Summer 8-1-2018

Tangible Teaching: The Effect of Physical Modeling on Community College Students’ Understanding of Conservation of Matter

Erick T. Moffett
University of Southern Mississippi

Tangible Teaching: The Effect of Physical Modeling on Community College Students' Understanding of Conservation of Matter

by

Erick Theon Moffett

A Dissertation
Submitted to the Graduate School,
the College of Science and Technology
and the Center for Science and Mathematics Education
at The University of Southern Mississippi
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy

Approved by:

Dr. Sherry S. Herron, Committee Chair
Dr. Deborah Booth
Dr. Richard Mohn
Dr. Kyna Shelley
Dr. Bridgette L. Davis

Dr. Sherry Herron Committee Chair	Dr. Sherry Herron Department Chair	Dr. Karen S. Coats Dean of the Graduate School

August 2018

Published by the Graduate School

THE UNIVERSITY OF SOUTHERN MISSISSIPPI®

ABSTRACT

Students' weak understanding of conservation of matter is well documented; however, there is a paucity of research that provides science educators with actual examples of empirically proven curricula employing physical modeling that can be used in the chemistry classroom to teach this fundamental concept. An intervention (three sequential physical modeling activities) was developed and evaluated. The intervention was administered to two sections of a General Chemistry I course at a community college in the southeastern United States, and pre-test/post-test data using a published instrument were collected to evaluate the physical model's effectiveness in developing students' understanding of conservation of matter compared to traditional teaching approaches. Because cognitive ability is theorized to play a significant role in understanding abstract concepts such as conservation of matter, student logical thinking ability was also measured using the abbreviated Group Assessment of Logical Thinking (GALT).

The results of a two-way mixed analysis of variance (2x2 ANOVA) revealed that statistically significant growth in understanding of conservation of matter and conceptual understanding occurred from pre-test to post-test for the treatment group only. In general, overall student understanding of conservation of matter was low with an average pre-test score of 39% and average post-test score of 47%. Initially, 87% of the students operated below the formal operational level, which decreased to 68% by the end of the study. The findings suggest that the physical model not only significantly enhanced students' understanding of conservation of matter, but also develop their conceptual understanding.

ACKNOWLEDGMENTS

This research project was by no means a lone effort. A tremendous amount of support was received from many individuals. The following paragraphs describe only some of the support I received.

My sincere thanks go to each of my committee members, especially Dr. Sherry Herron for serving as my committee chair and research advisor throughout this long endeavor. As my first professor in the doctoral program, Dr. Herron kept me productive and motivated with her continued guidance, patience, encouragement, and numerous edits. Without a doubt, I am extremely fortunate to have had such a seasoned faculty member as my committee chair.

From the very first day that we met seven years ago at the beginning of my graduate career, Dr. Bridgette Davis has been the driving force behind my professional and academic growth. Even though we did not initially hold the same educational philosophy, she took me under her wing and revolutionized what I thought I knew about education through tremendous mentorship. Countless times, she made unselfish investments in me and my family without hesitation. For that, I will always admire and respect her as a mentor, colleague, and friend.

Dr. Deborah Booth was the first teacher to show me the power of modeling in chemistry. I will never forget her tour of “Moleville”, which made mole conversions and stoichiometry so much more bearable for a first-semester college student. Dr. Booth is a continual reminder of the kind of science educator I strive to be.

Dr. Richard Mohn served as my most invaluable local statistician. I am extremely grateful for his help, support and continued input in data analysis and study design.

Having listened to his recorded lectures countless times, I have a newfound appreciation for his remarkable intelligence and sense of humor. *So concludes my prepared remarks on Dr. Mohn*.

Dr. Kyna Shelley also served as an invaluable statistician, but in a different capacity. Everything I know about the intricacies of study design and methodology in general is a result of her phenomenal instruction. Dr. Shelley was my very first statistics professor, and got the ball rolling as I began to frame my dissertation study. Her advice to not allow myself to lose momentum served me well.

I am very fortunate to have had a strong network of family, church family, neighbors, coworkers, and friends as a source of continued encouragement and support. Nothing was more motivating than the desire to make them all proud. I would also like to thank my doctoral colleagues: Adam Smith, Kelly Byrd, and Antoine Gates. They kept me in good spirits with their humor, kind words, friendship, and advice. I'll be waiting for you all at the finish line.

Finally, I would like to acknowledge Dr. Ansley Abraham of the State Doctoral Scholars Program. Five years ago, I was in a tough situation with a difficult decision to make. Thus, I will never forget the moment I received his call. His question, "Are you still interested", was music to my ears and God's answer to my prayers. Thank you, Dr. Abraham, for opening the door for me towards a Ph.D.

DEDICATION

I dedicate ***Chapter 1*** to my mother, Dr. Moffett, who gave me life, and through her incredible example ***introduced*** me to the idea of what hard work really is.

I dedicate ***Chapter 2*** to my father who has been a light guiding me through the ***difficulties*** of life and is a loving example of the man I wish to one day become.

I dedicate ***Chapter 3*** to my daughter who arrived right in the ***middle*** of this endeavor and provided the motivational push to keep me going as she sat in my arms at the computer before bedtime.

I dedicate ***Chapter 4*** to the Doctoral Scholars Program. I am a ***result*** of its incredible vision to increase the number of minority students earning doctoral degrees.

I dedicate ***Chapter 5*** to my wife with whom I intend to ***conclude*** the last chapter of my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER I – INTRODUCTION

Overview

The history of education in the United States is a story of constant change and reform. Over the years, policymakers and other stakeholders have proposed a variety of curricular and pedagogical ideas in an effort to remedy deficiencies that undermine the academic performance of students from all over the country. Science education reform is no exception. In 2012, the National Research Council published their *Framework for K – 12 Science Education (Framework)*. From this *Framework*, the *Next Generation Science Standards (NGSS)* were born in 2013. The release of these newly developed science standards signified the start of a paradigm shift in science education in the United States. Suddenly, science and engineering practices were considered an important part of student engagement in authentic scientific inquiry (Bybee, 2013). The *Framework* states that true science is much more than just the memorization of a collection of facts about the natural world, but that it is also "a set of practices used to establish, extend, and refine that knowledge" (NRC, 2012, p. 26). The assumption made is that if scientists use a specific set of practices to refine their knowledge and understanding, then students should also use these same practices to develop their own learning.

The *Framework* and the *NGSS* identify eight science and engineering practices used by scientists that should be emulated by students in the science classroom. These practices are: (1) asking questions (for science) and defining problems (for engineering), (2) developing and using models, (3) planning and carrying out investigations, (4) analyzing and interpreting data, (5) using mathematics and computational thinking, (6) constructing explanations (for science) and designing solutions (for engineering), (7) engaging in argument from evidence, and (8) obtaining, evaluating, and communicating information. The focus of this study is on the second scientific practice – developing and using models – and its implications for teaching abstract concepts in chemistry, namely, conservation of matter.

This study sought to explore the effect of a novel instructional manipulative on students' understandings of conservation of matter. This effect was examined within chemistry classrooms at a public community college in a southeastern state of the United States. This chapter presents an overview of the study and is organized according to the following sections: (a) background of the problem, (b) theoretical framework, (c) statement of the problem, (d) the purpose of the study, (e) research questions, (f) hypotheses, (g) delimitations, (h) assumptions, (i) definitions of terms, and (j) summary.

Background of the Problem

The phrase "seeing is believing" is well known by many around the world. It suggests a certain degree of skepticism to those things that are invisible to the naked eye and implies that one cannot truly accept the unseen. This skepticism of the intangible can be extended to the realm of chemistry, in which students are often faced with the difficult task of learning concepts that cannot always be readily observed. It is unsurprising to discover that students tend to struggle with the glut of abstract concepts chemistry has to offer, especially when mental manipulation on the molecular level is involved (Copolo & Hounshell, 1995). Related research has suggested that traditional teaching methods in the chemistry classroom are ineffective (Ozman & Ayas, 2003; Plass, Homer, & Hayward, 2009) and calls for instructional strategies that deepen students' conceptual understanding instead of promoting rote memorization.

Fortunately, there exists a countermeasure that every chemistry teacher has access to in their arsenal of instructional tools that can help students overcome the challenges that abstract concepts present; the use of instructional manipulatives as physical models. Unlike model cars or model cities, chemistry's models are big representations of small (often invisible) things. As external representations of mental concepts, models provide scientists a way to visualize and understand certain phenomena (Krajcik & Merritt, 2012). Using instructional manipulatives as physical models, students are also able to physically interact with concrete representations of natural phenomena (Berk, 1999; Carbonneau, Marley, & Selig, 2013).

The general consensus regarding the role of instructional manipulatives in the science classroom is that they effectively bridge the gap between the concrete and the abstract, thereby promoting students' conceptual understandings of scientific concepts that are notoriously difficult for students to learn (Bruner, 1964; Marley & Carbonneau, 2014; Piaget & Inhelder, 1969), a claim that has been verified by decades of research in educational contexts (Chiu & Linn, 2014; Salta & Tzougraki, 2011; Sowell, 1989; Wise & Okey, 1983; Za).

Conceptual versus Algorithmic Understanding

How, though, is student understanding of science gauged? One of the most important measures of success in science education is whether students have achieved conceptual understanding – as opposed to algorithmic understanding - of scientific concepts (Doucerain & Schwartz, 2010; Slavings, Cochran, & Bowen, 1997; Vosniadou, 2007). Being able to evaluate the degree of student conceptual understanding, however, is much more difficult than it may seem. Often, science educators unknowingly make the mistake of assuming that students have a conceptual understanding of an idea simply because they can solve a related algorithmic problem (Nakhleh & Mitchell, 1993; Nurrenbern & Pickering, 1987). Related research has revealed that over time many students' understanding of science concepts remains weak, while their ability to apply algorithms to solve problems improves significantly. This is probably so because algorithmic understanding is the primary beneficiary of the often-used traditional approach to teaching (Abraham & Williamson, 1994; Pfundt & Duit, 2000; Sawrey, 1990). Researchers have called for a focus on instructional strategies that develop students' conceptual understanding of conservation of matter (Agung & Schwartz, 2007).

Conservation of Matter

This study focuses on the law of conservation of matter, which states that matter can neither be created nor destroyed. Thus, the amount of matter in a closed system is conserved and remains the same regardless of other processes that may be active within the system (Doucerain & Schwartz, 2010). Conservation of matter in physical and chemical processes is not only described as a central crosscutting concept in the *Framework* and the *NGSS*, but it is also integral in understanding more advanced concepts in chemistry (Ozmen & Ayas, 2003) and in other sciences (Hartley, Wilke, Schramm, D'Avanzo, and Anderson, 2011; Thomson & Lotter, 2014). Therefore, a sound understanding of the law of conservation of matter is an especially important factor in a student's ability to understand the world around them (Pyke & Ochsendorph, 2004). Most students, however, show a weak understanding of this law (Agung & Schwartz, 2007; Benjaoude & Barakat, 2000) and hold numerous misconceptions (Andersson, 1986; Barkar & Millar, 1999; Ben-Zvi et al., 1987; Driver et al., 1984; Andersson, 1984;

Hesse & Anderson, 1992; Ozmen & Ayas, 2003; Ramsden, 1997; Yarroch, 1985). Formal instruction has helped students perform better on algorithmic problems such as balancing equations and stoichiometry but does not seem to improve students' conceptual understanding of the law (Gomez, Pozo, & Sanz, 1995). Considering its significance and challenging nature, conservation of matter is perfect for studying the impact of instructional manipulates on science learning.

Instruction

Since the creation of universities in Western Europe over 900 years ago, the predominant form of instruction has been lecturing (Freeman, Eddy, McDonough, Smith, Okoroafor, Jordt & Wenderoth, 2014). The traditional style of instruction relies heavily on the textbook, lectures, and worksheets. Unfortunately, this approach to teaching seldom influences students' misconceptions, as they tend to retain any preconceptions after instruction has ended (Driver, 1985; Perkins, 1992; Osborne, 1985, Harlen, 1985). Recently, a large-scale meta-analysis on STEM (Science, Technology, Engineering, and Mathematics) education revealed that in comparison to the traditional, instructor-focused approach to teaching, a constructivist approach (i.e., allowing students to construct their own knowledge as more active learners) not only decreases the likelihood that students will fail, but also contributes to an increase in exam scores by an average of 6 percentage points (Freeman et al., 2014). This study sought to extend these findings to the concept of conservation of matter, which is both fundamental to the sciences and notoriously difficult for students to learn.

Theoretical Framework

Over the past few decades, constructivism has been at the forefront of the paradigm shift in science education. In support of this framework, research has shown that student learning is maximized when students take control of their own learning (National Research Council, 2000). The integration of hands-on activities and particularly the use of instructional manipulatives as aids in student learning are hallmark of Jerome Bruner's (1966) constructivist theory of learning.

Constructivism calls for students to take an active role in their own learning as they construct new knowledge through a process involving the acquisition of new information and its assimilation into preexisting knowledge schemas (Bruner, 1966). Essentially, all students come already preconfigured with their own repertoires of knowledge that are dependent upon personal experiences and prior knowledge. There are three major principles of the constructivist approach. First, instruction must provide an environment in which students are willing and able to learn. Second, instruction must be structured in a way in which students are able to understand. Lastly, instruction should be preconfigured in a way that allows for students to explore beyond the information presented. Essentially, the main goal of the instructor should be the active engagement of all students (Schlechty, 2002).

In contrast, the traditional method of teaching places students in a much more passive role in the learning process, as the teacher takes center stage in the classroom and teaches in a way that generally requires students to use rote memorization to master concepts. Naturally, this method of teaching leaves very little opportunity for students to become actively engaged in the learning process.

Following a constructivist approach to teaching does not suggest, however, that the learning process will be easy or that the acquisition of knowledge is guaranteed. On the contrary, teachers and students face many challenges on the road towards the construction of knowledge that are often specific to the various academic disciplines. As previously discussed, chemistry is notoriously difficult for both teachers and students. Educational research in chemistry has identified several reasons contributing to complexities in the teaching and learning of chemistry. Johnstone's (1982) triangle which attempts to shed light on why chemistry is often difficult for students offers a popular theoretical framework in understanding how chemistry concepts are often represented. Johnstone describes three interconnected levels of chemical representation: the macroscopic level, the submicroscopic level, and the symbolic level. An overview of the three levels are shown in Figure 1, with examples of each of the three levels of chemical representation of matter in Figure 2.

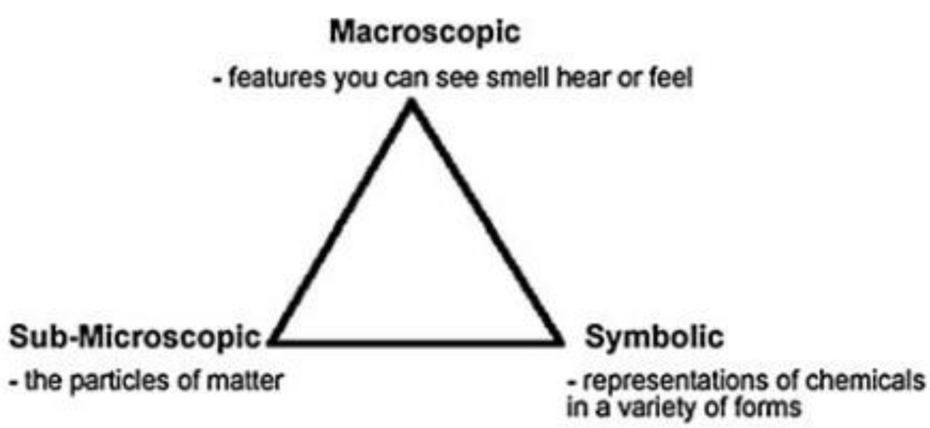


Figure 1. Three levels of chemical representation of matter (Johnstone, 1982).

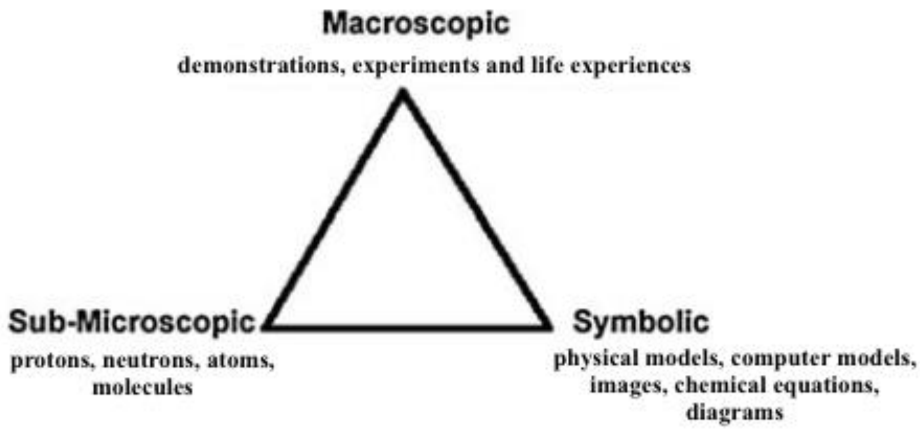


Figure 2. Examples of each of the three levels of chemical representation of matter (Johnstone, 1982).

Harrison and Treagust (2002) discovered that for the majority of 8th grade students, and even for many junior high and high school science teachers, their understanding of the submicroscopic level of matter is deficient. Furthermore, subsequent research has revealed that students ranging from secondary school to college as well as teachers have trouble shifting from one level to another, which suggests a dire need for teachers to know how and facilitate the process of transferring students from one level to another (Treagust and Chittleborough, 2001).

As suggested by Philips (1995), there are many constructivist theories of learning. This study focused on the constructivist theory of Piaget, as it is among the most commonly referenced in the realm of science education. Jean Piaget's (1973) cognitive development theory provides additional support for the study as an explanation of how students can successfully construct abstract concepts such as conservation of matter through the manipulation of models as symbolic representations within Johnstone's hierarchal level (Herron, 1975; Johnstone, 1993). Piaget describes the process of cognitive development in terms of four sequential stages: sensorimotor, preoperational, concrete operational, and formal operational. The stages most relevant to this study are the concrete operational and the formal operational stages. Each stage is associated with an age range in which it is most likely to be observed. According to this theory, we would expect children to enter the formal operational stage at the age of 12 and essentially complete their cognitive development by the age of 15. However, much research has been done that contradicts the age progression outlined by Piaget. Huddle and Pillay (1996) and Smith (1978) found that most of the high school chemistry students in their study lacked the capacity for formal operational thought and therefore struggled. These findings have been corroborated in other studies done around the world such as England (Lovell, 1961) and Australia (Dale, 1970). Even more striking is what science education researchers have discovered about attainment of formal operational thought by students at the college level. A study of 131 college freshman enrolled at the University of Oklahoma revealed that 50% of the sample were functioning at the concrete operational level as defined by Piaget and only 25% met the criteria for formal operational thinking established by the researchers (McKinnon & Renner, 1971). The implications of these studies become plain once a clear distinction is made between a student at the concrete operational level of development and a student at the formal operation level of development.

Recently, new research has been conducted specifically in the field of science education that corroborates the story being told by the classic studies of students' logical thinking ability presented above. Bird (2010) conducted a study surveying 466 students who were enrolled in the second semester of General Chemistry. Approximately 59% of the students fell below the formal operational level. Considering that the logical reasoning skill test was administered at the start of the second semester of General Chemistry, it can be assumed that the percentages do not include students who failed or withdrew during the first semester of the General Chemistry course. Therefore, it would be sensible to speculate that the excluded students falling below the formal operational level could further inflate the percentage value of 59%. McConnel, Steer, Owens, & Knight (2005) had comparable findings in their study of 741 students enrolled in an introductory geoscience course. These students were found that 57% of these students were functioning below the formal operational level based on the same assessment and score ranges used in Bird's (2010) study.

The nature of concrete operational thinking is revealed in the name of the stage itself – concreteness. The thinking process of a student at the concrete operational stage of cognitive development is completely oriented towards concrete things. This student is unable to understand abstract concepts that stray from concrete reality and can be readily observed (Herron, 1975). On the other hand, a formal operational thinker can think abstractly and has the capacity for higher-order reasoning (Inhelder and Piaget, 1958).

Figure 3 presents the theoretical framework for this study, which attempts to combine Piaget's stages of cognitive development and Johnstone's three levels of representation in chemistry to show how physical modeling can be used as a bridge to transition students unable to move from macro level of representation (concrete operational stage) to the submicro level of representation (formal operational stage) on their own. In theory, this is accomplished because the physical modeling represents the unobservable in a way that allows students to not only visualize but manipulate it while it is being taught.

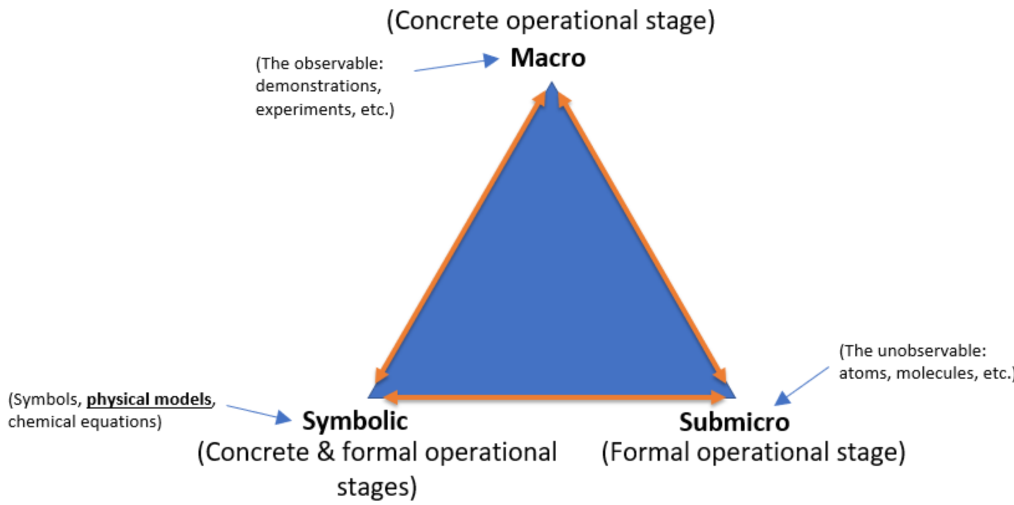


Figure 3. Transitioning through the three types of representations – the combination of Piaget's stages of cognitive development and Johnstone's three levels of representation in chemistry.

Note. Triplet component image modified from "The role of submicroscopic and symbolic representations in chemical explanations" by D.F. Treagust, G. Chittleborough, and T. L. Mamiala (2003), International Journal of Science Education, 25, p. 1354.

Statement of the Problem

Conservation of matter is a concept that is not well understood by students of various ages, a trend that persists in all around the globe, from the United States to Indonesia (Agung & Schwartz, 2007). As the law of conservation of matter is central in chemistry (Ozmen & Ayas, 2003), a study investigating how students grasp this concept while evaluating the use of various teaching strategies would be an important benchmark in science education research. While research shows that instructional manipulatives as physical models work to improve student learning, there is a paucity of research that provides science educators with actual examples of a proven curriculum employing physical modeling that can be directly implemented in the chemistry classroom to teach conservation of matter (Chiu & Linn, 2014; Salta & Tzougraki, 2011; Sowell, 1989; Wise & Okey, 1983). To help fill the gap in the literature, this study drew from current research on student learning in science education to create a hands-on lesson designed to help students master this difficult concept.

Purpose of the Study

While previous studies have examined student understanding of conservation of matter and called for the increased use of modeling as a tool to promote the growth of knowledge in the chemistry classroom, none have specifically examined the extent to which the implementation of an actual physical model affects students' understanding of conservation of matter (Chittleborough and Treagust, 2007; Coll and Treagust, 2001; Gabel & Sherwood, 1980; Harrison & Treagust, 1996). The purpose of the study was to investigate whether the physical modeling lesson developed by the researcher helped community college students better understand the concept of conservation of matter as reflected by improved achievement on the Conservation of Matter questionnaire (Sadler & Schwartz, 2004). Because cognitive ability is theorized to play a significant role in understanding abstract concepts such as conservation of matter, student logical thinking ability was also measured (Inhelder and Piaget, 1958; Roadranka, Yeany, & Padilla, 1983).

Research Questions

1. How does the use of physical modeling compare to traditional teaching methods regarding differences made in students' overall understanding of conservation of matter?
2. How does the use of physical modeling compare to traditional teaching methods regarding differences made in students' logical thinking abilities?

3. How does the use of physical modeling compare to traditional teaching methods regarding differences made in students' conceptual and algorithmic understanding of conservation of matter?

Research Hypotheses

This study was designed based off the research questions listed above, to investigate the following hypotheses:

- H_1: There is a difference between students who receive instruction that incorporates physical modeling and students who receive instruction that does not regarding their overall understanding of conservation of matter.
- H_2: There is a difference between students who receive instruction that incorporates physical modeling and students who receive instruction that does not regarding their logical thinking abilities.
- H_3: There is a difference between students who receive instruction that incorporates physical modeling and students who receive instruction that does not regarding their conceptual and algorithmic understanding of conservation of matter.

Delimitations

Delimitations of this research study are as follows. For practical and logistical purposes, this study was delimited to students enrolled at a public community college in the southeastern United States. Therefore, findings of this study may not be generalizable to students in other areas, due to cultural and societal differences. Justification of this delimitation exists in the way in which researchers in the reviewed literature often delimited their studies to single schools or a small group of schools.

While this delimitation is effective in increasing internal validity, generalizability suffers. Nevertheless, the potential differences between participants from different schools, cities, or even states may hamper internal validity by introducing confounding variables or variables not relevant to this study.

Justification for using a public community college located in the South also exists in the fact that positive findings from this study can have the greatest impact in the South, where student achievement in the sciences is lower than anywhere else in the country. The latest National Assessment of Educational Progress (NAEP) results have serious implications about student performance in science in the South compared to the rest of the country and beyond. In 2015, achievement-level results for twelfth-grade students assessed in NAEP Science was lower for the South than any other region of the U.S. 81% of tested students performed below the proficient level on the science assessment in the South while percentages for students in the same achievement range for the remaining regions of the U.S. was 74% for the Northeast, 74% for the Midwest, and 80% for the West (NCES, 2016). Assumptions from these data are even further corroborated by comparing the percentage of students that met science benchmark scores on the ACT in southern states. In twelve of sixteen Southern states, the percent of students meeting the science benchmark score was well below the national average of 37%. Of these twelve Southern states, Mississippi had the lowest percentage of 20% (ACT, Inc., 2017).

The sample population of students enrolled in General Chemistry I, chosen because naïve participants were available, may further reduce generalizability. In addition, this study was delimited to the data gathered over the course of one month in the Spring 2018 academic semester. Therefore, the researcher was unable to fully examine the longitudinal effects of the proposed intervention. Furthermore, it is possible that the sample was not be wholly representative since participation is voluntary.

This study was delimited to the inclusion of just two instructional strategies. Future research may explore other types of instructional strategies that have proven successful in various disciples outside of the sciences, such as discourse initiated through argumentation and simulations.

Assumptions

Two assumptions were made for this study. First, it was assumed that all participants responded honestly to the Group Assessment of Logical Thinking (GALT), and the Conservation of Matter questionnaires (Roadranka, Yeany, & Padilla, 1983; Sadler & Schwartz, 2004). Second, it is assumed that all parties involved in the study (i.e., students, teachers, and researchers) adhered to the intervention with fidelity.

Definition of Terms

The following definitions are provided to clarify terms identified as pertinent to the study:

1. *Abstract concept* – concepts that exist solely as thoughts or ideas and cannot be conceived at the physical level.
2. *Algorithmic understanding* – algorithmic understanding describes a level of understanding that involves simply being able to complete mathematical functions to correctly solve problems.
3. *Balanced chemical equation* – a chemical equation that shows an equal number and kinds of atoms between reactants and products.

4. *Chemical formula* – the written representation of a substance's chemical makeup.
5. *Chemical reaction* – the process by which a new substance is formed.
6. *Coefficient* – the number before the formula of a substance in a chemical equation.
7. *Conceptual understanding* – conceptual understanding reflects a deep level of understanding that goes beyond the rote memorization of ideas.
8. *Concrete operational* – the third stage in Piaget's theory of cognitive development that involves the ability to use logical thought but can only apply logic to concrete objects.
9. *Conservation of matter* – a fundamental chemistry concept that describes how matter is neither created nor destroyed in a chemical reaction, despite changes in appearance or state of matter. Thus, matter is conserved and simply reshuffled to form new substances.
10. *Formal operational* – this final stage in Piaget's theory of cognitive development involves fully mature logical thinking and the ability to think abstractly without reliance on the concrete.
11. *Manipulative* – an object designed to allow a learner to perceive some concept through manipulation.
12. *Matter* – traditionally defined as anything that has mass and takes up space. In addition, matter is made up of atoms.
13. *Physical model* – used to describe unobservable phenomena by providing a physical representation that can be visualized and manipulated.

14. *Products* – the substances that are formed in a chemical reaction.
15. *Reactants* – the substances that are consumed in a chemical reaction.
16. *Subscript* – quantifies the number of atoms of each element in a molecule (ex: O_2).
17. *Traditional instruction* – the predominant form of teaching which incorporates instructor-led lectures and promotes heavy use of textbooks and worksheet practice.

Summary

This study investigated how understanding of conservation of matter was impacted as a result of two separate instructional techniques used to teach students taking General Chemistry I at a public community college in the southeastern United States. This study is presented in five chapters. Chapter I has provided an introduction and overview of the study and is followed by an exhaustive review of the literature relating to science education in Chapter II. Chapter III outlines the research methodology followed by a discussion of the findings in Chapter IV. A summary of overall conclusions for this study is presented in Chapter V.

CHAPTER II – REVIEW OF THE LITERATURE

Overview

To provide a rationale for this study, this chapter presents a review of the literature, which includes the key components for understanding the many variables related to increasing the conceptual understanding of conservation of matter in high school chemistry students. Chapter two contains three major sections. The first section provides an historical background of science curriculum reform in the United States. The second section focuses on recommendations of instructional techniques related to the proposed intervention that have been proven to theoretically facilitate conceptual change in understanding abstract science concepts. The third section presents literature evaluating students' understanding of conservation of matter.

Foundations of Science Curriculum Reform in the United States

During the 1950s, the many criticisms of science education in America had become an important topic of discussion and reflection among scientists. Initially backed by various professional organizations and the National Science Foundation, groups of scientists pondered ways in which science education in the United States could be revitalized with a healthy injection of scientific rigor. This initiative finally received governmental support and financial backing with the launch of the Russian satellite, *Sputnik I*, into orbit in 1957 (DeBoer, 1991, p. 147). In the following year, Congress passed the *National Defense Education Act*, which provided substantial funding to boost science education, ushering the U.S. into a Golden Age of Science and federal involvement in science teaching. During this time, many national science programs were started to further encourage science education in the United States. Such programs included the Biological Science Curriculum Study (BSCS), which played a large part in redefining science instruction and revolutionizing the content of early science textbooks. With the release of A Nation at Risk in 1983, recommendations for science education were made that focused on a broader approach to achieving science literacy, as opposed to a narrower focus on content knowledge. This reform movement of inducing positive change within the educational system continued into subsequent years, as shown by the inception of Project 2061: *Science for all Americans* (Rutherford & Ahlgren, 1990) and the *National Science Education Standards* (NSES) (National Research Council, 1996). Most recently, the *Next Generation Science Standards* (NGSS) have emerged as the latest reform effort in science education.

Scientific Practices in Science Learning

The *Framework* and *NGSS* are constructed around the three dimensions of crosscutting concepts, disciplinary core ideas, and scientific practices. The reasoning behind this structure exists in the idea that allowing students to explore crosscutting concepts and core ideas through engagement in scientific practices will result in an increased achievement and a deeper understanding of the nature of science itself (NRC, 2012). These scientific practices involve a specific set of behaviors and activities in which scientists are constantly engaged.

There are eight science and engineering practices identified in the *Framework* and *NGSS*. These practices are: (1) asking questions (for science) and defining problems (for engineering), (2) developing and using models, (3) planning and carrying out investigations, (4) analyzing and interpreting data, (5) using mathematics and computational thinking, (6) constructing explanations (for science) and designing solutions (for engineering), (7) engaging in argument from evidence, and (8) obtaining, evaluating, and communicating information. The focus of this study is on the second scientific practice – developing and using models. Therefore, a brief description of this practice only is provided below.

Developing and Using Models. The *NGSS* assert that as soon as the eighth grade, students should be able to (1) develop a model to predict and/or describe phenomena; and (2) develop a model to describe unobservable mechanisms (p. 56). Lehrer and Schauble (2006) argue the importance of model-based reasoning to science practice asserting, "Scientific ideas derive their power from the models that instantiate them, and theories change as a result of efforts to invent, revise, and stage competitions among models" (p. 371). Selley (1981) goes so far as to say that familiarity with models and their role in the development of scientific ideas should be a key component to every chemistry educator's philosophy of education and pedagogy beliefs. This idea of the importance of models to science supports the use of modeling as practice. Specifically related to the focus of this study, the *NGSS* performance expectation MS-PS1-5 calls for the use of physical models to describe how mass is conserved in chemical reactions. Figure 4 depicts the *NGSS* performance expectations for the middle school physical science relevant to this study.

<table>
<tr><td colspan="2">Performance Expectations
● MS-PS1-1. Develop models to describe the atomic composition of simple molecules and extended structures.
● MS- PS1-5 Develop and use a model to describe how the total number of atoms does not change in a chemical reaction and thus mass is conserved. [Clarification Statement: Emphasis is on law of conservation of matter and on physical models or drawings, including digital forms, that represent atoms.]</td></tr>
<tr><td colspan="2">Disciplinary Core Idea
PS1.B: Chemical Reactions
● Substances react chemically in characteristic ways. In a chemical process, the atoms that make up the original substances are regrouped into different molecules, and these new substances have different properties from those of the reactants. (MS-PS1-2),(MS-PS1-3),(MS-PS1-5)
● The total number of each type of atom is conserved, and thus the mass does not change. (MS-PS1-5)</td></tr>
<tr><td>Dimension</td><td>NGSS code/citation</td></tr>
<tr><td>Science and Engineering Practices</td><td>Developing and Using Models</td></tr>
<tr><td>Disciplinary Core Ideas</td><td>PS1.B: Chemical Reactions
In a chemical process, the atoms that make up the original substances are regrouped into different molecules, and these new substances have different properties from those of the reactants. (MS-PS1-2),(MS-PS1-3),(MS-PS1-5)
The total number of each type of atom is conserved, and thus the mass does not change. (MS-PS1-5)</td></tr>
<tr><td>Crosscutting Concept</td><td>Scale, Proportion, and Quantity
Time, space, and energy phenomena can be observed at various scales using models to study systems that are too large or too small (MS-PS1-1)
Energy and Matter
Matter is conserved because atoms are conserved in physical and chemical processes. (MS-PS1-5)</td></tr>
</table>

Figure 4. Examining the *Next Generation Science Standards* (NGSS Lead States, 2013) at the middle school level.

In keeping with the three-dimensional style of teaching described by the *NGSS Framework,* this model included appropriate *NGSS* middle school standards and an integration of the three *NGSS* dimensions (science and engineering practices [SEP], disciplinary core ideas [DCI], and crosscutting concepts [CC]) into both student lessons and assessments.

The physical model curriculum directly addresses the DCI, PS1.B: Chemical Reactions, that provides the foundation knowledge base for understanding conservation of matter and the balancing of chemical equations (NGSS Lead States, 2013). Mastering this basic concept in middle school is essential in preparation for higher-level science courses offered in high school and college. Considering previously aforementioned research that indicates that undergraduate students do not have a strong fundamental understanding of conservation of matter, it would make sense to address middle school standards that focus on the concept even when dealing with college students.

The *NGSS* focus on students' abilities to engage with science and engineering practices by developing and using models to demonstrate and explain abstract phenomena or unobservable mechanisms as within balancing chemical reactions. Using this model allows students to manipulate and visualize the abstract recombination of atoms within a chemical reaction, helping them to better grasp conservation of matter.

This model also addressed *NGSS* CC 2: Scale, Proportion, and Quantity, as students use tangible items (colored chips) and mathematical applications to concretely demonstrate conservation of matter. In addition, CC 5, Energy and Matter, is the most evident connection, as students were able to visualize the phenomena of atoms being conserved in physical and chemical processes.

To further assess the ability that manipulative models have in relaying abstract concepts to students, a simple curriculum comprised of three sequential physical modeling activities used to scaffold the learning tasks related to conservation of matter was developed by the researcher. These activities paired the *NGSS* crosscutting concepts of *Energy and Matter* with the science and engineering practice of *Developing and Using Models* within the disciplinary core idea of *Chemical Reactions*. Table 1 shows how the proposed intervention connects to the *NGSS*.

Table 1

Connecting to the Next Generation Science Standards.

Standard: MS-PS1: Matter and Its Interactions		
Performance Expectation: MS-PS1-5: Develop and use a model to describe how the total number of atoms does not change in a chemical reaction and thus mass is conserved. [Clarification Statement: Emphasis is on law of conservation of matter and on physical models or drawings, including digital forms that represent atoms.]		
Dimension	***NGSS*** **code or dimension name**	**Link to the standard**
Disciplinary Core Idea	Chemical Reactions	The total number of each type of atom is conserved, and thus the mass does not change. (MS-PS1-5)
Science and Engineering Practices	Developing and Using Models	Develop a model to describe unobservable mechanisms. (MS-PS1-5)
Crosscutting Concept	Energy and Matter	Matter is conserved because atoms are conserved in physical and chemical processes. (MS-PS1-5)

Note. (*NGSS* Lead States, 2013).

Teaching Chemistry

Long before the inception of initiatives to reform science education on the national level, research had been conducted in almost every scientific discipline to maximize student learning of science curricula. Instead of a more typical general approach to improving instructional strategies, research in science education seems to focus more specifically on ways to improve science learning.

Constructivism is at the forefront of this revolution, as it without a doubt has a positive impact in student learning. Curriculum based on constructivism tends to actively involve students in the learning process as opposed to traditional lecture-style teaching.

Teaching Chemistry Using Different Types of Representations

It is well known that one of the main reasons explaining why learning science can be so difficult for students is that scientific knowledge can be represented on several different levels. Not all these levels are readily observable to students, and therein exists the problem (Handbook of Science Research, p.382). Chittleborough and Treagust (2007) purport that just as scientists must be able to represent knowledge to conduct research, teachers must also do the same to teach students. This is especially relevant in the realm of chemistry, in which students are often faced with the difficult task of learning concepts that cannot always be readily observed (Johnstone, 1991; Nakhleh, 1992; Taber, 2002). Understandably, visualization cannot be realized in the traditional classroom environment that relies solely on textbooks and lecturing. To improve students' cognitive and affective outcomes, several external representations have been integrated into science education over the past few decades. These external representations have been heavily researched and documented in the science education literature across a wide variety of content areas. Treagust and Tsui (2014) categorize these external representations into six major overlapping categories: *analogies* (e.g., Dagher, 1994; Treagust, Harrison, & Venville, 1998) and *metaphors* (e.g., Aubusson, Harrison, & Ritchie, 2006; Martins & Ogborn, 1997); *visualization* (e.g., Gilbert, Reiner, & Nakhleh, 2008); *models* and *model-based learning* (e.g., Buckley, 2000; Clement & Rae-Mamirez, 2008; Gilbert, 2004; Gilbert & Boulter, 1998); *multilevel representations* (e.g., Chandrasegaran, Treagust, & Mocerino, 2011; Gilbert & Treagust, 2009); *multimodal representations* (e.g., Waldrip & Prain, 2012; Waldrip, Prain, & Carolan, 2010); and *multiple external representations* (eg., Ainsworth, 1999; de Jong et al., 1998; Treagust & Tsui, 2013; Tsui & Treagust, 2003). Each type of representation contributes in its own unique way to the development of students' understanding of the science concept being studied (Coll and Treagust, 2001). This study is grounded in the use of models as external representations, which overlaps with most of the previously mentioned methods and strategies.

Models and Modeling. The release of the National Research Council's (NRC) *Framework for K-12 Science Education* in 2012 and the *NGSS* in 2013 marked a shift in the dialogue surrounding science education reform. This shift acknowledged the idea that science is more than just the rote memorization and recitation of knowledge, but also "a set of practices used to establish, extend, and refine that knowledge" (NRC, 2012, p. 26). As listed in the *NGSS*, one of these essential practices is that of modeling. A model is simply comprised of actual objects representative of some natural phenomena. For example, modeling was used by the scientists James Watson and Francis Crick to elucidate the correct structure of DNA. Even today, modeling is used across a wide variety of disciplines to help scientists investigate and predict phenomena. The premise for incorporating modeling into science education is that if scientists use modeling to further their own scientific understanding, why should students not also use modeling to enhance their own learning?

The use of models has distinct advantages. One main advantage is the fact that models can incite discussions and explanations that encourage students to evaluate the logic of their own thinking and understanding (Raghavan & Glaser, 1995). Considering the heavy amount of abstract content in chemistry, the use of models and modeling is commonplace as it allows students to circumvent the problem of unobservable phenomena by helping them to fabricate their own mental models of chemical compounds. In fact, Coll and Treagust (2001) state that it is nearly impossible to explain chemical phenomena without the use of modeling.

While the use of models has been shown to improve students' understanding of chemical concepts (Gabel & Sherwood, 1980; Harrison & Treagust, 1996), they can present a unique problem within science education in which students often confuse the reasons for using models and modeling (Renstrom et al., 1990; Treagust, Chittleborough & Mamiala, 2001). To elaborate, students often do not regard models differently from the natural phenomenon that the model represents (Harrison & Treagust, 1996) and see them through a simplistic lens as mere copies of the scientific phenomena they represent (Grosslight, Unger, Jay, & Smith, 1991). Lehrer and Schauble (2006) note that novice interpretation of physical models by students typically focuses on literal similarity, a finding shared by other researchers (Grosslight, Unger & Jay, 1991; Schwarz et al., 2009). To confront this issue, Hardwicke (1995) suggests that a solution lies in discussing the strengths and limitations of each model so that students can assess its value. Gilbert, Boulter, & Elmer (2000) and Harrison & Treagust (1996) place responsibility on the teacher, suggesting that educators be trained to use models in the classroom in a more scientific way.

Manipulative Instruction

Manipulatives are known to be physical items such as markers, blocks, or puzzles that are often used in education. Manipulatives are not always concrete, however and can be virtual as well in the form of simulations and games. The use of manipulatives in an educational environment is regarded as a constructivist approach, as students are using the items for active engagement as they learn the concepts taught to them. In this approach, the educator supplies the students with the items that they require to accomplish specific learning objectives and may provide them with a basic direction for pursuit or guide them more directly in a 'step-by-step' format of guidance. Many students may prefer this approach to traditional lecture-based approaches, and the use of manipulatives has potential to be effective through the providence of a multi-sensory experience, representation capacity, facilitating communication between students, and increasing understanding (Green, Piel, & Flowers, 2008; Marley & Carbonneau, 2014).

Several analysts across a variety of disciplines have discussed or studied the effects of instruction with a manipulatives emphasis, recommending considerations for practice while confirming the practical benefits of using the approach (Gire et al., 2010; Klahr & Williams, 2007; Olympiou & Zacharia, 2012). These recommendations generally imply that the methods and techniques could be used in educational processes, such as science classrooms that teach conservation laws. More specifically, Burns (2007) suggests that manipulative models help students to make sense of abstract concepts, allow students to evaluate ideas, are useful in problem solving, and increase student motivation to learn by making it more engaging. Furthermore, Chiu and Linn (2014) point out the fact that manipulative models allow students the opportunity to physically interact with chemical reactions on the molecular level through the visualization of atoms, invisible to the naked eye.

Green, Piel, & Flowers (2008) studied the potential for manipulatives to address common misconceptions regarding arithmetic, and completed two research efforts, targeting 50 and 39 students, respectively. In the first study involving 50 participants, the researchers assessed their capacity to solve basic math problems (i.e. regarding whole numbers and fractions) using concrete and representational manipulatives across five sessions. In the second study, the researchers attempted to reproduce the results with a slightly smaller sample of 39 people. Studying the results, the researchers found that the participants in the first study had been able to use the manipulatives to successfully substantially increase their knowledge, while the participants in the second study generated results that led the researchers to conclude that manipulatives can facilitate both an improvement in general knowledge and a reduction of misunderstandings or misconceptions regarding a given topic. Discussing the relevant theory in further detail, Green, Piel, & Flowers (2008) explained that while the impacts of learning with this approach vary based on the context, the approach is generally associated with increased learning outcomes and improved conceptual comprehension in mathematics and other subjects such as chemistry. They also questioned why they are generally not more commonly used for math instruction if they are known to tend to facilitate improved learning outcomes; they hypothesized that it may be due to a common belief that manipulatives function well as attractors of interest and stimulation during parts of a lecture or traditional teaching method, but that they cannot be used on any type of regular basis to optimize learning. Considering all of this, they recommended that educators consider using manipulatives on more of a regular basis, especially in courses of a more quantitative nature such as mathematics and chemistry, claiming that they can even be effective for teaching in a subject with a qualitative emphasis.

Marley and Carbonneau (2014) discussed a range of empirical evidence and theory regarding their general use, focusing on how evidence supports the theory, and the implications for adopting the strategies in classrooms. They explained that technology has increasingly been used as a way of applying manipulatives to education, and that virtual manipulatives have great potential to be useful in a wide range of classroom settings. Marley and Carbonneau (2014) offer the following thoughts about manipulatives for consideration:

> If one accepts that manipulatives are core instructional materials in educators' toolkits, this question goes beyond head-to-head comparisons of manipulatives vs. no manipulatives conditions. One might even consider summative contrasts of this nature uninteresting, as there are many dimensions that symbolic or iconic control conditions can differ from enactive treatment conditions that are unrelated to the presence or absence of manipulatives. (p. 3).

While a great deal of literature has been concerned with whether teaching through manipulative use is better than teaching through the more traditional methods, the analysts recommended that ongoing research and analyses focus on the capacity for manipulatives to be used in conjunction with traditional approaches to optimize educational objectives and learning outcomes. Meanwhile, educators and those involved with designing educational programs are recommended to consider literature demonstrating the effectiveness of manipulatives in guiding learning, in teaching science, in improving retention, and in generally improving the educational process in the range of specific settings that have been targeted. Marley & Carbonneau (2014) concluded that while there is a consensus that manipulatives are effective in instruction, and therefore play an important role in it, many aspects of the nature of the role and potential mechanics in the process as a teaching technique are not fully understood, and therefore experts should continue to work to understand them. There is much that experts can attempt to do to do this, including designing research studies to strategically target which techniques are most effective in a specific setting, target variables for compatibility or affinity in potential students, and strategically assess the underlying theory of what can make the use of manipulatives most effective in teaching processes in general. Moreover, they can assess the existing research in attempt to isolate patterns that had not been noticed, and to attempt to further extrapolate meaning from the conclusions that had been reached in these studies.

Much of the underlying science involved in the effectiveness of using manipulatives is cognitive science and underlying processes that might be involved with a person believing that they are learning more (or are learning more effectively) from the use of manipulatives (as opposed to any combination of traditional auditory and visual approaches of conveying information). Conducting an analysis in this area, Pouw, Gog, & Paas (2014) examined a range of literature regarding the capacity for manipulatives to increase the general 'richness' of learning, disagreeing with a so-called 'moderate view' that the use of manipulatives can actual compromise the effectiveness of learning through its demand of so much attention and restricting the development of inferences believed to be better facilitated through traditional methods. The researchers referred to the underlying concepts in the Embedded Embodied perspectives on cognition, and argued that the ‘richness’ of interaction can actually reduce the net amount of ‘cognitive load’ proposed to be high in manipulative use, and that transferring information may be better facilitated by the range of sensorimotor experiences the student has when using manipulatives (Pouw, Gog, & Paas, 2014). The latter statement may be true when considering that there are more mediums for potential memorization when manipulatives are used, and further research may help to facilitate better understanding of this aspect of the cognitive science. Concluding, Pouw, Gog, & Paas (2014) wrote:

> Manipulatives have specific properties that make only certain actions possible…a mouse-based virtual interface that only allows for unimanual manipulation, or a pie-wedge that only allows for re-arranging parts in preset wholes. However, manipulative perceptual properties also determine which behavior given the possibilities is likely to be solicited. (p. 59).

They explained that while children requested to use blocks, they were observed to struggle with instructions to use one hand for one block at a time. The automatic nature of engagement is generally fundamentally inherent in people, which suggests potential for more engagement than would occur from simply listening or watching as instruction is received in the traditional learning environment.

Loong (2014) further reported on potential for a combination of physical and virtual manipulatives to benefit teaching, focusing on mathematics. According to this analyst, “physical manipulatives aid deep conceptual understanding because they present alternative representations that help reconstruct concepts…and aid concrete thinking…Students who have worked with manipulatives tend to perform better in maths" (Loong, 2014, p. 3). They also explained that using virtual manipulatives in conjunction with traditional ones could assist students in making substantial improvements over students that use only one type. They reported that educators and educational developers are recommended to consider mathematical fidelity, cognitive fidelity, and pedagogical fidelity when choosing manipulatives to use for math course instruction, and to consider comparable variables when choosing them for instruction of another course subject. They concluded that the importance of virtual manipulatives should not be ignored, but that traditional ones should not be ignored either, as a combination of both is recommended for use.

Satsangi & Bouck (2015) studied the impact of virtual manipulative use in online education, explaining "secondary students with a learning disability in mathematics often struggle with the academic demands presented in advanced mathematics courses, such as algebra and geometry. With greater emphasis placed on problem solving and higher-level thinking skills in these subject areas, students with a learning disability in mathematics often fail to keep pace with their general education peers" (p. 174). Their research sample was small, comprised of only three participants, but they were able to make an original empirical contribution to literature in documenting their finding that the virtual manipulatives used were able to improve learning outcomes. They further reported, "results from this study provide new evidence showing virtual manipulatives to be a viable and accessible technology to teach students with learning disabilities advanced mathematical concepts" (Satsangi & Bouck, 2015, p. 174).

Student Understanding of Conservation of Matter

First named by Antoine Lavoisier, the law of conservation of matter states that matter cannot be created or destroyed in a chemical reaction (Ebbing, 1996). As such, the total amount of mass also remains unchanged in the same system. Numerous scientists since Lavoisier have shared his sentiment on the importance of this law in being able to understand chemistry in general (Ozmen & Ayas, 2003). Piaget was the first to conduct a relevant research study on how children learn conservation of matter while studying children's ideas on solutions (Piaget and Inhelder, 1974). The researchers quizzed children between the ages of 4 and 12, on the phenomenon of sugar dissolving in water. The gathered responses were categorized into three specific stages ranging from no conservation (for example, "the sugar disappears") to full conservation (related to both quantitative and qualitative observations). Naturally, due to the quantitative aspect of the full conservation stage, it is only possible when children can grasp the abstract nature of matter on the atomic level. Piaget & Inhelder (1974) concluded that basic acquisition occurs at the concrete operational stage (ages 7-11).

In addition to studies done on how to improve student learning of science concepts, much research has also been done on student understanding of the concepts themselves. The focus of this study was on student understanding of the law of conservation of matter.

The concept of conservation of matter is not unique from other fundamental science concepts in the sense that it seems to be difficult for students to grasp. Several studies have discovered that students at the high school and even college levels have a fragmented understanding of conservation of matter (Agung & Schwartz, 2007;

Boujaoude & Barakat, 2000). More shocking is the fact that this trend of weak understanding of conservation of matter plagued with misconceptions extends beyond students to preservice chemistry teachers (Haidar, 1997).

Ozmen & Ayas (2003) identified many misconceptions that students harbor regarding the law of conservation of matter. The two most common misconceptions among the 10^{th} grade students in the study were that “the total mass increases in a precipitation reaction because the precipitate produced is solid and is heavier than a liquid”, and “when chemical combustion occurs in a closed system, the total mass decreases” (p. 279). Stavy’s (1991) claim that misconceptions are especially difficult to extinguish with traditional teaching methods is further supported by Ozmen and Ayas.

Agung and Schwartz (2007) examined Grade 12 chemistry students’ understanding of conservation of matter, balancing equations, and stoichiometry. They found that in general, student understanding of this fundamental principle in chemistry was low, with an average score of 41%. Results also suggested that students performed better on algorithmic-based problems (i.e., stoichiometry). Algorithmic concepts such as stoichiometry and balancing chemical equations are deeply based upon the law of conservation and are commonly taught in the chemistry curriculum; however, as stated by Mason, Shell, and Crawley (1997), student success with either of these algorithmic concepts does not necessarily correlate with student conceptual understanding of this fundamental principle. Thus, Agung and Schwartz (2007) call for future research to be done that focuses on teaching practices and curricula that focus specifically on developing students’ conceptual understanding of conservation of matter.

The importance of conceptual understanding was emphasized in a study involving 436 chemistry professors at 205 separate institutions in the United States (Slavings, Cochran, & Bowen, 1997). Faculty from various areas of chemistry identified "conceptual understanding" as one of the two most significant outcomes for student learning.

Doucerain and Schwartz (2010) decided to tackle the challenge of finding teaching strategies that enhance conceptual understanding. The researchers conducted a study in which the impact of two teaching strategies on students' conceptual understanding of the conservation of matter was probed. The two strategies existed in guided inquiry and argumentation, respectively. It was found after a pre- and post-test analysis that guided inquiry was especially effective in improving conceptual understanding.

Gomez, Pozo, and Sanz (1995) conducted a more detailed study that examined students' understanding of conservation of matter and how it correlated with age, level of instruction, and level of expertise in chemistry. The researchers discovered that understanding of conservation of matter was strongly correlated with age group. However, when the researchers compared subjects of similar age, the influence of level of expertise in chemistry was lower than predicted from other novice-expert studies. More interesting is the finding that level of instruction had a more significant influence on understanding of conservation of matter when quantitative reasoning was introduced, suggesting that formal instruction enables students to solve algorithmic problems better, but does not really enhance their conceptual understanding. These findings are support differences observed between algorithmic and conceptual understanding in other studies

(Agung & Schwartz, 2007; Nurrenbern & Pickering, 1987; Sawrey, 1990), which find that only a small number of students in secondary school and the first year of college could solve conceptual chemistry problems.

Considering the widespread deficient level of understanding of conservation of matter across age, educational level, and location, it seems important to focus on interventions that have proven to be successful in teaching conservation of matter. A small handful of studies have reported substantial gains in students' understanding of conservation of matter because of some special teaching strategy or instructional unit (Campanario, 1995; Doucerain and Schwartz, 2010; Garnet, Oliver, & Hackling, 1998; Johnston & Scott, 1991; Lynch, Kuipers, Pyke, & Szesze, 2005; Paixao & Cachapuz, 2000; Stavy, 1991). More specifically, Kimberlin and Yezierski (2016) devised an intervention composed of two inquiry-based activities for high school chemistry students that was found to have a profound effect on student achievement.

As much of the relevant literature describes conservation of matter as a cornerstone of chemistry, it is reasonable to assume that student mastery of this concept is imperative to future success. Unfortunately, the literature also identifies issues with students' understanding of this fundamental concept. Furthermore, the literature does not present a specific example of a proven curriculum employing physical modeling that can be directly implemented in the chemistry classroom to teach conservation of matter.

Summary

Students are expected to use models when learning conservation of mass in a chemical reaction as early as middle school according to the *NGSS*. However, research has shown that students often have trouble understanding this concept (Lempinen, 2010).

Students' understanding of fundamental concepts such as conservation of mass has a profound impact on their ability to grasp more advanced science concepts as they progress from middle school to high school and college. Science educators must try to utilize the best instructional strategies to teach these big ideas in science. One such instructional strategy is the use of instructional manipulatives as physical models.

Potential implications of this study could be tremendous. While research discussed in this chapter shows that physical modeling works to improve student learning, there is a paucity of research that provides science educators with actual examples of curriculum employing modeling practices that can be directly implemented in the classroom to teach conservation of matter, specifically.

CHAPTER III - RESEARCH DESIGN AND METHODOLOGY

Overview

In this chapter, the research questions and research hypotheses are stated. Also, the research design, sampling and experimental procedures, data collection instruments, and the analytic approach employed are discussed.

Research Questions

To explore the effect of physical modeling in the classroom on students' understanding of conservation of matter, the following research questions were examined:

1. How does the use of physical modeling compare to traditional teaching methods regarding differences made in students' overall understanding of conservation of matter?
2. How does the use of physical modeling compare to traditional teaching methods regarding differences made in students' logical thinking abilities?
3. How does the use of physical modeling compare to traditional teaching methods regarding differences made in students' conceptual and algorithmic understanding of conservation of matter?

Research Hypotheses

This study was designed based off the research questions listed above, to investigate the following hypotheses:

- H_1: There is a difference between students who receive instruction that incorporates physical modeling and students who receive instruction that does not regarding their overall understanding of conservation of matter.

- H_2: There is a difference between students who receive instruction that incorporates physical modeling and students who receive instruction that does not regarding their logical thinking abilities.
- H_3: There is a difference between students who receive instruction that incorporates physical modeling and students who receive instruction that does not regarding their conceptual and algorithmic understanding of conservation of matter.

Research Design

This study employed a type of quasi-experimental research design commonly used in social research called the nonequivalent groups design. The nonequivalent groups design is structured similarly to a pre- and post-test randomized experiment but lacks a crucial component of randomized experimentation - random assignment. This means that for the study, the assignment to the treatment or control groups was not controlled via random assignment. Instead, classes of students that were already intact were randomly assigned to either treatment or control groups. Ideally, each class consisted of students that had much in common and were quite similar. However, groups may have been significantly different prior to the start of the study. For this reason, pre-test measures taken on the same outcome variables as the two post-tests were employed. Shadish, Cook, and Campbell (2002) describe the usefulness of pre-tests in this type of design, stating that they help to identify initial differences between the groups being compared, which allows for stronger causal inferences to be made. Figure 5 depicts a diagram summarizing the overall research design.

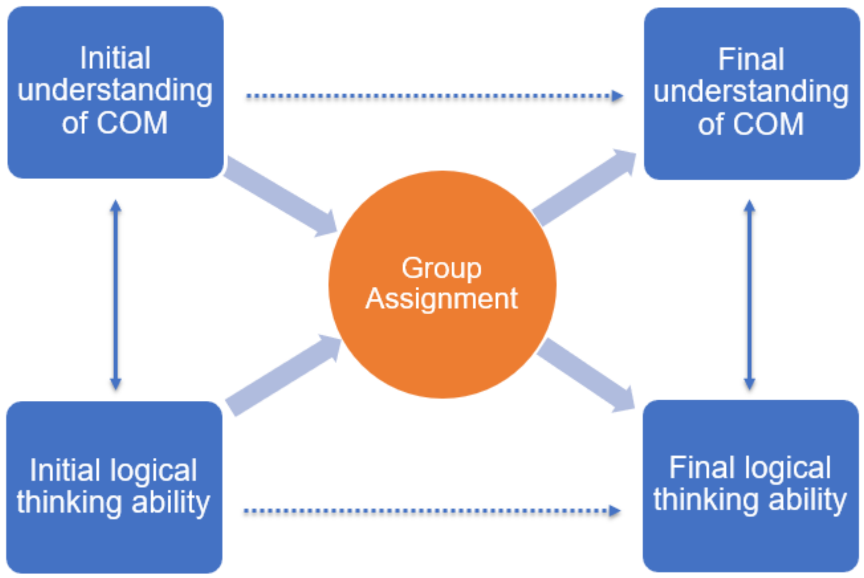


Figure 5. Research design diagram.

Variables

Independent variables. This study follows a quasi-experimental research approach with a nonequivalent control group design using pre-tests and post-tests. For this study, a 2 x 2 mixed design approach was employed. Mixed designs are designs that include at least one between-subjects variable (group assignment condition) in addition to at least one within-subjects variable. In this study, participants were divided into two groups. Thus, treatment condition (control and treatment groups) was the between-subjects factor in this study. Participants were also given pre-tests and post-tests, and these together served as the within-subjects factor.

Dependent variables. In the first research question, the dependent variable is understanding of conservation of matter. In the second research question, the dependent variable is logical thinking ability. The third research question probes specific types of understanding of conservation of matter – conceptual and algorithmic understanding. Figure 6 highlights the variables of the study.

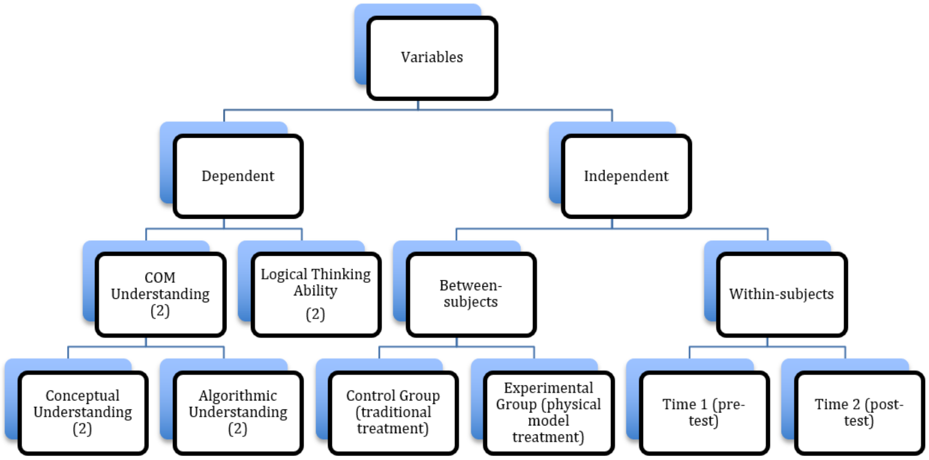


Figure 6. A hierarchical diagram highlighting variables of the study.

Note. (2) – two testing periods (*i.e.*, pre-test and post-test)

Materials Design

Conservation of Matter Lesson

Every participant, regardless of their group assignment condition received initial and follow-up lectures on Chemical Reactions and Reaction Stoichiometry (Chapter 3 of the course textbook), in which the idea of matter conservation and its application in chemistry was introduced and covered. Lecture material was comprised of the chapter PowerPoint provided by the publisher of the adopted textbook and was used in all classes. The treatment in this study were two types of enrichment activities and depended on group placement. These two treatments are discussed below.

Traditional Treatment

Traditional teaching methods are often teacher-directed and involve the use of standard materials such as worksheets, group work, textbooks, and demonstrations.

Traditional activities used in this study included one worksheet on counting atoms and one worksheet on balancing chemical equations. An overview of the traditional curriculum can be found in Appendix H.

Physical Modeling Treatment

To probe the incredible potential that physical models have in relaying abstract concepts to students in science, a simple physical model comprised of three sequential activities used to scaffold the difficult learning tasks of counting atoms and balancing chemical equations was utilized in the treatment group (see Appendix G). During development of the physical modeling activities, a panel of 14 elementary, junior high, high school, and college science educators participated in a pilot study to help ensure the validity of the model and correct usage of terminology.

The physical model used colored chips to represent atoms of different elements that were placed in designated areas on a template. Each activity was followed up with informal assessments and discussion questions to incite thought and clear misconceptions that may have been introduced. As Levy Nahum, Hofstein, Mamlok and Bar-Dov (2004) suggest, 'in chemistry, almost all models are metaphorical models' (p. 303). Therefore, special effort was made to make sure that the features of the physical model used in this study were not interpreted literally but symbolically. Through discussion, students may have gained further appreciation for the use of models in science while also becoming aware of the limitations.

In activity 1 (25 minutes), students were presented with chemical formulas and were challenged with counting the total number of atoms, taking into consideration coefficients and subscripts, Figure 7.

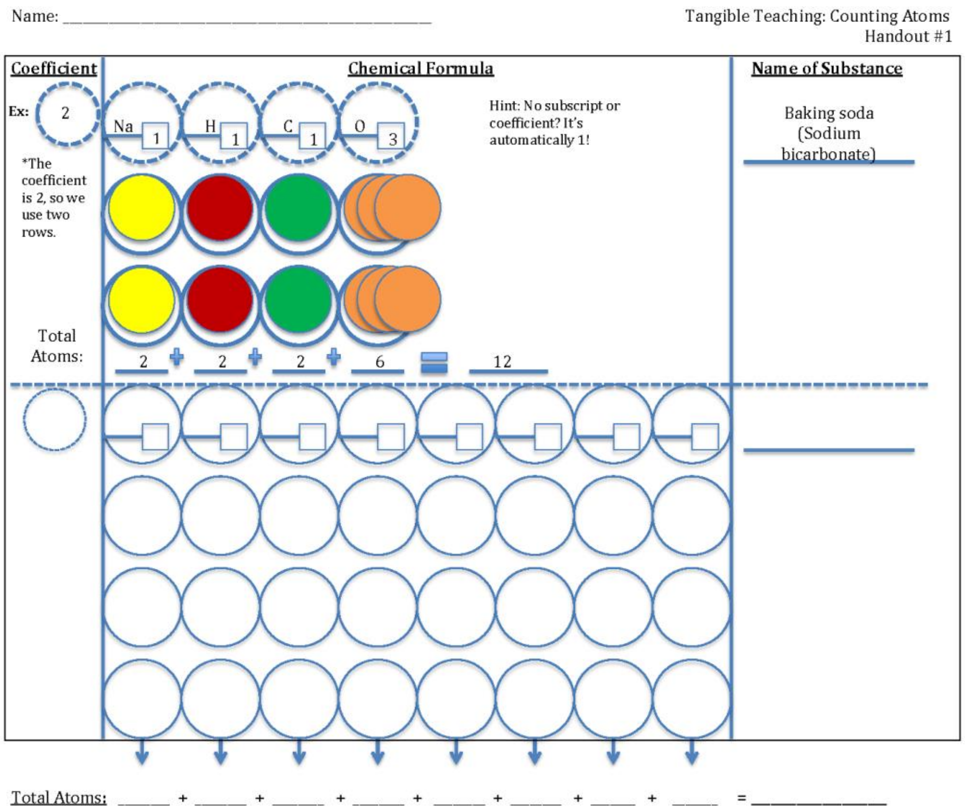


Figure 7. Template used for Activity 1.

Activity 2 (25 minutes) required students to count atoms for both sides of an entire chemical equation, as opposed to a simple chemical formula. Activity 2 presented more of a challenge because students were now presented with multiple substances present on both sides of a chemical reaction. Activity 2 ended with students being able to recognize if a chemical equation was balanced or not but did not require them to take any further steps. In Activity 3, students brought together everything that had been introduced and reinforced in Activities 1 and 2 to take on the final task of balancing entire chemical equations. The template used for Activity 2 and Activity 3 are shown in Figure 8.

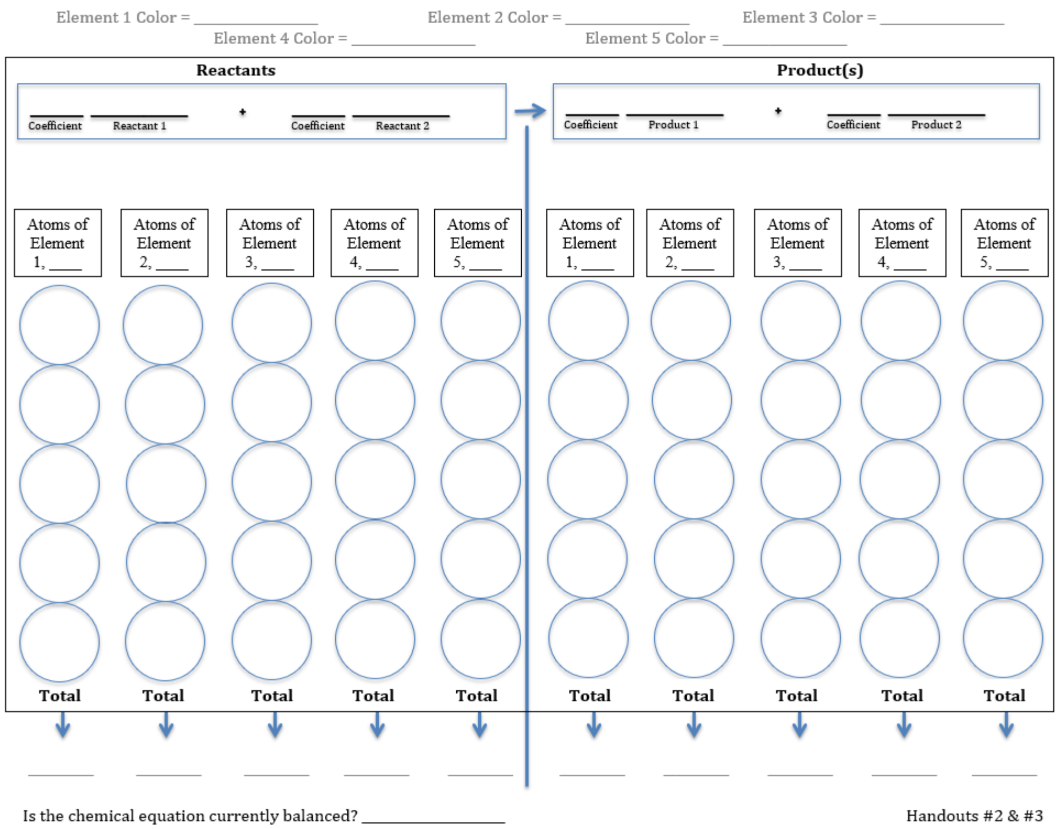


Figure 8. Template used for Activities 2 and 3.

In addition, formative assessments were integrated into each activity to guide the instruction and feedback of the teacher. The use of these formative assessments was backed by research conducted on learning progressions and sequencing that investigate how students develop their understanding and ability to use practices and disciplinary core ideas with time (NRC, 2012; Merritt & Krajcik, 2013). Teachers are encouraged to create formal assessments appropriate for their classroom. Recently published literature (Bybee, 2013; Lyon 2013; NRC, 2014) recommends that assessments be developed along each of the three dimensions of the *NGSS* and that students have numerous opportunities to demonstrate their knowledge.

Procedures

Sampling Procedure

Following approval from the Institutional Review Board at the University of Southern Mississippi, sixty-eight undergraduate students currently taking General Chemistry I at a public community college in a southeastern state of the United States participated in this study.

To attain the sample, the researcher solicited the participation of the community college by contacting the institutional research offices with a letter of introduction and completed all applications for study approval. The researcher determined when to administer the pre-tests and post-tests. Following consent, any student enrolled in General Chemistry I in the Spring 2018 academic semester was eligible for participation.

Regarding anonymity and confidentiality, each student was assured that their real name would not be collected nor revealed in any presented form of the research results. In addition, as participation in the study was not part of the curriculum, each student was informed that participation was not mandatory.

Experimental Procedure

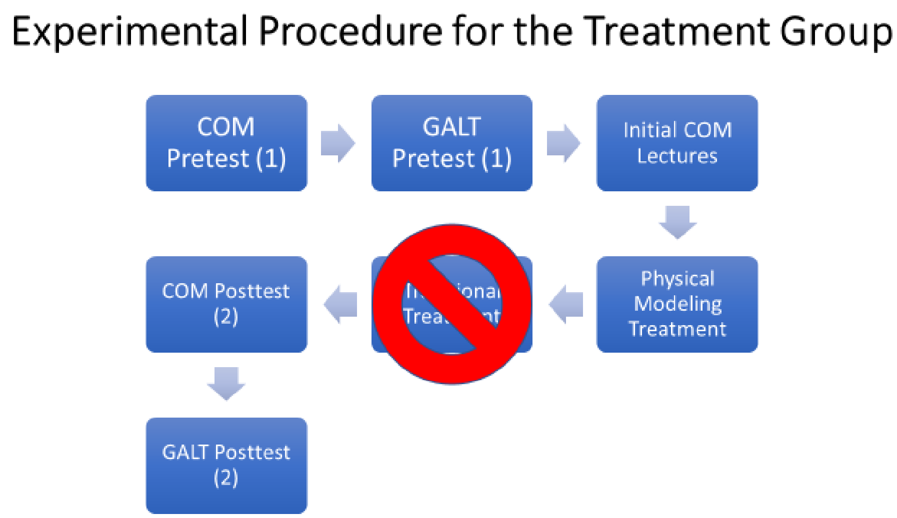


Figure 9. Chain diagram showing the experimental procedure for the treatment group.

Note. COM = Conservation of Matter; GALT = Group Assessment of Logical Thinking; (1) – pre-test; (2) – post-test.

Two separate learning sessions per group were held, which took place over the course of a week during the regular laboratory time. Rather than employing random assignment of students to determine placement in either the treatment or control groups, intact groups of students were used according to their respective scheduled lab times. Therefore, the design of this experiment fell under the scope of a non-equivalent groups design which does not employ random assignment as a tactic in which to safeguard equivalence between the two groups. The fidelity of the treatment was maintained through fully structured physical modeling activities.

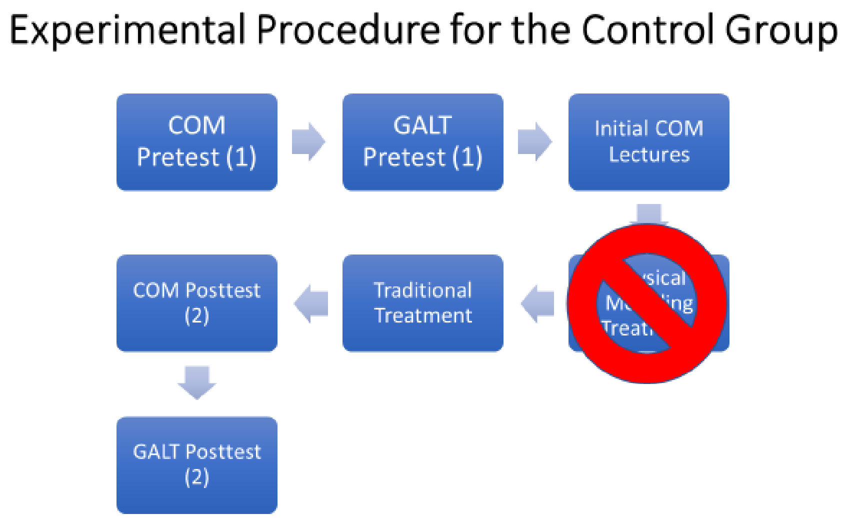


Figure 10. Chain diagram showing the experimental procedure for the control group.

Note. COM = Conservation of Matter; GALT = Group Assessment of Logical Thinking; (1) – pre-test; (2) – post-test.

Figures 9 and 10 summarize the experimental flow for the treatment and control groups, respectively. After consenting and completing a short demographic questionnaire, participants completed the COM pre-test (1) and GALT pre-test (1) to assess their initial understanding of conservation of matter and their current level of logical thinking ability, respectively. Afterwards, participants in both groups received instruction on learning tasks related to conservation of matter via lectures. Following instruction, participants in the treatment group received additional practice and enrichment through physical modeling while participants in the control group received additional practice and enrichment through traditional activities in the form of practice worksheets. Next, all participants took a COM post-test (2) in addition to a GALT post-test (2).

Instrumentation

Two instruments were used for data collection in this study. They included the Group Assessment of Logical Thinking (GALT) questionnaire (see Appendix E), and a Conservation of Matter questionnaire (see Appendix D).

Group Assessment of Logical Thinking (GALT)

The logical thinking ability of each student was assessed with the Group Assessment of Logical Thinking (Roadranka, Yeany, & Padilla, 1983). The GALT has been used in a number of research studies to successfully measure logical thinking ability of students (Bird, 2010; Bitner, 1991; Brunce et. al., 1993; Fah, 2009; McConnell et. al, 2005).

To account for limitations found in other logical thinking tests, Roadranka et al. (1983) developed the GALT, a Piagetian test of logical thinking that integrates the six reasoning nodes identified as essential abilities for success in advanced science and mathematics courses. One node of reasoning is concrete operational. The remaining five nodes of reasoning are formal operational in nature (Brunce et al., 1993). The GALT does not measure logical thinking ability below the concrete operational level, but instead identifies subjects as either concrete, transitional, or formal operational thinkers. Roadranka et al. (1983) suggested that the original 21 items in the GALT be used instead as a pool of items, so that shorter tests could be utilized in consideration of time constraints. Their suggestion, which was used in this study, includes the following recommended 12 items and 6 subscales: conservation of matter, items 1 and 4; proportional reasoning, items 8 and 9; controlling variables, items 11 and 13;

probabilistic reasoning, items 15 and 16; correlational reasoning, items 17 and 18; and combinatorial reasoning, items 19 and 20.

The test format for the GALT, apart from the combinatorial reasoning questions, consists of an illustration of the problem and a corresponding multiple-choice response for both the correct answer and justification. Each combinatorial item is open response, requiring students to provide their own answers. The answers to the GALT items 1 to 18 were considered correct only if the student selected the appropriate answer and justification. For item 19, Roadrangka et al. (1983) explain that students "must (1) show a pattern and (2) have no more than one error or omission". Item 20 follows similar guidelines, requiring students to show a pattern while not having more than two errors or omissions. For the abbreviated version of the GALT used in this study, classification of students as concrete, transitional, or formal thinkers was dependent on the following score criteria: (a) 0-4, concrete; (b) 5-7, transitional; (c) 8-12, formal.

The GALT was chosen as a measure of formal reasoning mainly due to validity and reliability results obtained by Roadranka et al. (1983) on a sample of 628 students from various grade levels ranging from sixth grade to college. The original 21-item GALT has a reported Cronbach's reliability coefficient for internal consistency of $\alpha = .85$ (Bitner, 1991). For the abbreviated GALT, Bunce and Hutchinson (1993) report the Cronbach's reliability coefficient for internal consistency as $\alpha = .62 - .70$.

Conservation of Matter Concept Questionnaire

To measure students' understanding of conservation of matter, a multiple-choice instrument developed at the Harvard-Smithsonian Center for Astrophysics was used. The construction of this 25-item instrument relied on numerous student interviews and answers to open-ended essay questions (Sadler & Schwartz, 2004). As a result, each item included distractors that aimed at identifying common misconceptions that emerged from the data gathered during the development phase. During the development of the instrument, a pilot study using Rasch analysis was conducted to assure one-dimensionality of certain constructs (Agung & Schwartz, 2007). Agung and Schwartz (2007) state that this one-dimensionality examination was used to ensure the reliability and validity of the instrument. Numerous rounds of student interviews, analysis of open-ended essay questions, question refinement, and reviews from subject matter experts including chemistry teachers and graduate students helped to provide validity checks for each question of the questionnaire. Over the years, a few studies have used this instrument to measure students' understanding of conservation of matter, further validating this instrument (Agung & Schwartz, 2007; Asghar, 2005; Doucerain, 2009; Doucerain & Schwartz, 2010).

The conservation of matter questionnaire consists of 17 conceptual questions exploring conservation of matter through various chemical and physical changes (items 1, 3-14, 16-17, 24-25), 3 questions regarding the nature of matter (item 15, 18-19), and 5 questions algorithmic in nature that are related to stoichiometry and balancing chemical equations (items 2, 20-23). For the purposes of this study, only the items that tested conceptual and algorithmic understanding were used to answer research question three. Items 15, 18 and 19 were omitted on the premise that their content was neither conceptual nor algorithmic in nature and, therefore, said items would not accurately assess students' conceptual and algorithmic understanding of conservation of matter.

Data Collection

Although no random selection was possible in this setting, all grouped laboratory sections of students were randomly assigned to treatment conditions. To eliminate having to also consider the many differences that may exist between instructors, the researcher was the sole administrator of treatment for of both groups in this study. The treatment group made use of the physical model to explore conservation of matter, while the control group was presented with traditional practice materials.

This study took place over the course of the Spring 2018 academic semester. Each student participating in the study completed the initial pre-tests during their General Chemistry I lecture class time at the beginning of the unit. Each COM and GALT pre-test established a baseline for the participants.

For the physical model curriculum, the researcher carried out all experimental protocols with the groups receiving the proposed treatment by following the activity steps and procedures for the physical model activities as seen in Appendix G. Since the researcher took the place of the administrator for both the treatment and control groups, the need to control for individual instructors was eliminated, and lent itself to the ability of the researcher to make stronger causal inferences from the results.

In the treatment group sessions, each student participating in the study explored the idea of conservation of matter through physical modeling. The physical model activities engaged students in hands-on experiences that allowed them to observe how atoms are conserved and rearranged in chemical reactions. The two major components of the treatment intervention were a) a physical model that students can visualize and b) hands-on activities. The activities were developed by the researcher in alignment with the recently adopted *NGSS* with the goal of enhancing students' understanding of conservation of matter. In addition, the model was piloted to ensure clarity and appropriateness and further evaluated by subject matter experts.

For the laboratory sections that served as the control group, the students were enriched with traditional materials. In general, the researcher asked them to apply their knowledge by completing worksheet assignments. Students were permitted to work in small groups while doing so.

At the end of the study, each participant completed the questionnaires a second time (as post-tests) to measure individual changes in both cognitive level and understanding of conservation of matter in addition to group differences across the treatment and control groups.

Analytic Approach

After data screening, descriptive statistics such as standard deviations, students' mean scores, and frequencies of student responses were computed using SPSS 20 and organized in tables. Furthermore, the final scores of the GALT and COM measures were tallied by computing raw scores and the percentage of correct answers.

For data analysis, a 2 x 2 mixed analysis of variance (ANOVA) was used to address the research hypotheses. Group assignment conditions (treatment and control) served as the between-subjects factor, and time (pre-test and post-test) was the within-subjects factor in this study. To satisfy a desire to understand the nature of any interaction between the two factors on the dependent variables, follow up simple effects tests were conducted on significant interactions. Scores on the GALT and COM questionnaires served as dependent variables. All data was analyzed using SPSS 20.

Summary

In summary, this chapter describes the methodology used for the study which examined the extent to which the incorporation of physical modeling affected students' understanding of conservation of matter. The students were enrolled at a public community college in the southeastern United States and were currently taking Chemistry I. The researcher employed a nonequivalent groups pre-test post-test quasi-experimental research design which utilized questionnaire scores to investigate the three identified hypotheses. A quantitative approach was most appropriate in comparing levels of content knowledge between participants receiving two different types of treatment. Power analysis using G*Power dictated that a minimum sample size of 68 was required to detect a large effect size with the given variables, assuming an alpha of .05.

Participants were obtained from a public community college. While participation did not depend on program of study or major, participants needed to be currently enrolled in General Chemistry I. The experimental procedure consisted of the two pre-tests, lectures on the chapter which introduced conservation of matter and its application in chemical reactions, supplemental practice with either traditional activities or physical modeling, and the two post-tests. Each participant took part in the study which took place over the course of a month. During the final phase of the instructional unit, participants were randomly divided into two groups, each receiving additional instruction via either traditional methods or through physical modeling, and then took the post-tests.

Quantitative data generated by the pre-tests and post-tests were analyzed through a 2 x 2 mixed-model ANOVA with follow up simple effects tests, using SPSS 20. The experimental design employed was most appropriate for comparing the effectiveness of two instructional strategies.

Results of these analyses are presented in the next chapter. Final participants are discussed first, followed by the quantitative and descriptive results by each research question and respective hypothesis.

CHAPTER IV – RESULTS

Research Questions

This study sought to answer three questions. These three questions are listed below and are followed by a discussion of the instruments used to answer each of the research questions:

1. How does the use of physical modeling compare to traditional teaching methods regarding differences made in students' overall understanding of conservation of matter?
2. How does the use of physical modeling compare to traditional teaching methods regarding differences made in students' logical thinking abilities?
3. How does the use of physical modeling compare to traditional teaching methods regarding differences made in students' conceptual and algorithmic understanding of conservation of matter?

Description of the Instruments

Two instruments were used to measure students' understanding of conservation of matter and logical thinking ability: (a) the Conservation of Matter (COM) questionnaire (see Appendix A) and (b) the abbreviated Group Assessment of Logical Thinking (GALT) questionnaire (see Appendix B). The Conservation of Matter questionnaire comprised of 25 multiple-choice items was used to assess understanding of conservation of matter (Sadler & Schwartz, 2004). Following the item analysis of Doucerain (2009), the questionnaire contains 17 conceptual questions that probe conservation of matter through scenarios involving physical and chemical changes (items #1, 3-14, 16-17, and

24-25), 3 questions about the nature of matter (items #15, 18-19), and 5 algorithmic questions related to stoichiometry and balancing equations (items #2, 20-23). Relevant to this study are the conceptual and algorithmic subscales, which were used to answer the third research question.

A 12-item paper-and-pencil abbreviated Group Assessment of Logical Thinking (GALT) was used to provide a direct measure of logical thinking abilities (Roadranka, Yeany, & Padilla, 1983). The test format for the GALT consists of an illustration of the problem and a corresponding multiple-choice response for both the correct answer and justification. There are also open response questions, requiring students to provide their own answers. Based on raw score, students were categorized as either concrete (0-4), transitional (5-7), or formal operational (8-12).

Both instruments were administered as a pre-test and post-test during the beginning and end of the Chemical Reactions and Reaction Stoichiometry unit of the Spring semester of 2018 at a community college in a southeastern state in the United States. One month passed between the administration of the pre-tests and post-tests. Raw scores and average percent correct scores on each pre-test and post-test for each group were computed.

Description of the Sample

Seventy-five individuals consented to participate in the study, mostly aged between 18 and 19 years old. Participants missing key data (i.e., pre-test or post-test) were excluded from the study. Thus, only 62 of the 75 total participants had a sufficiently complete data set for inclusion in data analysis. Of the 62 students, the majority were female (n = 40, 65%) and the minority were male (n = 22, 35%). All participants were enrolled in General Chemistry I. The composition of the control group (n = 30) was also more female (n = 19, 63%) than male (n = 12, 37%). This trend was also observed in the treatment group (n = 32) which also had more females (n = 22, 69%) than males (n = 10, 31%).

Description of the Variables

The overall average percent correct score for all participants in the control group on the COM questionnaire was 37.5% (SD = 14%) on the pre-test and 41.7% (SD = 17.2%) on the post-test. COM scores increased within the control group from 37.5% to 41.7% from pre-test to post-test, but not significantly (p = .104). The treatment group performed better overall. For the treatment group, overall average score on the pre-test was 39.8% (SD = 14.7%) and 53.1% on the post-test (SD = 19%), a significant difference (p <.001). Despite the apparent group differences, both groups still performed poorly on the COM questionnaire. Table 2 presents an item analysis of the COM questionnaire and overall scores for both groups.

Of the twenty-five items on the COM questionnaire, 17 items tested conceptual understanding (Items 1, 3 – 14, 16 -17 and 24 – 25), and 5 items tested algorithmic understanding of conservation of matter (Items 2 and 20-23). The overall average percent correct scores for students were very low for both subscales and are shown in Table 3. For the 17 conceptual questions, the control group obtained an average score of 38.2% on the pre-test and 42.5% on the post-test and was not significantly different (p = .134). For the treatment group, average scores significantly increased from 39.2% to 51.8% (p < .001). For the 5 questions on the algorithmic understanding subscale, average percent correct scores for the control group rose ten percentage points from 28% to 38%, but not significantly (p = .075). For the treatment group, average scores also increased, from 32.5% to 47.5%, which was a significant change (p = .007).

In general, correlations between post-tests across each subscale were weak. However, a significant correlation was found when looking at how conceptual understanding post-test scores correlated with algorithmic understanding post-test scores (p = .011). This was surprising, considering the literature describes a disconnect between algorithmic and conceptual learning, suggesting that they are not correlated. Therefore, a second correlation analysis was conducted for each group separately, to further explore this significant correlation observed that is not present in the literature.

Follow-up correlations between post-test scores across each of the two subscales in the control group were weak. However, a similar analysis for the treatment group uncovered a strong correlation between post-test scores for each subscale. This finding suggests that the physical model treatment is likely the factor explaining the significant correlation between algorithmic and conceptual understanding. The results of the Pearson correlations are presented in Tables 4 and 5.

The GALT was used to measure and assess changes in logical thinking ability in the student sample. At the beginning of the study, the general student population was found to be 40.7% concrete, 46.5% transitional and 12.8% abstract. Table 6 displays the percentages of the students' various cognitive levels as measured by the GALT for pre-tests and post-test in both groups.

All test results were normally distributed. The COM pre-test had a skewness of .437 and kurtosis of -.231, and the COM post-test had a skewness of .150 and kurtosis of -.392, both well within the acceptable bounds for skewness of kurtosis. The GALT pre-test had a skewness of .517 and kurtosis of .327, while the post-test had a skewness of .298 and kurtosis of -.990, also well within the acceptable bounds.

Table 2

Percent Correct for All Items for COM Pre-test and Post-test

	Control Group (*N* = 30)		**Treatment Group (*N* = 32)**	
Question number	**Pre-test (%)**	**Post-test (%)**	**Pre-test (%)**	**Post-test (%)**
1 [c]	57	73	66	66
2 [a]	40	60	59	75
3 [c]	40	43	62	53
4 [c]	53	63	75	69
5 [c]	43	67	41	53
6 [c]	53	60	56	75
7 [c]	30	43	28	66
8 [c]	47	47	41	34
9 [c]	23	27	22	47
10 [c]	37	53	38	69
11 [c]	23	30	22	34
12 [c]	33	37	19	53
13 [c]	60	43	72	78
14 [c]	17	20	19	34
15	83	80	91	94
16 [c]	40	40	28	53
17 [c]	37	40	38	44
18	17	17	19	31
19	50	33	56	75
20 [a]	47	37	34	44
21 [a]	20	27	28	44
22 [a]	17	37	28	41
23 [a]	17	30	12	34
24 [c]	20	17	9	16
25 [c]	33	20	31	47
Overall Average Score	37.5	41.7	39.8	53.1

Note. [c]Conceptual understanding subscale; [a]Algorithmic understanding subscale.

Table 3

Overall Average Score by Subscale.

	Control Group (*N* = 30)		Treatment Group (*N* = 32)	
Subscale	**Pre-test (%)**	**Post-test (%)**	**Pre-test (%)**	**Post-test (%)**
Conceptual	38.2	42.5	39.2	51.8
Algorithmic	28.0	38.0	32.5	47.5

Note. As measured by the Conservation of Matter (COM) Questionnaire.

Table 4

Pearson Correlation on Subscales

		Conceptual pre-test	**Conceptual post-test**	**Algorithmic pre-test**	**Algorithmic post-test**
Conceptual pre-test	Pearson correlation	1	.680**	.065	.189
	p (two-tailed)		.000	.613	.141
Conceptual post-test	Pearson correlation	.680**	1	.115	.415**
	p (two-tailed)	.000		.372	.001
Algorithmic pre-test	Pearson correlation	.065	.115	1	.220
	p (two-tailed)	.613	.372		.086
Algorithmic post-test	Pearson correlation	.189	.415**	.220	1
	p (two-tailed)	.141	.001	.086	

Note. **Correlation is significant at the 0.01 level (two-tailed), *Correlation is significant at the 0.05 level (two-tailed).

Table 5

Pearson Correlation on Subscales by Group

Group			**Conceptual pre-test**	**Conceptual post-test**	**Algorithmic pre-test**	**Algorithmic post-test**
Control	Conceptual pre-test	Pearson correlation	1	.740**	.010	.240
		p (two-tailed)		.000	.956	.201
	Conceptual post-test	Pearson correlation	.740**	1	.002	.341
		p (two-tailed)	.000		.991	.065
	Algorithmic pre-test	Pearson correlation	.010	.002	1	.215
		p (two-tailed)	.965	.991		.254
	Algorithmic post-test	Pearson correlation	.240	.341	.215	1
		p (two-tailed)	.201	.065	.254	
Treatment	Conceptual pre-test	Pearson correlation	1	.657**	.117	.132
		p (two-tailed)		.000	.522	.473
	Conceptual post-test	Pearson correlation	.657**	1	.172	.443*
		p (two-tailed)	.000		.347	.011
	Algorithmic pre-test	Pearson correlation	.117	.172	1	.192
		p (two-tailed)	.522	.347		.291
	Algorithmic post-test	Pearson correlation	.132	.443*	.192	1
		p (two-tailed)	.473	.011	.291	

Note. **Correlation is significant at the 0.01 level (two-tailed), * Correlation is significant at the 0.05 level (two-tailed).

Table 6

Student Logical Thinking Ability Domain Percentages

Domain	**Control Group (*N* = 30)**		**Treatment Group (*N* = 32)**	
	Pre-test (%)	**Post-test (%)**	**Pre-test (%)**	**Post-test (%)**
Concrete Operational	53.3	50.0	28.1	40.6
Transitional	36.7	26.7	56.3	18.8
Formal Operational	10.0	23.3	15.6	40.6

Note. As measured by the Group Assessment of Logical Thinking (GALT).

Data Analysis Summary

Analysis Overview

A Two-Way Mixed ANOVA was conducted to examine the effects of physical modeling strategy on understanding of conservation of matter and logical thinking ability in community college chemistry students. Scores for both measures at various times before (pre-test) and after (post-test) treatment were compared for two grouping assignments (control and treatment). Thus, the design is a 2 x 2 factorial design where Group is a between-subjects factor and Time is a within-subjects factor. A mixed-model ANOVA was chosen over other individual analyses due to increased protection against Type I error, and to thoroughly examine any significant interactions that may be present in the data. To avoid conflict, a separate mixed-model ANOVA was run while investigating the third research question since the subscales involved were a part of a measure already included the model. Simple effects tests were run as follow-up analyses to further examine significant interactions. A summary of these results is show in in Table 7.

Due to the lack of random assignment and the inclusion of two separate measures as dependent variables, a Box's Test of Equality of Equality of Covariance Matrices was conducted. The Box's Test was not significant, which indicated a reduced risk of Type I error. In addition, to satisfy the assumption of homogeny of variance of each group at each level of the within-subjects variable (pre-test and post-test) Levene's Test was conducted. Levene's test was not significant for any of the dependent variables.

Table 7

Summary of Results of Mixed Design Analysis of Variance (ANOVA)

Within-Subjects Effects	**Measure**	**df**	**Mean Square**	***F***	***p***	**Partial Eta Squared**
Time	COM	1/60	150.594	24.060	.000	.286
	GALT	1/60	1.539	.878	.353	.014
	Conceptual	1/60	64.643	18.466	.000	.235
	Algorithmic	1/60	12.097	10.557	.002	.150
Time * Group	COM	1/60	40.153	6.414	.014	.097
	GALT	1/60	.249	.142	.708	.002
	Conceptual	1/60	15.675	4.478	.038	.069
	Algorithmic	1/60	.484	.422	.518	.007

Between-Subjects Effects	**Measure**	**df**	**Mean Square**	***F***	***p***	**Partial Eta Squared**
Group	COM	1/60	90.487	3.195	.079	.051
	GALT	1/60	40.216	3.432	.069	.054
	Conceptual	1/60	23.316	1.241	.270	.020
	Algorithmic	1/60	3.794	2.228	.141	.036

Note. Time - Pre-test & Post-test; Group - Control Group & Treatment Group; COM - Conservation of Matter Questionnaire; GALT - Group Assessment of Logical Thinking; Conceptual - Conceptual Understanding subscale; Algorithmic - Algorithmic Understanding subscale.

Analysis of Data

The researcher explored three research questions and three corresponding research hypotheses. Each of the three research questions were quantitative in nature and dealt with differences in overall understanding of conservation of matter, logical thinking ability, and conceptual and algorithmic understanding of conservation of matter, respectively. The data were collected in relation to each of the questions and hypotheses posed below.

Research Question 1: Differences Made in Overall Understanding

Regarding this area of examination, the researcher asked the following question concerning differences made in overall understanding of conservation of matter:

How does the use of physical modeling compare to traditional teaching methods regarding differences made in students' overall understanding of conservation of matter?

Hypothesis 1 predicted that there is a difference between students who receive instruction that incorporates physical modeling and students who receive instruction that does not regarding their overall understanding of conservation of matter.

The results of the Two-Way Mixed ANOVA showed that there was no significant main effect of Group ($F(1, 60) = 3.195$, $p = .079$, $\eta p^2 = .051$) on Conservation of Matter raw scores, with the control group ($M = 9.90$) and treatment group ($M = 11.609$) performing similarly overall.

There was, however, a significant main effect of Time for the Conservation of Matter questionnaire ($F(1, 60) = 24.060$, $p < .001$, $\eta p^2 = .286$), with students having significantly higher post-test scores ($M = 11.857$) over pre-test scores ($M = 9.652$).

There was a significant interaction between Group and Time ($F(1, 60) = 6.414$, $p = .014$, $\eta p^2 = .097$). As depicted in Figure 11, examination of the cell means indicated that while the control group showed improvement in raw scores from pre-test ($M = 9.367$, $SD = 3.557$) to post-test ($M = 10.433$, $SD = 4.376$) following treatment; the treatment group showed even larger overall score gains in response to treatment from pre-test ($M = 9.938$, $SD = 3.741$) to post-test ($M = 13.281$, $SD = 4.814$).

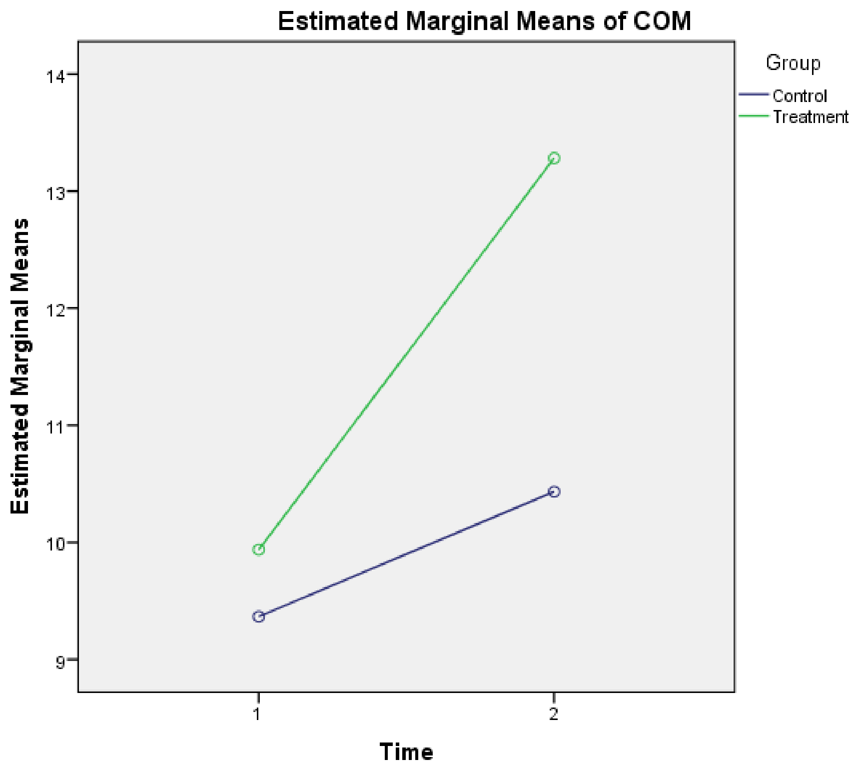


Figure 11. Time x Group interaction graph for the Conservation of Matter (COM) measure.

A follow up simple effects test was conducted to further determine the nature of the significant interaction. Simple main effects analysis showed that conservation of matter score changed significantly in response to treatment for the treatment group ($p < .001$), but not the control group ($p = .104$).

Simple effects tests also showed that before treatment there were no significant group differences, $F(1, 60) = .38$, $p = .541$, but after treatment there was a statistically significant difference between groups, $F(1, 60) = 5.92$, $p = .018$.

These findings support the notion that physical modeling can be used as a powerful tool to effectively teach abstract concepts such as conservation of matter.

Research Question 2: Differences Made in Logical Thinking Ability

Regarding this area of examination, the researcher asked the following question concerning differences made in logical thinking ability:

How does the use of physical modeling compare to traditional teaching methods regarding differences made in students' logical thinking abilities?

Hypothesis 2 predicted that there is a difference between students who receive instruction that incorporates physical modeling and students who receive instruction that does not regarding their logical thinking abilities.

The results of the Two-Way Mixed ANOVA showed that there was no significant main of effect of Group ($F(1, 60) = 3.432$, $p = .069$, $\eta p^2 = .054$) on logical thinking ability scores, with the control group ($M = 4.767$) and treatment group ($M = 5.906$) comprised of mostly transitional thinkers.

In addition, there was no significant main effect of Time on logical thinking ability scores ($F(1, 60) = .878$, $p = .353$, $\eta p^2 = .014$), with students showing similar average GALT scores for the pre-test ($M = 5.225$) and post-test ($M = 5.448$).

As with each main effect, the interaction between Group and Time was not significant, $F(1, 60) = .142$, $p = .708$, $\eta p^2 = .002$. As depicted in Figure 12, neither the control nor treatment group attained significant gains in GALT scores from pre-test to post-test.

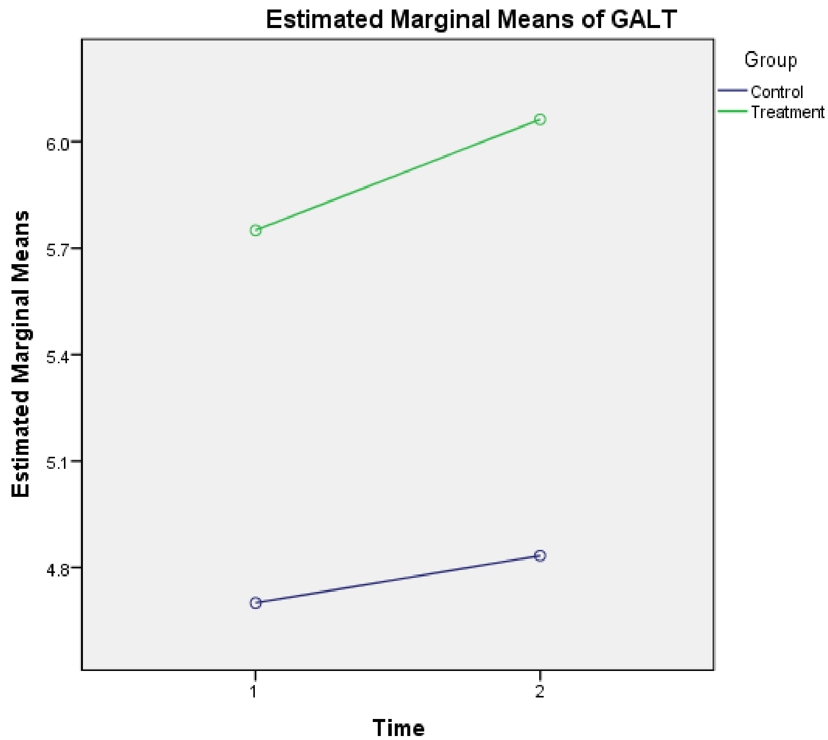


Figure 12. Time x Group interaction graph for the Logical Thinking Ability (GALT) measure.

Because no significant interaction was found, no follow-up tests were run. These findings suggest that the physical model used in this study was not effective in moving students towards formal operational thought. A possible explanation may exist in the relatively short amount of time between the implementation of the physical model and the measurement of logical thinking ability.

Research Question 3: Differences Made in Conceptual and Algorithmic Understanding

Regarding this area of examination, the researcher asked the following question concerning differences made in conceptual and algorithmic understanding of conservation of matter:

How does the use of physical modeling compare to traditional teaching methods regarding differences made in students' conceptual and algorithmic understanding of conservation of matter?

Hypothesis 3 predicted that there is a difference between students who receive instruction that incorporates physical modeling and students who receive instruction that does not regarding their conceptual and algorithmic understanding of conservation of matter.

When comparing mean raw scores in pre-test and post-test conditions for the conceptual and algorithmic subscales of the conservation of matter questionnaire, significant differences were found in the pre-test (M = 6.578) and post-test raw scores (M = 8.023) for the seventeen conceptual questions ($F(1, 60) = 18.466$, $p < .001$, $\eta p^2 = .235$) and in the pre-test (M = 1.513) and post-test raw scores (M = 2.137) for the five algorithmic questions ($F(1, 60) = 10.557$, $p = .002$, $\eta p^2 = .150$). Thus, a significant main effect of Time was observed.

In addition, there was not a significant main effect of group assignment condition for neither the conceptual ($F(1, 60) = 1.241$, $p = .270$) nor the algorithmic subscales ($F(1, 60) = 2.228$, $p = .141$).

A significant Time x Group interaction effect was found for the conceptual subscale ($F(1, 60) = 4.478$, $p = .038$, $\eta p^2 = .069$), but not for the algorithmic subscale ($F(1, 60) = .422$, $p = .518$).

As depicted in Figure 13, examination of the cell means shows larger increases in scores on conceptual questions for the treatment group from pre-test ($M = 6.656$) to post-test ($M = 8.813$), in comparison to increases in scores for the control group from pre-test ($M = 6.500$) to post-test ($M = 7.233$). Before treatment, there was not much difference in conceptual understanding raw scores between control ($M = 6.500$) and treatment groups ($M = 6.656$). However, after treatment, the control group ($M = 7.233$) had much lower conceptual understanding raw scores than the treatment group ($M = 8.813$). Figure 14 shows the nonsignificant interaction graph for the algorithmic subscale of the COM questionnaire.

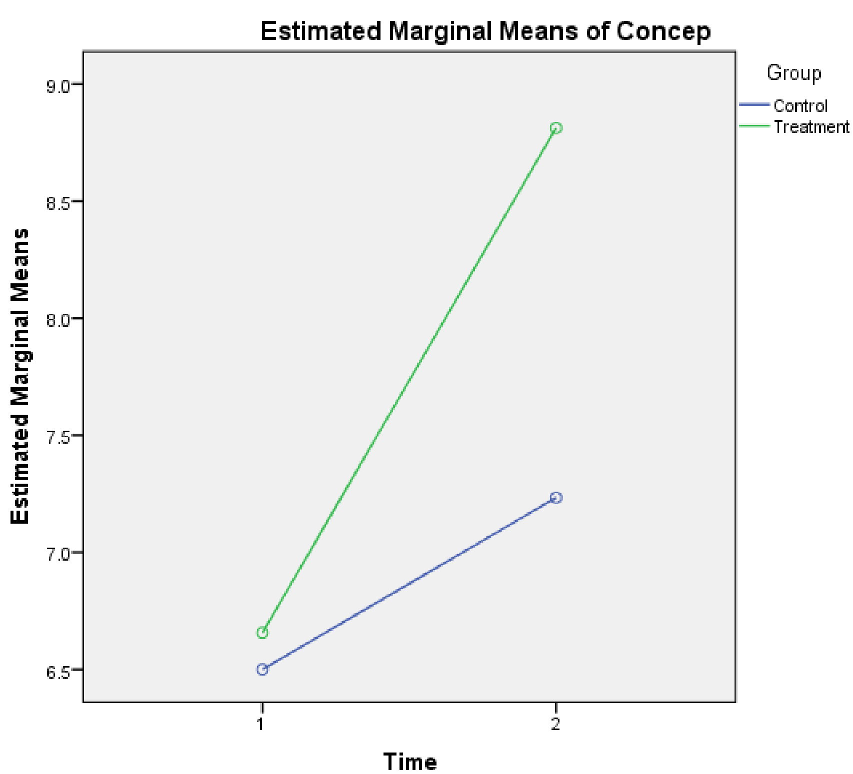


Figure 13. Time x Group interaction graphs for the conceptual subscale of the Conservation of Matter (COM) questionnaire.

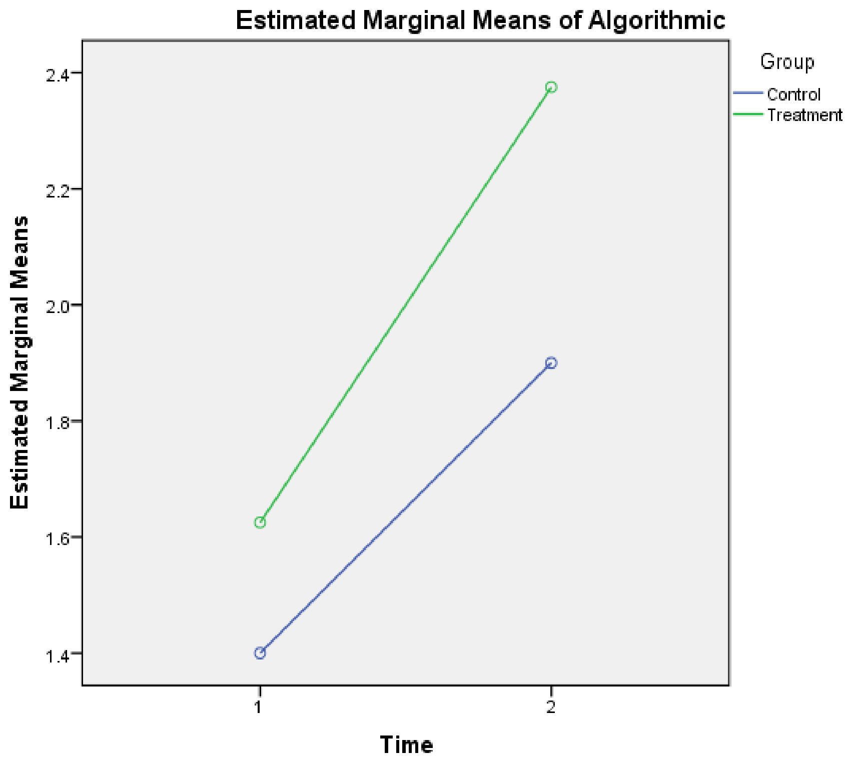


Figure 14. Time x Group interaction graphs for the algorithmic subscale of the Conservation of Matter (COM) questionnaire.

Follow-up simple effects tests were conducted to further explore the nature of the significant interaction regarding conceptual understanding of conservation of matter. Simple effects analysis showed that conceptual understanding changed significantly as a result of treatment for the treatment group ($F(1, 60) = 21.25$, $p < .001$), but not for the control group ($F(1, 60) = 2.30$, $p = .134$). Furthermore, although no significant group differences were observed before treatment ($F(1, 60) = .04$, $p = .845$), group differences after treatment were approaching significance ($F(1, 60) = 3.08$, $p = .084$).

These findings support the notion that physical modeling could be used to fill gaps identified by previous researchers concerning the paucity of empirically proven instructional interventions that can be used to increase the often-neglected conceptual understanding goal of student learning.

The following chapter concludes with a discussion of the findings and their implications on integrating physical modeling into the chemistry classroom. Limitations and suggestions for further research to help build and expand on the findings of this research study are also given.

CHAPTER V – DISCUSSION

This chapter discusses the findings, conclusions and implications of this study. It is organized into four major sections: (a) a summary of the study, (b) conclusions and discussions of the findings that can be drawn from the results of each research question, (c) implications for practice, (d) limitations of the study, and (e) suggestions for further research.

Summary of the Study

Learning chemistry often requires students to imagine and visualize structures and reactions on particulate and sub-particulate levels. This can be difficult for many students, particularly those who learn best in a visual or tactile way. Instructional manipulatives used as physical models of unobservable phenomena can be a great solution for students to study chemistry concepts that require them to 'see' the structure or process to fully understand it.

Recent educational reform has shifted the pedagogy of science education away from traditional instruction towards a more interactive approach. Because conservation of matter is so fundamental in understanding a wide array of concepts across multiple science disciplines, educators need empirical evidence to further determine what type of instruction leads to the greatest growth in content knowledge of this concept. This need led to the formulation of the following three research questions:

1. How does the use of physical modeling compare to traditional teaching methods regarding differences made in students' overall understanding of conservation of matter?

2. How does the use of physical modeling compare to traditional teaching methods regarding differences made in students' logical thinking abilities?
3. How does the use of physical modeling compare to traditional teaching methods regarding differences made in students' conceptual and algorithmic understanding of conservation of matter?

The study took place at a public community college in the southeastern United States. The researcher conducted all research sessions in person. The pre-test and post-test questionnaires used in data collection were paper-based. All necessary materials including pens, pencils, and dry-erase markers were provided for students who did not have or did not wish to use their own.

The researcher recruited community college students currently enrolled in General Chemistry I via in-person solicitation. No incentive was offered for participation in this study.

Seventy-five students participated, and 13 were eliminated on the basis that each participant had to be concurrently enrolled in the lecture and laboratory sections of General Chemistry I for the duration of the entire study. The remaining 62 students provided usable data sets for analysis. Of the 62 students, 40 were female and 22 were male.

To preserve anonymity, participants were assigned a random code to be used as a pseudonym. While general instruction and the completion of the pre-tests and post-tests occurred during normal lecture times, the intervention was administered during the designated laboratory period. Each of the four laboratory sections was randomly assigned to either the control or treatment groups. At the beginning of the Chemical

Reactions and Reaction Stoichiometry unit that introduced conservation of matter, the researcher explained the study to the participants and distributed information sheets about the study and a consent form, with the researcher's and the chair's contact information (see Appendix C). Following consent, the COM and GALT pre-tests were administered. After the pre-test, instruction continued as usual over the course of two weeks. During the third week, participants in the control and treatment groups were exposed to the intervention for the entire 1 hour and 50 minutes of the laboratory period. The COM and GALT post-tests were administered during each students' lecture class a week later. Although extra time outside of the 50-minute lecture class time allotment was not needed, students were assured that they would not be time constricted while taking the pre-test and post-tests. No major difficulties, such as campus closures due to bad weather occurred during the data collection process.

The researcher analyzed the results of the pre-tests and post-tests using SPSS 20, through a mixed-model ANOVA with follow-up simple effects tests on significant interactions. A mixed-model ANOVA compares the mean differences between two groups that are split on two independent variables, where one independent variable (or factor) is referred to as "between-subjects" factor and the other is referred to as 'within-subjects". In this study, Group was the between-subjects factor and Time was the within-subjects factor. COM, GALT, conceptual understanding, and algorithmic understanding scores for both pre-tests and post-tests served as the dependent variables. These analyses revealed answers for the proposed researched questions.

Research Question One: How does the use of physical modeling compare to traditional teaching methods regarding differences made in students' overall understanding of conservation of matter? Initial data analysis revealed that there were no overall differences in COM mean scores between the control and treatment groups. However, the researcher did find significant differences in mean scores between pre-tests and post-tests in general, suggesting that both interventions were effective in increasing students' understanding of conservation of matter. Further analysis revealed a statistically significant interaction effect between the effects of Time and Group on conservation of matter mean scores. Simple main effects analysis revealed the nature of this significant integration; there was a significant increase in COM scores for the treatment group, but not for the control group. In addition, it was observed that before the intervention, there were no significant group differences between groups, but after treatment there was a statistically significant difference observed between groups. This lack of significant group differences before the implementation of the intervention supports the internal validity of the study and suggests that despite the lack of random assignment, the groups were both normally distributed.

Research Question Two: How does the use of physical modeling compare to traditional teaching methods regarding differences made in students' logical thinking abilities? No significant differences existed in logical thinking ability between participants in the control and treatment groups. Thus, the physical model intervention was not effective in developing students' logical thinking abilities.

Research Question Three: How does the use of physical modeling compare to traditional teaching methods regarding differences made in students' conceptual and algorithmic understanding of conservation of matter? When comparing mean raw scores in pre-test and post-test conditions for the conceptual and algorithmic subscales of the conservation of matter questionnaire, significant differences were found in the pre-test and post-test scores for both the seventeen conceptual questions and five algorithmic questions. Further analysis showed a significant change in conceptual understanding scores from pre-test to post-test for the treatment group, but not for the control group.

Conclusions and Discussions

The results of the quantitative analysis led to several possible conclusions. In terms of a simple score analysis, there was a significant difference between pre-test and post-test scores overall for the COM questionnaire, but not for the GALT. This suggests that the collective experiences of the traditional and physical modeling treatments increased students' understanding of conservation of matter but were not effective in developing students' logical thinking abilities.

Furthermore, a significant Time x Group interaction effect was observed for the COM questionnaire, but not for the GALT. Follow-up up simple effects tests further supported the superior effectiveness of the treatment intervention by showing that COM raw scores changed significantly because of treatment for the treatment group but not for the control group. In consideration of the lack of random assignment in this study, the researcher was also pleased to see that before treatment (on the pre-test), there was not a significant difference between raw scores of the control (M = 9.367) and treatment groups (M = 9.938). After treatment (on the post-test), however, a statistically significant difference between the control group raw scores (M = 10.433) and treatment group raw scores (M = 13.281) was observed. Thus, further supporting the claim that the observed score increases could be confidently attributed to differences in treatment and not other variables. These findings also corroborate the meta-analysis findings of Freeman et. al

(2014) which concluded that STEM students engaged in active learning (such as the physical modeling intervention examined in this study) achieve higher student performance over students who are strictly taught via traditional lecturing.

Agung and Schwartz (2007) suggest that future studies focus on instructional strategies that focus on the development of students' conceptual understanding of conservation of matter. This study provides a promising solution to this unique issue. Data obtained in this study suggests that the physical modeling treatment was successful in presenting a targeted approach in which educators can increase the often-neglected conceptual understanding of conservation of matter. When comparing mean raw scores in pre-test and post-test conditions for the seventeen conceptual and five algorithmic questions of the conservation of matter questionnaire, significant differences were found in the pre-test and post-test raw scores for both the conceptual and algorithmic subscales. This indicates that treatment was effective for both group assignment conditions in increasing understanding on both fronts. While there was no significant difference in the mean raw scores of each subscale between the control and treatment groups, a significant Time x Group interaction was observed for conceptual understanding of conservation of matter, but not for algorithmic understanding. In other words, conceptual understanding changed significantly because of treatment for the treatment group, but not for the control group. The observed tremendous increase in conceptual understanding for the treatment group was evidence that physical modeling was indeed successful in increasing students' conceptual understanding of conservation of matter, as measured by the COM questionnaire. However, as Harrison and Treagust (2001) indicated, the attainment of high scores on a questionnaire does not always truly mirror students' actual conceptual understanding. Therefore, the observed increased in students' conceptual understanding scores must be considered with caution.

In line with the findings of Agung and Schwartz (2007), a Pearson correlation showed that post-test scores for the conceptual and algorithmic subscales in the control group were not significantly correlated ($p = .065$). However, the researcher was surprised to find that post-test scores for the conceptual and algorithmic subscales in the treatment group were significantly correlated ($p = .011$). This was surprising, considering the literature describes a disconnect between algorithmic and conceptual learning, reporting that they are not correlated. As Nurrenbern and Pickering's (1987) study first showed with quantitative certainty, the ability to solve algorithmic problems is not equivalent to a conceptual understanding. Simply put, due to the general lack of correlation between conceptual and algorithmic understanding, achieving one does not necessarily mean achievement in the other. This issue manifests itself as a significant issue for science educators. One of the most difficult hurdles to overcome in science education is the reality that the best instructional approach is one in which one type of understanding is not neglected over the other. Sawrey (1990) follows up the work of Nurrenbern and Pickering with a bold proposition for further researchers. Realizing that teaching strategies that give attention to algorithmic understanding often do not address conceptual understanding simultaneously, he petitions for ways in which science educators can remedy this problem. This discrepancy between the study findings and the literature suggests that the physical modeling intervention was able to address both types of understanding simultaneously.

Implications

This study may have varying implications for classrooms instruction and pedagogy. The most obvious implication from this research is that four weeks of instruction including a single session of physical modeling is not enough to move students forward in terms of logical thinking ability. As Ashgar (2004) proposed, a more prolonged experience using concrete physical modeling as a scaffolding strategy to facilitate understanding of abstract scientific principles is needed to help bridge the gap between concrete and formal operational thought. While the physical modeling curriculum employed in the treatment group used scaffolding strategies, its use was limited to just one session, and not repeated sessions over a long period of time. This likely explains the lack of significant growth in logical thinking ability, as measured by the GALT. The exact length of time required to observe changes in logical thinking ability, however, is difficult to determine. In their study of college chemistry students, Bunce and Hutchinson (1993) determined that even a 13-week science course was not enough exposure to yield measurable improvement in logical thinking ability.

One important discovery by the researcher is that as McConnell (2003) pointed out, many instructors may wrongly assume that their undergraduate students operate at the highest order of logical thinking ability simply because they have not taken measures to know the cognitive levels of their students. In this study, the cognitive level of students was far below what one would expect. Overall, 41% of the sample population were functioning at the concrete operational level, 46% were transitional, and 13% were at the formal operational level at the time that the GALT pre-test was administered. At the time the GALT post-test was administered, 45% of the sample were concrete, 23% transitional, and 32% were found to be at formal operational thought.

When analyzing student logical thinking ability across groups, the findings were even more interesting. While overall differences in logical thinking ability over time were not significant ($p = .107$), a clear shift from transitional thought to formal operational thought was apparent. During pre-test administration, a large percentage of the study population were in fact labeled as transitional (37% of the control group and 56% of the treatment group), which describes the domain between concrete and formal operational thought. While the control group saw a small reduction in transitional thinkers from pre-test to post-test (37% to 27%) and a small increase in formal operational thinkers from pre-test to post-test (10% to 23%), the treatment group saw percentages of transitional thinkers reduced from 56% to 19%. Unfortunately, not all students were moved forward towards formal operational thought. While the treatment group did see a substantial increase in formal operational thinkers (16% to 41%), the percentage of concrete thinkers also went up (28% to 41%). These findings suggest that the physical modeling helped to move some students to formal operational thought but may have also handicapped other students and caused them to regress back to the concrete domain due to an overreliance of concrete examples. Perhaps these students did not have the required prerequisite knowledge necessary to move forward. Considering these findings, a crucial need becomes apparent. Science educators must devise effective curriculum that caters to not one but all forms of thinkers in the classroom. While the physical model used in this study was effective, it appears to have not only boosted logical thinking ability from the concrete to formal operational stages for some students, but also promoted a transition away from formal operational thought back to the concrete level for others. A plan in which to incorporate the physical model effectively must be devised if it is to be used to bridge the transition between concrete and formal operational thought. Therefore, further research needs to be conducted to explore the cause of this bidirectional shift in logical thinking ability over time.

In addition, this research supported the notion that a more constructivist approach to teaching plays an important role in building knowledge. While the specific aspect of the physical model intervention that contributed to the significant changes in students' understanding of conservation of matter may be unknown, social interaction cannot be prematurely ruled out as a possible contributor considering the amount of discussion between students the physical model intervention incited. A study done by Doucerain and Schwartz (2010) highlighted the importance of discourse and argumentation in shaping cognitive processes. However, the authors acknowledge that any conclusions suggesting that language and dialog play a role in shaping students' ideas must be made with caution pending supporting empirical evidence regarding mechanisms underlying how exactly discourse enhances learning.

College educators may find that effective use of physical models within the classroom can present some inherent difficulties. One difficulty exists in student resistance to using the physical model itself. Gray, Owens, Steer, McConnell, and Knight (2011) agree that using physical models is not a common experience for most undergraduate students. As a result, students either refused to engage with the models or played with them inappropriately. However, the researchers found that these initial observations did not tell the entire story. Repeated use of the modeling activities resulted in the students eventually accepting and actively engaging in the hands-on activities. Gray et. al. (2011) make an interesting point regarding some students' preference for a more traditional instructional approach that may prove useful as future studies are designed:

> Preference for a traditional didactic lecture may reflect a desire to remain within a familiar pedagogy that has worked for them in the past, but our data suggest that continued use coupled [with] clear instructions and formative feedback will eventually reduce or eliminate such resistance. (p. 21)

Another hinderance is that educators may find that the use of physical models during lecture takes up too much time to do within the confines of the allotted lecture period. Fortunately, in this study the intervention was able to be administered during the much longer laboratory period.

As Gray et. al (2011) point out, educators must be sure to explain the model and clearly describe its use to students in order to minimize not only resistance but also other complications such as the reinforcement of misconceptions. One common problem experienced when using the models was students' tendency to incorrectly assume certain qualities of the physical models. For example, students interpreted the models literally and assumed that all atoms were spherical in shape, equal in size, or bonded to other atoms solely in a linear fashion. Students must be reminded that while models are useful as tools to visualize unobservable phenomena, they are not always perfect representations are not meant to be interpreted literally.

Limitations

Each participant received treatment only once within a 1 hour and 50-minute time block during his or her scheduled chemistry laboratory section. This proved to be a serious limitation because it severely restricted the amount of exposure students had to either the traditional or physical modeling instructional strategies.

Suggestions for Further Research

While there is no research in literature supporting differences in gender or ethnicity regarding understanding of conservation of matter, a follow-up study should be conducted to further assist in evaluating if the intervention had an effect that differs by various demographic variables, such as gender, ethnicity, or socioeconomic status.

Other age groups should also be explored to determine if age is a factor in the effectiveness of using physical modeling in the learning of conservation of matter. The incorporation of various age groups across different areas of the country could uncover interesting results and add to the findings of this study.

In addition, considering the importance of conservation of matter as a core concept in the sciences, future studies should be conducted that administer a delayed conservation of matter post-test at least thirty days after treatment to probe how well students in different treatment conditions are able to retain and apply their newly gained understanding over a prolonged period.

Suggestions for future research also involve students' logical thinking ability. The researcher suggests future studies of a longitudinal nature, to help determine how long-term exposure to physical modeling in teaching unobservable concepts affects logical thinking ability. Furthermore, by using GALT pre-test scores, the researcher may be able to isolate transitional students to more accurately assess the impact of physical modeling as a bridge to formal operation thought.

The physical model intervention was limited in a few ways. Most notable were the limitations to the size and coefficients of the chemical formulas in Activity 1 and the limitation of a maximum of two reactants and two products for each chemical equation used in Activity 2 and Activity 3. While the model was effective in increasing understanding of conservation of matter, the physical model could not address some difficulties experienced when balancing chemical equations such as the presence of numerous reactants and products and the use of fractional coefficients to balance chemical equations. Considering the inability of the physical model to handle complex chemical formula and equations, a redesign would be beneficial to increase usability for more advanced students.

Researchers may also find value in using qualitative approaches to compare teaching strategies used to teach conservation of matter. Considering the argument made by Harrison and Treagust (2001) that test scores are not always a perfect indicator of conceptual understanding, a qualitative study may provide a richer picture of students' understanding of conservation of matter. Furthermore, as shown in a mixed methods study done by Ashgar (2004), the use of case studies to evaluate various chemistry curricula have the potential to uncover truths and insight that quantitative methods cannot.

Last, the findings of this study are based upon students from one public community college in a southeastern state of the United States. A replication of this study that includes more institutions of higher learning in different parts of the country is needed to produce results that can be confidently generalized.

Summary

This chapter presented the findings, conclusions, and discussion of this study as well as implications for science educators and suggestions for future research. The purpose of this study was to investigate whether the proposed version of physical modeling helped community college students better understand the concept of conservation of matter, reflected by improved achievement on the Conservation of Matter questionnaire. In addition, because cognitive ability is theorized to play a significant role in understanding abstract concepts such as conservation of matter (Inhelder and Piaget, 1958), student logical thinking ability was also measured. Results from the quantitative analysis led to the following conclusions. First, compared to traditional methods, physical modeling made a difference in students' understanding of conservation of matter. Second, the physical model as used, did not make a difference in students' logical thinking ability. Third, compared to traditional methods, there was also a difference made in students' conceptual understanding of conservation of matter.

APPENDIX A – IRB Approval Letter (1)

THE UNIVERSITY OF
SOUTHERN MISSISSIPPI

INSTITUTIONAL REVIEW BOARD
118 College Drive #5147 | Hattiesburg, MS 39406-0001
Phone: 601.266.5997 | Fax: 601.266.4377 | www.usm.edu/research/institutional.review.board

NOTICE OF COMMITTEE ACTION

The project has been reviewed by The University of Southern Mississippi Institutional Review Board in accordance with Federal Drug Administration regulations (21 CFR 26, 111), Department of Health and Human Services (45 CFR Part 46), and university guidelines to ensure adherence to the following criteria:

- The risks to subjects are minimized.
- The risks to subjects are reasonable in relation to the anticipated benefits.
- The selection of subjects is equitable.
- Informed consent is adequate and appropriately documented.
- Where appropriate, the research plan makes adequate provisions for monitoring the data collected to ensure the safety of the subjects.
- Where appropriate, there are adequate provisions to protect the privacy of subjects and to maintain the confidentiality of all data.
- Appropriate additional safeguards have been included to protect vulnerable subjects.
- Any unanticipated, serious, or continuing problems encountered regarding risks to subjects must be reported immediately, but not later than 10 days following the event. This should be reported to the IRB Office via the "Adverse Effect Report Form".
- If approved, the maximum period of approval is limited to twelve months.
  Projects that exceed this period must submit an application for renewal or continuation.

PROTOCOL NUMBER: 17100507
PROJECT TITLE: Tangible Teaching: The effect of physical modeling on conceptual understanding of conservation of matter
PROJECT TYPE: Doctoral Dissertation
RESEARCHER(S): Erick T. Moffett
COLLEGE/DIVISION: College of Science and Technology
DEPARTMENT: Center for Science and Mathematics Education
FUNDING AGENCY/SPONSOR: N/A
IRB COMMITTEE ACTION: Exempt Review Approval
PERIOD OF APPROVAL: 11/21/2017 to 11/20/2018
Lawrence A. Hosman, Ph.D.
Institutional Review Board

APPENDIX B – IRB Approval Letter (2)

PEARL RIVER
COMMUNITY COLLEGE

Pearl River Community College
Institutional Review Board Decision Letter

The Institutional Review Board (IRB) has completed its review of the following project:

Principal Investigator: Erick T. Moffett
Pearl River Community College

On-Campus Liaison*: Dr. Jennifer Seal
Telephone: 601-403-1146 Email: jseal@prcc.edu

Project Title: Tangible Teaching: The Effect of Physical Modeling on Conceptual Understanding of Conservation of Matter

Funding Agency: N/A

Proposal Number (if applicable): N/A

The determination of the board is that:

√ This project **complies** with the institution's policy and procedures regarding use of human subjects in a grant-funded research project (Common Rule Section 101, subsection b). The project may be conducted as planned subject to continuing review as outlined in the Board's procedures.

You are authorized to implement this study as of the date of the final approval. This approval is valid for one year. If the project continues beyond this time frame, the IRB will request continuing review and update of the project. After receiving initial approval of your project, any proposed changes that may affect the format, implementation, or exempt status of your project and/or any unanticipated or serious adverse events involving risk to the participants requires that notification must be received, and approval must be granted, by the PRCC IRB before the implementation of the project.

As stated and agreed upon in your petition to the Pearl River Community College Institutional Review Board, **study findings must be shared with the PRCC IRB.** Copies of the report may be shared with both internal and external personnel associated with the College who have an interest in the topic and results. In addition, **permission of the PRCC IRB is required prior to publication of your study.** Approval is at the sole discretion of the Board.

*Please note that you must discuss your project and implementation plans with the On-Campus Liaison (indicated above) to ensure that your plans can be implemented as you requested.

____This project **does not comply** with the institution's policy and procedures regarding use of human subjects in a grant-funded research project. Concerns of the Institutional Review Board are outlined in an attached document. The Principal Investigator has the right to modify and re-submit the proposal for another review.

Jennifer Seal
Chair, Institutional Review Board

11-17-17
Date

6/13

101 Highway 11 North • Poplarville, Mississippi 39470 • 601-403-1000 • www.prcc.edu

APPENDIX C – Consent Form

ORI | Office of Research Integrity

INSTITUTIONAL REVIEW BOARD

STANDARD (SIGNED) INFORMED CONSENT

STANDARD (SIGNED) INFORMED CONSENT PROCEDURES

This completed document must be signed by each consenting research participant.

- The Project Information and Research Description sections of this form should be completed by the Principal Investigator before submitting this form for IRB approval.
- Signed copies of the consent form should be provided to all participants.

Last Edited July 20th, 2017

Today's date: September 28, 2017		
PROJECT INFORMATION		
Project Title: Tangible Teaching: The effect of physical modeling on conceptual understanding of Conservation of Matter		
Principal Investigator: Erick T. Moffett	Phone: 601-909-0654	Email: erick.moffett@usm.edu
College: The University of Southern Mississippi	Department: Department of Science and Mathematics Education	
RESEARCH DESCRIPTION		

1. **Purpose**:

This research has one main objective. The purpose of the proposed study is to investigate whether the proposed version of manipulative modeling helps junior college students enrolled in Biology I, Principles of Biology I, or Chemistry I better understand the concept of conservation of matter, reflected by improved achievement on the Conservation of Matter Questionnaire.

The overarching hypothesis in this study is formulated as such, "if the treatment group was able to receive instruction about conservation of matter that implements physical modeling, then they will outperform the comparison group in a test of conceptual understanding of conservation of matter."

Furthermore, because an extensive literature review uncovered that cognitive ability is theorized to play a significant role in understanding abstract concepts, student logical thinking ability was also measured.

2. **Description of Study:**

Upon approval from the IRB, course professors who teach Biology I, Principles of Biology I, or Chemistry I will be contacted by the researchers to ask for their support in recruiting their students to voluntarily participate in this study. Each participant will complete three sessions that will occur during their respective class time on the campus of the junior college. Sessions will not last longer than 50 minutes. Session 1 will involve a short pre-test of cognitie thinking ability and conceptual understanding of conservation of matter. Session 2 will involve the implementation of the physical model (for experimental groups) or traditional instruction (comparison group) to teach and reinforce the idea of conservation of matter. Session 3 will involve a post test of the two questionnaires administered during Session 1.

The consent form will be provided at the beginning of the interview informing participants that their participation is voluntary and he/she may decline without any penalty or prejudice.

Participation will only begin once the participants sign the general consent form.

3. **Benefits:**

While participants will not receive direct benefits, it is hoped that this research will contribute to the literature and practice in the discipline. By examining students' conceptual understanding and retention rate of conservation of matter, the researcher can begin to propose possible ways that may serve to enhance how fundamental scientific concepts are taught. Furthermore, as participants participate in the study, another benefit would be a particular focus on (often difficult) topics already in the curriculum of the courses they are already in.

4. **Risks:**

As it is with any study, there is always some risk associated. For this particular study, participants may find the academic material taught and refinfored during the interactions frustrating or stressful as they strive to gain a complete understanding of the science concepts

5. **Confidentiality:**

After the analysis is complete the original data sources will be destroyed or deleted. The digital data files will be secured and protected by the researchers for future analytical comparisons or evaluation. After a period of one year, digital and physical data sets will be deleted and destroyed.

The researcher will make every effort to protect the anonimity and confidentiality of the participants. Students' names will not be used on data questionnaires, but instead a random, unique identifier code assigned to each student that will be used for both the pretests and posttests.

6. **Alternative Procedures:**

N/A

7. **Participant's Assurance:**

This project has been reviewed by the Institutional Review Board, which ensures that research projects involving human subjects follow federal regulations.

Any questions or concerns about rights as a research participant should be directed to the Chair of the IRB at 601-266-5997. Participation in this project is completely voluntary, and participants may withdraw from this study at any time without penalty, prejudice, or loss of benefits.

Any questions about the research should be directed to the Principal Investigator using the contact information provided in Project Information Section above.

CONSENT TO PARTICIPATE IN RESEARCH

Participant's Name: ____________________

Consent is hereby given to participate in this research project. All procedures and/or investigations to be followed and their purpose, including any experimental procedures, were explained to me. Information was given about all benefits, risks, inconveniences, or discomforts that might be expected.

The opportunity to ask questions regarding the research and procedures was given. Participation in the project is completely voluntary, and participants may withdraw at any time without penalty, prejudice, or loss of benefits. All personal information is strictly confidential, and no names will be disclosed. Any new information that develops during the project will be provided if that information may affect the willingness to continue participation in the project.

Questions concerning the research, at any time during or after the project, should be directed to the Principal Investigator with the contact information provided above. This project and this consent form have been reviewed by the Institutional Review Board, which ensures that research projects involving human subjects follow federal regulations. Any questions or concerns about rights as a research participant should be directed to the Chair of the Institutional Review Board, The University of Southern Mississippi, 118 College Drive #5116, Hattiesburg, MS 39406-0001, 601-266-5997.

Include the following information only if applicable. Otherwise delete this entire paragraph before submitting for IRB approval: The University of Southern Mississippi has no mechanism to provide compensation for participants who may incur injuries as a result of participation in research projects. However, efforts will be made to make available the facilities and professional skills at the University. Participants may incur charges as a result of treatment related to research injuries. Information regarding treatment or the absence of treatment has been given above.

______________________________	______________________________
Research Participant	**Person Explaining the Study**
______________________________	______________________________
Date	**Date**

APPENDIX D – Conservation of Matter Pre- and Post-test

Conservation of Matter Questionnaire 1

Please answer all of the questions on this test using a #2 pencil. In some cases, there may be more than one correct answer. However, each question has only one best answer. Choose the single best answer from the five choices for each question. Mark your answer sheet by completely filling in the circle on the sheet that matches your choice. If you change an answer, be sure to thoroughly erase your original choice.

1. Some smelly yellow sulfur is mixed with brown magnetic iron filings in a ceramic dish. The mixture is heated until a black solid is formed. This solid has no odor and is not magnetic. What do you think happened?
   a. A new substance was formed, with new properties.
   b. Some matter was destroyed, along with its properties.
   c. All substances lose their basic properties when they are heated.
   d. A crust formed, sealing in the sulfur smell and magnetism.
   e. This is the type of phenomenon scientists are trying to understand.

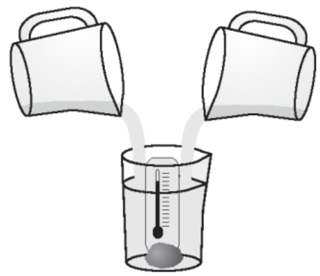
2. Which of the equations below represents a balanced equation for the conversion of oxygen to ozone?
   a. $6O_{2(g)} \longrightarrow 2O_{3(g)}$
   b. $6O_{2(g)} \longrightarrow 9O_{3(g)}$
   c. $6O_{2(g)} \longrightarrow 12O_{6(g)}$
   d. $6O_{2(g)} \longrightarrow 4O_{3(g)}$
   e. $6O_{2(g)} \longrightarrow 6O_{3(g)}$

3. When red food coloring is added to a cup of water, the water looks red. A scientist's explanation for the color change of the water is that:
   a. the water molecules were dyed red.
   b. the particles of red coloring spread out in the water.
   c. a chemical reaction occurred between the food coloring and the water.
   d. the particles of food coloring were pulled inside the water molecules.
   e. the food coloring particles and water molecules have similar properties and repelled each other.

4. Ako collects shiny pennies and saves them in an open glass jar. After a couple of years he observes that the pennies have turned grayish green. What do you think has happened?
   a. The pennies have combined with gases in the air to form a new chemical.
   b. The pennies are the same; they only changed color.
   c. Grayish green material was present in the air and it deposited on the pennies.
   d. Sunlight changed the color of the pennies.
   e. Something from the glass jar was transferred to the pennies.

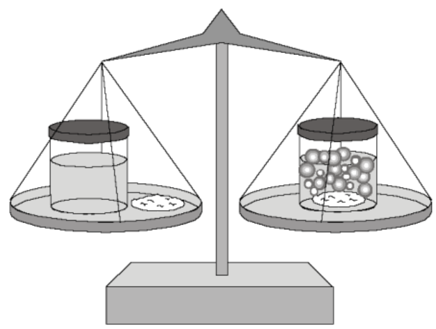
NEXT

5. Imagine that when two different clear liquids are mixed together in a beaker, the temperature in the beaker rises and a blue-colored solid forms and sinks to the bottom. The total amount of matter that now exists in the beaker is:
   a. greater than before because a new substance was formed.
   b. less than before because some matter was destroyed.
   c. less than before because some matter was turned into energy.
   d. the same as before.
   e. It is impossible to determine an answer using the information given.

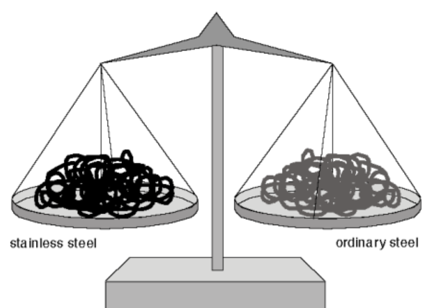


6. Ashley combines 45 grams of vinegar and 5 grams of baking soda in a cup on a scale. After 10 minutes, the scale reads only 48 grams. The change in mass is the result of:
   a. a gas being produced and escaping into the air.
   b. a gas being produced and causing the cup to float.
   c. the baking soda dissolving in the vinegar.
   d. the substances being together instead of two different containers.
   e. the substances combining with the air.

7. A scale is balanced with two sealed jars. The left pan has a sealed jar containing vinegar and 5 grams of baking soda is lying outside. The right pan has a sealed jar containing vinegar and the same amount of baking soda inside the jar. As the baking soda fizzes, what will happen to the pan with the fizzing baking soda?
   a. The pan will move up.
   b. The pan will not move.
   c. The pan will move down.
   d. The pan will first move up and then down.
   e. The pan will first move down and then up.

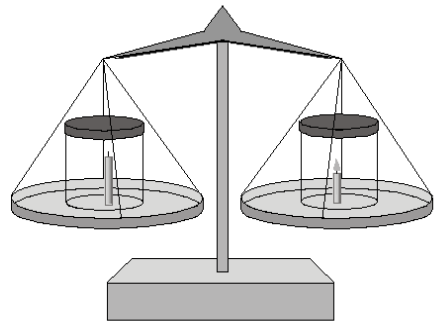


8. Sophia balances a pile of stainless steel wire and ordinary steel wire on a scale. After a few days the ordinary wire on the right pan starts rusting. What will happen to the pan with the rusting wire?
   a. The pan will move up.
   b. The pan will not move.
   c. The pan will move down.
   d. The pan will first move up and then down.
   e. The pan will first move down and then up.

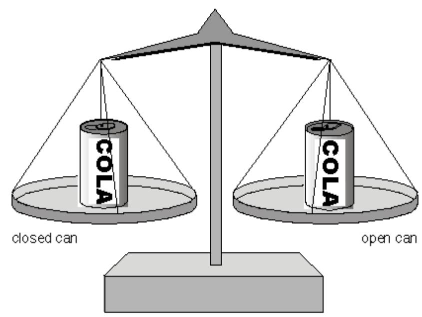


NEXT

9. Nico balances two sealed containers—one with an unlit birthday candle and one with a lit candle—on opposite sides of a scale. As the lit candle burns halfway down, what will happen to the side holding the burning candle?
   a. The pan will move up.
   b. The pan will not move.
   c. The pan will move down.
   d. The pan will first move up and then down.
   e. The pan will first move down and then up.



10. Fatima balances two cans of cola, one closed and one just opened on a scale. As the cola in the open can goes flat, what will have happened to the pan holding the open can?
    a. The pan will move up.
    b. The pan will not move.
    c. The pan will move down.
    d. The pan will first move up and then down.
    e. The pan will first move down and then up.

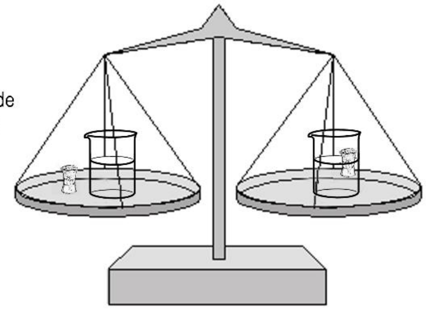


11. Two astronauts are sealed inside a test chamber. One of them is celebrating his birthday and is given a cake with 40 lit candles. What do you think will happen to the test chamber after the astronaut blows out the candles and the chamber is filled with smoke?
    a. The test chamber will get lighter by a tiny amount.
    b. The test chamber will get heavier by a tiny amount.
    c. The test chamber will stay exactly the same weight.
    d. The test chamber will first get lighter and then heavier.
    e. There is not enough information to answer the question.

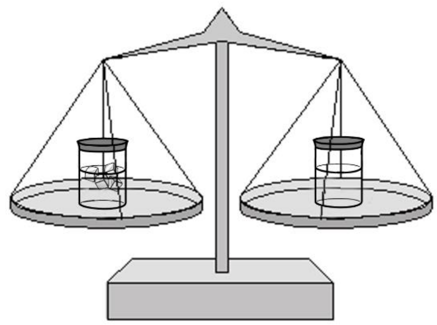
12. One gram of water is sealed in a strong tube. The tube and the water together weigh 25 grams. The tube is heated until all the water boils and it is no longer visible. How much will the sealed tube now weigh?
    a. Less than 25 grams.
    b. 25 grams.
    c. Between 25 and 26 grams.
    d. 26 grams.
    e. More than 26 grams.

NEXT

13. Mike balances a beaker of water and a piece of cork on each side of the balance. On the left side the cork remains on the pan. Mike places the cork on the right side into the beaker. What will happen to the pan holding the beaker with the floating cork?
    a. The pan will move up.
    b. The pan will not move; the keys are still balanced.
    c. The pan will move down.
    d. The pan will first move up and then down.
    e. The pan will first move down and then up.



14. Silvia balances two beakers. The left side has ice and water. The right side has just water. What will happen to the pan on the left once all the ice as melted?
    a. The pan will move up.
    b. The pan will not move; the keys are still balanced.
    c. The pan will move down.
    d. The pan will first move up and then down.
    e. The pan will first move down and then up.

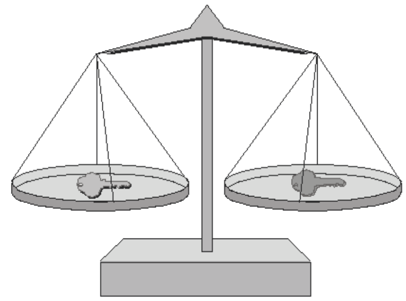


15. Which of the following would a scientist say is an element?
    a. Oxygen
    b. Water
    c. Flour
    d. More than one of the above.
    e. None of the above.

16. What is the weight of sugar water when 1 pound of sugar is dissolved in 30 pounds of water?
    a. 29 Pounds.
    b. 30 Pounds.
    c. Between 30 and 31 pounds.
    d. 31 pounds.
    e. More than 31 pounds.

NEXT

17. Sonya balances a metal key with a key of a different metal. After a few weeks the key on the right pan has changed color. The key changed color because oxygen in the air combined with the metal. What will happen to the pan holding this key?
    f. The pan will move up.
    g. The pan will not move; the keys are still balanced.
    h. The pan will move down.
    i. The pan will first move up and then down.
    j. The pan will first move down and then up.



18. Which of the following processes would change one kind of element into another kind of element?
    a. Rusting
    b. Boiling
    c. Oxidizing
    d. All of the above
    e. None of the above.

19. Which of the following processes would result in a new compound?
    a. Rusting
    b. Boiling
    c. Melting
    d. All of the above
    e. None of the above

20. According to the reaction below, when 4 grams of hydrogen gas react with sufficient oxygen gas, 36 grams of water will result. How many grams of water would result when 3 grams of hydrogen gas are fully reacted with oxygen?

$$2H_{2\,(g)} + O_{2\,(g)} \longrightarrow 2H_20_{\,(l)}$$

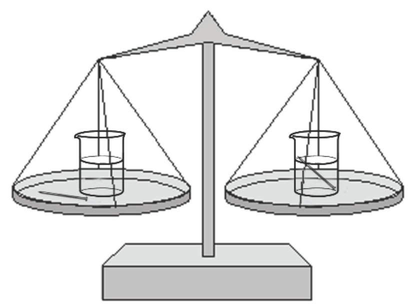
    a. 33 g
    b. 32 g
    c. 27 g
    d. 6 g
    e. 3 g

21. If you react equal volumes of hydrogen gas and oxygen gas according to the above equation, then you can expect that:
    a. hydrogen gas will be left over.
    b. oxygen gas will be left over.
    c. no gases will be left over.
    d. no reaction will take place.
    e. impossible to determine.

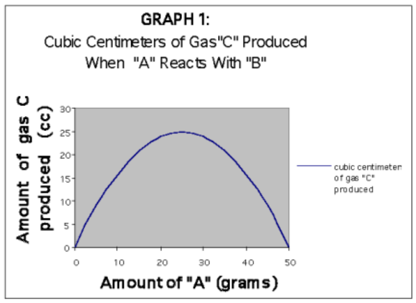
22. Which of the equations below represents a balanced equation for a reaction where aluminium and oxygen form aluminium oxide?
    a. $4Al_{(s)} + 3O_{2(g)} \rightarrow 2Al_2O_{3(s)}$
    b. $2Al_{2(s)} + 3O_{2(g)} \rightarrow 2Al_2O_{3(s)}$
    c. $Al_{4(s)} + O_{6(g)} \rightarrow Al_4O_{6(s)}$
    d. $4Al_{(s)} + 6O_{(g)} \rightarrow Al_4O_{6(s)}$
    e. $Al_{4\,(s)} + 3O_{2(g)} \rightarrow 2Al_2O_{3(s)}$

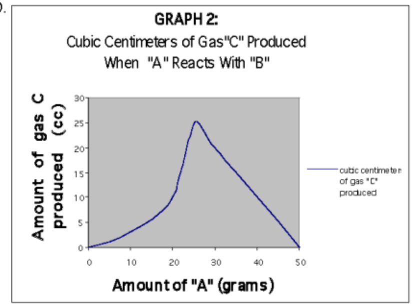
23. Which of the equations below represents a balanced equation for a reaction where zinc and oxygen form zinc oxide?
    a. $Zn_{2(s)} + 2O_{(g)} \rightarrow 2ZnO_{(s)}$
    b. $Zn_{2(s)} + O_{2(g)} \rightarrow Zn_2O_{2(s)}$
    c. $2Zn_{(s)} + 2O_{(g)} \rightarrow 2ZnO_{(s)}$
    d. $2Zn_{(s)} + O_{2(g)} \rightarrow 2ZnO_{(s)}$
    e. $2Zn_{(s)} + O_{2(g)} \rightarrow Zn_2O_{2(s)}$

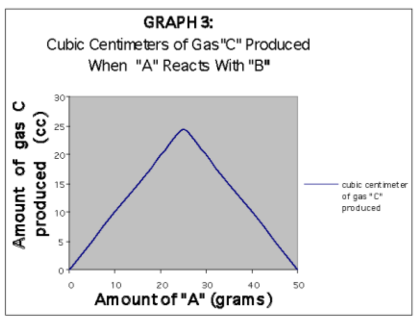
24. Farouq balances a nail and beaker of water on each side of a balance. He then places the nail in the beaker on the right side. What will happen to the pan with the nail in the beaker of water?
    f. The pan will move up.
    g. The pan will not move; the keys are still balanced.
    h. The pan will move down.
    i. The pan will first move up and then down.
    j. The pan will first move down and then up.

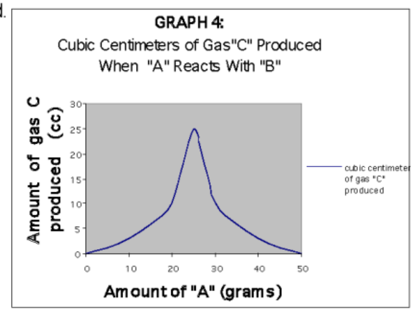


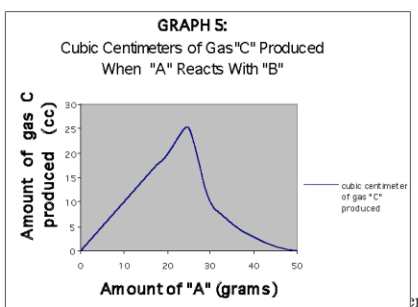
25. Alexandra decides to mix compounds "A" and "B", and discovers she can make gas "C". She tries a variety of mixtures always making sure the total weight of both A and B equals 50 grams. She knows that 0 grams of A and 50 grams of B will not produce any gas (the coordinates 0,0 on all the graphs below). She also knows that mixing 50 grams of A and 0 grams of B will not produce any gas (the coordinates 0,50 on all the graphs). The following graphs show five possibilities for the amount of gas C produced with different mixtures of A and B. Which of the graphs do you think best predicts the production of gas C between (0,0) and (0,50)? [Note: The x- axis only shows the amount of compound "A" that Alexandra uses.]

a.


b.


c.


d.


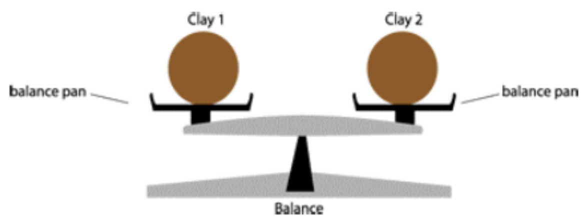
e.


FINISH

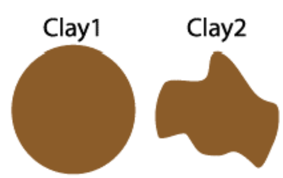
APPENDIX E – Group Assessment of Logical Thinking Ability Pre- and Post-test

Item 1. Piece of Clay

Tom has two balls of clay. They are the same size and shape. When he places them on the balance, they weigh the same.



The balls of clay are removed from the balance pans. Clay 2 is flattened like a pancake.



WHICH OF THESE STATEMENTS IS TRUE?

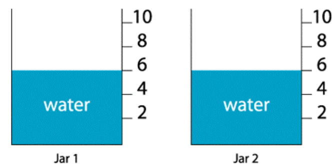
a. The pancake-shaped clay weighs more.
b. The two pieces weigh the same.
c. The ball weighs more.

REASON

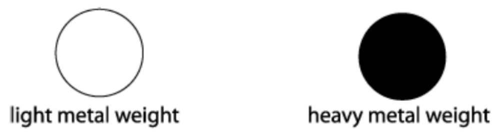
1. You did not add or take away any clay.
2. When clay 2 was flattened like a pancake, it had a greater area.
3. When something is flattened, it loses weight.
4. Because of its density, the round ball had more clay in it.

Item 4. Metal Weights

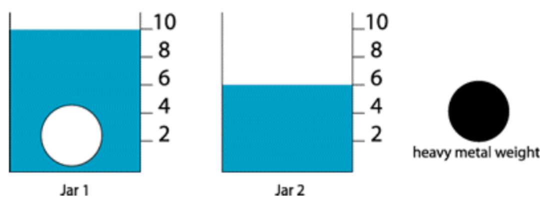
Linn has two jars. They are the same size and shape. Each is filled with the same amount of water.



She also has two metal weights of the same volume. One weight is light. The other is heavy.



She lowers the light weight into jar 1. The water level in the jar rises and looks like this:



IF THE HEAVY WEIGHT IS LOWERED INTO JAR 2, WHAT WILL HAPPEN?

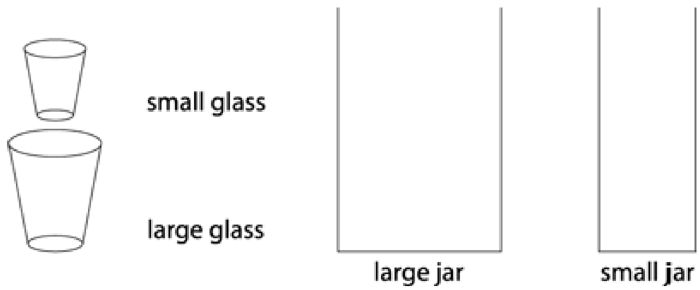
a. The water will rise to a higher level than in jar 1.
b. The water will rise to a lower level than in jar 1.
c. The water will rise to the same level as in jar 1.

REASON

1. The weights are the same size so they will take up equal amounts of space.
2. The heavier the metal weight, the higher the water will rise.
3. The heavy metal weight has more pressure, therefore the water will rise lower.
4. The heavier the metal weight, the lower the water will rise.

Item 8. Glass Size #2

The drawing shows two glasses, a small one and a large one. It also shows two jars, a small one and a large one.



It takes 15 small glasses of water or 9 large glasses of water to fill the large jar. It takes 10 small glasses of water to fill the small jar.

HOW MANY LARGE GLASSES OF WATER DOES IT TAKE TO FILL THE SAME SMALL JAR?

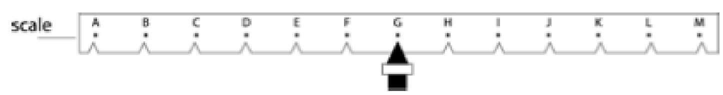
a. 4
b. 5
c. 6
d. other

REASON

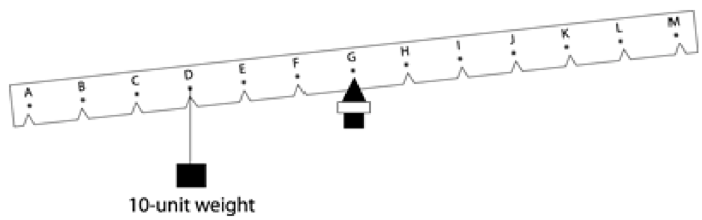
1. It takes five less small glasses of water to fill the small jar. So it will take five less large glasses of water to fill the same jar.
2. The ratio of small to large glasses will always be 5 to 3.
3. The small glass is half size of the large glass. So it will take about half the number of small glasses of water to fill up the same small jar.
4. There is no way of predicting.

Item 9. Scale #1

Joe has a scale like the one below.



When he hangs a 10-unit weight at point D, the scale looks like this:



WHERE WOULD HE HANG A 5-UNIT WEIGHT TO MAKE THE SCALE BALANCE AGAIN?

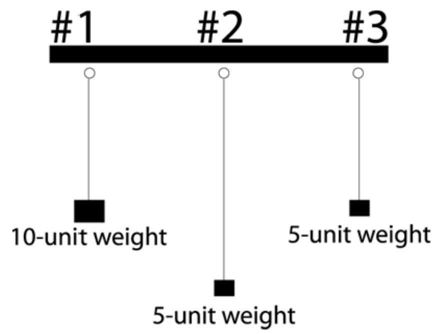
a. at point J
b. between K and L
c. at point L
d. between L and M
e. at point M

REASON

1. It is half the weight so it should be put at twice the distance.
2. The same distance as 10-unit weight, but in the opposite direction.
3. Hang the 5-unit weight further out, to make up its being smaller.
4. All the way at the end gives more power to make the scale balance.
5. The lighter the weight, the further out it should be hung.

Item 11. Pendulum Length

Three strings are hung from a bar. String #1 and #3 are of equal length. String #2 is longer. Charlie attaches a 5-unit weight at the end of string #2 and at the end of #3. A 10-unit weight is attached at the end of string #1. Each string with a weight can be swung.



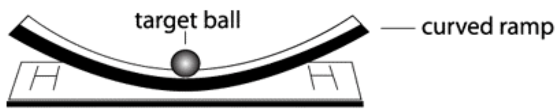
Charlie wants to find out if the length of the string has an effect on the amount of time it takes the string to swing back and forth.

WHICH STRING AND WEIGHT WOULD HE USE FOR HIS EXPERIMENT?
a. string #1 and #2
b. string #1 and #3
c. string #2 and #3
d. string #1, #2, and #3
e. string #2 only

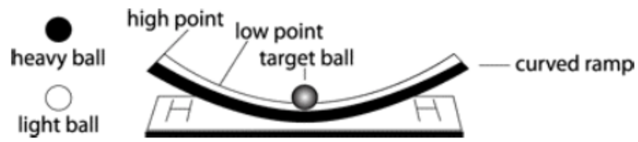
REASON
1. The length of the strings should be the same. The weights should be different.
2. Different lengths with different weights should be tested.
3. All strings and their weights should be tested against all others.
4. Only the longest string should be tested. The experiment is concerned with length not weight.
5. Everything needs to be the same except the length so you can tell if length makes a difference.

Item 13. Ball #1

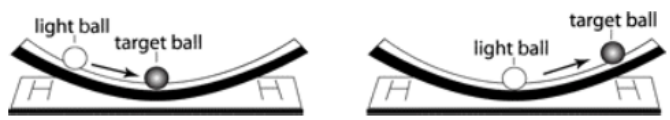
Eddie has a curved ramp. At the bottom of the ramp there is one ball called the target ball.



There are two other balls, a heavy and a light one. He can roll one ball down the ramp and hit the target ball. This causes the target ball to move up the other side of the ramp. He can roll the balls from two different points, a low point and a high point.



Eddie released the light ball from the low point. It rolled down the ramp. It hit and pushed the target ball up the other side of the ramp.



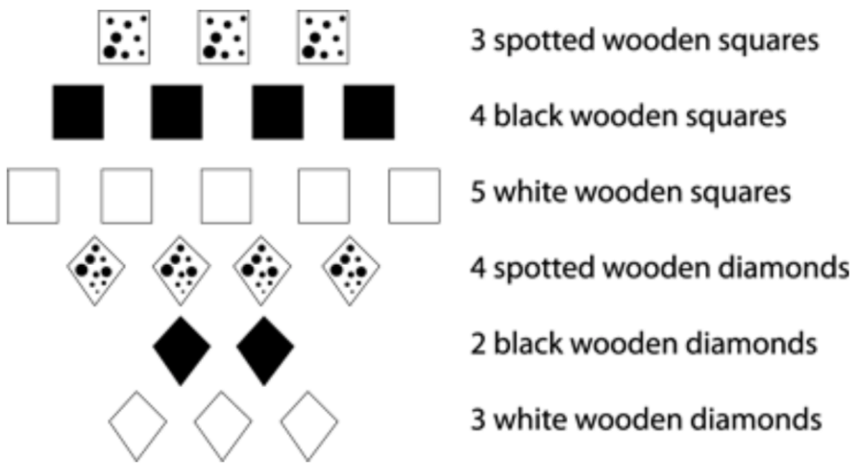
He wants to find out if the point a ball is released from makes a difference in how far the target goes.

TO TEST THIS WHICH BALL WOULD HE NOW RELEASE FROM THE HIGH POINT?
a. the heavy ball
b. the light ball

REASON
1. He started with the light ball he should finish with it.
2. He used the light ball the first time. The next time he should use the heavy ball.
3. The heavy ball would have more force to hit the target ball farther.
4. The light ball would have to be released from the high point in order to make a fair comparison.
5. The same ball must be used as the weight of the ball does not count.

Item 15. Square and Diamonds #1

In a cloth sack, there are



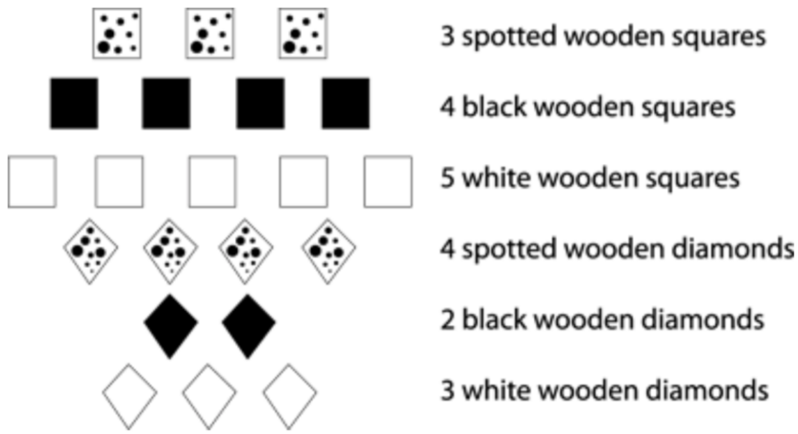
All of the square pieces are the same size and shape. The diamond pieces are also the same size and shape. One piece is pulled out of the sack.

WHAT ARE THE CHANCES THAT IT IS A SPOTTED PIECE?
a. 1 out of 3
b. 1 out of 4
c. 1 out of 7
d. 1 out of 21
e. Other

REASON
1. There are twenty-one pieces in the cloth sack. One spotted piece must be chosen from these.
2. One spotted piece needs to be selected from a total of seven spotted pieces.
3. Seven of the twenty-one pieces are spotted pieces.
4. There are three sets in the cloth sack. One of them is spotted.
5. ¼ of the square pieces and 4/9 of the diamond pieces are spotted.

Item 16. Square and Diamonds #2

In a cloth sack, there are



All of the square pieces are the same size and shape. The diamond pieces are also the same size and shape. Reach in and take the first piece you touch.

WHAT ARE THE CHANCES OF PULLING OUT A SPOTTED DIAMOND OR A WHITE DIAMOND?

a. 1 out of 3
b. 1 out of 9
c. 1 out of 21
d. 9 out of 21
e. Other

REASON

1. Seven of the twenty-one pieces are spotted or white diamonds.
2. 4/7 of the spotted and 3/8 of the white are diamonds.
3. Nine of the twenty-one pieces are diamonds.
4. One diamond piece needs to be selected from a total of twenty-one pieces in the cloth sack.
5. There are 9 diamond pieces in the cloth sack. One piece must be chosen from these.

Item 17. The Mice

A farmer observed the mice that live in his field. He found that the mice were either fat or thin. Also, the mice had either black tails or white tails.

This made him wonder if there might be a relation between the size of a mouse and the color of its tail. So he decided to capture all of the mice in one part of his field and observe them. The mice that he captured are shown below.

DO YOU THINK THERE IS A RELATION BETWEEN THE SIZE OF THE MICE AND THE COLOR OF THEIR TAILS (THAT IS, IS ONE SIZE OF MOUSE MORE LIKELY TO HAVE A CERTAIN COLOR TAIL AND VICE VERSA)?

a. Yes
b. No

REASON

1. 8/11 of the fat mice have black tails and ¾ of the thin mice have white tails.
2. Fat and thin mice can have either a black or a white tail.
3. Not all fat mice have black tails. Not all thin mice have white tails.
4. 18 mice have black tails and 12 have white tails.
5. 22 mice are fat and 8 mice are thin.

Item 18. The Fish

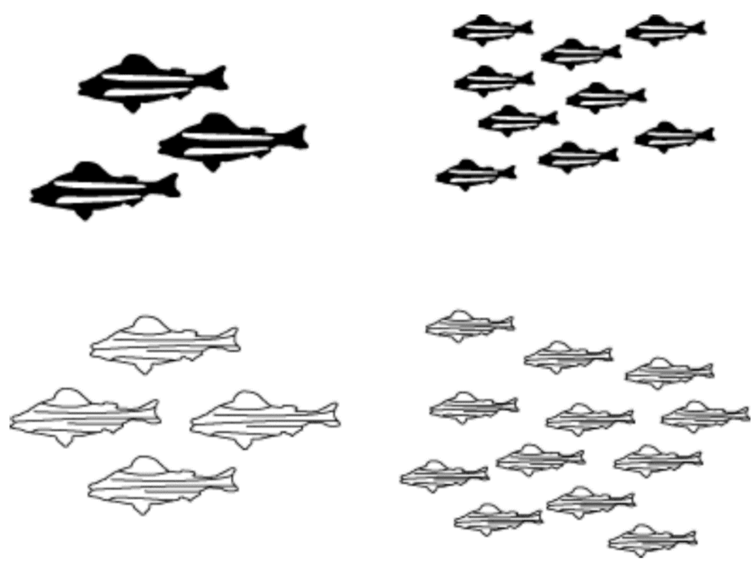
Some of the fish below are big and some are small. Also some of the fish have wide stripes on their sides. Others have narrow stripes.

IS THERE A RELATIONSHIP BETWEEN THE SIZE OF THE FISH AND THE KIND OF STRIPES IT HAS (THAT IS, IS ONE SIZE OF FISH MORE LIKELY TO HAVE A CERTAIN TYPE OF STRIPES AND VICE VERSA)?

a. Yes
b. No

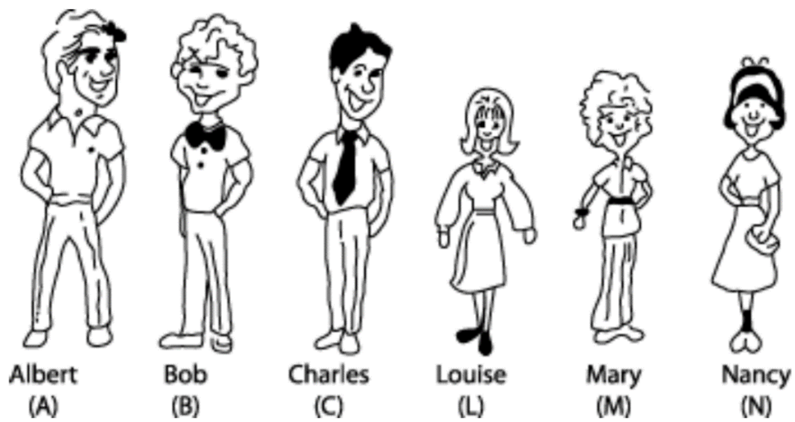
REASON

1. Big and small fish can have either wide or narrow stripes.
2. 3/7 of the big fish and 9/21 of the small fish have wide stripes.
3. 7 fish are big and 21 are small.
4. Not all big fish have wide stripes and not all small fish have narrow stripes.
5. 12/28 of fish have wide stripes and 16/28 of fish have narrow stripes.



Item 19. The Dance

After supper, some students decide to go dancing. There are three boys: ALBERT (A), BOB (B), and CHARLES (C), and three girls: LOUISE, (L), MARY (M), and NANCY (N).

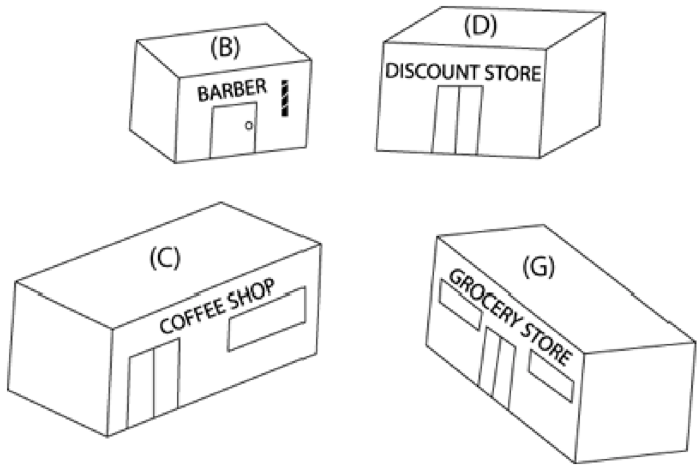


One possible pair of dance partners is A-L, which means ALBERT and LOUISE.

LIST ALL OTHER POSSIBLE COUPLES OF DANCERS. BOYS DO NOT DANCE WITH BOYS, AND GIRLS DO NOT DANCE WITH GIRLS.

Item 20. The Shopping Center

In a new shopping center, 4 stores are going to be placed on the ground floor. A BARBER SHOP (B), a DISCOUNT STORE (D), a GROCERY STORE (G), and a COFFEE SHOP (C) want to locate there.



One possible way that the stores could be arranged in the 4 locations is BDGC. Which means the BARBER SHOP first, the DISCOUNT STORE next, then the GROCERY STORE and the COFFEE SHOP last.

LIST ALL OTHER POSSIBLE WAYS THAT THE STORES CAN BE LINED UP IN THE FOUR LOCATIONS.

Answer Sheet

Pseudonym: ___________________

Instructions: For each item you are to choose the best answer and reason for selecting that answer. Record your answer in the space provided according to the test item.

ITEM	BEST ANSWER	REASON
1. Piece of Clay	_______________	_______________
4. Metal Weights	_______________	_______________
8. Glass Size #2	_______________	_______________
9. Scale #1	_______________	_______________
11. Pendulum Length	_______________	_______________
13. Ball #1	_______________	_______________
15. Square and Diamonds #1	_______________	_______________
16. Square and Diamonds #2	_______________	_______________
17. The Mice	_______________	_______________
18. The Fish	_______________	_______________

Item 19. "The Dance"

Place your answers below:

A-L	_____	_____	_____	_____	_____
_____	_____	_____	_____	_____	_____
_____	_____	_____	_____	_____	_____
_____	_____	_____	_____	_____	_____
_____	_____	_____	_____	_____	_____

Item 20. “The Shopping Center”

Place your answers below:

BDGC	_____	_____	_____	_____	_____
_____	_____	_____	_____	_____	_____
_____	_____	_____	_____	_____	_____
_____	_____	_____	_____	_____	_____
_____	_____	_____	_____	_____	_____

APPENDIX F – Letter of Support from the Author of the COM Questionnaire

From: Schwartz, Marc SCHWARMA@uta.edu
To: Erick Moffett erick.moffett@eagles.usm.edu

Dear Erick,

I regret the length of time it has taken me to respond. Hopefully the attachment is still useful to you.

This version of the questionnaire is the last one on which I worked with Dr. Sadler. However I would like to point you to follow-on work with another colleague, Dr. Theo Dawson. While the questionnaire is a multiple choice test using insights from research on misconceptions (as distractors), the work of Dr. Dawson is based on the complexity of student answers. In this case the questions are re-framed as open-answer questions, and Dr. Dawson uses a rubric to measure the complexity of the answer. If this approach interests you I'm sure Dr. Dawson would be happy to discuss what we attempted to do a few years back. This approach remains of great interest to me, so I would be happy to explore this with you and Dr. Dawson (if she has the time).

Best, Marc

Marc Schwartz
Professor and Director
SW Center for Mind, Brain and Education
University of Texas
Arlington, Texas 76019

phone: 817-272-5641

APPENDIX G – An Overview of the Physical Modeling Curriculum and its Activities

The physical model curriculum has been divided into three separate activities, scaffolding the material that leads up to balancing chemical equations.

Initial Instruction (same for both treatment conditions):

<u>Lecture, Reading, Review Questions</u>

- Lecture: PowerPoint
- Reading: Textbook
  - Ch*emistry: The Central Science (14th edition)* by Brown *et. al.*
    - Chapter 3 – Stoichiometry
- Key vocabulary:
  - Reactants
  - Products
  - Coefficients
  - Subscripts
  - Law of conservation of matter
  - Balanced equation
  - Physical change
  - Chemical change
- Sample review questions:
  - Describe what a chemical reaction is.
  - What are reactants and products?
  - What are key differences between physical and chemical changes?
  - What does it mean to say a chemical equation is balanced?

Activities:

<u>Activity 1: Counting atoms (25 minutes)</u>

Before doing Activity 1, students should already be familiar with basic atomic structure and have a fundamental knowledge of the Periodic Table of Elements. At this point in the model, students are presented with chemical formulas and are challenged with counting the total number of atoms in a compound or molecule, taking into consideration coefficients and subscripts.

Activity 2: Modeling conservation of matter (25 minutes)

In the second activity, students should have an understanding of conservation of matter, and be able to identify the reactants and products in a chemical equation. Activity 2 presents more of a challenge because students are now presented with multiple molecules or compounds present on both sides of a chemical reaction.

Activity 3: Learning to balance chemical equations (40 minutes)

In the third activity, students will bring together everything that has been introduced and reinforced in Activities 1 and 2 to take on the final task of balancing chemical equations. It is important that students master Activity 2 before moving on to Activity 3.

APPENDIX H – An Overview of the Traditional Curriculum and its Activities

This curriculum follows a traditional lecture approach which utilizes the textbook and practice worksheets. Thus, ideas are mainly transferred via lectures and limited student interaction.

Initial Instruction (same for both treatment conditions):

Lecture, Reading, Review Questions

- Lecture: PowerPoint
- Reading: Textbook
  - Ch*emistry: The Central Science (14th edition)* by Brown *et. al.*
    - Chapter 3 – Stoichiometry
- Key vocabulary:
  - Reactants
  - Products
  - Coefficients
  - Subscripts
  - Law of conservation of matter
  - Balanced equation
  - Physical change
  - Chemical change
- Sample review questions:
  - Describe what a chemical reaction is.
  - What are reactants and products?
  - What are key differences between physical and chemical changes?
  - What does it mean to say a chemical equation is balanced?

Activities:

Worksheets

REFERENCES

Abraham, M. R., Williamson, V. M. and Westbrook, S. L. (1994), A cross-age study of the understanding of five chemistry concepts. *J. Res. Sci. Teach., 31*: 147–165. doi:10.1002/tea.3660310206

ACT, Inc. (2017). The Condition of College & Career Readiness 2017. Iowa City, IA: Author. Retrieved from www.act.org/condition2017

Agung, S., & Schwartz, Marc S. (2007). Students' understanding of conservation of matter, stoichiometry and balancing equations in Indonesia. International Journal of Science Education, 29(13), 1679-1702. doi: 10.1080/09500690601089927

Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education, 33*(2/3), 131-152.

Andersson, B. (1984). Chemical reactions. EKNA Report No: 12, University of Gˆteborg, Gˆteborg.

Andersson, B. (1986). Pupils' explanation of some aspects of chemical reactions. *Science Education, 70,* 549-563.

Asghar, A. (2005). *Exploring children's cognitive and emotional development related to conservation of mass.* Montreal, Canada: American Educational Research Association (AERA).

Aubusson, P. J., Harrison, A. G., & Ritchie, S. M. (Eds.). (2006). *Metaphor and analogy in science education*. Dordrecht, the Netherlands: Springer.

Barker, V. & Millar, R. (1999). Students' reasoning about chemical reactions: What changes occur during a context-based post-16 chemistry course? *International Journal of Science Education, 21*, 645-665.

BouJaoude, S., & Barakat, H. (2000). Secondary school students' difficulties with stoichiometry. *School Science Review*, *296*, 91–98.

Ben-Zvi, R., Eylon, B. & Silberstein, J. (1987). Students' visualization of a chemical reaction. *Education in Chemistry, 24*, 117-120.

Berk, E.G. (1999). *Hands-on science: Using manipulatives in the classroom*. Hatfield, PA.

Bird, L. (2010). Logical reasoning ability and student performance in general chemistry. *Journal of Chemical Education*, *87*(5), 541 546.

Bitner, B. (1991). Formal Operational Reasoning Modes: Predictors of Critical Thinking Abilities and Grades Assigned by Teachers in Science and Mathematics for Students in Grades Nine Through Twelve. *Journal of Research in Science Teaching*, *28*(3), 265-274.

Bruner, J. S. (1960). *Process of Education (Revised Education)*. Cambridge, MA, USA: Harvard University Press.

Bruner, J. (1966). *Toward a theory of instruction*. Cambridge, MA: Harvard University Press.

Bunce D.M. and Hutchinson K.D., (1993), The use of the GALT (Group Assessment of Logical Thinking) as a predictor of academic success in college chemistry, *Journal of Chemical Education*, *70*, 183-187.

Burns, M. (2007). *About teaching mathematics: A K-8 resource* (3rd Edition). Sausalito, CA: Math Solutions.

Bybee, R. (2013). *Translating the NGSS for classroom instruction*. Arlington, VA: NSTA Press.

Bybee, R.W., & Van Scotter, P. (2006). Reinventing the Science Curriculum. *Educational Leadership*, *64*(4), 43-47.

Buckley, C.B. (2000). Interactive multimedia and model-based learning in biology. *International Journal of Science Education*, *22*(9), 895-935.

Campanario, J. M. (1995). Automatic "balancing" of chemical equations. *Computers Chemical, 19*(2), 85–90.

Carbonneau, K.J., Marley, S.C., & Selig, J.P. (2013). A Meta-Analysis of the Efficacy of Teaching Mathematics with Concrete Manipulatives. *Journal of Educational Psychology*, *105*(2), 380-400.

Chandrasegaran, A. L, Treagust, D. F., & Mocerino, M. (2011). Facilitating high school students' use of multiple representations to describe and explain simple chemical reactions. *Teaching Science*, *57*(4), 13-20.

Chittleborough, G.D. and Treagust D. F. (2007). The Modelling Ability of Non-Major Chemistry Students And Their Understanding Of The Sub-Microscopic Level. *Chem. Educ. Res. Pract.*, *8*, 274-292.

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

Clement, J. J., & Rae-Mamirez, M. A. (Eds.). (2008). *Model based learning and instruction in science*. Dordrecht, The Netherlands: Springer

Coll, R. K. & Treagust, D. F., (2001). Learners' Mental Models Of Chemical Bonding, *Research in Science Education*, *31*, 357–382.

Committee, O. C. F. F. T., & National, R. C. (2012). Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas. Washington, DC, USA: National Academies Press.

Copolo, C.F., & Hounshell, P.B. (1995). Using Three-Dimensional Models to Teach Molecular Structures in High School Chemistry. *Journal of Science Educations and Technology*, *4*(4), 295-305.

Dagher, Z.R., (1994) Does the Use of Analogies Contribute to Conceptual Change?, *Science Education*, *78*(6), 601-614

Dale, L. G. (1970). The growth of systematic thinking: Replication and analysis of Piaget's first chemical experiment. *Australian Journal of Psychology*, *22*, 277-286.

DeBoer, G. E. (1991). A history of ideas in science education. New York: Teachers College Press.

De Jong, T., Ainsworth, S., Dobson, M., Van der Hulst, A., Levonen, J., Reimann, P., . . . Swaak, J. (1998). Acquiring knowledge in science and mathematics: The use of multiple representations in technology based learning environments. In M.W.V. Someren, P. Reimann, H.P.A. Boshuizen, & T. D. Jong (Eds.), *Learning with multiple representations* (pp. 9-40). London: Elsevier.

Doucerain, M., & Schwartz, M.S. (2010). Analyzing learning about conservation of matter in students while adapting to the needs of a school. *Mind, Brain, and Education*, *4*(3), 112-124.

Driver, R., Child, D., Gott, R., Head, J., Johnson, S., Worsley, C. & Wylie, F. (1984). *Science in schools, age 15*. Research Report No: 2, Assessment of performance unit. Department of Education and Science, London.

Driver, R. (1985). *Beyond appearances: The conservation of matter under physical and chemical transformations*. In: Driver, R., Guesne, E., Tiberghien, A. (Eds.) Children's Ideas in Science. Open University Press, Milton Keynes , 145-169.

Driver, R., & Oldham , V. (1985). A constructivist approach to curriculum development in science. *Studies in Science Education, 13*, 105-122.

Ebbing, D. D. (1996). *General chemistry* (5th ed.). Boston, MA: Houghton Mifflin.

Fah, L. Y. (2009). Logical thinking abilities among form 4 students in the interior division of Sabah, Malaysia. *Journal of Science and Mathematics Education in Southeast Asia, 32*(2), 161-187.

Freeman, S., Eddy, S., McDonough, M., Smith, M., Okoroafor, N., Jordt, H., Wenderoth, M. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, *111*(23), 8410-8415.

Gabel, D., & Sherwood, R. (1980). The effect of student manipulation of molecular models on chemistry achievement according to Piagetian level. *Journal of Research in Science Teaching*, *17*(1), 75-81.

Garnet, P., Oliver, R., & Hackling, M. (1998). *Designing interactive multimedia materials to support concept development in beginning chemistry classes.* Edith Cowean University, Australia. Retrieved November 24, 2005, from http://elrond.scam.ecu.edu.au/oliver/docs/98/ICCE.pdf

Gilbert, J.K. (2004). Models and modelling: Routes to more authentic science education. *International journal of Science and Mathematics Education*, *2*(2), 115-130.

Gilbert, J. K., & Boulter, C. J. (1998). Learning science through models and modeling. In B. J. Fraser & K. G. Tobin (Eds.), *International handbook of science education* (pp. 53–66). Dordrecht, The Netherlands: Kluwer.

Gilbert, J. K., Boulter, C. J., & Elmer, R. (2000). Positioning models in science education and in design and technology education. In J. K. Gilbert & C. J. Boulter (Eds.), *Developing Models in Science Education* (pp. 3-18). Dordrecht: Kluwer.

Gilbert, J. K.; Reiner, M. & Nakhleh, M. (2008). Introduction. In J. K. Gilbert, M. Reiner, & M. Nakhleh (Eds.) Visualization: Theory and Practice in Science Education. (pp. 1-2). Dordrecht: Springer.

Gilbert, J. K. and D. F. Treagust (2009). *Multiple Representations in Chemical Education*. Dordrecht, the Netherlands: Springer.

Gire, E., Carmichael, A., Chini, J. J., Rouinfar, A., Rebello, S., & Puntambekar, S. (2010). The effects of physical and virtual manipulatives on students" conceptual learning about pulleys. *International Conference of the Learning Sciences*. Chicago.

Gomez, M., Pozo, J., Sanz, A. (1995). Students' ideas on conservation of matter: Effects of expertise and context variables. *Science Education*, *79*(1), 77-93.

Gray, K., Owens, K., Steer, D., McConnell, D. & Knight, C. (2011). An Exploratory Study Using Hands-On Physical Models in a Large Introductory, Earth Science Classroom: Student Attitudes and Lessons Learned. Electronic Journal of Science Education, 12(2).

Green, M., Piel, J., & Flowers, C. (2008). Reversing education majors' arithmetic misconceptions with short-term instruction using manipulatives. *Journal of Educational Research, 101*(4), 234-242.

Grosslight, L., Unger, C., Jay, E., & Smith, C. L. (1991). Understanding models and their use in science: Conceptions of middle and high school students and experts. *Journal of Research in Science Teaching*, *28*(9), 799-822.

Haidar, A.H. (1997). Prospective chemistry teachers' conceptions of the conservation of matter and related concepts. *Journal of Research in Science Teaching*, *34*(2), 181-197.

Hardwicke, A. J. (1995) Using molecular models to teach chemistry: Part 1 using models. *School Science Review*, *77*(278), 59-64.

Harlen, W. (Ed.). (1985). *Primary science: Taking the plunge*. Oxford: Heinemann Educational.

Harrison, A.G. & Treagust, D.F. (1996). Secondary students' mental models of atoms and molecules: Implications for teaching chemistry. *Science Education*, *80*, 509-534.

Harrison A.G. and Treagust D.F., (2002), The particulate nature of matter: challenges in understanding the submicroscopic world, In J. K. Gilbert, O. De Jong, R. Justi, D. F. Treagust and J. H. Van Driel (Eds.), *Chemical education: towards research-based practic*e, 213-234, Dordrecht: Kluwer Academic Publishers.

Hartley L.M., Wilke B.J., Schramm J.W., D'Avanzo C., Anderson C.W. (2011). College students' understanding of the carbon cycle: contrasting principle-based and informal reasoning. *BioScience*, *61*, 65-75.

Herron J.D. (1975). Piaget for chemists: explaining what ‘good’ students cannot understand. *Journal of Chemical Education, 52*, 146-150.

Hesse, J. J., & Anderson, C. W. (1992). Students’ conceptions of chemical change. *Journal of Research in Science Teaching*, *29*, 277-299.

Huddle, P.A. & Pillay, A.E. (1996). An In-Depth Study of Misconceptions in Stoichiometry and Chemical Equilibrium at a South African University. *Journal of Research in Science Teaching, 33*(1), 65-77.

Inhelder, B. and Piaget, J. (1958). *The Growth of Logical Thinking from Childhood to Adolescence*. Basic Books, New York, NY.

Instructional Design Learning Theories. (n.d.) http://www.instructionaldesign.org/theories/

Johnston, K., and Scott, P. (1991). Diagnostic teaching in the classroom: teaching/learning strategies to promote development in understanding about conservation of mass on dissolving. *Research in Science and Technological Education, 9*, 193-212.

Johnstone, A. H. (1982). Macro- and micro-chemistry. *School Science Review, 64*, 377–379

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.

Johnstone, A. H. (1993). The development of chemistry teaching: A changing response to changing demand. *Journal of Chemical Education*, *70*(9), 701–705.

Kimberlin, S., and Yezierski, E. (2016). Effectiveness of Inquiry-Based Lessons Using Particulate Level Models To Develop High School Students' Understanding of Conceptual Stoichiometry. *Journal of Chemical Education*, *93*(6), 1002-1009.

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

Krajcik, J., & Merritt, J. (2012). Engaging students in scientific practices: what does constructing and revising models look like in the science classroom? Understanding a Framework for K-12 Science Education. *Science and Children*, *49*(7), 10-13.

Lehrer, R., & Schauble, L. (2006). Scientific thinking and science literacy: Supporting development in learning in contexts. In W. Damon, R. M. Lerner, K. A. Renninger & I. E. Sigel (Eds.), *Handbook of child psychology, 6th ed*. (Vol. 4). Hoboken, NJ: John Wiley and Sons.

Levy Nahum, T., Hofstein, A., Mamlok-Naaman, R., & Bar-Dov, Z. (2004). Can final examinations amplify students' misconceptions in chemistry? *Chemistry Education: Research and Practice in Europe*, *5*(3), 301-325.

Loong, E. (2014). Fostering mathematical understanding through physical and virtual manipulatives. *Australian Mathematics Teacher*, *70*(4), 3-10.

Lovell, K. (1961). The Growth of Basic Mathematical and Scientific Concepts in Children. Philosophical Library: New-York.

Lynch, S., Kuipers, J., Pyke, C., & Szesze, M. (2005). Examining the effects of a highly rated science curriculum unit on diverse students: Results from a planning grant (PDF). *Journal of Research in Science Teaching*, *42*(8), 912-946.

Lyon, E.G. (2013). Next generation science assessment: Putting research into classroom practice. In M.S. Khine & I.M. Saleh, *Approaches and strategies in next generation science learning*, (Eds.), (pp. 247- 264). Hershey, PA: IGI Global.

Marley, S., & Carbonneau, K. (2014). Theoretical perspectives and empirical evidence relevant to classroom instruction with manipulatives. *Educational Psychology Review, 26*(1), 1-7.

Martins, I., & Ogborn, J. (1997). Metaphorical reasoning about genetics. *International Journal of Science Education*, *19*, 47-63.

Mason S.D., Shell D.F. and Crawley F.E. (1997), Differences in problem solving by nonscience majors in introductory chemistry on paired algorithmic-conceptual problems. *Journal of Research in Science Teaching*, *34,* 905-923.

McConnell, D.A., Steer, D.N., and Owens, K.D. (2003). Assessment and active learning strategies for introductory geology courses. *Journal of Geoscience Education, 51*, 205-216.

McKinnon, J. W., & Renner, J. W. (1971). Are colleges concerned about intellectual development? *American Journal of Physics, 39*, 1047-1052.

Merritt, J., & J. Krajcik. (2013). Learning progression developed to support students in building a particle model of matter. In G. Tsaparlis and H. Sevian, (Eds.), *Concepts of matter in science education*, (pp.11-45). Netherlands: Springer.

Nakhleh, M., (1992). Why Some Students Don't Learn Chemistry: Chemical Misconceptions, *Journal of Chemical Education*, *69*(3), 191-196.

Nakhleh, M.B., & Mitchell, R.C. (1993). Concept learning versus problem solving: There is a difference. *Journal of Chemical Education*, *70*(3), 190-192.

National Center for Educational Statistics (2016). *The Nation's Report Card: 2015 Science at Grades 4, 8, and 12*. National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, Washington, D.C. Retrieved from www.nationsreportcard.gov

National Research Council. (1996). *National science education standards*. Washington, DC: National Academy Press.

National Research Council. (2000). *Inquiry and the National Science Education Standards: A Guide for Teaching and Learning*. Retrieved from http://www.nap.edu/catalog.php?record_id=9596

National Research Council. (2012). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas.* Washington, DC: The National Academics.

National Research Council. (2014). *Developing Assessments for the Next Generation Science Standards.* Washington, DC: National Academies Press.

NGSS Lead States. (2013). *Next Generation Science Standards: For state, by states.* Washington, DC: National Academies Press. www.nextgenscience.org/next-generation-science-standards.

Nurrenbern, S.C., & Pickering, M. (1987). Concept learning versus problem solving: Is there a difference? *Journal of Chemical Education*, *64*(6), 508.

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

Osborne, R. (1985). Children's own concepts. In Primary science: Taking the plunge. W. Harlen. (ed)., 75-91. London: Heinemann Educational.

Ozmen, H., & Ayas, A. (2003). Students' Difficulties in Understanding of the Conservation of Matter in Open and Closed-system Chemical Reactions. *Chemistry Education: Research and Practice*, *4*(3), 279-290.

Perkins, D.N. (1992). Smart schools: From training memories to educating minds: New York: The Free Press.

Pfundt, H., and Duit, R. (2000). Bibliography: Students' alternative frameworks and science education, 5th ed. Kiel, Germany: Institute for Science Education, University of Kiel.

Phillips, D.C. (1995). The good, the bad, and the ugly: The many faces of constructivism. *Educational Researcher*, *24* (7), 5-12

Piaget, J., & Inhelder, B. (1969). *The psychology of the child*. New York: Basic Books.

Piaget, J. (1973). *To understand is to invent: The future of education* (G. A. Roberts, Trans.). New York: Grossman Publishers.

Piaget, J., & Inhelder, B. (1974). *The child's construction of quantities.* London: Routledge & Kegan Paul.

Paixo, M. F. & Cachapuz, A. (2000). Mass conservation in chemical reactions: the development of an innovative teaching strategy based on the history and philosophy of science. *Chemistry Education: Research and Practice in Europe*, *1*, 201-215. [http://www.uoi.gr/cerp]

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

Plass, J.L., Homer, B.D., & Hayward, E.O. (2009). Design Factors for Educationally Effective Animations and Simulations. *Journal of Computing in Higher Education*, *21*(1), 31-61.

Pyke C., Ochsendorf R. (2004). Conservation of matter assessment manual. Unpublished manuscript, The George Washington University.

Raghavan, K., & Glaser, R. (1995). Model-based analysis and reasoning in science: The MARS curriculum. *Science Education*, *79*, 37-61.

Ramsden, J. M. (1997). How does a context-based approach influence understanding of key chemical ideas at 16+? *International Journal of Science Education, 19*, 697-710.

Renström, L., Andersson, B., & Marton, F. (1990). Student's conceptions of matter. *Journal of Educational Psychology*, *82*(3), 555–569.

Roadranka, V. Yeany, R.H. & Padilla, M.J. (1983). *The construction and validation of GALT*. A paper presented at the annual meeting of the National Association of Research in Science Teaching. Dallas Texas.

Rutherford, F. J., & Ahlgren, A. (1990). *Science for all Americans*. Oxford: Oxford University Press.

Sadler, P. S., & Schwartz, M.S. (2004). *Research tools to evaluate student understanding.* Harvard-Smithsonian Center for Astrophysics, Science Education

Department, Cambridge, MA.

Salta, K., & Tzougraki, C. (2011). Conceptual versus Algorithmic Problem-Solving: Focusing on Problems Dealing with Conservation of Matter in Chemistry. *Research in Science Education*, *41*(4), 587-609.

Satsangi, R., & Bouck, E. (2015). Using virtual manipulative instruction to teach the concepts of area and perimeter to secondary students with learning disabilities. *Learning Disability Quarterly*, *38*(3), 174-186.

Sawrey, B. A., (1990). Concept learning versus problem solving: Revisited. *Journal of Chemical Education*, *67*, 253-254.

Schlechty, P. C. (2002). *Working on the work: An action plan for teachers, principals, and superintendents*. San Francisco: Jossey-Bass.

Schwarz, C., Reiser, B., Davis, E., Kenyon, L., Acher, A., Fortus, D., Swartz, Y., Hug, B., & Krajcik, J. (2009). Developing a learning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners. *Journal for Research in Science Teaching*, *46*(6), 632-654.

Selley, N.J. (1981). The Place of Alternative Models in School Science. *School Science Review*, *63*(223), 252-259.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. Boston, MA, US: Houghton, Mifflin and Company.

Slavings, R., Cochran, N., & Bowen, C.W. (1997). Results of a national survey on college chemistry faculty beliefs and attitudes of assessment-of-student-learning practices. *The Chemical Educator, 2*, 1-28.

Smith, P. (1978). Piaget in High School Instruction. *Journal of Chemical Education*, *55*, 115-118.

Sowelll, E. (1989). Effects of manipulative materials in mathematics instruction. *Journal for Research in Mathematics Education*, *20*, 498-505.

Stavy, R. (1991). Children's Idea About Matter. *School science and mathematics, 91*(6), 240.

Taber, K. S. (2009). College students' conceptions of chemical stability: The widespread adoption of a heuristic rule out of context and beyond its range of application. *International Journal of Science Education*, *31*(10), 1333-1358.

Thompson, S. & Lotter, C., (2014). Conservation of Matter in the Life Sciences. *Science Scope*, *38*(2), 57-69.

Treagust, D. F., Chittleborough, G. and Mamiala, T. (2001) Learning introductory organic chemistry: secondary students' understanding of the role of models and the development of scientific ideas. Paper presented at AERA 2001, Seattle, WA.

Treagust, D. F., Harrison, A. G. & Venville, G. J. (1998). Teaching science effectively with analogies: An approach for preservice and inservice teacher education. *Journal of Science Teacher Education*, *9*, 85-101.

Treagust, D.F., & Tsui, C.-Y. (Eds.). (2013). *Multiple representations in biological education*. Dordrecht, the Netherlands: Springer.

Treagust, D.F. & Tsui, C.-Y. (2014). General Instructional Methods and Strategies. In: N.G. Lederman & S.K. Abell (Eds.), *Handbook of Research on Science Education. Volume III.*, 303–320. New York: Routledge

Tsui, C.-Y., & Treagust, D.F. (2003). Genetics reasoning with multiple external representations. *Research in Science Education*, *33*(1), 111-135.

Vosniadou, S. (2007). The cognitive-situative divide and the problem of conceptual change. *Educational Psychologist*, *42*(1), 55-56.

Waldrip, B., & Prain, V. (2012). Learning from and through representations in science. In B.J. Fraser, K. Tobin, & C.J. McRobbie (Eds.), *Second international handbook of science education* (Vol. 40, pp. 145-154). Dordrecht, the Netherlands: Springer.

Waldrip, B., Prain, V., & Carolan, J. (2010). Using Multi-Modal Representations to Improve Learning in Junior Secondary Science. *Research in Science Education*, *40*, 65-80.

Wise, K.C., & Okey, J.R. (1983). A meta-analysis of the effects of various science teaching strategies on achievement. *Journal of Research in Science Teaching*, *20*(5), 419-435.

Yarroch, W. L. (1985). Student understanding of chemical equation balancing. *Journal of Research in Science Teaching*, *22*, 449-459.

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