S., & Veermans, K. (2011). A comparison of students' conceptual understanding of electric circuits in simulation only and simulation-laboratory contexts. *Journal of Research in Science Teaching*, 48, 1, 71-93.

Johnson-Glenberg, M. C., Birchfield, D. A., Tolentino, L., & Koziupa, T. (2014). Collaborative embodied learning in mixed reality motion-capture environments: Two science studies. *Journal of Educational Psychology*, *106*(1), 86.

Johnson-Glenberg, M. C., Koziupa, T., Birchfield, D., & Li, K. (2011, June). Games for learning in embodied mixed-reality environments: principles and results. In *Proceedings of the 7th international conference on Games+ Learning+ Society*

Conference (pp. 129-137). ETC Press.Johnstone, A. H. (1991). Why is science difficult to learn? Things are seldom what they seem. Journal of Computer Assisted Learning, 7(2), 75–83.

Kaufmann, H., & Dünser, A. (2007). Summary of usability evaluations of an educational augmented reality application. In *Virtual Reality* (pp. 660-669). Springer Berlin Heidelberg.

Kerawalla, L., Luckin, R., Seljeflot, S., & Woolard, A. (2006). “Making it real”: exploring the potential of augmented reality for teaching primary school science. *Virtual Reality*, *10*(3-4), 163-174.

Kim, M. C., & Hannafin, M. J. (2011). Scaffolding problem solving in technology-enhanced learning environments (TELEs): Bridging research and theory with practice. *Computers & Education*, *56*(2), 403-417.

Klahr, D., Triona, L. M., & Williams, C. (2007). Hands on what? The relative effectiveness of physical versus virtual materials in an engineering design project by middle school children. *Journal of Research in Science Teaching*, *44*(1), 183-203.

Klahr, D., Zimmerman, C., & Jirout, J. (2011). Educational interventions to advance children’s scientific thinking. *Science*, *333*(6045), 971-975.Klahr,

Korakakis, G., Boudouvis, A., Palyvos, J., & Pavlatou, E. A. (2012). The impact of 3D visualization types in instructional multimedia applications for teaching science. *Procedia-Social and Behavioral Sciences*, *31*, 145-149.

Kozma, R. B. (2000). The use of multiple representations and the social construction of understanding in chemistry. *Innovations in science and mathematics education:*

Advanced designs for technologies of learning, 11-46.

Kozma, R. B., & Russell, J. (1997). Multimedia and understanding: Expert and novice responses to different representations of chemical phenomena. *Journal of research in science teaching*, *34*(9), 949-968.

Kreijns, K., Kirschner, P. A., & Jochems, W. (2003). Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: a review of the research. *Computers in human behavior*, *19*(3), 335-353.

Lakoff, G., & Johnson, M. (1980). Metaphors we live by. Chicago: University of Chicago Press.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought*. Basic books.

Larkin, M. J. (2002). *Using scaffolded instruction to optimize learning*. Arlington, VA: ERIC Clearinghouse on Disabilities and Gifted Education.

Lee, H. S., Linn, M. C., Varma, K., & Liu, O. L. (2010). How do technology ‐ enhanced inquiry science units impact classroom learning?. *Journal of Research in Science Teaching*, *47*(1), 71-90.

Lee, H., Plass, J. L., & Homer, B. D. (2006). Optimizing cognitive load for learning from computer-based science simulations. *Journal of educational psychology*, *98*(4), 902.

Levy, S. T., & Wilensky, U. (2009). Students' learning with the Connected Chemistry (CC1) curriculum: navigating the complexities of the particulate world. *Journal of Science Education and Technology*, *18*(3), 243-254.

Levy, S. T., Novak, M., & Wilensky, U. (2006). Students' foraging through the complexities of the particulate world: Scaffolding for independent inquiry in the connected chemistry (MAC) curriculum. In *Annual Meeting of the American Educational Research Association, San Francisco, CA*.

Lindgren, R., & Johnson-Glenberg, M. (2013). Emboldened by Embodiment Six Precepts for Research on Embodied Learning and Mixed Reality. *Educational Researcher*, *42*(8), 445-452.

Lindgren, R., & Moshell, J. M. (2011, June). Supporting children's learning with body-based metaphors in a mixed reality environment. In *Proceedings of the 10th International Conference on Interaction Design and Children* (pp. 177-180). ACM.

Linn, M. C., & Eylon, B. (2006). Science education. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of Educational Psychology* (2nd ed.) (pp.511-544). Mahwah, NJ: Lawrence Erlbaum Associates.

Linn, M. C. & Eylon, B.-S. (2011). *Science Learning and Instruction: Taking Advantage of Technology to Promote Knowledge Integration.* New York: Routledge.

Linn, M. C. (2006). The knowledge integration perspective on learning and instruction. In R. K. Sawyer (Ed.), Cambridge handbook of the learning sciences (pp. 265–286). Cambridge: Cambridge University Press.

Linn, M. C., Chang, H. Chiu, J. L., Zhang, Z. H., & McElhaney, K. (2010). Can desirable difficulties overcome deceptive clarity in scientific visualizations? In A. Benjamin (Ed.), *Successful remembering and successful forgetting: a Festschrift in honor of Robert A. Bjork* (pp. 239-262). New York: Routledge.

Liu, O. L., Lee, H. S., Hofstetter, C., & Linn, M. C. (2008). Assessing knowledge integration in science: Construct, measures, and evidence. *Educational Assessment*, *13*(1), 33-55.

Liu, X. (2006). Effects of combined hands-on laboratory and computer modeling on student learning of gas laws: A quasi-experimental study. *Journal of Science Education and Technology*, *15*(1), 89–100. doi:10.1007/s10956-006-0359-7

Liu, W., Cheok, A. D., Mei-Ling, C. L., & Theng, Y. L. (2007, September). Mixed reality classroom: learning from entertainment. In *Proceedings of the 2nd international conference on Digital interactive media in entertainment and arts* (pp. 65-72). ACM.Lowe, R. (2004). Interrogation of a dynamic visualization during learning. *Learning and instruction*, *14*(3), 257-274.

Lowe, R. (2004). Interrogation of a dynamic visualization during learning. *Learning and instruction*, *14*(3), 257-274.

MacFarlane, S., Sim, G., & Horton, M. (2005, June). Assessing usability and fun in educational software. In *Proceedings of the 2005 conference on Interaction design and children* (pp. 103-109). ACM.

Marshall, P. (2007, February). Do tangible interfaces enhance learning?. In *Proceedings of the 1st international conference on tangible and embedded interaction* (pp. 163-170). ACM.

Marshall, P. Cheng, P.C.-H., and Luckin, R. (2010). Tangibles in the balance: a discovery learning task with physical or graphical materials. In *Proceedings of the 4th international conference on tangible, embedded, and embodied interaction* (pp. 153-160). ACM.

Martín-Gutiérrez, J., Contero, M., & Alcañiz, M. (2010, January). Evaluating the usability of an augmented reality based educational application. In *Intelligent tutoring systems* (pp. 296-306). Springer Berlin Heidelberg.

McBride, D. L., Murphy, S., & Zollman, D. A. (2010). Student understanding of the correlation between hands-on activities and computer visualizations of NMR/MRI. *AIP Conference Proceedings* (Vol. 225, pp. 225–228).

McElhaney, K. W., & Linn, M. C. (2011). Investigations of a complex, realistic task: Intentional, unsystematic, and exhaustive experimenters. *Journal of Research in Science Teaching*, *48*(7), 745-770.

Milgram, P., & Kishino, F. (1994). A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems*, *77*(12), 1321-1329.

Minstrell, J., & Kraus, P. (2005). Guided inquiry in the science classroom. *How people learn: History, mathematics, and science in the classroom*, 475-12.

Mostmans, L., Vleugels, C., & Bannier, S. (2012). Raise Your Hands or Hands-on? The Role of Computer-Supported Collaborative Learning in Stimulating Intercreativity in Education. *Educational Technology & Society*, *15*(4), 104-113.

Mullet, K., & Sano, D. (1995). *Designing visual interfaces: Communication oriented techniques* (Vol. 2550). Englewood Cliffs NJ: SunSoft Press.

n.a. (n.d.). *The Kinetic Molecular Theory*. Retrieved from http://chemed.chem.purdue.edu/genchem/topicreview/bp/ch4/kinetic4.html

National Center for Education Statistics. (2012). The nation's report card: Science 2011 (NCES 2012-465). Institute of Eduation Sciences, U.S. Department of Education: Washington, D.C.

National Research Council (NRC). (2010). Rising above the gathering storm, revisited: Rapidly approaching Category 5. Washington, DC: The National Academies Press.

National Research Council (NRC). (2005). *America's lab report: Investigations in high school science*. Washington, DC: The National Academies Press.

Nelson, T. H. (1990). The right way to think about software design. *The Art of Human-Computer Interface Design, Addison-Wesley*, 235-243.

NGSS Lead States. (2013). *Next generation science standards: For states, by state*s. Washington, DC: The National Academies Press.

Norman, D. A. (1988; 1993). *The design of everyday things*. New York, NY: Basic Books.

Northrup, P. (2001). A framework for designing interactivity into web-based instruction. *Educational Technology. Englewood Cliffs, NJ*, *41*(2), 31-39.

Novellis, F., & Moher, T. (2011). How real is "real enough"? Designing artifacts and procedures for embodied simulations of science practices. In Proceedings of the 10th International Conference on Interaction Design and Children - IDC '11 (pp. 90–98).

Novick, S., & Nussbaum, J. (1981). Pupils' understanding of the particulate nature of matter: A cross‐age study. *Science education*, *65*(2), 187-196.

National Research Council. (2011). A framework for K-12 science education: Practices, crosscutting concepts, and core ideas. Washington, D.C.: The National Academies Press.

Olsson, T., Kärkkäinen, T., Lagerstam, E., & Ventä-Olkkonen, L. (2012). User evaluation of mobile augmented reality scenarios. *Journal of Ambient Intelligence and Smart Environments*, *4*(1), 29-47.

Olympiou, G., & Zacharia, Z. (2012) Blending physical and virtual manipulatives: An effort to improve students' conceptual understanding through science laboratory experimentation. *Science Education*, 96(1), 21-47. doi: 10.1002/sce.20463

Oviatt, S., Arthur, A., & Cohen, J. (2006, October). Quiet interfaces that help students think. In *Proceedings of the 19th annual ACM symposium on User interface software and technology* (pp. 191-200). ACM.

Pillat, R., Nagendran, A., & Lindgren, R. (2012, December). A mixed reality system for teaching STEM content using embodied learning and whole-body metaphors. In *Proceedings of the 11th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry* (pp. 295-302). ACM.

Pillay, S. (2010). *Sequencing instruction in chemistry: simulation or traditional first?* (Doctoral dissertation, University of British Columbia).

Pouw, W. T., van Gog, T., & Paas, F. (2014). An embedded and embodied cognition review of instructional manipulatives. *Educational Psychology Review*, *26*(1), 51-72.

Price, S., & Rogers, Y. (2004). Let's get physical: the learning benefits of interacting in digitally augmented physical spaces. *Computers & Education*, *43*(1), 137-151.

Price, S., Sheridan, J. G., Falcão, T. P., & Roussos, G. (2008). Towards a framework for investigating tangible environments for learning. International Journal of Arts and Technology, 1(3/4), 351. doi:10.1504/IJART.2008.022367

Provasnik, S., Kastberg, D., Ferraro, D., Lemanski, N., Roey, S., and Jenkins, F. (2012). Highlights from TIMSS 2011: Mathematics and science achievement of U.S. fourth- and eighth- grade students in an international context (NCES 2013-009 Revised). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, D.C.

Romeo, G., Edwards, S., McNamara, S., Walker, I., & Ziguras, C. (2003). Touching the screen: Issues related to the use of touchscreen technology in early childhood education. *British Journal of Educational Technology*, *34*(3), 329-339.

Rosenbaum, E., Klopfer, E., & Perry, J. (2007). On location learning: Authentic applied science with networked augmented realities. *Journal of Science Education and Technology*, *16*(1), 31-45.

Rossman, G. B., & Rallis, S. F. (2011). *Learning in the field: An introduction to qualitative research*. Sage publications.

Ryan, G. W., & Bernard, H. R. (2003). Techniques to identify themes. *Field methods*, *15*(1), 85-109

Sadler, T. D., Romine, W. L., Stuart, P. E., & Merle-Johnson, D. (2013). Game-Based Curricula in Biology Classes: Differential Effects Among Varying Academic Levels. *Journal of Research in Science Teaching*, *50*(4), 479-499.

Sadler, P. M., & Tai, R. H. (2001). Success in introductory college physics: The role of high school preparation. *Science Education*, *85*(2), 111-136.

Schneider, K. G. (1996). Children and information visualization technologies. *Interactions*, *3*(5), 68-73.

Sims, R. (1997). Interactivity: A forgotten art?. *Computers in Human Behavior*, *13*(2), 157-180.

Sipitakiat, A., Blikstein, P., & Cavallo, D. P. (2004, June). GoGo board: augmenting programmable bricks for economically challenged audiences. In *Proceedings of the 6th international conference on Learning sciences* (pp. 481-488). International Society of the Learning Sciences.

Smith, G. W., & Puntambekar, S. (2010). Examining the combination of physical and virtual experiments in an inquiry science classroom. In *Proceedings of the Conference on Computer Based Learning in Science, Warsaw, Poland*.

Squire, K. D., & Jan, M. (2007). Mad City Mystery: Developing scientific argumentation skills with a place-based augmented reality game on handheld computers. *Journal of Science Education and Technology*, *16*(1), 5-29.

Squire, K., & Klopfer, E. (2007). Augmented reality simulations on handheld computers. *The Journal of the Learning Sciences*, *16*(3), 371-413.

Srinivasan, M. A., & Basdogan, C. (1997). Haptics in virtual environments: Taxonomy, research status, and challenges. *Computers & Graphics*, *21*(4), 393-404.

Theng, Y. L., Mei-Ling, C. L., Liu, W., & Cheok, A. D. (2007). Mixed reality systems for learning: a pilot study understanding user perceptions and acceptance. In *Virtual Reality* (pp. 728-737). Springer Berlin Heidelberg.

Tobin, K. (1990). Research on science laboratory activities: In pursuit of better questions and answers to improve learning. *School Science and Mathematics*, *90*(5), 403-418.

Tolentino, L., Birchfield, D., Megowan-Romanowicz, M. C., Johnson-Glenberg, M. C., Kelliher, A., & Martinez, C. (2009). Teaching and learning in the mixed reality science classroom. Journal of Science Education and Technology, 18(6), 501–517. doi:10.1007/s10956-009-9166-2

Trundle, K. C., Atwood, R. K., Christopher, J. E., & Sackes, M. (2010). The effect of guided inquiry-based instruction on middle school students' understanding of lunar concepts. *Research in Science Education*, *40*(3), 451-478.

van der Meij, J., & de Jong, T. (2006). Supporting students' learning with multiple representations in a dynamic simulation-based learning environment. *Learning and Instruction*, *16*(3), 199-212.

Varelas, M., Pieper, L., Arsenault, A., Pappas, C. C., & Keblawe-Shamah, N. (2014). How science texts and hands-on explorations facilitate meaning making: Learning from Latina/o third graders. *Journal of Research in Science Teaching*, *51*(10), 1246-1274.

Wagner, E. D. (2005, June). Interaction strategies and experience design: Guidelines for technology-mediated learning. In *World Conference on Educational Multimedia, Hypermedia and Telecommunications* (Vol. 2005, No. 1, pp. 4118-4119).

Wagner, E. D. (2006). On designing interaction experiences for the next generation of blended learning. *The Handbook of Blended Learning: Global perspectives, local designs*, 41-55.

Weiser, M. (1991). The computer for the 21st century. *Scientific American*, *265*(3), 94-104.

Wilensky, U. (1999). NetLogo (and NetLogo User Manual), Center for Connected Learning and Computer Based Modeling, Northwestern University. http://ccl.northwestern.edu/netlogo/

Wilensky, U., & Resnick, M. (1999). Thinking in levels: A dynamic systems approach to making sense of the world. *Journal of Science Education and Technology*, *8*(1), 3-19.

Windschitl, M., & Andre, T. (1998). Using computer simulations to enhance conceptual change: The roles of constructivist instruction and student epistemological beliefs. *Journal of Research in Science Teaching*, 35, 2, 145-60.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, *9*(4), 625-636.

Winn, W., Stahr, F., Sarason, C., Fruland, R., Oppenheimer, P., & Lee, Y.-L. (2006) Learning oceanography from a computer simulation compared with direct experience at sea. *Journal of Research in Science Teaching*, 43, 25-42.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving*. *Journal of Child Psychology and Psychiatry*, *17*(2), 89-100.

Wu, H. K., Krajcik, J. S., & Soloway, E. (2001). Promoting understanding of chemical representations: Students' use of a visualization tool in the classroom. *Journal of research in science teaching*, *38*(7), 821-842.

Yannier, N., Koedinger, K. R., & Hudson, S. E. (2013, January). Tangible Collaborative Learning with a Mixed-Reality Game: EarthShake. In *Artificial Intelligence in Education* (pp. 131-140). Springer Berlin Heidelberg.

Yin, R. K. (2009). *Case study research: Design and methods* (Vol. 5). Sage Publications.

Zacharia, Z.C., & Anderson, O. R. (2003). The effects of an interactive computer-based simulation prior to performing a laboratory inquiry-based experiment on students' conceptual understanding of physics. *American Journal of Physics*, *71*(6), 618-629.

Zacharia, Z. C., & Constantinou, C. P. (2008a). Comparing the influence of physical and virtual manipulatives in the context of the Physics by Inquiry curriculum: The case of undergraduate students' conceptual understanding of heat and temperature. *American Journal of Physics*, *76*(4), 425-430.

Zacharia, Z. C., & Olympiou, G. (2011). Physical versus virtual manipulative experimentation in physics learning. *Learning and Instruction*, 21(3), 317-331.

Zacharia, Z. C., Olympiou, G., & Papaevripidou, M. (2008b). Effects of experimenting with physical and virtual manipulatives on students' conceptual understanding in heat and temperature. *Journal of Research in Science Teaching*, *45*(9), 1021-1035.

Zhang, Z. H., & Linn, M. C. (2011). Can generating representations enhance learning with dynamic visualizations? *Journal of Research in Science Teaching*, *48*(10), 1177-1198.

Zhang, Z. H., & Linn, M. C. (2013). Learning from Chemical Visualizations: Comparing generation and selection. *International Journal of Science Education*, *35*(13), 2174-2197.

Znaniecki, F. (1934). *The method of sociology* (p. 12). New York: Farrar & Rinehart.

Zuckerman, O., Arida, S., & Resnick, M. (2005, April). Extending tangible interfaces for education: digital montessori-inspired manipulatives. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 859-868). ACM.

APPENDIX A: GAS LAB CURRICULUM (TUI)

Name______________________________Teacher____________________ Block ______

Gas Lab with the Frame (TUI Version)

This lab combines real objects with a computer simulation of gas molecules. Your task is to connect these two levels – the everyday and the tiny – into an understanding of the behavior of gases. The "Frame" is like having super-magnifying glasses. You can observe and manipulate simulations of very, very small gas molecules that you wouldn't be able to see and interact with otherwise. Make sure that you have a Frame set up at your lab station. Check boxes have been provided for you throughout the steps to help you monitor your progress.

Behaviors of gas molecules
In this activity, you will investigate the behaviors of gas molecules.

Motion
☐ Double-tap the TUI icon on the desktop. The model shows a chamber with a piston filled with gas. The gas molecules are initially invisible. (Note: If the simulation is running, you may skip this step).
☐ Tap the Reveal button to show the gas molecules represented by small green dots.
☐ Tap the Run button to run the simulation.
Please describe the motion of the gas molecules.

____
____
____

Kinetic energy and speed
☐ Tap the Energy button to show colors that represent the kinetic energy of the molecules: red for high energy, pink for medium energy, and white for low energy.
☐ Tap the Velocity button to show arrows that represent the speed and direction of the gas molecules. Longer arrows mean faster speed.
☐ Tap the Run button again to pause the simulation.
How is the color of the gas molecules related to the length of the arrows? Explain why they have this relationship.

____
____
____

☐ Tap the Run button again to resume the simulation.
Are all gas molecules moving at the same speed and have the same amount of kinetic energy? Explain why or why not.

____
____
____

Collisions

☐ Tap the Trace button to highlight one gas molecule.
When and why does the highlighted gas molecule change direction?

____
____
____

☐ The horizontal arrows along the piston wall represent the impulses of collisions between the gas molecules and the piston wall. Longer arrows represent greater collision impulses (or more forceful collisions).
Why do these horizontal arrows have different lengths?

____
____
____

☐ Turn off these visual displays by tapping the trace, energy, and velocity buttons again.

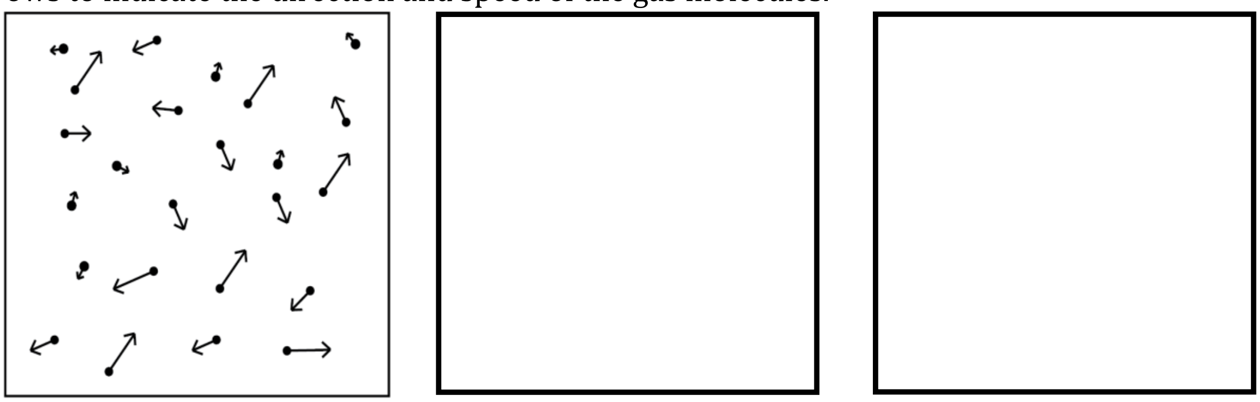
☐ Tap the Reset button to reset the simulation.

Temperature

In this activity, you will investigate the nature of gas temperature. Temperature (T) is a measure of the average kinetic energy of the molecules. An individual molecule's kinetic energy (KE) is proportional to its mass (m) and its velocity squared (v2). The formula for kinetic energy is:

KE=(1/2)mv2

Reset the simulation and DON'T run it yet. Answer this question first and you will have a chance to revise your answer later. The picture below shows gas molecules at room temperature. The length and the direction of the arrows represent molecules' speed and direction respectively. Draw how you think the molecules would look in a cold room and a hot room, assuming that the walls of the rooms are fixed. Include arrows to indicate the direction and speed of the gas molecules.



Room temperature   Cold room   Hot room

Run the simulation and reveal the molecules. Hold the piston in place using the spring to keep volume as close to constant as possible. Change the temperature of the Frame by touching hot or cold jars next to the temperature sensor (ask teacher for hot water and ice cubes). What happens to the average kinetic energy and speed of the molecules when you increase or decrease temperature? Why do you think that happens? (Press the Energy and/or Velocity buttons to help you visualize)

____

____

____

When you decrease the temperature, do all the molecules (Circle One)
spread out more.
gather together in the middle.
fill the chamber the same as before.
Explain your observation:

____

____

Based on your observation, would you change your drawing of molecules in a hot or cold room? ☐YES ☐NO

Explain why or why not and provide new drawings if needed.

Cold room

Hot room

Pressure

In this activity, you will investigate the nature of gas pressure. When a molecule collides with the piston wall, it exerts a force on it. The total force on the piston wall is the sum of forces exerted by individual molecules. Pressure (P) is defined as the force (F) on a given surface (A).

$P=F/A$

There is a constant atmospheric pressure outside the chamber (on the right side of the piston). Observe the gas molecules inside the chamber. How do the gas molecules inside the chamber exert pressure to counterbalance the atmospheric pressure?

____

____

Why is the piston wall fluctuating around a certain position?

____

____

Why don't you observe such fluctuation in real life (e.g., , membrane of an air balloon)?

____

____

Reset the frame. Look where the piston begins and check the number of molecules in the system. Pump 100 gas molecules into the chamber by pressing in the attached syringe. Push the piston back to its original position by pressing (exerting a force on) the attached spring. Feel the increased force on your hand. What happens to the collisions on the piston wall that makes the pressure increase?

____

____

Relationship between pressure and volume

In this activity, you will investigate the relationship between volume and pressure of a certain amount of gas at constant temperature. Make a hypothesis before conducting the experiment:
When you decrease the volume of the gas, the pressure…

____

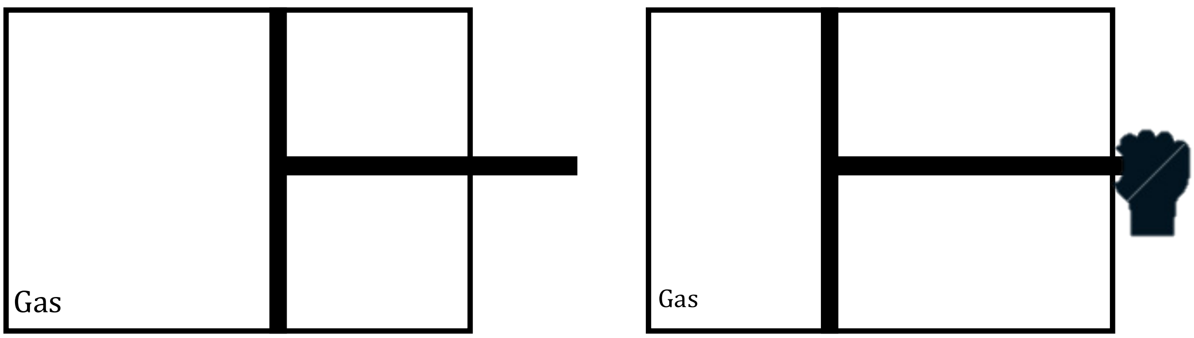
Explain your hypothesis based on what you know about the behaviors of the gas molecules. Draw on the following diagrams to help you explain.

____

____

____

____



Before pushing the piston   After pushing the piston

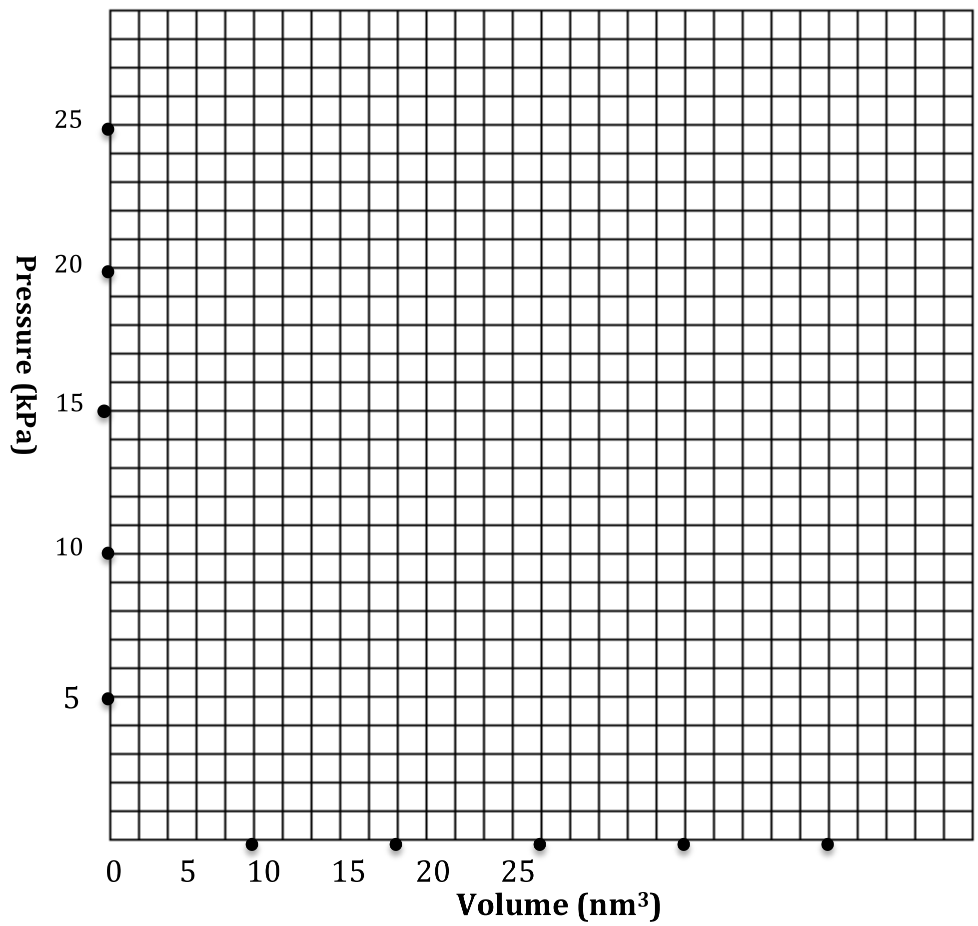
Experiment design and procedure
Reset and run the frame.
The table below lists out the control variables (CV), independent variable (IV), and dependent variable (DV) for this experiment. Follow the procedure and collect your data.

Design	Variable	Procedure
Control	Molecular mass	1. Use the same type of molecules throughout the experiment.
Control	Number of molecules	2. Do not add or remove molecules during the experiment.
Control	Temperature	3. Remove any heat or cold source away from the temperature sensor and wait until the temperature stabilizes.
Independent Variable	Volume	4. Press or depress the piston to five different positions that are relatively spread out. For each new position, wait for at least 10 seconds for the piston to stabilize. Record the volume in the table below.
Dependent Variable	Pressure	5. Record the pressure readout for each volume. You may pause the simulation (tap the Run button) to help you read the number.

Record your data below.

Trial #	Volume	Pressure
1		
2		
3		
4		
5		

Display your data in a line graph below.



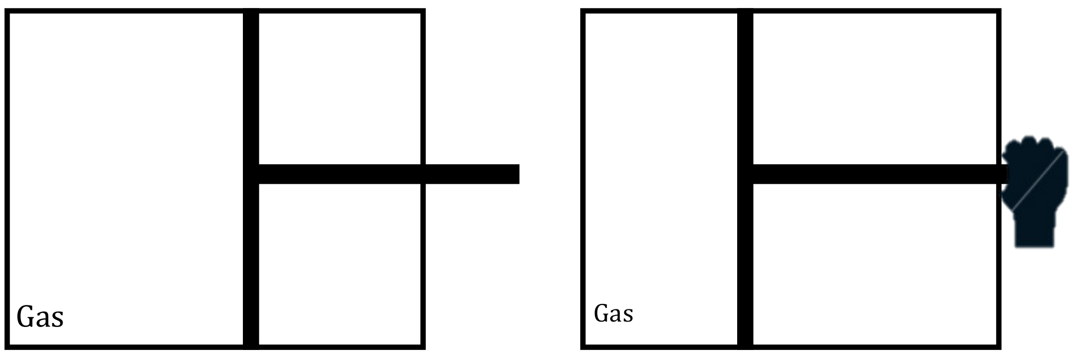
Do your data support your hypothesis? What is your conclusion about the relationship between volume and pressure?

____

Change or refine your explanation for the relationship between volume and pressure based on the behaviors of gas molecules. Draw on the following diagram to help you explain.

____

____

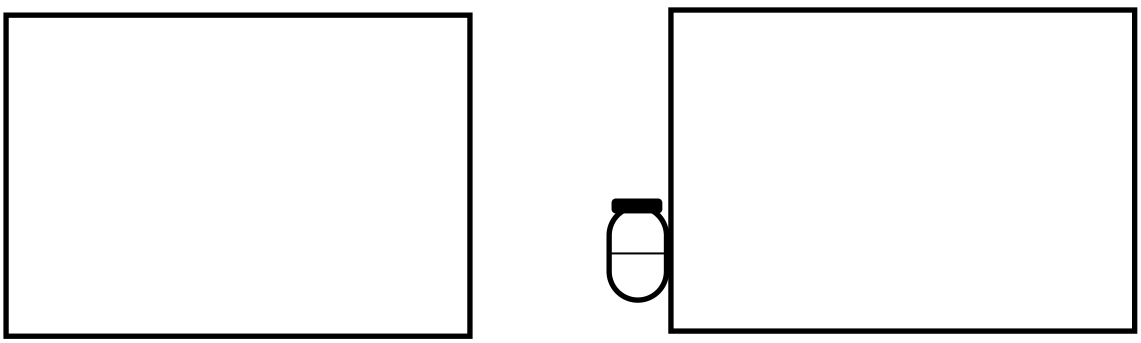


Before pushing the piston

After pushing the piston

Relationship between temperature and volume
In this activity, you will investigate the relationship between temperature and volume of a certain amount of gas at the same volume. Form a hypothesis before conducting the experiment.
When you increase the temperature of the gas, the volume will:

Explain your hypothesis based on what you know about the behaviors of the gas molecules. Draw to help you explain.



Before being heated

After being heated

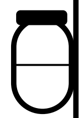
Pay close attention to the where the piston is before changing the temperature in the chamber. Use hot and cold jars to increase and decrease the temperature in the chamber and observe what happens. Be sure to let the temperature stabilize between each change.

When temperature increased, what happened to the volume of the gas?

As the temperature increased, what happened to the speed of the molecules and the length of the arrows when the molecules hit the container walls?

Change or refine your explanation for the relationship between temperature and volume based on the behavior of gas molecules. Draw in the diagrams below to help you explain.

Before being heated



After being heated

APPENDIX B: GAS LAB CURRICULUM (GUI)

Name________________________Teacher________________________ Block ______

Gas Lab with the Frame (GUI Version)

This lab combines uses computer simulation of gas molecules. Your task is to connect the everyday and the microscopic into an understanding of the behavior of gases. You can observe and manipulate simulations of very, very small gas molecules that you wouldn't be able to see and interact with otherwise. Make sure that you have a simulation set up at your lab station. Check boxes have been provided for you throughout the steps to help you monitor your progress.

Behaviors of gas molecules
In this activity, you will investigate the behaviors of gas molecules.
Motion

☐ If your machine is not currently running a simulation, please ask for instructions, otherwise proceed to the next step.

☐ The model shows a chamber with a piston filled with gas. The gas molecules are initially invisible. Click the Reveal button to show the gas molecules represented by small green dots.

☐ Click the Run button to run the simulation.

Using complete sentences, ldescribe the motion of the gas molecules.

____

____

Kinetic energy and speed

☐ Click the Energy button to show colors that represent the kinetic energy of the molecules: red for high energy, pink for medium energy, and white for low energy.

☐ Click the Velocity button to show arrows that represent the speed and direction of the gas molecules. Longer arrows mean faster speed.

☐ Click the Run button again to pause the simulation.

What is the relationship between the color of the gas molecules and the length of the arrows? Why do you think this relationship exists? (Make sure to answer both questions)

____

____

____

☐ Click the Run button once more to resume the simulation.

Are all gas molecules moving at the same speed and have the same amount of kinetic energy? Explain why or why not.

____

____

____

Collisions

☐ Click the Trace button to highlight one gas molecule.
When and why does the highlighted gas molecule change direction?

____
____
____

☐ The horizontal arrows along the piston wall represent the impulses of collisions between the gas molecules and the piston wall. Longer arrows represent greater collision impulses (or more forceful collisions).
Why do these horizontal arrows have different lengths?

____
____
____

☐ Turn off these visual displays by clicking the trace, energy, and velocity buttons again.

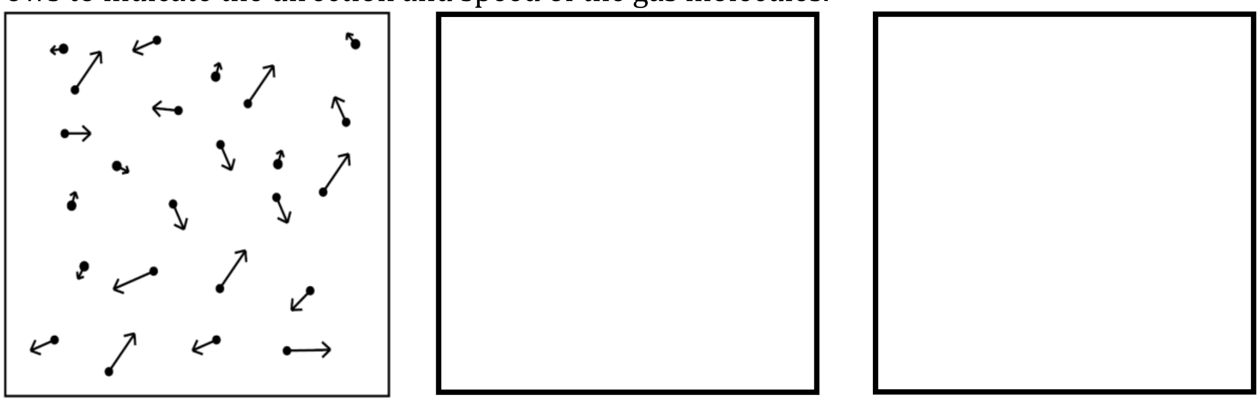
☐ Click the Reset button to reset the simulation.

Temperature

In this activity, you will investigate the nature of gas temperature. Temperature (T) is a measure of the average kinetic energy of the molecules. An individual molecule's kinetic energy (KE) is proportional its mass (m) and its velocity squared (v2). The formula for kinetic energy is:

KE=(1/2)mv2

Reset the simulation and DON'T run it yet. Answer this question first and you will have a chance to revise your answer later. The picture below shows gas molecules at room temperature. The length and the direction of the arrows represent molecules' speed and direction respectively. Draw how you think the molecules would look in a cold room and a hot room, assuming that the walls of the rooms are fixed. Include arrows to indicate the direction and speed of the gas molecules.



Room temperature   Cold room   Hot room

Run the simulation and reveal the molecules. Hold the piston in place using the force slider to keep the volume as close to constant as possible. Change the temperature of the Frame by adjusting the temperature slider to the left of the chamber. What happens to the average kinetic energy and speed of the molecules when you increase or decrease temperature? Why do you think that happens? (Press the Energy and/or Velocity buttons to help you visualize)

____

____

____

When you decrease the temperature, do all the molecules (circle one):
spread out more.
gather together in the middle.
fill the chamber the same as before.
Explain your observation:

____

____

Based on your observation, would you change your drawing of molecules in a hotter or colder room? ☐YES ☐NO

Explain why or why not, and provide new drawings if needed.

____

____

____

____

____

____

Cold room

Hot room

Pressure

In this activity, you will investigate the nature of gas pressure. When a molecule collides with the piston wall, it exerts a force on it. The total force on the piston wall is the sum of forces exerted by individual molecules. Pressure (P) is defined as the force (F) on a given surface (A).

P=F/A

There is a constant atmospheric pressure outside the chamber (on the right side of the piston). Observe the gas molecules inside the chamber. How do the gas molecules inside the chamber exert pressure to counterbalance the atmospheric pressure?

Why is the piston wall fluctuating around a certain position?

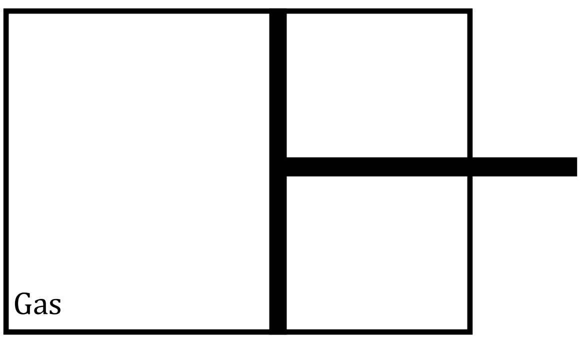
Why don't you observe such fluctuation in real life (e.g., , membrane of an air balloon)?

Reset the frame. Look where the piston begins and check the number of molecules in the system. Put 100 more gas molecules into the chamber by clicking the "Add Molecules" radio button at the bottom of the chamber. When enough molecules are in the chamber, stop adding molecules. Push the piston back to its original position by adjusting the force slider located to the right of the piston. What happens to the collisions on the piston wall that makes the pressure increase?

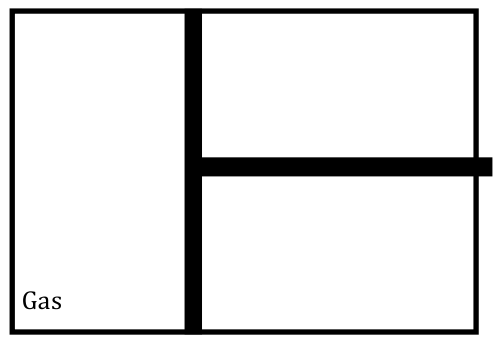
Relationship between pressure and volume

4.1 In this activity, you will investigate the relationship between volume and pressure of a certain amount of gas at constant temperature. Make a hypothesis before conducting the experiment--When you decrease the volume of the gas, the pressure…:

Explain your hypothesis based on what you know about the behaviors of the gas molecules. Draw on the following diagrams to help you explain.



Before moving the piston



After moving the piston

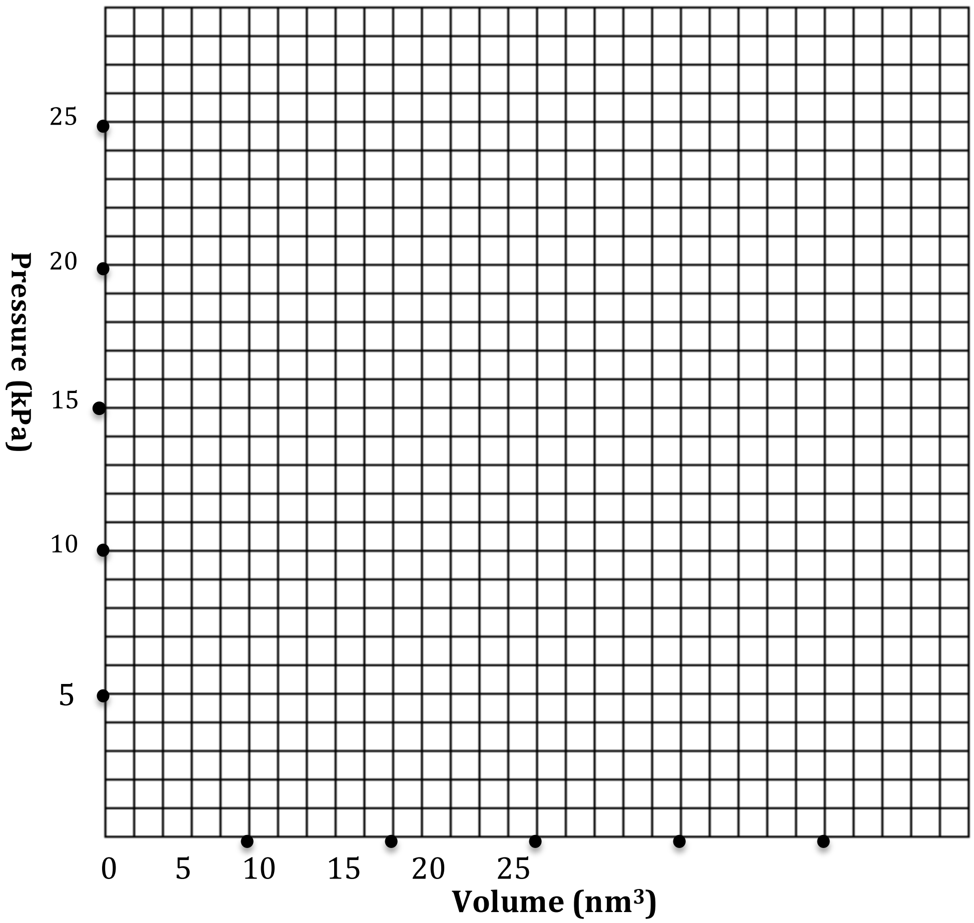
Experiment design and procedure
Reset and run the simulation.
The table below lists out the control variables (CV), independent variable (IV), and dependent variable (DV) for this experiment. Follow the procedure and collect your data.

Design	Variable	Procedure
Control	Molecular mass	1. Use the same type of molecules throughout the experiment.
Control	Number of molecules	2. Do not add or remove molecules during the experiment.
Control	Temperature	3. Using the slider, stabilize the temperature.
Independent Variable	Volume	4. Adjust the force slider to five different positions that are relatively spread out. For each new position, wait for at least 10 seconds for the piston to stabilize. Record the volume in the table below.
Dependent Variable	Pressure	5. Record the pressure readout for each volume. You may pause the simulation (Click the Run button) to help you read the number.

Record your data below.

Trial #	Pressure (kPa)	Volume (nm3)
1		
2		
3		
4		
5		

Display your data in a line graph below.



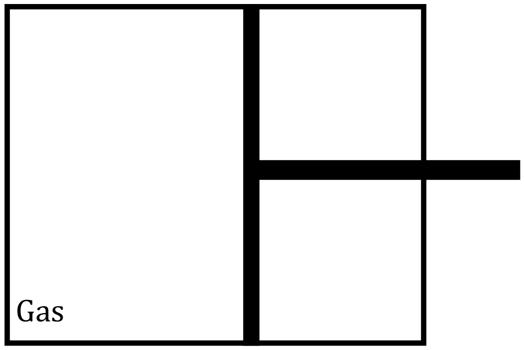
Do your data support your hypothesis? What is your conclusion about the relationship between volume and pressure?

____

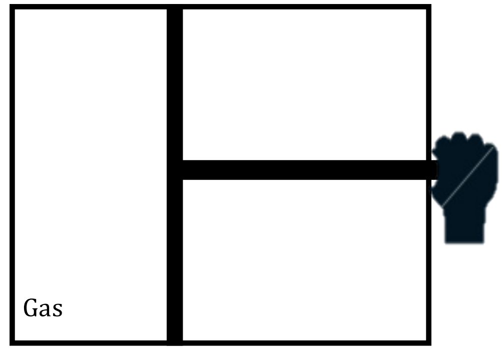
Change or refine your explanation for the relationship between volume and pressure based on the behaviors of gas molecules. Draw on the following diagram to help you explain.

____

____



Before pushing the piston

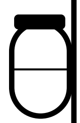


After pushing the piston

Relationship between temperature and volume

In this activity, you will investigate the relationship between temperature and volume of a certain amount of gas at the same volume. Form a hypothesis before conducting the experiment. When you increase the temperature of the gas, the volume will: ______________________________________

Explain your hypothesis based on what you know about the behaviors of the gas molecules. Draw to help you explain.



Before being heated

After being heated

Now click run to help you understand how temperature relates to volume. Using the temperature slider, collect 5 data points with 5 different temperatures. Record both the temperature and the volume in the table below. Be sure to keep the number of molecules constant and do not change the force slider.

Trial #	Temperature (K)	Volume (nm3)
1		
2		
3		
4		
5		

When temperature increased, what happened to the volume of the gas?

____

As the temperature increased, what happened to the speed of the molecules and the length of the arrows when the molecules hit the container walls?

Change or refine your explanation for the relationship between temperature and volume based on the behavior of gas molecules. Draw in the diagrams below to help you explain.

APPENDIX C: INTERVIEW PROTOCOL

MRL Gas Laws
Follow-up Interview Protocol (High School TUI/GUI)

State the purpose of the interview. Introduce yourself.

Remember: try to show strong curiosity about what students think, and make sure they perceive that; make sure those quieter students get chance to speak first.

Liking

1. In general, how do you like the lab activity you did yesterday? 1-10 scale
2. Tell me a couple of things you don't like about it.
3. Tell me a couple of things you liked about it.

TUI Interface Condition

4. Have you ever done any lab activities like this one before? What were they like?
5. You do something around the frame such as putting a hot jar or pressing a spring, then things happen in the computer, what do you think about this kind of interaction with computer? Does it feel natural, or strange, or need some time to get comfortable with?
6. Compare the frame with a computer simulation with button controls, which would you prefer using in classroom? Why?
7. If they prefer TUI, further ask them whether they think using physical controls would impact their understanding of the concepts (not just about engaging or entertaining).

GUI interface condition

8. Have you ever used a simulation like this before? What were they like?
9. Compare this simulation to a hands-on lab, which would you prefer using in the classroom? Why?
10. Ask them to imagine the buttons and scrollbars are replaced with physical controls, give them a concrete example (e.g., , syringe gas adder). Ask them whether they would like that and why.
11. Regarding the simulation, what was the easiest thing about using it?
12. Regarding the simulation, what was the most difficult thing to figure out?

Help learning

13. Do you think the activity helped you understand the properties (temperature/pressure/volume) of gas better?
14. Tell me which parts of the activities helped you most to understand these theory and laws?
15. Tell me which parts of the activities were not very helpful, or irrelevant, or even confusing.

Design change

16. What would you like to change if you were the designers of the activity?

Focus on Interactions (To be determined from video recording)

17. Using video clips from students' experiences with the frame in the classroom, focus on one or two interactions of interest and ask them what they were thinking about when they were interacting with the frame, given the video example.

APPENDIX D: PRE/POST/DELAYED ASSESSMENT

Pre/Post/Delayed Post Assessment
Name____________________ Teacher___________________ Block________ Date___________

Challenge Questions

Which of the following describes the motion of the gas molecules in a sealed container of gas?
a) The gas molecules are not moving.
b) The gas molecules are all moving at the same speed.
c) The gas molecules move only when the container of the gas is moving.
d) The gas molecules are moving at different speeds.

Compared to molecules of cold air, molecules of hot air...
a) are larger in size.
b) have more mass.
c) have more heat molecules mixed with them.
d) move faster on average.

An air mattress can hold weight because the air inside exerts pressure onto the walls of the mattress. Air is made of gas molecules, so how do the gas molecules cause pressure? Please explain.

When you pump your bicycle tires, the air pump feels harder and harder to push as you pump in more air. Please explain what happens to the gas molecules inside the tire and how they make it harder to push.

The picture below shows a syringe with its nozzle covered.

As the plunger is pushed into the syringe, the amount of force required to move the plunger increases. Explain this phenomenon by describing the behaviors of gas molecules.

As the plunger is pulled out of the syringe, the amount of force required to move the plunger also increases. Explain this phenomenon by describing the behaviors of gas molecules.

A scientist filled a steel container with air and sealed it so that no air would leak out. She measured the pressure inside the container in the morning when it was cold and then in the afternoon when it was hot. Which of the following results is most likely true?

The gas pressure in the cold morning was higher than that in the hot afternoon.

The gas pressure in the cold morning was lower than that in the hot afternoon.

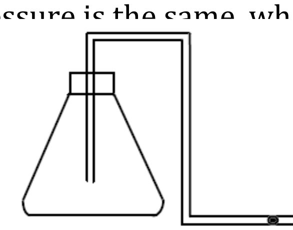
The gas pressure did not change from morning to afternoon.

It is difficult to tell.

Please explain your choice by describing what happened to the gas molecules. Feel free to draw to help you explain.

Morning	Afternoon

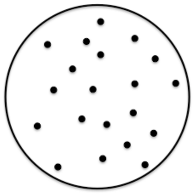
A flask is closed by a stopper connected to a glass tube. The glass tube is sealed by a drop of mercury as shown in the following figure. If we move the whole apparatus from a room with a temperature of 26 °C to an outdoor yard with a temperature of 5 °C, and the indoor air pressure and the outdoor air pressure is the same, what will happen to the mercury drop?



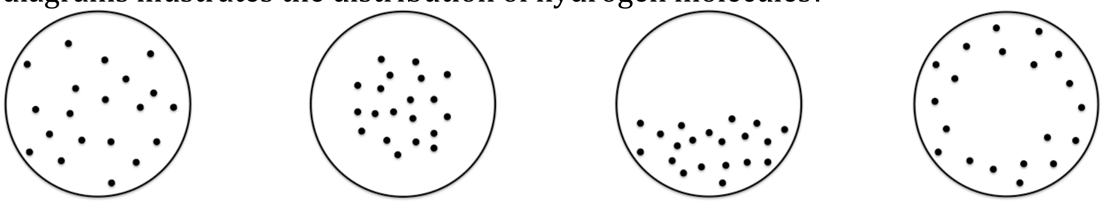
a) The mercury will not move.
b) The mercury will move to the right.
c) The mercury will move to the left.
d) It is impossible to predict.

Explain your prediction:

A steel tank is filled with hydrogen gas at a temperature of 20°C and a pressure of 3 atm. In the following diagram, the circle represents the tank, and the dots represent the distribution of hydrogen molecules.



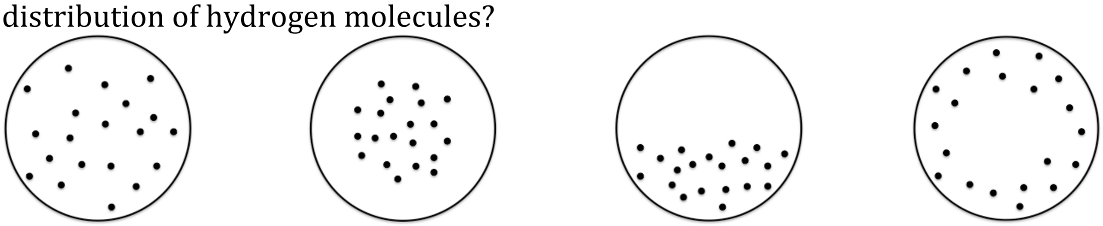
If the tank is cooled to 5°C (hydrogen still remains a gas), which of the following diagrams illustrates the distribution of hydrogen molecules?



a) b) c) d)

Please explain your choice:

If the tank is heated to 50°C, which of the following diagrams illustrates the distribution of hydrogen molecules?



a) b) c) d)

Please explain your choice:

APPENDIX E: QUALITATIVE MEMOS

Analytical Memo – P38 GUI

This GUI workgroup consisted of 4 members: 2 males and 2 females. Three of the group members are seated throughout the duration of the lab, while another one is standing. The young lady who is standing looks at her friend's paper often and rarely interacts with the computer simulation. She interacted with it at the very end, 41 minutes into the intervention.

This group, in general, is very focused on the directions provided in the packet. For example, when one student is running the simulation, another group member calls him out, looking at the paper, saying "we're not supposed to do that". The group members were regulating each others use of the simulation and also regulating themselves in terms of making progress. Although they did not finish the entire laboratory in the class period, it seems like this group of students worked well together and communicated with each other throughout the duration of the intervention without any trouble.

To navigate the simulation, both the mouse and the touchpad were used. The mouse was the preferred method of interaction; students in this group did not appear to notice that the monitor was a touch screen. Throughout the laboratory, there were several points where there appeared to be usability issues with the simulation. Some confusing regarding the controls was present—for example, another student took over the controls because the student who was previously controlling the simulation was having a bit of trouble adjusting the force slider. The force slider is the virtual representation of the spring, which is found in the TUI version. Likewise, there appeared to be some confusion with the use of the temperature slider later on in the lab; students in this group noted early on that the slider was small. When asked to decrease the temperature, one of the students clicked on the slider 9 distinct times (this can be heard and counted in the video) before getting the desired effect.

In the video, the student closest to the camera has a packet which is visible. His hypothesis is that molecules in a cold room bunch up together (almost in a ball) prior to conducting the investigation. Further inspection of the packet is warranted to check this out. As far as that whole activity about the hot/cold room investigation portion of the lab, students in this group had a lively discussion about the molecular distribution that they observed. All students were involved, and they revisited the simulation to try and support their points of view. One student thought that because the chamber size had changed (due to the compression of the piston) that the molecules were "bunching" together more. She was using this logic to support the idea that the hypothesis of molecules bunching together was correct. Another student challenged her, saying that the size of the chamber was irrelevant. Upon further review, it seems as though the phrase "spread out the same as before" which was intended to describe the molecular distribution within the container was interpreted differently by students within this group. One half of the group took this to mean what we had intended, while the others thought that because the chamber size changed due to not being able to control the external pressure effectively (a glitch in the simulation), the molecules did not spread out the same as before. Even after discussing this amongst themselves, referring to the simulation, and consulting a researcher, all students in this group did not seem to reach consensus on the behavior of molecules in a closed chamber when temperature increases or decreases. They agreed to disagree. This is consistent with some of the answers (multiple choice and their corresponding explanations) that seem to be mismatched on the assessment.

This group referenced the simulation components multiple times throughout the intervention. The references appeared to be used primarily in two ways (1) to explain the functionality of simulation controls to other group members and (2) to go back through activities in the packet, usually adjusting variables to double check observations. Towards the end of the lab, students became more distracted, as was evident by increased off-topic conversation and doodling on the packet. This lack of concentration appeared to lead them to not completely follow the directions for the last activity, which included hypothesizing about the relationship of pressure and volume. The drawing that they included in their packet represented what they observed, not what they predicted. One student remarked that they should go back and draw something else in the section instead, to stand for the prediction, commenting that it might throw off the data. Another group member suggested that they'd better not. Further investigation of the packet is needed to see what exactly happened with all of the groups packets, but regardless, the group did not complete the final activity in its entirety.

Analytical Memo – O37 GUI

RQ #3 What differences, if any, are there between students using the Frame and students using virtual labs on students' interactions with the technologies?

The qualitative analysis of purposefully selected cases in this study is intended to shed light upon the types of interactions that students have with technology in virtual and sensor augmented labs. As a classroom teacher, I am particularly interested in knowing how students make use of a technology to facilitate their understanding as well as how technology may not be as helpful as many hope. The selection of two high performing and two low performing groups in each condition was made in order to make comparisons between groups to seek out any existing commonalities or differences which may help or hinder students' performance and conceptual understanding of the content material. In this study, students are using technology in groups to investigate the molecular behaviors of gases in the context of state standardized chemistry curricula provided by their assigned instructors.

The process of video selection involved the comparison of overall KI gain scores from pre to post and from pre to delayed post across groups. The idea behind this decision was to get an idea of students' performance, and to avoid selecting only the "good" students. This selection method allowed me to also select groups who did not fare so well on the assessments. While there are several possibilities that exist for their performance, it is important to understand how groups differing in achievement made use of the technologies provided to them in this lab.

When investigating this video (and all subsequent videos) I observed two types of student interactions (1) interactions with technology and (2) interactions with group members and researchers (if present). I noticed in this particular video that the interactions with technology were minimal. The classroom setting was bright and not crowded; each group had their own lab station and was working at least one lab station away from other students (diagram can be drawn if necessary). There was plenty of natural light in the room, and the ceilings were high so the sounds were, at times, drowned out by the chatter of other students.

For this particular group, it wasn't difficult to understand what was being said. This was helped by the fact that this group really didn't communicate with each other much. There were three students arranged in front of a laptop computer running the virtual simulation, complete with a mouse to the right of the machine. To the right of the machine was one student, the middle another, and to the left, the final student of the group. For a group of three students, they were exceptionally quiet. The group consisted of 2 male students and 1 female student. Initially, the students expressed confusion about the components of the simulation, but by reading they seemed to figure it out. One student, the middle student, was the primary driver for the simulation. The other students did not touch or interact with the simulation at all, although they did refer to it on occasion by using hand gestures to seemingly assist in their verbal explanations. All students appear engaged in the laboratory activity, but most of their time seems to be spent bent over their packets and writing as opposed to interacting with the simulation. The middle student does the most talking and provides the most verbal explanation for the phenomena being examined as the lab goes on. Each student draws independently; they do not consult each other about drawings or written explanations. Once an explanation is verbalized by one student in the group, there are no questions from other group members. There is also no pushback, so the explanation seems based solely on one student's articulation of events. The delegation of tasks within the group is minimal; the middle student seems to be in charge of running the simulation and the other two students don't seem to mind at all. Discussions are minimized to "I got this" and a response that is "okay". At one point, there was an inconsistency between group members' graphs. The students sorted it out by using this fact-checking mechanism; in determining what point should be plotted on the graph based on the data, a student indicated "I got this number for pressure" and the other student acquiesced. Interestingly, even though the simulation was right in front of them, they did not go back and rerun the simulation to double check the data they had previously recorded. What I observed seemed like a fact-checking exchange more than a discussion.

Analytic Memo O77 GUI

This group was selected due to low performance on the assessments. This group consisted of 3 members; 2 females and 1 male. In the beginning of the intervention, following the instruction, identifying parts of the simulation seemed like a struggle. This could be due to a lack of explicit clarity in the packet or because students were not paying attention—it's hard to tell which. The general operating procedure for this group appears to be that one person reads aloud the details about the activity to the other group members, and the other group members work with the simulation and try and figure out what is going on. The boy in this group appears dominant, sure of himself, and confident in his knowledge of the subject material. The girl between the narrator and the boy appears to be lacking self-confidence and most of the time, she acquiesces when there is a disagreement. This may not be a reflection of what she knows, but rather a mechanism through which she can operate and not hear the boy tell her constantly that what she is saying or thinking is wrong. The boy is quick to point out what the group should do in the simulation before everyone is prepared to move forward; it appears that, while he is not mean-spirited, he lacks patience or is bored.

Immediately, the sound of the simulation became an issue for the boy in the group, who appeared to find it rather annoying. This group spent a lot of time in trying to address the initial part of the laboratory, which is basically a tutorial for students to get a feel for how to operate the simulation. It is possible that this group of students was uncomfortable with technology use, in general, and that made them proceed slower than other groups.

As the activities progressed, I noticed that the boy was frequently referring to Newton's Laws of motion, specifically the Third Law of Motion, in his explanations for why certain things were going on in the simulation. As the simulation has been constructed with underpinnings of Newtonian mechanics, this seems to be a good application of the student's existing knowledge to a novel situation/phenomena. He also linked phenomena observed in different phases of matter to the gas laws content (i.e. solid and liquid). He used his knowledge of water to inform and explain hypotheses. In spite of this knowledge, the students in this group struggled to understand the molecular mechanism that drives pressure until there were only a few minutes left in the class period. Students initially had the impression that pressure is a result of what the piston is doing and as a result, what the molecules are doing. At one point, they called on another classmate for an opinion about what they were thinking, asking them if they agreed with their understanding of the aspect of the phenomena under investigation.

Throughout the course of this activity, I noticed several of what I'll term "rule-following" behaviors. This was enforced by the male student, who sometimes insisted that the directions must be followed exactly, and anything contrary to that approach was not correct. This was tempered by general group curiosity once engaging with the simulation, as students exhibited some play behaviors (e.g. Can we get the molecule number to 0?)

Analytic Memo G82 GUI

This student group consisted of 2 members; 1 male and 1 female. The computer is situated in front of the male student; the female student is located to the left. The lab is, generally, pretty noisy. There are quite a few students here working in a small classroom space; most of the groups (with the exception of the video group) are working 2 groups to a station. The video group was given a station by themselves in the hope that some of the excess noise would be filtered out in the video recording.

The boy in the group has assumed a leadership role and is in control of the simulation for the duration of the lab. The pace of this group appears to be slower when compared to other groups that I have observed. There is not a whole lot of interaction in this group. It seems as though the boy is doing a majority of the work, while the girl is transcribing the answers that he dictates aloud. A couple of times, the boy asks her for her opinion, and she doesn't really have much to add—whether she just doesn't feel like it or just doesn't understand the activity is not immediately clear. The girl often accepts the boy's assumptions/assertions about the phenomena being observed as truths. I noticed that she make her own hypothesis nor do they confer as a group and discuss the hypothesis. What happens is the boy will think of something, write it down, occasionally ask her what she thinks, and she just agrees that what he has is sufficient. She never challenges him. She even seems to copy his drawings. It would be a good idea to review the interview and packet data for this group closely to see if I might glean any other information about the lab from that data.

Like the previously reviewed low performing group, this group calls on another classmate to check in with them for help. In this case, the boy asks another student how to change the mass of the molecules. This is explained in the packet, so either it was misunderstood or never read. I wouldn't be entirely surprised by either; the packet is a bit dense and text heavy, so I can understand how it would be cumbersome for a high schooler to get through. A researcher soon came by and helped with the molecular mass issue. I noted also in this video that the student opens the HTML prompt by clicking on the data box in an attempt to manipulate the numbers within the simulation; however, the simulation does not allow for that kind of adjustment to be made. Some of the explanations I heard in the video make it clear that they don't seem to understand the underlying mechanism of pressure.

Analytic_Memo_O19 TUI HP

RQ #3 What differences, if any, are there between students using the Frame and students using virtual labs on students' interactions with the technologies?

There are three students in this group; 2 females and 1 male. One of the females leaves approximately 7 minutes into the activity and never returns to the classroom. All students are standing over the computer screen. For the majority of the activity, students are focused on the screen, and each student in this group takes turns talking and sharing ideas related to the simulation and procedures outlined in the packet. They also take turns in using the interactive components of the device (i.e. syringe, spring, temp. sensor, and touch screen aspect of simulation). The roles of the group here are not very rigid; the students take turns using the technology and confer with each other regularly. Students improvised a bit during the water activity, getting water out of the faucet to have more testing points (as 5 were asked of them) than just the hot and cold water—they wanted warm water, too.

I noticed that the temperature related activities seemed to take the most time to complete in this particular group. Part of this seems to be due to the nature of planning. This group read the packet carefully, and they understood prior to beginning the temperature investigation that they needed to take five temperature readings for the data collection component of that activity. Before they ever had jars filled, they talked to each other about a plan of action. They also consulted with a researcher who was floating around to address technical issues about what they could use to measure temperature. After they planned what they would measure, they went ahead with the investigation.

During the investigation, I noticed that students would reset and stop the simulation. Then they would place the jar on the sensor and hit run. This was not necessarily an intended method for the observation, as it may make it difficult to see change, but it was the method the students decided would work for them. Whether this was because they had trouble understanding the purpose of the investigation or because they preferred this, it is not possible to say with any degree of certainty. The interview data may shed some light on the procedural aspects of this group.
It seemed like a lot of time was taken to record data—both numerical and written. It seems like there could have been more interaction with the device.

When students conducted the investigation that requests them to add more molecules and to investigate pressure, I noticed that the students were a bit perplexed by the representation of the syringe. The boy was looking at the numbers on the syringe and trying to figure out how that might correspond to the volume of the sealed container represented in the simulation. He was initially trying to match the numbers on the barrel of the syringe with the numbers of molecules being added. One reason he may have been confused is that the simulation is set to accommodate only 200 molecules. He noticed a discrepancy between that limitation and the numbering on the barrel of the syringe and asked for help. It was clarified by a researcher floating in the class to answer technical questions that the syringe did not directly correspond to the size of the chamber on the screen, nor did the numbering have anything to do with the molecules.

I also noted that initial interpretations of the visualization itself were very literal. For example, when asked what the significance was of the molecules, which had longer trailing arrows, the girl noted the following, focusing on angle of impact.

11:54 “Like, if the arrow is longer [on the piston wall], a molecule with more energy hit it”

12:00 “Yeah. Hit it more directly I guess, because if you’ll notice the way they are, like if they hit it at a very obtuse angle, it doesn’t do much. ((pause)) so if they hit it at a…if they hit it as close to head on as possible [they have more force]”

I also noted that when students were not completely sure about a concept, they revisited the visualization and made use of some of the external tools, in this case, the spring, to reinforce their understanding. The girl did this when trying to remember what the relationship was between volume and pressure. After reexamining this, she decided that less volume had less pressure and was able to use this knowledge to inform her hypothesis.

Analytic_Memo_G62 TUI HP

RQ #3 What differences, if any, are there between students using the Frame and students using virtual labs on students' interactions with the technologies?

This group consists of three students; 2 males and 1 female. They are located in the back half of the classroom/laboratory area. This class is not as noisy as previous classes, and the students are focused on working with the lab most of the time. There is an open dialogue throughout—no one demeans another member of the group by telling them they are "wrong". The observations exhibited by the students in this group take a fair amount of time and are very thorough. Throughout the investigations in this lab, they allow the computer ample time to adjust to changes in temperature and volume, waiting to record their data until they are sure that they can record it with some measure of precision and accuracy.

Like the previous group O77, they are a bit innovative in their approach to collecting data points for the temperature investigation. When this group was provided with two jars of water (hot and cold) they improvised beyond their fingertips, deciding to combine the water into one jar after each were recorded individually. This was not a practice that we recommended, nor was it a common practice throughout the implementation of the lab activity.

The girl in this group is directing activity and leading the group forward; she is occasionally narrating the questions to the rest of the group, but this task is shared by the group members in some capacity as well. She trades places with the boy on the far side in order to try the piston/spring. She is able to pull the spring back so far that the system registers a negative pressure.

I noticed that this group made use of the pause feature (pressing the Run button a second time) more frequently than other groups that I've observed. This group appeared to find this feature exceptionally useful when recording data. Other features, such as energy, were less used. I noticed in the temperature data collection phases, the colors were not displayed on the screen (a function of the energy button), and this group did not rely on that function to understand characteristics of gas molecules.

There were a couple of minor difficulties that this group had while working through the laboratory. (1) students thought that the plunger's position in the syringe was a direct correspondence to the volume of the chamber displayed in the visualization. When extracting molecules from the chamber, this group pulled the plunger out of the syringe entirely and needed assistance for resetting the model afterward. (2) They also lacked a bit of clarity in terms of the type of graph they were expected to make—a line graph or a line that connects most of the points (regression). In spite of these difficulties with the features and curriculum, they were able to make their own suggestions as they went. For example, I heard the girl say that it would be nice to have a display of the running average temperature instead of oscillating back and forth so much. Before they started using the pause button, it was a bit difficult to record the fluctuating data.

Students in this group found it fun to take a break in the middle of their lab to play with the simulation and the hands-on components. This occurred briefly when the girl went to ask a question, and all three were involved at the conclusion of the lab. They developed a game that basically pitted humans versus molecules in a simulated tug of war. The piston/spring represented the human, and the molecules from the syringe, while human controlled, seemed to represent molecules.

Analytic Memo - P28 TUI LP

This group consisted of four students arranged around the computer, almost in a circular formation. This group was located at the back of the chemistry laboratory next to the chemistry prep room. There are three males and one female in this group. Overall, the class has quite a few students, and the students in this class are working in larger groups, on average, than students in the other implementations. This class is relatively social and loud; students are generally doing a lot of talking.

This group experimented with the piston prior to collecting data for the temperature activity. When they were using the cold water with the temperature sensor, they noticed that a bar appears within the chamber (on the side of the temp. sensor) which they used to help them denote the change of temperature within the chamber. Throughout the investigations, the group made regular use of the simulation features, velocity and energy. This seemed particularly helpful to them in understanding what happens, on the molecular level, with regard to temperature change.

While this group was talkative, it was mostly related to the activity at hand. At one point, in between activities, students began talking about time machines, solar systems, and galaxies, but after one student examined the packet and noticed that there was a lot yet to be done in a relatively short amount of time, he encouraged them to move forward, and they did.

It is clear that students in this group have a good rapport with one another, and they get along socially. Overall, they worked very well, in cooperation with one another, and by distributing tasks within the group—specifically, students took turns collecting data points and using the spring/piston and syringe. There was a point when the girl was behind, and the group continued to move on to the next activity. In what seemed like an effort to catch up, she took another group member's paper and copied his drawing to represent her hypothesis (31:11).

The response to the question, "why don't you observe fluctuations like this in real life?" is kind of superficial, but I don't think that is uncharacteristic of participants. It seems as though the question could be more specific to drill down to a molecular explanation. Students normally just say that they can't see the reaction, and this group was no different.

The boys seemed to be more inclined to play with these features of the simulation when there was some downtime; examining the end of this video, there are several minutes where the students are just playing to see what they can do with the system (after they'd completed the lab). It was a interesting that at the end, during play, one student actually referred to the computer and asked, "Are there actual gas molecules in there?" No one from the group responded and he did not pursue this question further.

Analytic Memo – G57 TUI LP

This group consisted of two students, one male and one female. The lab station is located at the back of the classroom. The female, who narrates the questions and checks in with her partner about what he thinks, leads the group. He doesn't speak very much, but he does provide input from time to time when requested. They waited for each other to finish tasks before moving on to the next one. After reviewing the video, I wonder if this is due to their insecurity or lack of confidence in their subject matter knowledge.

There were a couple of points in the lab activity where they really struggled (1) understanding how the temperature sensor was supposed to work and (2) understanding the concept of partial pressure and learning how to use the syringe to add/remove molecules and collect data from the simulation. The hands on component appeared to be really challenging for these students; it may be possible that for these students, they would have had improved success in the GUI condition.
It was more obvious, to me, as an observer that the girl lacked confidence. She was consistently checking in with the boy seemingly because she was not sure of her answers. The boy provided answers, but he did not appear outwardly to be exceptionally confident in his subject matter knowledge either.

At some point the sound became annoying to the girl, and the boy, detecting this without it being verbalized, shut it off. Just as students are beginning the molecular mass activity, the boy gives his paper to the girl (it seems like she is copying something from him) and he leaves for about a minute. She does some work in the interim and they swap papers when he returns, and he proceeds to copy down something from her paper. It is unclear from the video what exactly is being copied.

The most challenging aspect of using the TUI version for this group involved the sensors. So the temperature sensor was tough, because the students did not initially realize where they should put the jar. The girl held it over the touchscreen first, and then she kept holding the jar and tried to adjust the temperature in the simulation by touching her fingers to the touchscreen. Ultimately, she asked for help and the location of the sensor was explained to her.

The boy then tried to do a temperature reading of his fingers following the researcher explanation. He grabbed onto the plastic tube connecting the syringe to the gas pressure sensor and held that for a few seconds. That's when his lab partner redirected the position of his hand. She physically moved his hand over the temperature sensor because he was struggling to do it himself. It is not clear if this was lack of understanding or lack of effort.

At this point, the reading focused solely on the numbers in the data display, and the visualization of molecules was not really used by the group in this portion of the investigation. The following investigation asked them to examine partial pressure.

When examining partial pressure, this group's activity came to a standstill because they had difficulty in understanding where the molecular mass was. Students ran the simulation and observed it for about a minute before. The girl even mentioned that molecular mass sounded like a familiar concept ("I feel like we've done this before in chemistry"). She got researcher assistance, and then it appeared that the gas sensor was not communicating with the simulation, once it was explained that the syringe was to be used to add/remove molecules. The computer was restarted and the students resumed collecting data operating on the information they were provided about how to use the device.