2015

Unpacking students' atomistic understanding of stoichiometry

John Ysrael Baluyut
Iowa State University

Recommended Citation

Baluyut, John Ysrael, "Unpacking students' atomistic understanding of stoichiometry" (2015). *Graduate Theses and Dissertations*. 14753.
https://lib.dr.iastate.edu/etd/14753

Unpacking students' atomistic understanding of stoichiometry

by

John Ysrael Baluyut

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Chemistry (Chemical Education)

Program of Study Committee:

Thomas A. Holme, Major Professor
Joseph W. Burnett
Thomas J. Greenbowe
Jacob W. Petrich
Theresa L. Windus

Iowa State University

Ames, Iowa

2015

Table of Contents

Acknowledgment

This dissertation is not the work of just one man. It is the fruit of many hours of hard work of a team who have directly or indirectly helped me see the accomplishment of a great milestone in my academic career. I wish to express my deepest gratitude to each of them:

Professor **Thomas A. Holme** – for the invaluable guidance and support, for showing me that indeed there can be so much fun even as we do so much work;

Professors Joseph Burnett, Thomas Greenbowe, Jacob Petrich, and **Theresa Windus** – for reminding me that there is still so much we can learn even after we have learned so much; Drs. Cristina Bonaccorsi, Kathy Burke, Jesudoss Kingston, Irmi Miller, and Javier Vela – for giving me access to my most important research resources: their students;

Drs. Allie Brandriet, Jenny Brigham, Sanjay Hari, Professors Kim Linenberger, Cindy Luxford, and Jeffrey Raker – for being mentors extraordinaire;

Drs. Chamila De Silva and Jessica Reed, Jack Polifka and Anna Prisacari – for always making me feel as young as they are;

The many students who participated in various phases of my study;

Jenne Minelli – for letting me into her home;

Precious Agnis, Dr. Vergel Bungay, Dr. Mel Cantorias, Dr. Kathlia De Castro, **Larry Malijan**, Dr. Danilo Ortillo, **Roy Parreno**, Dr. Lani Ramos, Dr. Celeste Regino, and Rannel Verdeflor - friends so far yet who are always so near;

And our **Heavenly Father** – from whom all things come, and to whom I offer everything.

Abstract

Despite the use by instructors of particulate nature of matter (PNOM) diagrams in the general chemistry classroom, misconceptions on stoichiometry continue to prevail among students tasked with conceptual problems on concepts of limiting and excess reagents, and reaction yields. This dissertation set out to explore students' understanding of stoichiometry at the microscopic level as they solved problems that using PNOM diagrams. In particular, the study investigated how students coordinated symbolic and microscopic representations to demonstrate their knowledge of stoichiometric concepts, quantified the prevalence and explained the nature of stoichiometric misconceptions in terms of dual processing and dual coding theories, and used eye tracking to identify visual behaviors that accompanied cognitive processes students used to solve conceptual stoichiometry problems with PNOM diagrams.

Interviews with students asked to draw diagrams for specific stoichiometric situations showed dual processing systems were in play. Many students were found to have used these processing systems in a heuristic-analytic sequence. Heuristics, such as the factor-label method and the least amount misconception, were often used by students to select information for further processing in an attempt to reduce the cognitive load of the subsequent analytic stage of the solution process.

Diagrams drawn by students were used then to develop an instrument administered over a much larger sample of the general chemistry student population. The robustness of the dual processing theory was manifested by response patterns observed with large proportions of the student samples. These response patterns suggest that many students seemed to rely on heuristics to respond to a specific item for one of two diagrams given for the same chemical context, and then used a more analytic approach in dealing with the same item for the other diagram. It was also found that many students incorrectly treated items dealing with the same chemical context independently of each other instead of using a more integrative approach.

A comparison of the visual behaviors of high-performing subjects with those of low-performers revealed that high performers relied heavily on the given diagrams to obtain information. They were found to have spent more time fixating on diagrams, looked between the chemical equation and the diagram for each problem more often, and used their episodic memory more heavily to collect information early on than low performers did. Retrospective think-alouds used with eye tracking also revealed specific strategies, such as counting and balancing of atoms and molecules across both sides of a diagram, as well as comparing ratios between atoms and molecules in a diagram with those given in a balanced equation, used by students to analyze PNOM diagrams.

CHAPTER 1. GENERAL INTRODUCTION

The Role of External Representations in Chemistry Education

The study of chemistry requires students to consider concepts and entities that are not visible to the naked eye. Students frequently have to deal with understanding processes that occur at the microscopic level and envisioning the components of these events as well as their interactions is often a challenge for novices. Frequently, this results in erroneous conceptions and poor course performances among beginning students of chemistry. The use of external representations, such as particulate nature of matter diagrams (PNOM), aims to help students understand unseen chemical processes (Gilbert, Reiner, & Nakhleh, 2008). External representations include physical symbols, objects, dimensions, as well as the rules, constraints, or relations within them (Zhang, 1997). Information from external representations are picked up, analyzed, and processed by perceptual systems alone. In chemistry, this is normally done by anchoring concepts with the help of characteristics such as color, size, and shape to help students visualize and understand abstract chemical concepts in terms of more familiar and concrete representations. Internal representations, on the other hand, are the knowledge and structure in memory. Information from internal representations is retrieved using cognitive processes, often with the help of external representations. Consider, for example, Figure 1, which is a PNOM diagram representing the complete reaction between three moles of methane and four moles of oxygen. From the diagram, students are supposed to infer that the oxygen molecules must be the limiting reagent because they were all used to form either carbon dioxide or water molecules, and that the methane molecule of the right side of the diagram must be an excess molecule. They are also supposed to determine that the reaction went to completion.

$$CH_3(g) + 2\ O_2(g) \rightarrow CO_2(g) + 2\ H_2O(g)$$

Figure 1. A particulate nature of matter (PNOM) diagram representing the complete combustion of three moles of methane in the presence of four moles of oxygen. In this diagram, each black sphere represents a carbon atom, red spheres represent oxygen atoms, and small light blue spheres represent hydrogen atoms.

Representational competence, or the understanding of how and when to use external representations (Kozma & Russell, 1997), is an important skill for students to gain as they study chemistry because representations do not automatically translate into learning. Students may not, at least initially, know how and when to use a representation in the domain. To completely understand how the use of representations can be most appropriately taught, it is important that instructors know the nature of interactions with representations which best support student learning, how a student's ability affects their interactions with the representations, and different kinds of tasks that are most suitable with the use of a representation (Hinze et al., 2013). Representational competence requires that students learn how to analyze, obtain information from, and explain different features of a representation (Kozma &

Russell, 2005). Experts, of course, possess these skills, usually from years of practice and experience. They are often able to smoothly transition between different types of representations, so that expert chemists, for example, can easily go obtain and integrate information from both a balanced chemical equation and the PNOM diagram that comes with it, and then decide whether the diagram accurately reflects the meaning of the equation or not. Students, on the other hand, usually focus only at the surface-level features of a representation and for the most part do not understand what these features mean (Kozma, 2003). It has been argued that the type and clarity of a representation is important to how students understand chemical concepts (Tasker, 2004) as representations have also been known to induce conceptual difficulties (Ametller & Pintó, 2002).

To use an external representation meaningfully, students must go through two developmental processes (Ainsworth, 2006). First, students must first learn the format and conventions of the representation including features and operations that come along with the representation’s use. Students need to know what kind of and how information from the representation is obtained. This means understanding how the different features of the representation are coordinated with each other to yield useful information. Second, students must understand exactly how information derived from the representation is used to generate inferences in relation to the representation’s domain.

Chemistry education is an especially interesting domain as far as the development of representational competency is concerned. The teaching of chemistry necessarily includes dealing with concepts and processes that are mostly unseen even as their implications are usually manifested in everyday human experiences. Chemistry instructors make use of visualizations to convey important information so that students may develop effective reasoning skills in chemistry. Still, integrating what is observed at the macroscopic level with what happens at the microscopic level and then translating these events in terms of symbols is a great challenge chemistry students face (Johnstone, 1993). External representations such as PNOM diagrams aim to make such an understanding more accessible to the beginning chemistry student.

PNOM Diagrams and Chemical Stoichiometry

Stoichiometry is often a difficult concept for many chemistry students to grasp. Proof of this comes from challenges students face as they try to understand concepts like excess and limiting reagents, and reaction yields, even if they can use algorithms effectively to solve numerical problems. A study that included responses to a pair of questions involving limiting reagents revealed that three out of five students could successfully determine which of two reactants was limiting as well as how much of the excess reagent was left over based on given numerical information (Nurrenbern & Pickering, 1987). However, less than 10% of students from the same group chose the correct form of the chemical equation that would symbolically represent a hypothetical reaction mixture depicted by a PNOM diagram or how changes in a reaction mixture can be appropriately depicted with a PNOM diagram. Even among honors students and declared chemistry majors from a large Midwestern university, less than half could correctly pick the limiting reagent given a mixture of a small numbers of hypothetical molecules that react with each other (Nakhleh, 1993). These examples reflect the gap between students' abilities to solve numerical problems on limiting reagents and their understandings of chemical change at the particulate level. Many previous studies have recommended giving emphasis to the use of visual approaches using PNOM diagrams when dealing with stoichiometric principles (Ben-Zvi, Eylon, & Silberstein, 1987; Sanger, 2005), yet the student difficulties persist. This begs the question: how exactly are students using (or failing to use) PNOM diagrams to help gain a better conceptual understanding of stoichiometric principles?

Purpose of the Study and Research Questions

The purpose of this study was to investigate first-year chemistry students' understandings of stoichiometry concepts of excess and limiting reagents, and yield, and to determine how students coordinate representations at the symbolic and microscopic levels as they solve problems involving these concepts. Specifically, the research questions that guided this dissertation include:

1. What are first-year chemistry students' understandings of excess and limiting reagents and yield?
    a. How are students' understandings of excess and limiting reagents and yield articulated using PNOM diagrams?
    b. What misconceptions about excess and limiting reagents and yield are manifested by how students coordinate information obtained from chemical equations and PNOM diagrams?

2. How prevalent are misconceptions on excess and limiting reagents and yield among first-year chemistry students?
3. How can students with high and low prior knowledge be distinguished from each other in terms of their visual behaviors when asked to solve stoichiometry problems that use PNOM diagrams?
4. What cognitive processes come with specific types of visual behaviors as students solve stoichiometry problems using PNOM diagrams?

Learning Theories Underlying This Study

Constructivism

The constructivist theory of learning sees students as being directly responsible for meaning and knowledge through active involvement in the learning process. Learners, therefore "construct" their own knowledge based on what they have previously known. This means students must have opportunities to express their own ideas, test those ideas with experiments and discourse in the classroom, and then think about connections between the chemical phenomena they might be investigating and other aspects of their lives (Wolffe & McMullen, 1995). Knowledge is then generally developed and transmitted within a social context (Crotty, 1998). The idea behind constructivism is that students have to make sense or assign meaning to information they obtain based on what they already previously know (Ferguson, 2007). Students are responsible for constructing their own knowledge instead of merely absorbing ideas talked about by their instructors (Lunenburg, 1998).

If students are to take an active role in building their own knowledge, then they must be able to modify knowledge that currently exists in their minds though a process of conceptual change (Posner, Strike, Hewson, & Gertzog, 1982). However, for conceptual change to occur, four conditions must first be met (Nussbaum & Novick, 1982): (1) learners must be dissatisfied with their existing conceptions; (2) learners must understand the new conception; (3) the new knowledge must somehow fit with other previously acquired knowledge; and (4) the new knowledge must be applicable and be useful to generate even newer knowledge in the future.

Meaningful Learning

Meaningful learning is the retention, understanding, and application of new information so that it can be used to interpret a situation that is different from the original context in which the information was obtained. Meaningful learning is diametrically opposed to rote learning, where information is merely memorized but is not understood well enough to be transferred to a different situation. To allow for meaningful learning to occur, students must be able to link new information with prior knowledge, give meaning to the new information, and then choose to incorporate the new information with prior knowledge (Ausubel, 2012; Bretz, 2001). It is, thus, important to identify prior knowledge among students because this has single-handedly the greatest influence on how students learn new information (Ausubel, 1963). Students' misconceptions, for example, often interfere with how students learn new material. Identification of prior knowledge among students, therefore, allows instructors to guide students in the assimilation of new information while at the same time breaking down misconceptions. This can be accomplished by explicitly showing how new material being presented in the classroom may be related with previous knowledge.

The Unified Learning Model

The unified learning model (ULM) is a composite of several principles obtained from different theories of learning that attempts to provide a singular model for learning (Kauffman & Shell, 2012). The ULM identifies a student's working memory, prior knowledge and motivation as the main drivers of the learning phenomenon. The working memory lies at the very center of the ULM. It is the part of the mind where storage and processing of information takes place, and it determines how learning occurs and what instructional methods or techniques might hinder or facilitate learning. A student's prior knowledge forms ULM's second core component. Given that new knowledge is the goal of this model, prior knowledge is used to influence just exactly how working memory operates. New information obtained from the teacher is sorted out and stored in the brain and influenced by the connections between ideas already in the learner's mind. Finally, motivation serves as the impetus for directing working memory to perform a task, specifically, the task of learning. ULM is primarily based on three principles of learning: (1) learning results principally from the allocation of working memory; (2) the capacity to allocate working memory is affected by the learner's prior knowledge; and (3) the manner in which working memory is allocated is directed by the learner's motivation. Using the framework of the ULM then, the role of the instructor is to help focus a student's attention to the concept being taught. The instructor needs to be aware of his or her students' prior knowledge so that students may be guided to make the most appropriate connections between what the students already know and what is being learned. For example, in teaching the concepts of excess and limiting reagents, and reaction yield using PNOM diagrams, the instructor must make sure his or her students understand exactly how amounts of products are determined from reactants in given stoichiometric quantities, how such a process may be appropriately depicted at the microscopic level using a PNOM diagram, and how the diagram can be connected to information students might obtain from the symbolic representation of the reaction. This framework was used to guide the analyses of how students coordinated information obtained from microscopic and symbolic representations as they responded to conceptual questions on stoichiometry.

Dual-Process Theory

Dual process theories suggest that there are two different modes by which information is cognitively processed and decisions are made (Evans, 2008). These two processing modes have been distinguished from each other based on their processing rates, capacity, and the manner in which information is processed by the decision-maker. The *System 1* mode of processing is characterized to occur unconsciously, rapidly, automatically, and is able to process large amounts of information (Kahneman & Frederick, 2002). On the other hand, *System 2* processes are described as those that occur consciously, slowly, and in a very deliberate manner. System 1 processing has also been described as being often influenced by the use of heuristics while System 2 is more analytical (Evans, 1996). This heuristic-analytical dichotomy describes the role of heuristics to account for biases towards a pragmatic and preconscious level of processing information before any analysis takes place. When students solve problems, for example, they often first use heuristics to selectively focus their attention on certain task features based on their prior knowledge. As a result, the analytical process that follows is applied only on these selected representations as well. The sequential application of these modes to process information may, sometimes, lead to relevant information being excluded from and irrelevant information being included in the analytical part of the information processing sequence. The processing sequence also often leads to selection of default responses provided by heuristics, except when the analytic system intervenes when a student is cued by some strong deductive reasoning instructions.

Dual Coding Theory

People have separate verbal and non-verbal information processing systems (Clark & Paivio, 1991). As a result, learning may be enhanced when both verbal and visual information are presented to students. The use of pictures with verbal information has been found to be superior because the verbal codes for pictures are more easily accessed by students than the image codes for words. Thus, pictures are more often than not dually coded. Still, sufficient verbal information to which the visual information may be referred is necessary for students to successfully perform problem-solving tasks (Mayer & Sims, 1994). In the absence of verbal information, students with insufficient prior knowledge may have difficulties executing problem-solving tasks using only the visual information they have. When both visual and verbal information are provided, students actively select, organize, and integrate information to come up with coordinated explanations in verbal and visual formats (Mayer, 2002). Students who display representational competence are able to use multiple representations to explain phenomena, make and support claims, and form predictions (Kozma & Russell, 2005). However, most college chemistry students are not proficient at transforming representations (Kozma & Russell, 1997). It is difficult for them to know what pieces of information to select, how to organize such information, and what form of integration makes the most sense. Students mostly have great difficulties both in understanding how representations are used to illustrate chemical concepts and how to articulate their own understanding of chemical processes using representations.

Mixed Methods Design

The objectives and research questions for this study directed the choice of methods used. Some of the research questions had a cause-and-effect nature that was better addressed using quantitative methods. These questions were tested using an assessment instrument that was administered online over a large sample of students. The findings from this part of the study can be generalized to a much larger population. On the other hand, some of the questions tended to be more exploratory and descriptive in nature. These required the use of more qualitative methodologies. In answering these questions, the main goal was to gain a deep understanding of what subjects experienced as they went through specific aspects of the study (Maykut & Morehouse, 1994). As a result, during the more exploratory parts of the study, sample sizes are limited and subjects are invited only to the point of *saturation* (Creswell &

Clark, 2007). In qualitative research, saturation is said to be reached when the collection of new data no longer sheds new light on the issue being investigated. While a few authors offer guidelines for sample sizes that may lead to saturation (Creswell & Clark, 2007; Morse, 1995, 2000), no empirical arguments are offered to explain the minimum number of subjects needed (Mason, 2010). In fact, work done by other researchers suggest that many do not strictly adhere to these guidelines (Thomson, 2004).

Once saturation is reached, data collected using qualitative methods may be broken down, examined, compared, and categorized in a process known as *coding* (Corbin & Strauss, 2014). Coding generally consists of steps such as labeling phenomena that took place during data collection, discovery of categories, and identification of properties and dimensions that describe the categories. The intent of coding is to break down collected data into conceptual components. As categories emerge from the analysis of initial sets of data, bits of information from succeeding sets are compared to help link and refined earlier defined categories, and even possibly define new categories. This is known as the *constant comparative method* (S. Kolb, 2012). The goal then is to generate a model about how categories defined using earlier obtained data fit with data that are analyzed later.

Eye Tracking

Eye tracking has been used to analyze how students view and perceive visual stimuli related to problem solving in chemistry (Stieff, Hegarty, & Deslongchamps, 2011; Tang, Kirk, & Pienta, 2014; Tang & Pienta, 2012; Tang, Topczewski, Topczewski, & Pienta, 2012; Williamson, Hegarty, Deslongchamps, Williamson Iii, & Shultz, 2013). This method provides insights in ways that no other technique can capture about how visual behavior may be related to the cognitive processes that go on as students solve problems (Havanki & VandenPlas, 2014). This section focuses on the use of eye tracking data first by describing the physiology of human eye movements, technology that exists in eye tracking, and how studies using eye tracking help relate visual behavior to cognitive processes.

The Physiology of Vision

Among the parts of the human eye that are most directly involved with receiving and processing visual information are the pupil, iris, cornea, and sclera (H. Kolb, Fernandez, & Nelson, 1995). The pupil allows light to go into the eye with the help of the iris, which regulates the size of the pupil. Meanwhile, the cornea is responsible for producing sharp images as it covers both the pupil and the iris. The sclera is the white portion of the eye that is supporting the wall of the cornea and is continuous with it.

The visual field is generally divided into three different regions (Figure 2): the fovea which is found near the middle of the retina; the parafovea, which is just outside of the fovea; and the peripheral , which is located just beyond the parafovea (Rayner, 1998). The fovea is densely covered by receptors which lead to higher acuity vision and covers two degrees of the visual field. The parafovea extends the vision by about five degrees on either side of fixation but the acuity is not as good. While the peripheral vision is generally characterized by greater acuity than the parafovea, people usually cannot see objects by using only the peripheral. To sufficiently see objects, the eyeball must be moved so that the object of visual interest appears directly on the fovea (Rayner, 1998).

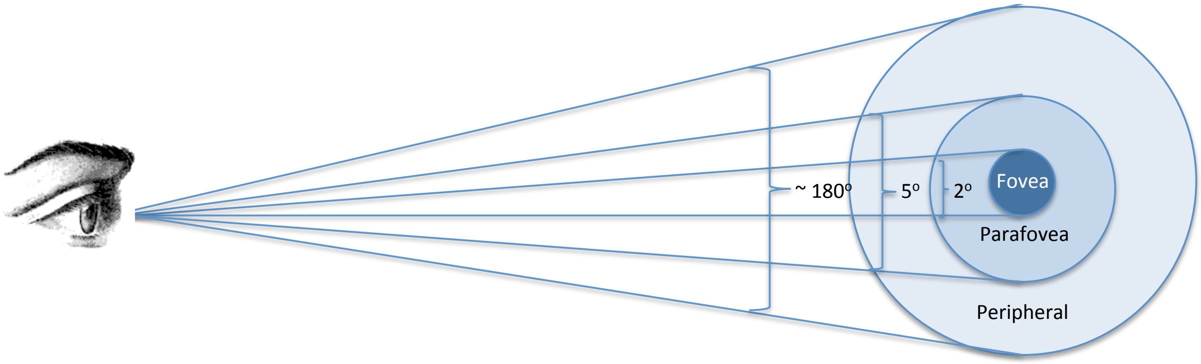


Figure 2. Schematic diagram of the human eye's vision field.

The eye generally does not move smoothly across the field covered by visual stimuli. The eye usually makes quick continuous movements known as *saccades*, and then remain relatively still during *fixations* that occur in between saccades. Most fixations last about 200 to 300 ms. Saccades can be described as quick jumps from one area of the visual stimulus to the next and may have angular velocities as quick as 500 degrees per second. Once a saccade starts, it is almost impossible to change either its destination or path. Often a saccade’s destination is chosen before the movement starts mostly with the use of peripheral vision. Saccades take between 30 and 120 ms and will cover up to 40 degrees of the visual field.

Fixations are periods of relative stillness of the eye and take place in between saccades. They are indications of visual attention, focusing on the object of the moment’s interest. The eye does not, of course, remain completely still during a fixation but goes through small motions of about one degree. Fixations usually last between 200 and 600 ms, and are followed by saccades.

Eye Movement and Cognitive Processes

There are many examples in the literature of how information from eye tracking has been used to understand cognitive processes that underlie visual behavior (Just & Carpenter, 1980; Rayner, 2009; Rayner, Chace, Slattery, & Ashby, 2006). Most eye tracking studies rely on two basic assumptions in trying to connect visual behavior to cognition. The *immediacy assumption* (Just & Carpenter, 1980) has the subject interpreting information from the referent before moving on to the next fixation. The subject decides how information fits with internal representations held in working memory and the eye does not move until this process is completed. The *eye-mind assumption*, on the other hand, suggests that there is no measurable time lag between eye fixation on a referent and the information obtained from processing. Thus, the amount of time spent fixating on a referent is a measure of the processing time. Based on these assumptions, researchers are able to make connections between visual behavior and how visual stimuli are processed (Havanki & VandenPlas, 2014).

Advantages, Disadvantages, and Limitations of Eye Tracking

Among the most frequently cited advantages of conducting an eye tracking study is that it gives researchers real-time access to some information about visual and cognitive processes going on within a subject in a sensitive yet unobtrusive way (Henderson & Ferreira, 2013). No special behavior is required from participants, and the more recent methods do not affect how subjects behave as they try to complete the task being studied. The researcher, thus, is able to directly observe natural viewing behaviors of subjects during task performance. Eye tracking also allows researchers to collect data that may be difficult for subjects to describe. Subjects may, for example, not remember every single fixation their eyes went through as they viewed a stimulus. Also, eye tracking provides large quantities of data that can be used for quantitative analysis. Even with the slowest eye trackers that have sampling frequencies of 25 Hz, a ten-minute eye tracking session provides 15,000 numerical data points about an individual subject that can be analyzed with a wide range of statistical tools.

These days, the extent of the time commitment required in carrying out a study seems to be the most significant disadvantage in doing eye tracking. Eye tracking studies need to be very carefully planned, from the design and creation of visual stimuli, review by a human studies board, selection of a sufficiently large number of appropriate subjects, time to run the actual eye tracking sessions, reduction of data into more manageable forms, and analysis (Havanki & VandenPlas, 2014).

Probably the most important limitation to remember in doing eye tracking is the fact that the data obtained directly reflects only what a subject viewed during the session, and not the cognition going on within the subject's mind (de Koning, Tabbers, Rikers, & Paas, 2010). It may very well be possible for subjects to spend a large chunk of time viewing a stimulus not because there is a lot of cognitive processing going on, but precisely because of the subject's own limitations to have any understanding of the stimulus. This limitation is often addressed by triangulation techniques such as interviews, testing, concept mapping, and physiological measurements.

Another limitation of eye tracking is that it is focused on foveal vision, which as earlier mentioned accounts for only two percent of a subject's visual field. None of the peripheral vision is measured by eye tracking even though significant visual events like color and shape recognition as well as decision making for subsequent fixations mostly take place using at the periphery (Holmqvist et al., 2011).

Visual behavior is greatly task dependent. How subjects exhibit visual behavior will depend to a great extent on the design of the tasks they are asked to perform, including the explicit instructions subjects receive from the researchers. This has major implications on the experimental design used in and the generalizability of the study. Subjects must, therefore, receive directions that are as identical to each other as possible, especially if they are expected to perform the same tasks in almost identical ways (Rayner, Rotello, Stewart, Keir, & Duffy, 2001).

Dissertation Outline

The dissertation follows the progression of the research project though its different phases. It reflects how a coherent attempt at understanding different aspects of students’ use of verbal and visual information when tasked with problems in stoichiometry that make use of PNOM diagrams. Chapter 2 presents qualitative information obtained from a pilot interview study that asked students to describe thought processes they went through as they drew PNOM diagrams to illustrate their concept of excess and limiting reagents and yield using specific chemical contexts. Students from honors programs as well as those who have declared majors in chemistry and closely-related fields were interviewed. Chapter 3 describes efforts to quantify the extent of the thinking processes identified earlier in terms of the prevalence of response patterns obtained using an instrument administered with a much larger sample of students. Chapter 4 highlights the use of eye tracking to identify visual behaviors that come with cognitive processes as students attempt to organize and integrate verbal and visual information to solve stoichiometry problems involving PNOM diagrams. This chapter also describes differences between the visual behaviors of high- and low-performing students. Chapter 5 gives a summary of major findings, their implications, and suggests future work.

REFERENCES

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

Ametller, J., & Pintó, R. (2002). Students' reading of innovative images of energy at secondary school level. *International Journal of Science Education, 24*(3), 285-312.

Ausubel, D. P. (1963). The psychology of meaningful verbal learning.

Ausubel, D. P. (2012). *The acquisition and retention of knowledge: A cognitive view*: Springer Science & Business Media.

Ben-Zvi, R., Eylon, B., & Silberstein, J. (1987). Students' visualisation of a chemical reaction. *Educ. Chem., 24*, 117-120.

Bretz, S. L. (2001). Novak's theory of education: Human constructivism and meaningful learning. *Journal of Chemical Education, 78*(8), 1107.

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational psychology review, 3*(3), 149-210.

Corbin, J., & Strauss, A. (2014). *Basics of qualitative research: Techniques and procedures for developing grounded theory*: Sage publications.

Creswell, J. W., & Clark, V. L. P. (2007). Designing and conducting mixed methods research.

Crotty, M. (1998). *The foundations of social research: Meaning and perspective in the research process*: Sage.

de Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2010). Attention guidance in learning from a complex animation: Seeing is understanding? *Learning and instruction, 20*(2), 111-122.

Evans, J. S. B. T. (1996). Deciding before you think: Relevance and reasoning in the selection task. *British Journal of Psychology, 87*(2), 223-240.

Evans, J. S. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annu. Rev. Psychol., 59*, 255-278.

Ferguson, R. L. (2007). Constructivism and Social Constructivism. In G. Bodner & M. Orgill (Eds.), *Theoretical Frameworks for Research in Chemistry/Science Education*. Upper Saddle River, NJ: Pearson Education, Inc.

Gilbert, J. K., Reiner, M., & Nakhleh, M. (2008). *Visualization: Theory and practice in science education*. United Kingdom: Springer.

Havanki, K. L., & VandenPlas, J. R. (2014). Eye Tracking Methodology for Chemistry Education Research. In D. M. Bunce & R. S. Cole (Eds.), *Tools of Chemistry Education Research*. Washington, D. C.: American Chemical Society.

Henderson, J., & Ferreira, F. (2013). *The interface of language, vision, and action: Eye movements and the visual world*: Psychology Press.

Hinze, S. R., Rapp, D. N., Williamson, V. M., Shultz, M. J., Deslongchamps, G., & Williamson, K. C. (2013). Beyond ball-and-stick: Students' processing of novel STEM visualizations. *Learning and Instruction, 26*, 12-21.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*: Oxford University Press.

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

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: from eye fixations to comprehension. *Psychological review, 87*(4), 329.

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. *Heuristics and biases: The psychology of intuitive judgment, 49*.

Kauffman, D., & Shell, D. F. (2012). The Unified Learning Model: How Motivational, Cognitive, and Neurobiological Sciences Inform Best Teaching Practices.

Kolb, H., Fernandez, E., & Nelson, R. (1995). Visual Acuity--Webvision: The Organization of the Retina and Visual System.

Kolb, S. (2012). Grounded Theory and the Constant Comparative Method: Valid Research Strategies for Educators. *Journal of Emerging Trends in Educational Research and Policy Studies, 3*(1), 83-86.

Kozma, R. B. (2003). The material features of multiple representations and their cognitive and social affordances for science understanding. *Learning and Instruction, 13*(2), 205-226.

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

Kozma, R. B., & Russell, J. (2005). Students becoming chemists: Developing representational competence *Visualization in science education* (pp. 121-145): Springer.

Lunenburg, F. C. (1998). Constructivism and Technology: Instructional Designs for Successful Education Reform. *Journal of instructional psychology, 25*(2), 75-81.

Mason, M. (2010). Sample Size and Saturation in PhD Studies Using Qualitative Interviews. *Forum: Qualitative Social Research, 11*(3).

Mayer, R. E. (2002). Cognitive Theory and the Design of Multimedia Instruction: An Example of the Two-Way Street Between Cognition and Instruction. *New directions for teaching and learning, 2002*(89), 55-71.

Mayer, R. E., & Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of educational psychology, 86*(3), 389.

Maykut, P., & Morehouse, R. (1994). *Beginning Qualitative Research: A Philosophic and Practical Guide*. London: Routledge-Falmer.

Morse, J. (1995). The significance of saturation. *Qualitative Health Research, 5*(3), 147-149.

Morse, J. (2000). Determining sample size. *Qualitative Health Research, 10*(1), 3-5.

Nakhleh, M. B. (1993). Are our students conceptual thinkers or algorithmic problem solvers? Identifying conceptual students in general chemistry. *Journal of Chemical Education, 70*(1), 52.

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

Nussbaum, J., & Novick, S. (1982). Alternative frameworks, conceptual conflict and accommodation: Toward a principled teaching strategy. *Instructional science, 11*(3), 183-200.

Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science education, 66*(2), 211-227.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological bulletin, 124*(3), 372.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The quarterly journal of experimental psychology, 62*(8), 1457-1506.

Rayner, K., Chace, K. H., Slattery, T. J., & Ashby, J. (2006). Eye movements as reflections of comprehension processes in reading. *Scientific Studies of Reading, 10*(3), 241-255.

Rayner, K., Rotello, C. M., Stewart, A. J., Keir, J., & Duffy, S. A. (2001). Integrating text and pictorial information: eye movements when looking at print advertisements. *Journal of Experimental Psychology: Applied, 7*(3), 219.

Sanger, M. J. (2005). Evaluating students' conceptual understanding of balanced equations and stoichiometric ratios using a particulate drawing. *Journal of Chemical Education, 82*(1), 131.

Stieff, M., Hegarty, M., & Deslongchamps, G. (2011). Identifying representational competence with multi-representational displays. *Cognition and Instruction, 29*(1), 123-145.

Tang, H., Kirk, J., & Pienta, N. J. (2014). Investigating the Effect of Complexity Factors in Stoichiometry Problems Using Logistic Regression and Eye Tracking. *Journal of Chemical Education, 91*(7), 969-975.

Tang, H., & Pienta, N. (2012). Eye-tracking study of complexity in gas law problems. *Journal of Chemical Education, 89*(8), 988-994.

Tang, H., Topczewski, J. J., Topczewski, A. M., & Pienta, N. J. (2012). Permutation test for groups of scanpaths using normalized Levenshtein distances and application in NMR questions. *Proceedings of the Symposium on Eye Tracking Research and Applications*, 169-172.

Tasker, R. (2004). Using multimedia to visualize the molecular world: educational theory into practice. *A chemist's guide to effective teaching, 16*, 256-272.

Thomson, B. (2004). Qualitative research: Grounded theory—Sample size and validity. Retrieved from http://www.buseco.monash.edu.au/research/studentdocs/mgt.pdf

Williamson, V. M., Hegarty, M., Deslongchamps, G., Williamson Iii, K. C., & Shultz, M. J. (2013). Identifying student use of ball-and-stick images versus electrostatic potential map images via eye tracking. *Journal of Chemical Education, 90*(2), 159-164.

Wolffe, R. J., & McMullen, D. W. (1995). The constructivist connection: Linking theory, best practice, and technology. *Journal of computing in teacher education, 12*(2), 25-28.

Zhang, J. (1997). The Nature of Eternal Representations in Problem Solving. *Cognitive Science, 21*(2), 179-217.

CHAPTER 2. STUDENTS’ VISUALIZATION OF LIMITING AND EXCESS REAGENTS, AND REACTION YIELDS

The study of chemistry can be seen as the observation of macroscopic phenomena, which can be explained in terms of properties of microscopic entities. Dealing with chemical phenomena, such as chemical reactions, often requires teachers and students to go across three levels of representations, namely the macroscopic, the microscopic, and the symbolic (Johnstone, 1993). Representations such as equations, concrete models, drawings, tables, and simulations are often employed to express the mental visualizations chemists have about molecules and their transformations (Harrison & Treagust, 2000). Students are, therefore, expected not only to learn chemistry from ideas that are expressed verbally but also from the creation and manipulation of visual representations both at the macroscopic and the microscopic levels (Bucat & Mocerino, 2009). For instance, the learning of limiting and excess reagents, as well as yields of chemical reactions might involve students’ ability to visualize how individual atoms and molecules of reactants may interact with each other (at the microscopic level) to form varying amounts of products (at the macroscopic level). Then they might compare their solutions with what they might obtain using a numerical algorithm.

Particulate nature of matter (PNOM) diagrams are used to convey information, explain, visualize, help predict, and form hypotheses about chemical phenomena. These diagrams are often static, although with animation technology, many are now also being presented in dynamic forms. PNOM diagrams represent molecules, atoms, and sub-atomic particles either as single particles or as arrays of particles schematically. Most chemistry experts have little difficulty with interpreting these diagrams. On the other hand, the interpretation of PNOM diagrams is often a great challenge to the novice (Johnstone, 1993; Treagust & Chittleborough, 2001). The use of PNOM diagrams has been shown to do little in overcoming student difficulties as far as gaining a conceptual understanding of stoichiometry and chemical equations is concerned (Ben-Zvi, Eylon, & Silberstein, 1987; Sanger, 2005). Probably an important reason for this difficulty among students in gaining competence with PNOM diagrams is students' lack of experience with the microscopic domain. For example, in a study that asked first-year college students in a midterm examination to balance the equation for the combustion of methane and then draw a corresponding PNOM diagram, it was found that while almost all of the students (96%) came up with a correctly balanced equation, only a little more than one in five (21.6%) drew PNOM diagrams that appropriately illustrate the same reaction (Nyachwaya, Warfa, Roehrig, & Schneider, 2014). This failure of many students to correctly coordinate symbolic and microscopic representations with each other has been pointed out and elaborated (Johnstone, 1993). A gap exists between the models chemists use to describe, explain, and predict properties of substances and chemical processes and understanding how chemical symbols are manipulated to represent and visualize the important components of such models (Talanquer, 2012).

Dual Coding Theory of Information Processing

The gap between models used by chemists and the coordination of chemical representations may be explained in terms of the differences by which verbal and visual information are processed (Clark & Paivio, 1991; Paivio, 2014). *Dual coding theory* assumes that most cognitive processes use both verbal and nonverbal representations, and that two independent yet connected systems are used during cognition (Figure 1). To process information effectively, connections have to be made between different representations both in the verbal and in the visual system. Two kinds of connections are described by the dual coding theory. Those occurring between mental visual representations within the visual system are called associative connections. The reading of a diagram forms a representation in the mental visual system. It is guided by previously existing mental visual representations in a student's mind. The formation of associative connections between mental visual representations is, thus, triggered by diagrams. The second kind of connections are those between mental visual representations and mental verbal representations. These are known as referential connections. Naming a diagram or drawing a diagram based on its name uses referential connections. When students, for example, study limiting reagents with diagrams, they need to describe how the limiting reagent is completely consumed by a complete reaction based on diagrams that show individual atoms and molecules of reactants and products involved in the reaction.

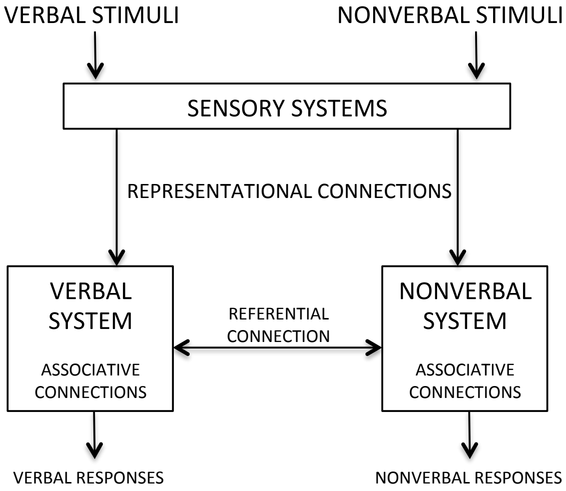


Figure 1. Verbal and nonverbal systems in the dual coding theory (Clark & Paivio, 1991).

The advantage of using visual cues in combination with verbal cues during chemistry instruction and problem solving has been pointed out in previous research (Cheng & Gilbert, 2014; Sanger & Greenbowe, 2000). Whereas words are coded verbally, students are more likely to code diagrams both verbally and visually. This usually results in better recall of pictures. Since learning chemical ideas often demands visual exactness, students should find it more effective when both visual and verbal representations are used.

Dual-Process Theory of Cognition

Aside from distinguishing between verbal and nonverbal information processing, researchers have suggested differentiating between two different kinds of cognitive processes (Evans, 2008; Kahneman & Frederick, 2002; Stanovich & Thompson, 2001). This *dual-process theory* suggests that both implicit and explicit learning work in the same mind underlying thinking and reasoning. *System 1* is said to account for the more implicit way of reasoning and is characterized by being automated, requiring little effort, having a large capacity, being fast, and being domain specific. Thinking that relies mostly on heuristics is said to be the result of System 1 processing. The use of System 1 processing also includes recognition of patterns, an overall perceptions of the situation, and orientation of the subject with respect to what is going on. On the other hand, thinking that seems to be more explicit, controlled, requires more effort and is slower is classified as having come from System 2 processes. This includes following rules, comparisons, and weighing of options. System 2 is, thus, more analytic in its approach and usually leads to improvement of the judgment calls made with System 1.

Many stoichiometry calculations students are asked to do in a general chemistry course use algorithms such as the factor-label method and, therefore, require System 1 processing, that is assuming students have already become adept at the use of common algorithms. Among these calculations are the determination of amounts of products that may be formed from specific amounts of reactants (or *vice versa*), amounts of reactants that will completely react with each other, and actual and theoretical yields of reactions. These heuristics often help students develop basic skills needed for successful chemistry problem solving, but the mastery of these skills do not necessary imply conceptual understanding (Nakhleh, Lowrey, & Mitchell, 1996; Nurrenbern & Pickering, 1987). Specifically, when the problems go even one step beyond the use of heuristics and require some analysis on their part, students more often than not encounter great difficulties.

PNOM Diagrams and Stoichiometry

PNOM diagrams allow students to visualize chemistry concepts and, thus, develop mental models (Gabel, 1998). It is quite common to have PNOM diagrams in textbooks drawn using circles of different sizes and colors to depict atoms, ions, and molecules although some diagrams may also use different shades of black, grey, and white with keys to help students interpret them. The use of PNOM diagrams, however, may also lead to misconceptions. For example, because PNOM diagrams represent reactions taking place at the microscopic level, students may be led to think that they are dealing with single particles when in fact several particles are actually being represented schematically (Ben-Zvi, Eylon, & Silberstein, 1988). How students link concepts with PNOM diagrams mainly depends on the diagrams' consistency with students' levels of understanding. Most students learn through actively choosing, organizing, and integrating the information obtained from diagrams used as visual inputs (Mayer, 2002). Many students fail to comprehend PNOM diagrams because they do not possess sufficient chemical knowledge and familiarity with representations of chemical concepts.

The three levels of representation, namely, the macroscopic, microscopic, and the symbolic (Figure 2), provide a framework for the understanding and teaching of chemistry (Johnstone, 1993). Chemistry experts easily go from one representation to another, but most students struggle with such transfer and some studies indicate the challenges are particularly important with the microscopic level.

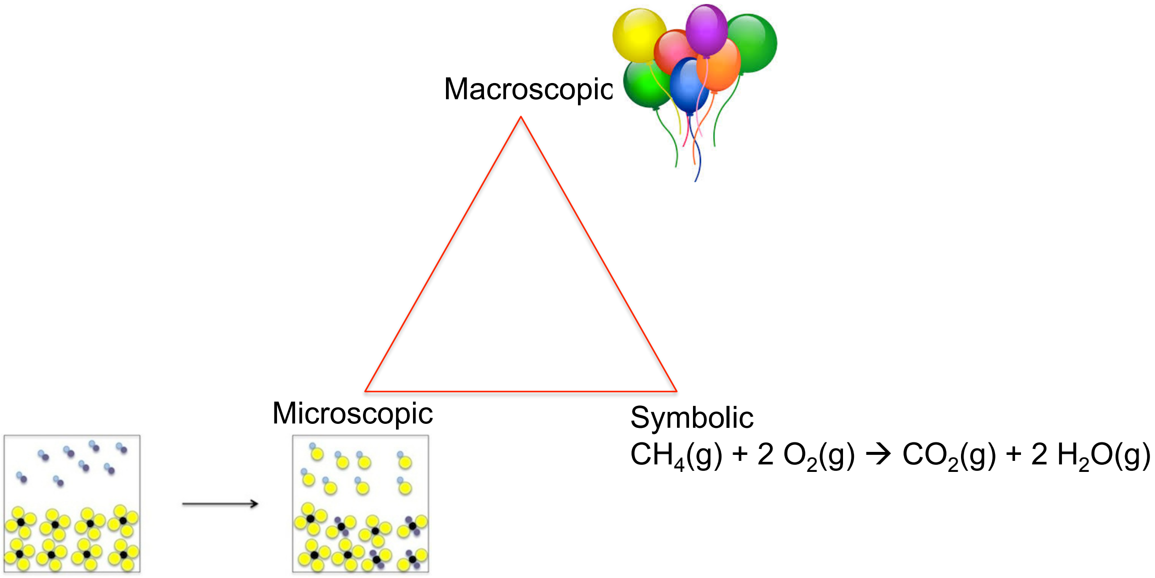


Figure 2. Johnstone's three level of representations.

It has been shown that even with the use of PNOM diagrams, students who can easily solve algorithmic chemistry problems often have great difficulties when asked to solve conceptual problems on the same topics (Nurrenbern & Pickering, 1987). Some may argue that this comes from a lack of development given by instructors to help students transition between the macroscopic and microscopic levels (Laugier & Dumon, 2004). Explanations for most chemical processes start from everyday experiences at the macroscopic level, but eventually go into the microscopic level, which makes use of invisible entities.

Chemical equations summarize only the net changes that take place in a reaction. They do not represent the microscopic nature of the species participating in the reaction, and they do not show details about how the changes occur or any species that, on the whole, do not participate in the process. Students usually see the balancing of chemical equations as the application a set of rules. When students fail to understand the implications of balanced chemical equations, it often arises from their failure to connect the symbolic representation to the actual changes occurring in the reaction (Laugier & Dumon, 2004).

The teaching of chemical equations and stoichiometry has mostly relied on the use of algorithms to solve different types of problems (Ault, 2001). Even with the increased use of PNOM diagrams in general chemistry textbooks, little is known about what student drawings reveal about how students understand chemical equations and stoichiometry at the microscopic level. Going through the literature also shows that students' difficulties in stoichiometry are generally sustained over long periods of time (Ben-Zvi et al., 1988). For instance, students have been found to misrepresent polyatomic ions as particles made up of single atoms, associate subscripts incorrectly to the wrong species (Smith & Metz, 1996), use subscripts outside of parentheses to indicate that polyatomic ions inside the parenthesis form a much bigger aggregate of ions (Roseman & McBride, 2011), form aggregates of product molecules, neglect excess reactants, or simply copy chemical equations as given in the questions (Davidowitz, Chittleborough, & Murray, 2010).

In this study, students were required to draw PNOM diagrams and explain how these diagrams relate to symbolic representations of specific chemical reactions. It was hoped that by using this approach, more insight into how students relate things from PNOM diagrams to stoichiometric concepts and see if there were ways to change the way questions are asked using PNOM diagrams. This chapter presents information about how students from two different types of general chemistry courses used PNOM diagrams to solve conceptual problems in stoichiometry that dealt with the concepts of limiting and excess reagents, and yield. In particular, this chapter attempts to address the following questions: (1) *What insights about student understanding of stoichiometry do student-drawn PNOM diagrams provide?*, and (2) *What challenges do students face when they are asked to draw their own PNOM diagrams in response to problems that deal with limiting and excess reagents, and yield?*

Study Participants

Students were recruited from two general chemistry courses, Chem A and Chem E, taught at a research intensive university for which the professors had given consent for this study. Chem A is the first part of a one-year course in general chemistry offered to physical and biological science majors, chemical engineering majors, as well as those intending to take 300-level chemistry courses. This course covers stoichiometry, parts of chemical equilibrium, acid-base chemistry, thermochemistry, rates and mechanism of reactions, changes of state, solution behavior, atomic structure, periodic relationships, chemical bonding. Chem E is a one-semester course aimed at providing students with an in-depth, broad-based view of modem chemistry. Chem E is also designed to introduce students to independent undergraduate research. Most students in Chem E have self-selected into the course after determining with their registration advisers before the start of the first semester that these students' chemistry preparation during high school is more than adequate compared to the average first-year student.

After obtaining approval from the Institutional Review Board (see Appendix F), the study was described to students by the researcher and students were asked to volunteer. Students from these classes were asked to fill out a demographics survey (Appendix A) that was used to identify volunteers to be interviewed. Those who did not wish to be interviewed were asked to return a blank survey.

Students were purposefully sampled from volunteers to achieve a relatively balanced representation based on sex, ethnicity, major, and course membership, as well as level of high school courses in chemistry and mathematics,. Since most students registered in Chem E indicated that they had taken honors or AP courses in chemistry and some specific mathematics courses listed in the survey, only students from Chem A who had similar academic backgrounds were invited for the interviews. A description of all students who participated in this study is given in Appendix B.

Data Collection

Students were each interviewed for about 60 minutes during the two weeks following their examination on stoichiometry in their respective course. Students used a Livescribe Echo smart pen to write notes and draw diagrams onto an accompanying Livescribe notebook (Livescribe, 2012) .The smart pen recorded each student's descriptions, notes, and illustrations. A total of 18 students participated during the think-aloud sessions using a common interview guide (Appendix C). The interviews were semi-structured to allow for deeper probing by the researcher if further exploration of the students' answers was warranted. Subjects were initially presented with a practice task that consisted of a problem that required a straightforward numerical solution to determine the percent yield obtained for a product based on the given amount of a reactant. The practice task was designed to be as simple as possible to allow students to get used to verbalizing their thoughts as they wrote into the notebook. This practice task also served as an opportunity for subjects to be familiar with being asked by the researcher to elaborate their responses. Although students were not specifically told that they were going through a practice task, responses collected at this point were not analyzed by the researcher.

The next problem, labeled Task 1 in the interview guide, asked subjects to draw a diagram that represents what would happen if given numbers of molecules of two gases were allowed to completely react with each other. In particular, the researcher was hoping to have an idea of how students understood the concept of limiting and excess reagents and how this translated in terms of PNOM diagrams drawn by students.

The last problem, labeled Task 2 in the interview guide, asked students to determine how a given diagram, representing complete reaction between two gases, would change if the percent yield of the reaction was reduced to half. This task puts the concept of yield on top of the limiting and excess reagents concept. The task aimed to determine how students coordinated these two concepts as indicated by changes they made to the given PNOM diagram. This problem was deliberately placed at the end of the session so that the complexity would increase as students went from one task to the next.

An important issue that needs to be addressed in conducting a qualitative study is the sample size required. Probably one of the most important goals in determining the adequacy of the sample is whether or not *saturation* is reached, *i.e.*, whether new ideas are no longer being found as more participants are brought into the study

(Creswell & Clark, 2007). The patterns of responses and thought processes described in the succeeding sections indicate that saturation was most likely achieved with the 18 participants in this study.

Data Analysis

All interviews were transcribed verbatim and the transcripts were managed using Dedoose (SocioCultural Research Consultants, 2014). All transcripts and documents coming from the student interviews were analyzed using the constant comparative method (Glaser, 1965), which classifies, describes, and connects data while looking for categories and common themes.

In general, student descriptions and solutions were first classified based on whether students ended up drawing correct, incorrect, or no diagram for each task. Each student's problem-solving strategy for each task was then carefully examined in order to classify them as having either a more numerical or pictorial approach. Numerical approaches were those that started from the use of ratios between numbers relating to species given in balanced chemical equations (Costu, 2010) similar to those found in most general chemistry textbooks (Brown, LeMay, Bursten, Murphy, & Woodward, 2014). These ratios may be in terms of amounts of substances given in moles or masses, or coefficients from the given or derived balanced equation. Solutions that relied more on drawings of PNOM diagrams or that went directly into the manipulation of such diagrams were considered to be pictorial in their approach. Specific instances of problem-solving strategies are discussed in later sections of this chapter. Diagrams submitted by students were also classified in terms of the types of representations used for atoms and molecules, *i.e.*, whether they used Lewis structures, differently sized and shaded circles, or not at all.

Student-Generated Diagrams in Response to Task 1

Task 1 required students to draw diagrams that would represent what would happen if three molecules of methane were completely reacted with four molecules of oxygen (Appendix B). Among the 18 students who voluntarily came for these interviews, 16 were actually able to draw a diagram in response to the task, while the remaining two did not draw a diagram at all.

The following subsections will, respectively, describe exemplars illustrating how solutions for Task 1 were classified as being numerically or pictorially based, describe in detail the more pictorial strategy followed by the more numerical strategy, discuss how some students came up with incorrect diagrams, and why others did not come up with diagrams at all.

General descriptions of student-drawn diagrams

Among the remaining 18 subjects, only 16 actually came up with diagrams of some sort in response to Task 1. Of these 16, eight came from Chem E, while the rest came from Chem A. Five of these students used circles of different sizes and colors to represent the different elements used in Task 1. The others drew Lewis or Lewis-like structures to represent their molecules. The Fisher exact test (see Table 1 and Table 2) gave no statistically significant correlation between either course or type of representation used and success or failure to come up with a correct diagram for this task. This is not surprising for a small-sample qualitative study. It appears that students in Chem E, who were assumed to have better mastery of chemistry concepts taught in high school chemistry, did not perform significantly better than Chem A students on Task 1. It was also initially thought that because most instructional materials in chemistry make use of different sized and colored circles in constructing PNOM diagrams, students would have preferred this type of schematic representation over other types. However, more students drew PNOM diagrams with Lewis structures than using circles. Almost the same fraction of students from each group turned in correct diagrams for Task 1. This indicates that when used properly, students should be able to solve conceptual problems in stoichiometry using Lewis or Lewis-like structures just as accurately as they do with circles.

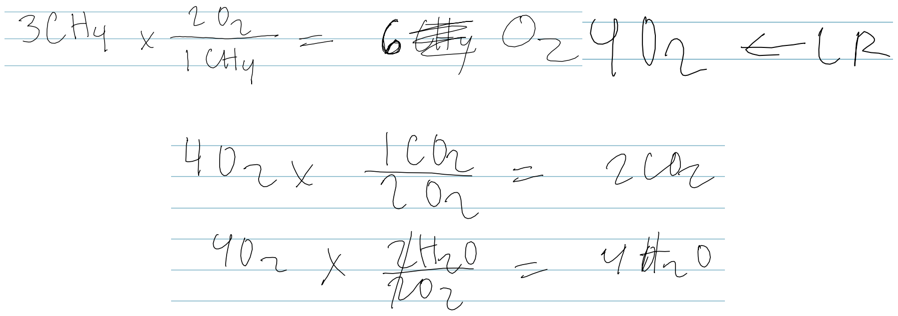
Table 1. Fisher exact test between success or failure to provide a correct PNOM diagram for Task 1 and course each participant came from.

Course	Correctness of Diagram		Row Total
	Correct	Incorrect	
Chem A	7	1	8
Chem E	6	2	8
Column Total	13	3	16
Cramer’s *V*			0.160
Fisher’s exact test *p*-value			0.500

Table 2. Fisher exact test between success or failure to provide a correct PNOM diagram for Task 1 and type of particulate representation used by each student.

Type of Representation	Correctness of Diagram		Row Total
	Correct	Incorrect	
Lewis structure	9	2	11
Labeled circles	4	1	5
Column Total	13	3	16
Cramer's *V*			0.022
Fisher's exact test *p*-value			0.705

On the other hand, the type of problem-solving strategy used by each student showed a significant correlation with the correctness of their initial diagram for this task (Table 3). As stated earlier, student strategies used in this tasked were classified as numerically based if it was shown that the student tended to reason first with stoichiometric ratios between the different species in the reaction. For the most part, students who used numerical strategies tended to rely on ratios between reactants and/or products as given by the coefficients in the balanced equation. Miley, who determined oxygen gas to be the limiting reagent after she saw that there was not enough oxygen available to completely react with the given number of molecules of methane, illustrates the numerical approach well (Figure 3): "*So I would need six molecules of* O_2 *to react with three moles, molecules of* CH_4*… And I only have four, so oxygen is limiting reagent*..." She then used her calculated numbers of molecules of products to guide her in drawing out her PNOM diagram (Figure 4).



$$3CH_4 \times \frac{2O_2}{1CH_4} = 6\,\cancel{CH_4}\,O_2 \quad 4O_2 \leftarrow LR$$

$$4O_2 \times \frac{1CO_2}{2O_2} = 2CO_2$$

$$4O_2 \times \frac{2H_2O}{2O_2} = 4H_2O$$

Figure 3. Miley's numerical approach to determining both the limiting reagent for the methane combustion reaction and determining the number of carbon dioxide and water molecules formed.

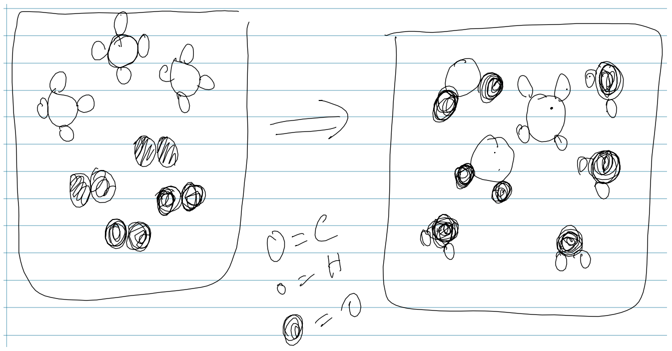


Figure 4. Miley's PNOM diagram for the methane combustion reaction task.

Other students used strategies that relied more on counting individual atoms and then drawing out product molecules in sets based on the given balanced equation to come up with their diagrams. Counting in itself may be thought of as an algorithm. However, these students did not explicitly rely on the use of ratios among species in the chemical equation, these strategies were tagged as being more pictorial than numerical in nature. Philip's approach started with him drawing a set of molecules consisting of one methane and two oxygen molecules (Figure 5). He saw in the balanced equation that from these, he should get one carbon dioxide and two water molecules as products of a complete reaction. He drew out these molecules based on the order of the species written in the balanced equation.

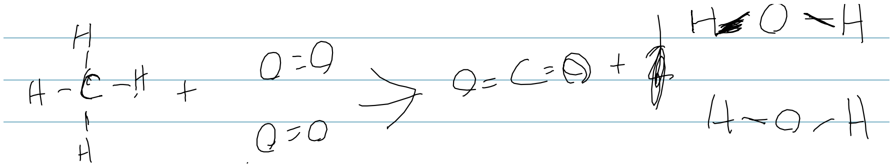


Figure 5. Philip's initial diagram showing how one set of reactant molecules react to form one set of product molecules.

Philip then realized that he still had two methane and two oxygen molecules and determined that he should have been able to make use of one set of reactant gas molecules to come up with a second set of products. He restarted drawing his product mixture, first, by drawing two sets of product molecules, and then by adding an unreacted molecule of methane (Figure 6).

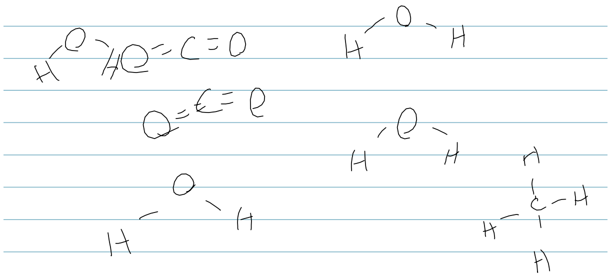


Figure 6. Philip's final diagram for Task 1.

Miley's and Philip's solutions for Task 1 each serve as an exemplars of the two general approaches used by students to come up with correct diagrams. One group of students who turned in what were considered to be correct diagrams first used calculations based on the ratios between coefficients in the balanced equation provided with Task 1, first, to correctly determine if methane and oxygen gas was the limiting reagent, and, second, to calculate the number of product and excess reactant molecules needed in their diagram. The second group, on the other hand, counted sets of reactant molecules, again based on the coefficients in the balanced equation, to form corresponding sets of products or to determine which molecule would be present in excess. These two approaches could be seen as going in relatively opposite directions as one solution led to the identification of the limiting reagent, while the other identified the excess reactant without explicitly stopping to determine which reagent would limit the reaction.

A Fisher exact test between the type of problem-solving strategy used and the correctness of the diagram for each subject yielded statistically significant result. This was accompanied by a strong effect size based on the Cramer's V value (Cramér, 1999). These numbers indicate that there might be reason to believe that the choice of problem-solving strategy used may be related to whether a participant draws the required PNOM diagram correctly or not. Table 3 shows that students who used pictorial approaches in completing Task 1 all drew correct diagrams, while those who used numerical approaches were split nearly equally between correct and incorrect. Detailed discussions of some of the more common features of these different strategies are given in the succeeding sections of this chapter.

Table 3. Fisher exact test between success or failure to provide a correct PNOM diagram for Task 1 and type of problem-solving strategy used by each student.

Type of Student Strategy	Correctness of Diagram		Row Total
	Correct	Incorrect	
Numerical	4	4	8
Pictorial	9	0	9
Column Total	13	4	17
Cramer's V			0.588
Fisher's exact test *p*-value			0.029

Student diagrams resulting from pictorial problem-solving strategies

Among students who did not explicitly use ratios between reactants and products in drawing their PNOM diagrams for Task 1, the most common approach was to draw product molecules in sets based on the coefficients of carbon dioxide and water given in the equation. Using this strategy, students drew one carbon dioxide and two water molecules as the initial products of the reaction. Some of these students crossed out reactant molecules during their thought processes to illustrate how they were keeping track of atoms used to form products. Then they realized that a second set of such molecules can still be obtained from the remaining reactant molecules. Finally, they saw that the third methane molecule had nothing to react with, and therefore, be left as an excess molecule.

Austin's and Billy's diagrams were examples of the strategy described previously. They both made the initial mistake of assuming that all of the methane molecules must be converted into carbon dioxides, an example of the *least amount assumption* (Davidowitz et al., 2010), where students incorrectly assume that the reactant present in the least amount must be the limiting reagent. However, even as Billy made this mistake (Figure 7), it was clear that he treated product molecules in terms of sets. Billy drew sets of products consisting of one carbon dioxide and two water molecules, lining each set horizontally in his diagram: "*So I'm gonna, if I have one* CO_2*... Then I will have, uhm, two* H_2O*'s... And then... so I used that correctly. Why did I... I'm not sure why I have a problem... Oh, ok, ok, ok. That would happen two more times.*" He realized after he had drawn his third set that he did not have enough oxygen atoms. After some moments of uncertainty, Billy saw that he should have only drawn two sets of product molecules and he finally determined that the third methane molecule must be an excess molecule: “*And then, uhm, there would still be one two three four... oh, ok... so this one (third set of molecules) would not happen and there would still be a C... ok, yeah, so that with this configuration, two carbon dioxides, four water molecules and a, uh,* CH_4 *molecule would be left over [in the product mixture]*.” Austin’s solution (Figure 8) reflected the same kind of thinking: “*There will only be two of the* CH_4 *molecules that are able to react.* [redraws diagram representing the reaction]*… So that will yield to us two carbon dioxides, uhm, along with four water molecules… And one* CH_4 *left over.* [counts atoms inaudibly] *Ok, now that’s right.*” It should be seen from their diagrams that the excess methane molecule was determined last by both students.

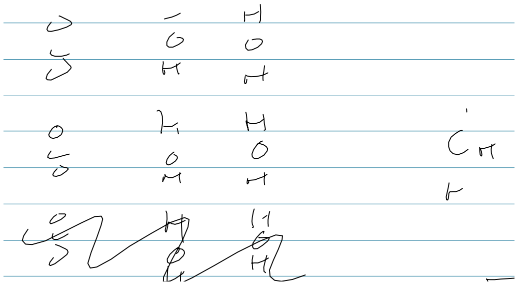


Figure 7. Billy's PNOM diagram for Task 1.

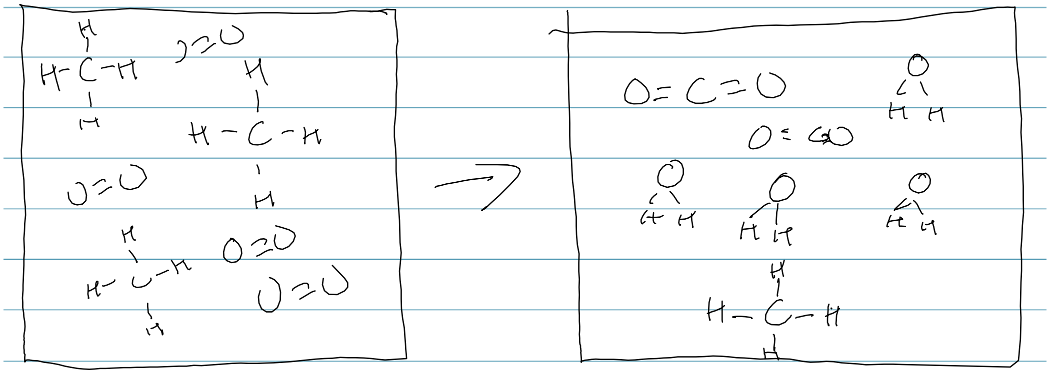


Figure 8. Ausin's diagram for Task 1.

Jason (Figure 9) also showed this kind of thought process in drawing out his product molecules in terms of sets of carbon dioxide and water molecules, although he did make the mistake of also initially viewing methane as the limiting reagent of the reaction. His second attempt at drawing his diagram showed clearly the initial formation of a set of products made up of one carbon dioxide and two water molecules.

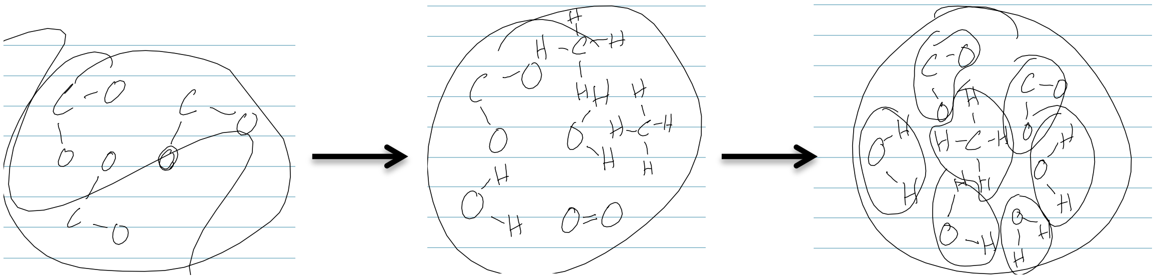


Figure 9. Jason's series of attempts in coming up with the correct diagram for Task 1.

Others, like BJ, Philip and Calvin, were quick to see that they should only have drawn two sets of product molecules and that the third methane molecule was left unreacted. Looking at Calvin’s PNOM diagram (Figure 10), one can spot the rough vertical alignment between the usual sets of product molecules, so that there an imaginary line

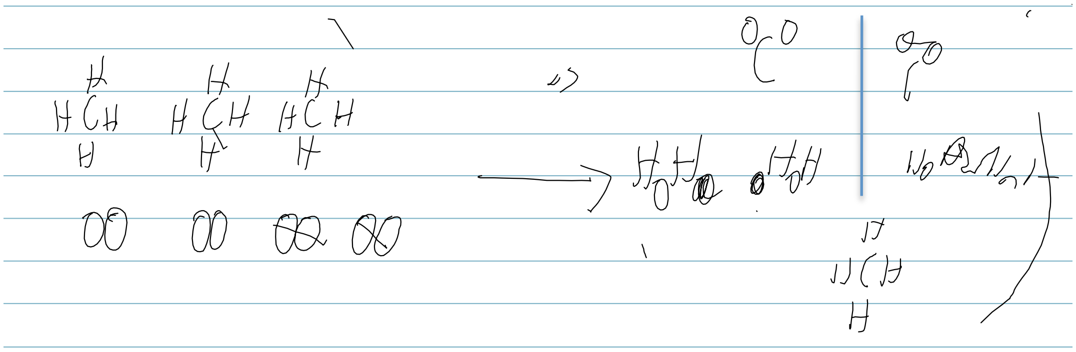


Figure 10. Calvin's PNOM diagram for Task 1 showing the imaginary line (added by the researcher) dividing his product molecules into two sets.

could be drawn between the two sets of molecules. The unreacted methane molecule is clearly seen to have been drawn as being the result of not having any oxygen molecules left to react with.

Psy’s diagram (Figure 11) also followed the same style of thinking although the line between his sets of product molecules is oblique. The relatively small size of the unreacted methane molecule resulted from having to fit this in the remaining space of his “After” box, since his product molecules have already taken up most of the right side of his diagram. It was pretty clear that Psy drew his products in terms of sets of carbon dioxide and water molecules coming from methane and oxygen. A unique feature of Psy’s illustration was his use of the wedge-and-dash symbolism to draw reactant molecules in the “Before” part of his diagram. Though completely unnecessary for this problem, this was quite illustrative of just how far advanced some of the students in Chem E were in terms of their prior knowledge.

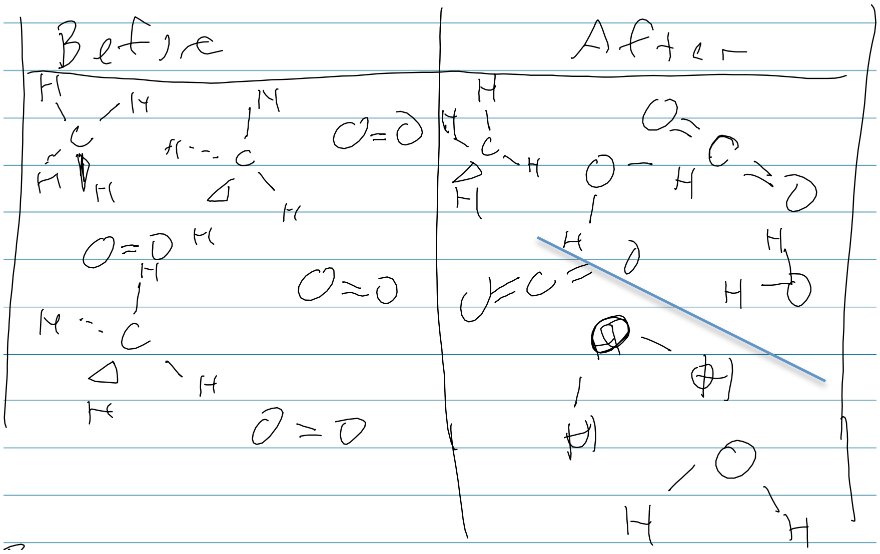


Figure 11. Psy’s PNOM diagram for Task 1 showing the imaginary line dividing his product molecules into two sets.

It is not always true though that imaginary lines could be drawn between sets of product molecules drawn by students. In Justin's solution (Figure 12), there were no distinct lines between sets of molecules that can be drawn across his diagram. However, it was still pretty clear that Justin's thought process involved forming the product molecules in sets in the same way the previous students did. It can be seen from his diagram that Justin took one methane molecule and combined it with two oxygen gas molecules to form one carbon dioxide and two water molecules. Justin even explicitly showed how he broke up each of the reactants and then recombined the atoms to form each of the product molecules. He then realized that with three methane and four oxygen molecules to start with, he could do this process twice as indicated by the multiplication step.

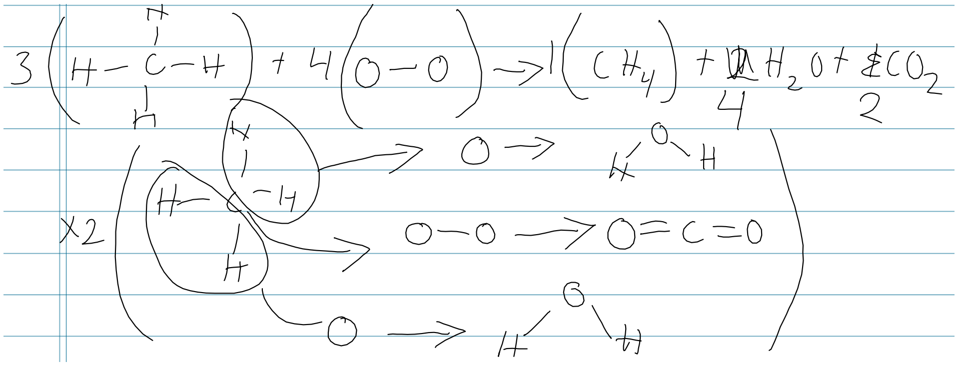
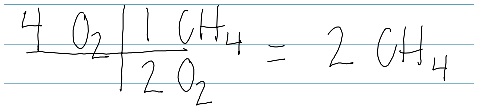


Figure 12. Justin's diagram for Task 1.

It is clear that an underlying theme among the different ways students described in this section constructed their diagrams for Task 1 is the treatment of reactant molecules as sets, which in turn form sets of product molecules. Several students who turned in correct responses drew diagrams treating reactants as sets of molecules that they needed to break apart and then recombine as guided by the product side of the balanced equation (Kelly & Jones, 2008). Students would continue this process as long as they could find sets of reactant molecules that could be used together based on the equation. Anything that did not have enough of the other type of reactant was then deemed to be the excess. The fact that these students go through sets of molecules in dealing with PNOM diagrams is not surprising given the fact that students selected for this interviews were all at the earliest stage of the undergraduate studies.

Student diagrams resulting from numerical problem-solving strategies

The remaining students used strategies that are modeled after quantitative algorithms for stoichiometry exercises to determine what the diagram they need for this task would look like. A common approach was to determine the limiting reagent based on reactant availability. Miley, Avril, and Eminem clearly used this method to determine that oxygen gas was the limiting reagent in the given problem. Whereas Miley calculated the number of oxygen gas molecules needed to completely react with the three methane molecules given in the task, Avril (Figure 13) and Eminem did the opposite. When either strategy is used correctly, students come to the conclusion that oxygen gas is the limiting reagent. Both students then used mole ratios between oxygen and each of the two products to correctly determine how many molecules of carbon dioxide and water they needed to draw.



$$4\ O_2 \left|\, \frac{1\ CH_4}{2\ O_2} \right. = 2\ CH_4$$

Figure 13. Avril's calculation of the number of molecules of methane that will react with the given number of oxygen gas molecules in Task 1.

Students may not always have explicitly shown how they used ratios from the reaction to decide which reactant was limiting. This was true for both Charice and Clark. In Clark's case (Figure 14), he made direct use of the ratio between oxygen and methane from the balanced equation in thinking about how they were going to react with each other: "*Ok, so now I'm gonna look at what the reaction says… And that says for every two moles of oxygen I have (a) mole of CH_4. So if I have four molecules of oxygen, then that can take care of only two molecules of CH_4. So, uhm, the oxygen is limiting, and there's actually one molecule of CH_4 left over.*"

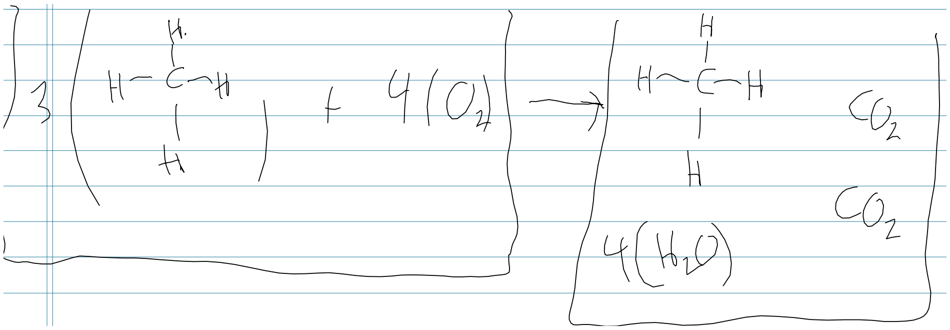


Figure 14. Clark's diagram in response to Task 1.

Charice (Figure 15) more explicitly showed that she used the ratio between methane and oxygen even though she did not use mathematical expressions. She used the 1:2 ratio to see that by the time she had used up two of her methane molecules, she would run out of oxygen gas molecules. Then she used the ratios between reactants and products to find that she would have two carbon dioxide and four water molecules as products and that a methane molecule would remain unreacted.

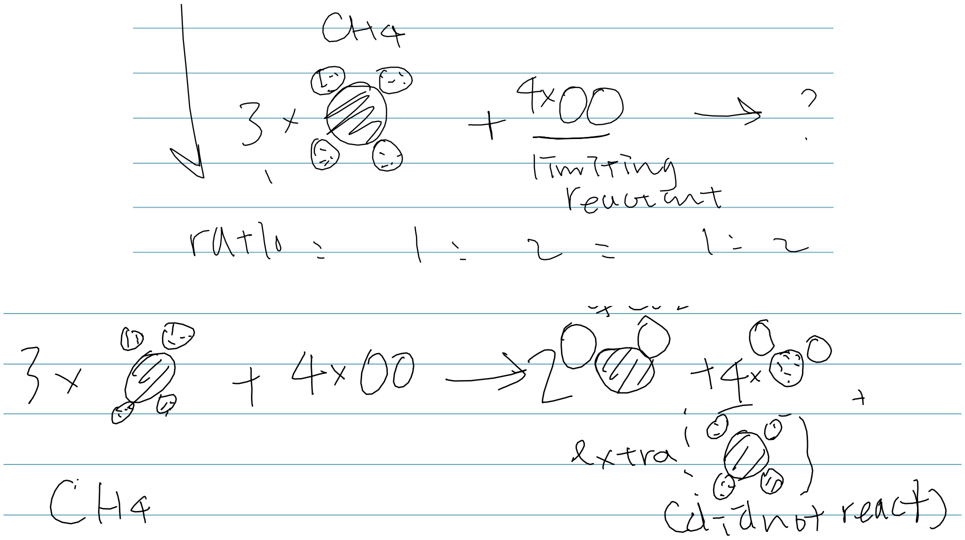
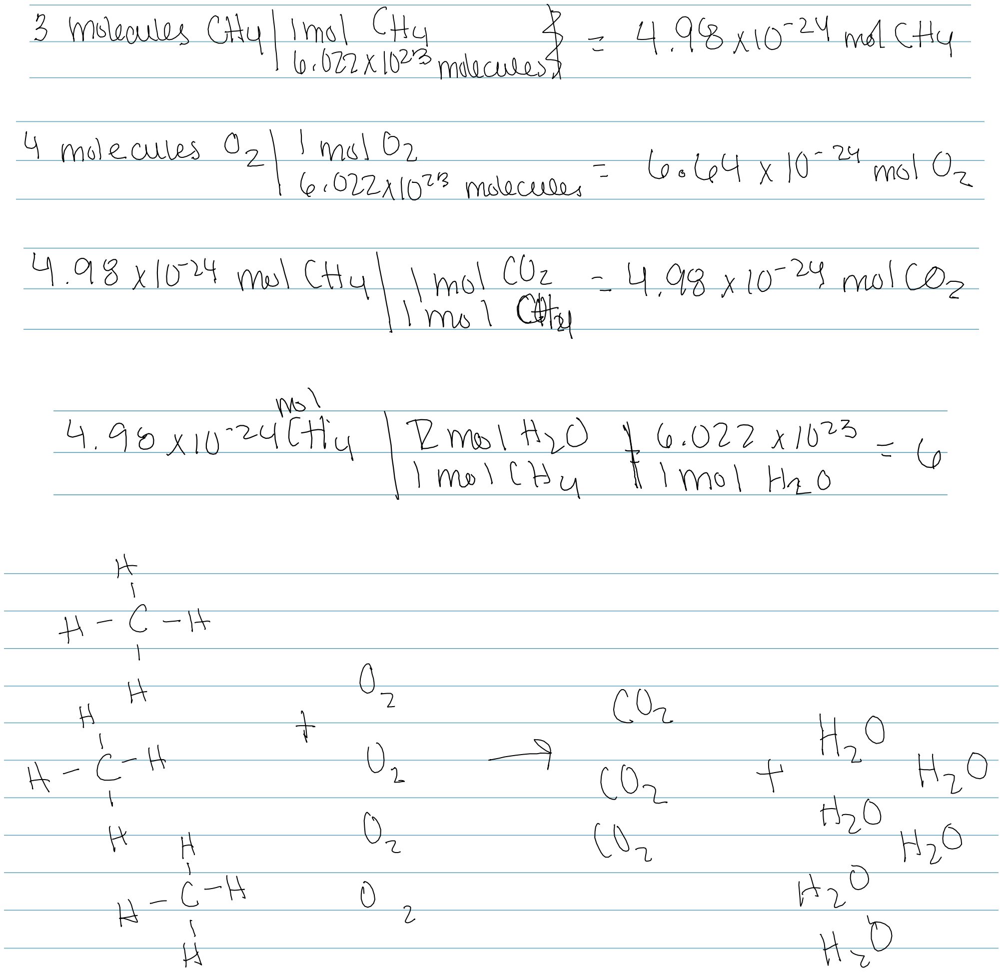


Figure 15. Charice's diagrams for Task 1.

Incorrect diagrams for Task 1

Similarly as what Kelly did for this task, Rihanna went through the unnecessary step of converting the numbers of molecules of reactants into moles using Avogadro's number (Figure 16). What led Rihanna, however, to turn in an incorrect diagram was her choice of methane as the limiting reagent based on its smaller number of moles compared to that she determined for oxygen gas. This incorrect choice of the limiting reagent for a reaction based on that which is present in the least amount was pointed out earlier with Austin's and Billy's cases. In any case, Rihanna used the amount of methane to calculate the number of moles of carbon dioxide and water produced, and then drew her diagram based on these calculations.



$3\text{ molecules }CH_4 \,\Big|\, \dfrac{1\text{ mol }CH_4}{6.022\times10^{23}\text{ molecules}} = 4.98\times10^{-24}\text{ mol }CH_4$

$4\text{ molecules }O_2 \,\Big|\, \dfrac{1\text{ mol }O_2}{6.022\times10^{23}\text{ molecules}} = 6.64\times10^{-24}\text{ mol }O_2$

$4.98\times10^{-24}\text{ mol }CH_4 \,\Big|\, \dfrac{1\text{ mol }CO_2}{1\text{ mol }CH_4} = 4.98\times10^{-24}\text{ mol }CO_2$

$4.98\times10^{-24}\text{ mol }CH_4 \,\Big|\, \dfrac{2\text{ mol }H_2O}{1\text{ mol }CH_4} \,\Big|\, \dfrac{6.022\times10^{23}}{1\text{ mol }H_2O} = 6$

Figure 16. Rihanna's calculations and diagram for Task 1.

Beyonce's solution (Figure 17) to this task also reflected her initial choice of methane as the limiting reagent. Based on the given numbers of molecules of methane and oxygen, and the ratios in the balanced equation, she listed down both what she was given in terms of reactants and what she initially thought she would get as products, using the ratios of both carbon dioxide and water with methane. Checking the balance between her atoms, she saw that there was more oxygen on the product side than on the reactant side. This led her to correctly think that she must somehow use the number of oxygen atoms to limit the number of product molecules she can obtain from the reaction. Unfortunately, she insisted on using the number of methane molecules given to determine the number of CO_2 molecules, which forced her to limit the number of water molecules in the product mixture. This seemed to be a case of making an incorrect choice between reagents to limit the number of product molecules which, anyway, leads to an incorrect diagram. Beyonce even went on to list what she would obtain as products using this approach of having both reactants limiting the amounts of products formed. When asked what would happen to hydrogen atoms she did not use in her diagram, she said that these would bond with each other because there is not enough oxygen present to form water molecules.

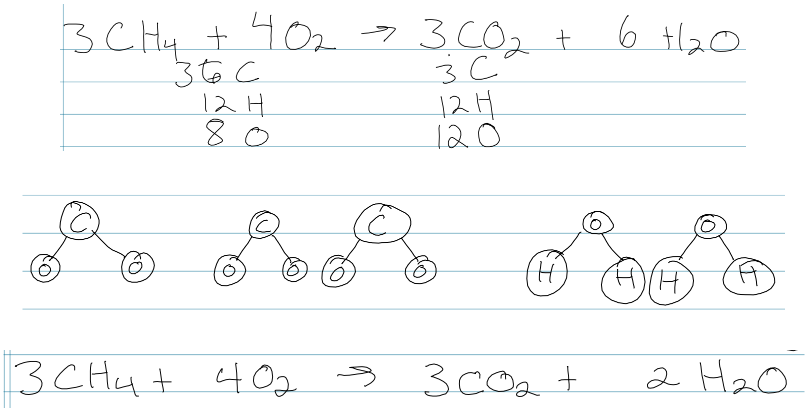


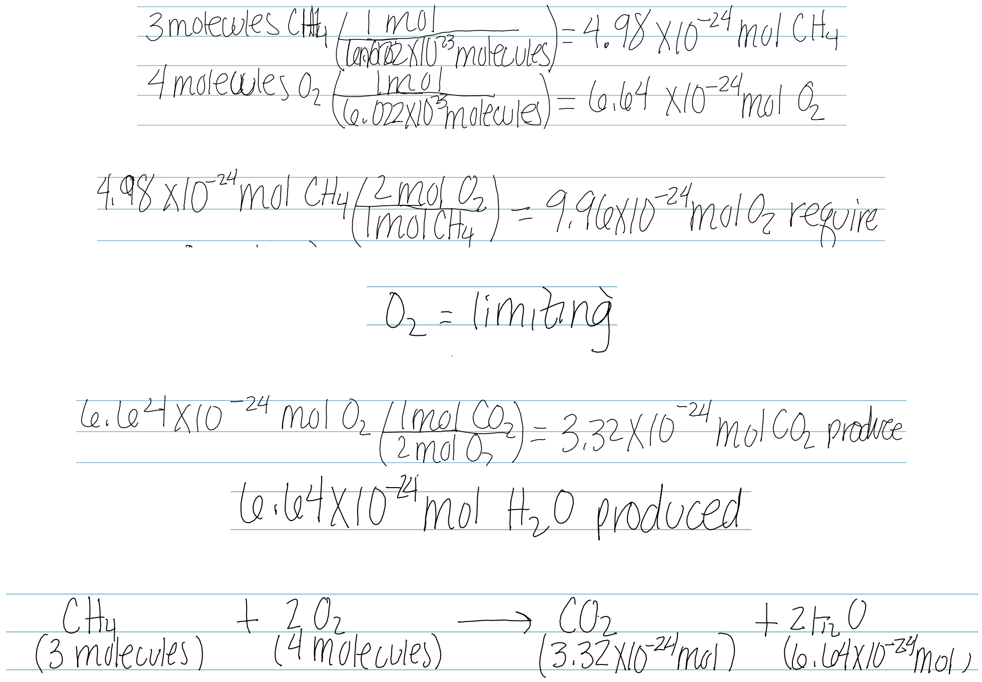
Figure 17. Beyonce's calculations and diagram for Task 1.

These two examples illustrate even further that a very likely source of student mistakes when drawing PNOM diagrams to illustrate reactions that have limiting reagents is the choice of the reactant given in a smaller amount as the one that determines how much of the reactants react and how much of the products were formed. This is typical of students classified as *commonsense learners* (Talanquer, 2006) who tend to automate their use of principles and strategies almost to the total disregard of other strategies or meanings. A typical approach of these students would be to use strategies simply because it has worked previously for a different problem. Common sense thinking often leads students to erroneous solutions and impedes problem solving. Both Beyonce and Rihanna illustrated how it was sometimes too difficult for students to let go of incorrect choices based on a heuristic that has probably frequently worked for them. Both students firmly held onto this strategy of using the reactant present in the least amount as their limiting reagent, even when it meant for them changing the chemical equation.

Students who failed to draw diagrams

Of the 18 students who voluntarily came for the interview, two students failed to draw a diagram for the first task. Both students came from the advanced general chemistry course, Chem E. Even with some prompting from the researcher, Adam did not understand exactly what it was he was being asked to do and how he was supposed to come up with diagrams of any sort from the information provided: "*I don't know if I understand what the question is asking… I don't know, I just don't know in general, I just don't know what it's asking with the diagram, or draw a diagram.*" When asked if he knew how to draw a diagram, Adam stated that he did not know what kind of diagram was required of him.

Kelly, who also did not draw a diagram for Task 1, seemed concerned about the kind of diagram she was supposed to draw: "*So, uhm, I don't know what type of a diagram you want… Uhm, uhm, ok. Uhm, I honestly don't know. Like were you, are you looking for what happens with the CO_2 and H_2O in this diagram?*" She was told that no specific type of diagram was needed from her as long as what would happen in the reaction was properly represented. Kelly then immediately went into calculations of the numbers of moles of reactants based on the number of molecules given for each of them (Figure 18). She used the *method of reactant availability*, with which a student calculates, for example the amount of oxygen gas needed to completely burn the amount of methane given in the problem. If the calculated moles or mass of oxygen was less than what was given in the problem statement, the student needed to pick out methane as the limiting reagent. Otherwise, the student should have seen that more oxygen than what was given in the problem was needed to combust the given amount of methane, and that oxygen was the limiting reagent. Although this was an entirely unnecessary set of calculations, Kelly did correctly determine oxygen gas to be the limiting one in the given problem.



$$3 \text{ molecules } CH_4 \left(\frac{1 \text{ mol}}{6.022 \times 10^{23} \text{ molecules}}\right) = 4.98 \times 10^{-24} \text{ mol } CH_4$$

$$4 \text{ molecules } O_2 \left(\frac{1 \text{ mol}}{6.022 \times 10^{23} \text{ molecules}}\right) = 6.64 \times 10^{-24} \text{ mol } O_2$$

$$4.98 \times 10^{-24} \text{ mol } CH_4 \left(\frac{2 \text{ mol } O_2}{1 \text{ mol } CH_4}\right) = 9.96 \times 10^{-24} \text{ mol } O_2 \text{ require}$$

$$O_2 = \text{limiting}$$

$$6.64 \times 10^{-24} \text{ mol } O_2 \left(\frac{1 \text{ mol } CO_2}{2 \text{ mol } O_2}\right) = 3.32 \times 10^{-24} \text{ mol } CO_2 \text{ produce}$$

$$6.64 \times 10^{-24} \text{ mol } H_2O \text{ produced}$$

$$\underset{(3 \text{ molecules})}{CH_4} + \underset{(4 \text{ molecules})}{2O_2} \longrightarrow \underset{(3.32 \times 10^{-24} \text{ mol})}{CO_2} + \underset{(6.64 \times 10^{-24} \text{ mol})}{2H_2O}$$

Figure 18. Kelly's calculations related to Task 1.

Kelly then determined the amounts of carbon dioxide and water produced based on the number of moles she determined for oxygen gas, which demonstrated that she knew that amounts of products formed by a reaction were based mainly on the given amount of the limiting reagent. She, however, ultimately revealed not having any idea about the diagram the problem was asking her to draw: "*I don't really know what you mean by draw a diagram. So I don't know where I would go from here, to be honest.*"

The fact that both Adam and Kelly came from the advanced course in general chemistry should not be too surprising with respect to their failure to draw diagrams for what is seemingly a straightforward stoichiometry problem. The failure of some students to deal with microscopic representations in chemistry in spite of their great abilities to solve algorithmic problems has been well documented in the literature (Agung & Schwartz, 2007; Nakhleh, 1993; Nakhleh & Mitchell, 1993; Pickering, 1990). It was possible that some students failed to draw or handle PNOM diagrams simply because these diagrams have not been included in the curricula of their previous chemistry courses. Others may have difficulty in transferring from one level of representation to another, especially if one of the representation levels is the particulate (Gabel, 1998). The new AP chemistry curriculum has taken steps to remedy this by requiring students to exhibit conceptual understanding of chemistry phenomena, including reactions, at the particulate level (Prilliman, 2014). The explanation of chemical reactions, like many other chemical phenomena, can be helped by using models to represent microscopic particles in substances being observed. It is true that many properties of matter observed in the macroscopic world are not simple extrapolations of the microscopic behavior of substances, as many students seem to think. Still, some explanations, when properly used, at the macroscopic level work just as well at the microscopic level, as in the case of stoichiometric concepts like limiting reagents. It appears that both Adam and Kelly could have benefitted from well-thought out illustrations of how PNOM diagrams work from their previous chemistry instructors.

Student-Generated Diagrams in Response to Task 2

Sixteen students from among those who volunteered came up with diagrams for Task 2. During this task, students were asked how they would change a diagram representing a reaction between four molecules of nitrogen gas and six molecules of hydrogen gas at 100% yield to represent 50% yield. An objective of this task was to see how students' diagrams would be affected by combining two closely-related chemistry principles in one task. It was determined that combining the concepts of limiting and excess reagents with yield resulted in the absence of a statistically significant relationship between the type of strategy used by each student in doing the task and the correctness of their diagrams (Table 4). Still, all students who went directly into manipulating the given diagram drew correct diagrams for this task. Diagrams from students who used numerical approaches were evenly split between correct and incorrect. So there may be a relationship going on between type of strategy used and accuracy of the diagram drawn that could be revealed with the use of a larger sample. Course membership was also not seen to be a factor as equal numbers of students from each course drew appropriate diagrams for this task (Table 5). All but one student used labeled circles to distinguish between their nitrogen and hydrogen atoms to draw diagrams for Task 2. This demonstrates the important effect of having an initial diagram on which students might be able to base their responses.

Table 4. Fisher exact test between success or failure to provide a correct PNOM diagram for Task 2 and type of problem-solving strategy used by each student.

Type of Student Strategy	Correctness of Diagram		Row Total
	Correct	Incorrect	
Algorithmic	4	4	8
Conceptual	8	0	8
Column Total	12	4	16
	Cramer's V		0.577
	Fisher's exact test *p*-value		0.077

Table 5. Fisher exact test between success or failure to provide a correct PNOM diagram for Task 2 and course each participant came from.

Course	Correctness of Diagram		Row Total
	Correct	Incorrect	
Chem A	6	1	7
Chem E	6	3	9
Column Total	12	4	16
	Cramer's V		0.218
	Fisher's exact test *p*-value		0.585

Students who turned in incorrect diagrams for Task 2

Four students, Charice, Clark, Eminem, and Rihanna, drew diagrams that varied slightly from each other drew for this task. Most instructors define the percent yield of a reaction, as the actual yield of a product divided by the same species' theoretical yield based on the limiting reagent. In the given diagram, four nitrogen molecules were reacted with six hydrogen molecules to produce ammonia. This meant hydrogen acted as the limiting reagent and should, therefore, be the only reactant used to calculate percent yield. Given that the yield of the reaction was 50%, then only three hydrogen molecules reacted with one nitrogen gas molecules. This leaeves behind three hydrogen and three nitrogen gas molecules unreacted, along with two ammonia molecules in the product mixture resulting from the reaction.

Rihanna's notes and diagram (Figure 19) for Task 2 were typical of what students who ended up turning in incorrect diagrams missed. She began by counting all of the atoms given on the reactant side of the given diagram (mistakenly labeling nitrogen atoms as "O"), and she found that there are 12 hydrogen and 8 nitrogen atoms.

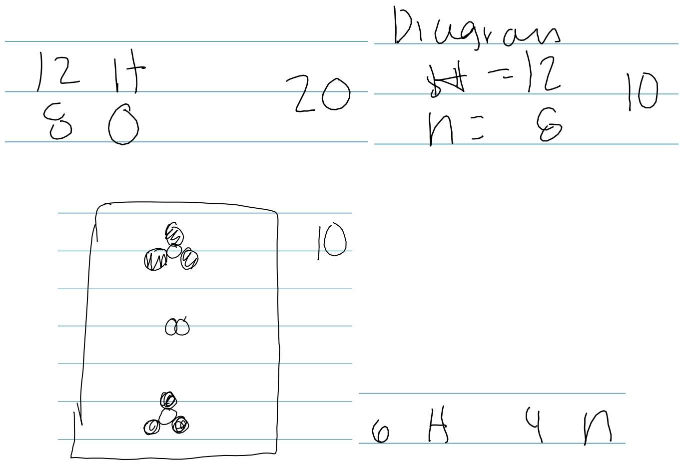


Figure 19. Rihanna's notes and diagram for Task 2.

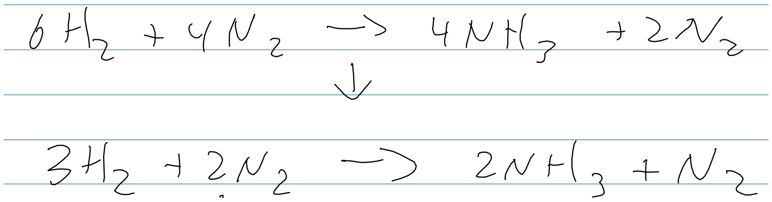
Since the given yield of the reaction was 50%, Rihanna divided the total number of atoms by two, and figured that she would need 6 hydrogen and 4 nitrogen atoms to make the necessary changes. As a result, she drew two ammonia molecules with one nitrogen molecule left behind. This application of the given percent yield of a reaction to all molecules of the reactants showed how some students failed to realize that the yield referred only to the fraction of the limiting reagent that reacts successfully. This divide-everything-in-half strategy, when improperly used by students, also points to a neglect of the law of conservation of mass. Students in this case failed to explicitly show all reactants that did not react (three hydrogen and two nitrogen molecules) as part of their diagram of the product mixture. For some students, the additional task complexity coming from having to address the percent yield led to a neglect of the role of the limiting reactant. This may have consequences on how instructors might make students understand the importance of laboratory tasks such as proper disposal of waste and isolation of desired products free from excess reactants, both of which are commonly dealt with in more courses like organic chemistry.

Students who drew correct diagrams for Task 2

Students who drew correct diagrams for Task 2 were divided into two groups. Four of them also focused on the dividing the molecules in the product mixture in half, using one part to form back reactant molecules. Others made use of some form of the definition of the percent yield. In this case, students divided either the number of product molecules formed or the number of molecules of the limiting reagent that successfully reacted in half. This was followed by either redrawing unused reactant molecules or the number of product molecules formed based on the reduced number of limiting reagent molecules that actually reacted, and then filling in the remaining molecules to balance atoms on both sides of the diagram.

Psy's work (Figure 20) for this task best represents how students who divided the number of molecules in the given diagram in half turned in correct diagrams for Task 2.

He started doing his work by first focusing on the compositions of both the reactant and product mixtures as drawn in Figure 21. He proceeded to divide everything in half.



$$6H_2 + 4N_2 \rightarrow 4NH_3 + 2N_2$$

$$\downarrow$$

$$3H_2 + 2N_2 \rightarrow 2NH_3 + N_2$$

Figure 20. Initial steps of Psy's solution for Task 2.

This got Psy stuck for a while after he saw that there were nitrogen molecules on both sides of his equation. He then proceeded to subtract a nitrogen molecule from each side of the equation and figured out that this nitrogen molecule must not have reacted. It became clear to Psy that this heuristic was not working, and it forced him to re-evaluate his mental model. He redirected his attention back to the product side of the diagram, and this is when he divided the product mixture in the provided diagram into halves (Figure 21): *"So I would, uhm, let's say you have n number of moles, n number of moles of, er,* H_2*. Which means, I must be having limiting reagents. Hmmm. What my gut answer tells me is if I just say it's only fifty percent, so you cut across in half… And what you do is, I guess, you mentally disassemble these…* [half of the ammonia molecules in the product mixture]."

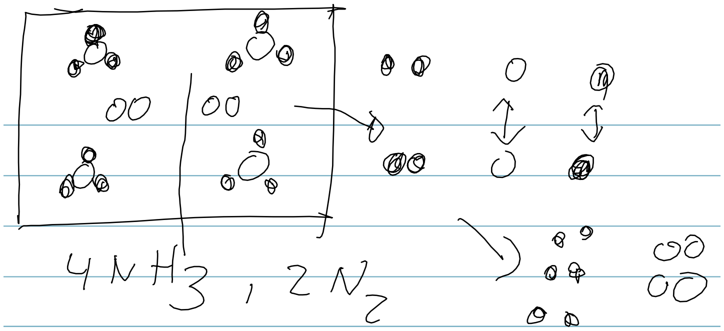


Figure 21. Psy's divide-in-half approach as applied to the product molecules drawn in the original diagram.

Psy showed how he recombined the *disassembled* atoms to form back nitrogen and hydrogen gas molecules, and then checked the balance of his atoms after he drew his final diagram.

Jason (Figure 22), on the other hand, focused more on the hydrogen gas molecules, which he said were the limiting reagent. Since there were six drawn in the given diagram, Jason said that the balanced equation must be multiplied by two (unnecessarily), which theoretically gives him four ammonia molecules. This meant that at 50% yield he should only be getting two ammonia molecules as products, and then Jason went on to "*pair up*" unused atoms.

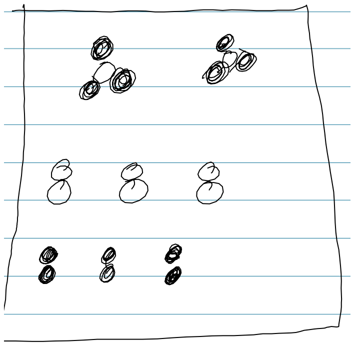


Figure 22. Jason's final diagram for Task 2.

Note from Jason’s diagram how he also lined up his unreacted nitrogen and hydrogen molecules after drew his ammonia molecules. This shows how Jason organized his thoughts about how accounting for all atoms used in this task.

Six students who turned in correct diagrams for Task 2 were, like Jason, more focused on the determining the number of product molecules formed or the number of molecules of the limiting reagent that reacted based on the given yield of the reaction. Avril, one of the six, immediately determined that at 50% yield, only two ammonia molecules would be formed by the reaction, which she drew at the top of her diagram (Figure 23). She then went on to draw the unreacted molecules to complete her diagram. This approach seemed to be the most direct as far as determining the composition of the product mixture since it made use of the percent yield’s definition as being a reaction’s actual yield divided by the theoretical yield as determined from the amount of the limiting reagent. BJ thought about the yield in terms of the fraction of the limiting reactant that actually formed products. He started his diagram (Figure 23) by drawing the unreacted gases on the left before finishing it up with the ammonia molecules formed: *“So with a fifty percent yield, how would I change it? Hmm, well I guess I would have some of the... three* H_2 *fooling around* (left over) *still and one more* N_2 *fooling around...”*

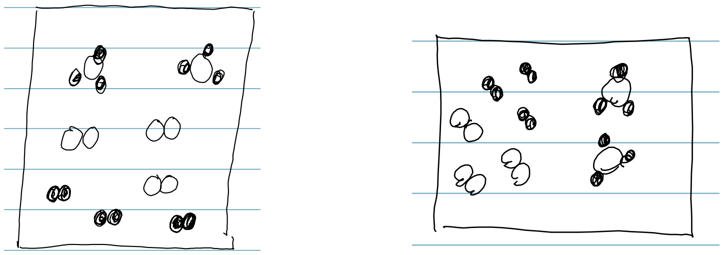


Figure 23. Avril's (left) and BJ's (right) diagrams of the product mixture for Task 2.

Themes Arising from Student Diagrams

It is seen from diagrams drawn by students in Task 1 that the choice of the reactant present in the least amount was a common student misconception. These students were focused on methane forming carbon dioxide and neglected to balance the remaining elements. In one variation of this approach, students made use of the “limiting reagent” to determine the amounts of all products based on ratios determined from the balanced equation. This is the case with Rihanna’s diagram, where in addition to three carbon dioxide molecules there were six molecules of water. Another variant of this approach makes use of the remaining atoms to form as many of the other products as possible (as in the case of Beyonce’s diagram). In either case, there was no indication at all that the diagram was checked against the law of conservation of mass.

The absence of a diagram with which to immediately connect verbal cues from the problem statements must have been a source of difficulty for some students. The lack of visual information to code in Task 1 made some of them to either resort to calculations based on previously learned algorithms, fail to come up with their own diagrams, or both. The fact that more students who had challenges with this task came from the advanced course may imply that while they may have greater competence with algorithmic problems compared to the typical first-year chemistry student, their expertise does not necessarily extend to conceptual problems.

The most common approach among those who came up with the correct diagram for Task 1 involved counting reactant molecules in sets as they were converted into product molecules. This meant using reactant molecules in sets based on the given balanced equation to form corresponding sets of product molecules. Students then checked whether more reactant molecules were available to form products. This strategy relied less on the factor-label heuristic that is most familiar to general chemistry students. This might be evidence that when forced to do so, some students illustrated use of System 2 processing to complete Task 1. If indeed this is so, then these students made the necessary connection between what they saw as the complete use of oxygen gas molecules to come up with the decision that this was the limiting reagent for the reaction and that the formation of more carbon dioxide and water was no longer possible.

That students who came up with correct diagrams for Task 1 used one of two approaches or both may be an example of the dual process learning model at work (Evans, 2008). Some students were forced by the lack of a diagram to reason using heuristics that they are familiar with. The use of ratios obtained from coefficients of species in the balanced equation for the combustion of methane to determine a limiting reagent is a manifestation of System 1 thinking. Because this kind of reasoning is more specific in its domain, concrete and contextualized, students who have not quite matured yet in the domain they are being tested on may get lost in their problem-solving processes. As a result, some students who have used numerical strategies to do Task 1 were not quite able to make the necessary connections between their drawings and the underlying calculations. Still, some students who started working on this task using heuristics did succeed in drawing the required diagram. This indicates that information flow between the verbal and the visual systems, at least in the context of Task 1, may be bidirectional.

It appears that the same kind of analysis may be applied to how students coordinated their thoughts about limiting reagents and yields for Task 2. Students also used one of two general approaches to come up with their modified diagrams. One group divided either reactant or product molecules into two groups of the same composition based on the given percent yield of the reaction. This sometimes led students to neglect the balance of atoms in their diagrams. The other group directly applied the definition of the percent yield in coming up with the diagram. All students who used this approach came up with a correct response. The more systematic approach that the second group of students used with their diagrams seems to be a manifestation of System 2 processes at work, while others let System 1 prevail.

It should be noted that across tasks, the use of numerical algorithms to start solutions only took place with Task 1, which did not provide students with a diagram to begin with. The use of numerical approaches definitely requires less effort than immediately going sketching a diagram, especially if students do not even know what it is supposed to look like. On the other hand, giving students a beginning diagram in Task

2 led students away from “tried-and-tested” numerical approaches. Students definitely used the given diagram as a cue that this was somewhat closer to what was being asked from them, even if it meant having to coordinate two different chemistry concepts together.

Implications

This study revealed how it is often difficult for students to go from symbolic to microscopic representations when asked to solve conceptual problems in stoichiometry. It is quite a challenge for students to visualize atoms and molecules when the initial information they have is a balanced chemical equation. This was shown to a great extent by the number of students who failed to draw appropriate diagrams for Task 1. Several students chose to focus their attention on the numerical relationships that could be inferred from the coefficients in the equation. Even students who used pictorial strategies in forming their responses to Task 1 relied heavily on the relationships among the coefficients in the equations to form their sets of reactant and product molecules (*i.e.,* one methane and two oxygen gas molecules, or one carbon dioxide and two water molecules). While there is nothing inherently wrong with these approaches, instructors need to point out to students the limitations of these strategies. Common pitfalls such as neglecting to balance atoms between the reactant and product sides of a PNOM diagram, choosing the reactant present in the least amount as the limiting reagent, or applying the percent yield across all molecules, may be addressed directly. It might help for instructors to actually model in front of their students how PNOM diagrams are constructed using the symbolic level mainly to guide such drawings. The use of ready-made PNOM diagrams to illustrate chemical concepts might have to be postponed until instructors see clearly for themselves that students understand how chemists build and interpret the pictures.

In light of the findings on this study, a teaching and learning sequence of the concepts of stoichiometry might go along the following:

1. Probe into students' prior knowledge of the particulate nature of matter. Identify how students explain the structure of matter at the microscopic level.
2. Illustrate the use of PNOM diagrams. Demonstrate the conventions used in preparing such diagrams. Identify the limitations of PNOM diagrams as far as exhibiting the structure of matter is concerned. Explain that while these diagrams are meant to make chemical relationships more explicit to students, diagrams are not supposed to be taken as general representations. State explicitly that often the representation of atoms and molecules with PNOM diagrams is partial.
3. Illustrate the relationship between PNOM diagrams (microscopic representations) and balanced chemical equations (symbolic representations) in dealing with stoichiometric concepts. This might include counting of atoms and molecules of reactants in sets based on the balanced equation to correctly determine the number of sets of product and unreacted molecules.
4. Make the concepts of limiting reagents and reaction yields more concrete for students. These are two highly abstract ideas for the beginning chemistry student. Illustrate exactly how the limiting reagent is completely converted into

products (as opposed to being *used up*) in a reaction that goes to completion using PNOM diagrams. Show that the yield of a reaction may be determined based on the number of limiting reagent molecules that have been transformed into products as shown in the PNOM diagram.

5. Illustrate the limitations of heuristics with well-defined examples. Show how selecting the reactant present in the least amount often fails to correctly identify the limiting reagent.

REFERENCES

Agung, S., & Schwartz, M. S. (2007). Students’ understanding of conservation of matter, stoichiometry and balancing equations in Indonesia. *International Journal of Science Education, 29*(13), 1679-1702.

Ault, A. (2001). How to say how much: Amounts and stoichiometry. *Journal of Chemical Education, 78*(10), 1347.

Ben-Zvi, R., Eylon, B., & Silberstein, J. (1987). Students’ visualization of a chemical reaction. *Education in chemistry, 24*(4), 117-120.

Ben-Zvi, R., Eylon, B., & Silberstein, J. (1988). Theories, principles and laws. *Education in chemistry, 25*(3), 89-92.

Brown, T. E., LeMay, H. E. H., Bursten, B. E., Murphy, C., & Woodward, P. (2014). *Chemistry The Central Science 13th Edition*: Prentice Hall.

Bucat, B., & Mocerino, M. (2009). Learning at the sub-micro level: Structural representations *Multiple representations in chemical education* (pp. 11-29): Springer.

Cheng, M. M. W., & Gilbert, J. K. (2014). Students' visualization of metallic bonding and the malleability of metals. *International Journal of Science Education, 36*(8), 1373-1407.

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational psychology review, 3*(3), 149-210.

Costu, B. (2010). Algorithmic, conceptual and graphical chemistry problems: A revisited study. *Asian Journal of Chemistry, 22*(8), 6013-6025.

Cramér, H. (1999). *Mathematical methods of statistics* (Vol. 9): Princeton university press.

Creswell, J. W., & Clark, V. L. P. (2007). Designing and conducting mixed methods research.

Davidowitz, B., Chittleborough, G., & Murray, E. (2010). Student-generated submicro diagrams: A useful tool for teaching and learning chemical equations and stoichiometry. *Chemistry Education Research and Practice, 11*(3), 154-164.

Evans, J. S. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annu. Rev. Psychol., 59*, 255-278.

Gabel, D. (1998). The complexity of chemistry and implications for teaching. *International handbook of science education, 1*, 233-248.

Glaser, B. G. (1965). The constant comparative method of qualitative analysis. *Social problems*, 436-445.

Harrison, A. G., & Treagust, D. F. (2000). A typology of school science models. *International Journal of Science Education, 22*(9), 1011-1026.

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

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. *Heuristics and biases: The psychology of intuitive judgment, 49*.

Kelly, R. M., & Jones, L. L. (2008). Investigating students' ability to transfer ideas learned from molecular animations of the dissolution process. *Journal of Chemical Education, 85*(2), 303.

Laugier, A., & Dumon, A. (2004). The equation of reaction: a cluster of obstacles which are difficult to overcome. *Chemistry Education Research and Practice, 5*(3), 327-342.

Livescribe, I. (2012). http://www.Livescribe.com*.*

Mayer, R. E. (2002). Cognitive Theory and the Design of Multimedia Instruction: An Example of the Two-Way Street Between Cognition and Instruction. *New directions for teaching and learning, 2002*(89), 55-71.

Nakhleh, M. B. (1993). Are our students conceptual thinkers or algorithmic problem solvers? Identifying conceptual students in general chemistry. *Journal of Chemical Education, 70*(1), 52.

Nakhleh, M. B., Lowrey, K. A., & Mitchell, R. C. (1996). Narrowing the gap between concepts and algorithms in freshman chemistry. *Journal of Chemical Education, 73*(8), 758.

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

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

Nyachwaya, J. M., Warfa, A.-R. M., Roehrig, G. H., & Schneider, J. L. (2014). College chemistry students' use of memorized algorithms in chemical reactions. *Chemistry Education Research and Practice, 15*(1), 81-93.

Paivio, A. (2014). *Mind and its evolution: A dual coding theoretical approach*: Psychology Press.

Pickering, M. (1990). Further studies on concept learning versus problem solving. *Journal of Chemical Education, 67*(3), 254.

Prilliman, S. G. (2014). Integrating Particulate Representations into AP Chemistry and Introductory Chemistry Courses. *Journal of Chemical Education, 91*(9), 1291-1298.

Roseman, R., & McBride, D. (2011). *Exploring Student Understanding of Chemical Equations through Representative Drawings.* Paper presented at the National Conference on Undergraduate Research, Ithaca, New York.

Sanger, M. J. (2005). Evaluating students' conceptual understanding of balanced equations and stoichiometric ratios using a particulate drawing. *Journal of Chemical Education, 82*(1), 131.

Sanger, M. J., & Greenbowe, T. J. (2000). Addressing student misconceptions concerning electron flow in aqueous solutions with instruction including computer animations and conceptual change strategies. *International Journal of Science Education, 22*(5), 521-537.

Smith, K. J., & Metz, P. A. (1996). Evaluating student understanding of solution chemistry through microscopic representations. *Journal of Chemical Education, 73*(3), 233.

SocioCultural Research Consultants, L. (2014). Dedoose.

Stanovich, K. E., & Thompson, V. A. (2001). [Who Is Rational: Studies of Individual Differences in Reasoning]. *Canadian Journal of Experimental Psychology, 55*(1), 77.

Talanquer, V. (2006). Commonsense chemistry: a model for understanding students' alternative conceptions. *Journal of Chemical Education, 83*(5), 811.

Talanquer, V. (2012). Chemistry education: Ten dichotomies we live by. *Journal of Chemical Education, 89*(11), 1340-1344.

Treagust, D. F., & Chittleborough, G. (2001). Chemistry: A matter of understanding representations. *Advances in research on teaching, 8*, 239-268.

CHAPTER 3. PREVALENCE OF GENERAL CHEMISTRY STUDENTS' MISCONCEPTIONS ON EXCESS AND LIMITING REAGENTS AND YIELD

Stoichiometry deals with the quantitative aspects of chemical formulas and reactions. It lies in the very core of any first-year college chemistry course. The literature, though, is rich with examples that illustrate the extent of difficulties students have with stoichiometry concepts (Agung & Schwartz, 2007; BouJaoude & Barakat, 2000; Davidowitz, Chittleborough, & Murray, 2010; de Astudillo & Niaz, 1996; Kern, Wood, Roehrig, & Nyachwaya, 2010; Olmsted, 1999; Sanger, 2005). A potential reason for the difficulties that students have with stoichiometry may be that the concepts are too abstract and seem to be unreal to novices (Upahi & Olorundare, 2012). Stoichiometry problems are also quite complex for many students. Problems in stoichiometry often require students to write correct and balanced chemical equations, apply principles of ratios and proportions with respect to amounts of reactants and products, identify limiting reagents, and then find yields.

It has also been shown that success with stoichiometry problems that use algorithmic strategies may not necessarily imply conceptual understanding, especially when students are asked to solve problems that are somewhat different from those presented in the classroom (BouJaoude & Barakat, 2000). Students have been shown to generally lack the ability to solve transfer problems involving situations that are different from those used by instructors in the classroom (Bodner & Herron, 2003). To alleviate this difficulty, visual approaches using diagrams that illustrate the particulate nature of matter (PNOM) at the microscopic level have been suggested (Ben-Zvi, Eylon,

& Silberstein, 1987). PNOM diagrams may be used as tools with which students might visualize chemical concepts and build mental models (Gabel, 1998). They usually include representations of molecular, atomic, and sub-atomic particles shown either as single particles or as arrays of particles. Many authors of general chemistry text books use PNOM diagrams to complement pictures that illustrate macroscopic properties of substances and symbols used in chemical and mathematical equations (Davidowitz et al., 2010). PNOM diagrams are now found in the body of the text, among end-of-chapter problems, and in many ancillary materials that accompany text books. Many conceptual questions found in the various types of examinations released by the American Chemical Society Examinations Institute make use of PNOM diagrams as well (Luxford et al., 2014). These diagrams most commonly use spheres of different colors and sizes to represent different types of atoms or ions.

Expert chemists can readily interpret PNOM diagrams, but many students have a weak grasp of the theory of the particulate nature of matter (Gabel, 1999; Johnstone, 1993; Treagust & Chittleborough, 2001). Thus, students are often forced by their circumstances to fall back on macroscopic properties and everyday expressions to predict the structure of matter (Ben-Zvi, Eylon, & Silberstein, 1988). While there has been an increase in the use of PNOM diagrams to help explain stoichiometry concepts, little has been done to understand exactly how students interact with these illustrations (Davidowitz et al., 2010). All that is known for sure is that even with the use of these diagrams, student understanding of concepts in stoichiometry such as limiting reagents and yield still lag farther behind their abilities to solve numerical problems (Nakhleh, 1993; Nakhleh, Lowrey, & Mitchell, 1996; Nurrenbern & Pickering, 1987).

Dual thinking processes

Human cognition has been viewed as being made up of two underlying systems (Evans, 2008), each of which has distinct roles, differ in the kind of information being processed, differ in terms of the level of knowledge expressed, and have different responses. The first system, simply referred to as System 1, mostly makes use of heuristic processes and mainly chooses representations relevant to a particular problem space. These heuristics are usually used in an unconscious manner and are characterized by having fast processing rates, high capacity, and somewhat automated in their execution. The heuristics may be based on prior experiences, beliefs, and background knowledge. On the other hand, System 2 uses more analytical processes, which often require deliberate and more explicit thinking. The analytic processes might operate on representations determined by the heuristics used by System 1 to generate inferences and form judgments (Evans, 1996). System 2 thinking usually follows a sequence, is more controlled, and requires more from working memory to operate. It also does not follow what some would consider to be logical conventions, but is quite capable of arriving at solutions to a wide range of problem types.

Beginning chemistry students often use heuristics to answer a wide range of chemistry questions because heuristics tend to simplify reasoning by reducing the amount of information that needs to be processed (Maeyer & Talanquer, 2010). The implied rules of thumb for how and when to search for information, as well as how to handle the results, from heuristics may not always lead to the correct solution, but often give answers that are quite reasonable. The limited time students have during high-stakes assessment motivates many of them to rely, sometimes misguidedly, on the efficiency with which algorithms can help students arrive at answers that agree well with their prior knowledge. What students mostly fail to realize, however, is that most heuristics are task-specific procedures. Many college-age students do not quite have the ability to discern the limitations of the algorithms they use to solve chemistry problems. This gives rise to many student misconceptions.

The data and analyses in this chapter address the following questions: (1) *How do general chemistry students interpret particulate nature of matter diagrams when solving problems on excess and limiting reagents and yield,* (2) *What misconceptions do students have and how are they inferred from the ways general chemistry students interpret PNOM diagrams in relation to excess and limiting reagents, and yield,* (3) *How prevalent are these misconceptions among general chemistry students*? To answer these questions, an instrument consisting of 30 items (see Appendix D) measuring six different chemical contexts with three different chemical reactions was developed and administered online. The sample included students from three different types of general chemistry courses (Chem A, Chem D, and Chem E) during the fall semester of 2013 at a large state university in the Midwest. Chem A is the first part of a one-year course in general chemistry offered to physical and biological science majors, chemical engineering majors, as well as those intending to take 300-level chemistry courses. This course covers stoichiometry, parts of chemical equilibrium, acid-base chemistry, thermochemistry, rates and mechanism of reactions, changes of state, solution behavior, atomic structure, periodic relationships, and chemical bonding. Chem D is an accelerated course designed for students with excellent preparation in math and science and is a terminal course intended for engineering students who do not plan to take additional courses in chemistry. It covers principles of chemistry and properties of matter explained in terms of modern chemical theory with emphasis on topics of general interest to the engineer. Topics discussed in Chem D usually include nomenclature, chemical reactions, stoichiometry, atomic structure, periodic properties, chemical bonding, thermodynamics, chemical kinetics, chemical equilibrium, and electrochemistry. Chem E is a one-semester course aimed at providing students with an in-depth, broad-based view of modem chemistry. Chem E is also designed to introduce students to independent undergraduate research. Professors in chemistry, chemical engineering, and biochemistry are invited to present the scopes of their research activities in an attempt to encourage advanced undergraduate students to join these research groups.

Instrument Development and Validation

The items on the instrument were grounded on information obtained from student interview data discussed in Chapter 3, *i.e.*, student drawings and descriptions. The instrument was developed to measure students' understandings of the concepts of excess and limiting reagents, and yield. Two of the chemical reactions used (the combustion of methane gas and the synthesis reaction of ammonia) for the online instrument were identical to those used during interviews with students during the previous fall semester. For each of these two reactions an incorrect diagram selected from among those drawn by students during the interviews or illustrating misconceptions most commonly committed by the same group of students. Items for a third chemical reaction (disubstitution of chlorine atoms in carbon tetrachloride in the presence of hydrogen fluoride) were written to determine the effect of increasing the visual complexity of the accompanying PNOM diagram on item difficulty and discrimination. Thus, a 30-item instrument was written.

The instructors of the student participants were asked to examine the face and content validity of the items included on the instrument. The most common concern among faculty members was the proper illustration of molecules of gases and liquids using PNOM diagrams. In particular, faculty members stressed that PNOM representations of molecules of liquids must be shown to occupy more compact volumes than gases while having no identifiable repeating spatial patterns. Molecules of gases, on the other hand, must be shown to fully occupy the remaining volume of the container represented by the diagram. These suggestions were all taken into account in writing a final form of the instrument.

Online Administration of the Instrument

The instrument was administered as an online survey, first, during the two weeks immediately following each course’s examination on stoichiometry covering the topics of limiting and excess reagents as well as yield. This instrument was given in four different versions to account for ordering effects among two of the three chemical reactions (methane and ammonia) as well as ordering effects of the diagrams on student responses. Items for the CCl_4 reaction were kept at the end of each version of the instrument to delay the onset of any instrument fatigue students might experience as a result of responding to questions about diagrams with increased visual complexity. Version 1 of the instrument is given in Appendix D. For the purposes of discussion in this chapter, items will be referred to as numbered in Version 1.

A total of 1225 students participated during this initial survey round. Only the responses from 1126 students were retained for the purposes of this study. Eliminated from the study were responses from students who were below 18 years of age, those who denied consent for the study, those who spent less than two minutes going through all parts of the instrument, as well as those who spent more than two hours responding to the instrument's questions. The instrument was administered a second time to students who participated in the instrument's first administration during the two weeks immediately before thanksgiving break of the same semester. A total of 1084 students participated during the instrument's second administration. Of these, responses from 211 students were removed from the study for the same reasons that they were deleted from the first set of responses. Students who participated during both rounds of the instrument's administration received extra credit from their respective instructors.

Descriptive Statistics

Individual student scores for the first round of testing ranged from 0 to 29 with a mean of 15.0 ± 0.30 at 95% confidence (Table 1) for all student participants. Ferguson's δ was determined to be 0.978, which suggests that this sample was distributed over

97.8% of the possible range of total scores using participants from all courses included. The skewness of the scores was at -0.143 while the kurtosis was at 2.626, indicating a relatively normal distribution of scores. However, the statistically significant result from the Kolmogorov-Smirnov *z* test (Table 3) indicated that the sample distribution of first-round instrument scores were significantly different from a normal distribution.

Second-round scores ranged from 2 to 30 with a mean of 16.6 ± 0.35 at 95% confidence. Ferguson's δ of 0.980, while the skewness and kurtosis of scores were at -0.027 and -0.373, respectively. However, similar to the first-round scores, the Kolmogorov-Smirnov *z* test gave a statistically significant result indicating significant deviation from a normal distribution. Thus, nonparametric analyses were used to test hypotheses.

Table 1. Descriptive statistics for each round of testing using scores of all participants.

Round	*N*	Mean	Standard Deviation	Min.	Max.	Ferguson's δ	Skewness	Kurtosis
First	1126	15.0	5.11	0	29	0.978	-0.143	2.626
Second	873	16.6	5.36	2	30	0.980	-0.054	2.550

Table 2. Descriptive statistics for each group of students after each round of testing.

Course	Round	*N*	Mean	Standard Deviation	Min.	Max.
Chem A	First	618	15.4	5.28	0	28
	Second	468	15.9	5.62	2	30
Chem D	First	474	14.4	4.75	3	29
	Second	372	17.1	4.88	5	29
Chem E	First	34	16.4	5.92	4	27
	Second	33	20.3	4.62	9	27

Table 3. Kolmogorov-Smirnov *z* statistic for each round using scores of all participants.

Round	*Z*	Degrees of Freedom	*p* value
First	0.0720	1125	<0.001
Second	0.0596	872	0.004

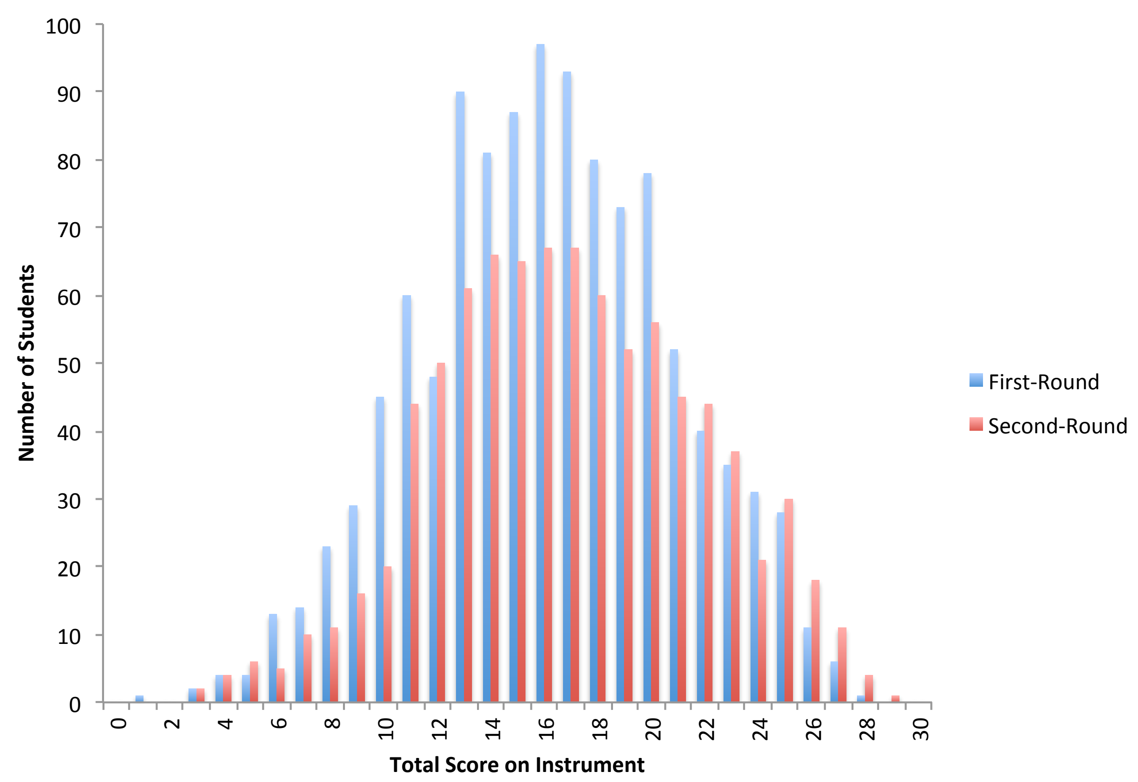


Figure 1. Distribution of total scores on both rounds of testing including students from all courses (*N* = 1126 for the first round and 873 for the second round).

Scores obtained by students in each course from both rounds of testing are summarized in Table 2. Total scores obtained by students from all courses also increased from the first to the second round of testing. The most noticeable difference is the much narrower range of scores obtained by Chem E students after the second round of testing. Increases in total scores most likely resulted from instruction given the six-week gap between rounds of testing. Pairwise comparisons of total scores from among the different groups of students were done using a *t* test at $\alpha = 0.05$ with a Bonferroni correction (Table 4). All first-round score differences had small effect sizes with students from Chem A obtaining significantly higher scores than those coming from Chem D, but not compared to those coming from Chem E. The difference between first-round scores of students from Chem D and Chem E was also not statistically significant. Second-round test score differences between groups of students were all found to be statistically significant. The difference between Chem A and Chem D had a small effect size. Second-round score differences between Chem A and Chem E had a large effect size, while the effect size of those between Chem D and Chem E was moderate.

Table 4. Pairwise comparison of total score on the instrument during each round of testing by each student group using *t*-test at $\alpha = 0.05$, with Bonferroni correction.

Pair Comparison	First Round			Second Round		
	Contrast	*p* value	Cohen's *d*	Contrast	*p* value	Cohen's *d*
Chem A - Chem D	+0.951*	0.002	-0.188	-1.236*	0.001	-0.233
Chem A - Chem E	-1.035	0.248	-0.195	-4.412*	0.001	-1.149
Chem D - Chem E	-1.986	0.028	-0.410	-3.176*	<0.001	-0.654

*Statistically significant at $\alpha = 0.05$.

Kruskal-Wallis tests on instrument scores (Table 5) for each group of students were conducted to evaluate ordering effects among chemical contexts as given in the four different versions of the instrument. For both rounds of testing, no ordering effect among the different chemical contexts was determined to be statistically significant.

Table 5. Kruskal-Wallis *p* values on instrument scores obtained by each student group across different versions of the instrument after each round of testing (α = 0.05).

Student Group	*p* value	
	First Round	Second Round
Chem A	0.885	0.748
Chem D	0.134	0.734
Chem E	0.622	0.806

Item Analysis

Item difficulty indices

The item difficulty index is the fraction of students who correctly answered an item. It is, therefore, a measure of the easiness of the item. Desired values for item difficulty index are between 0.25 and 0.80. Items with difficulty indices less than 0.25 are considered to be too difficult while those with indices greater than 0.80 are deemed too easy (Ding & Beichner, 2009). Item difficulty indices were determined for each group of participants after every round of testing. For students from Chem A item difficulty indices ranged from 0.26 to 0.70 after the first run of the instrument and from 0.29 to 0.68 after the second run (Figure 2).

Item difficulties ranged from a low of 0.26 to a high of 0.69 after the first run and from a low of 0.25 to a high of 0.75 after the second run for students coming from Chem D (Figure 3). These numbers are all well within the desired difficulty levels. One difference between these two sets of item difficulty indices, however, is that a greater number of items showed slight increases for students from Chem D. Item difficulty indices for students from Chem A, on the other hand, showed minimal changes between each round of testing.

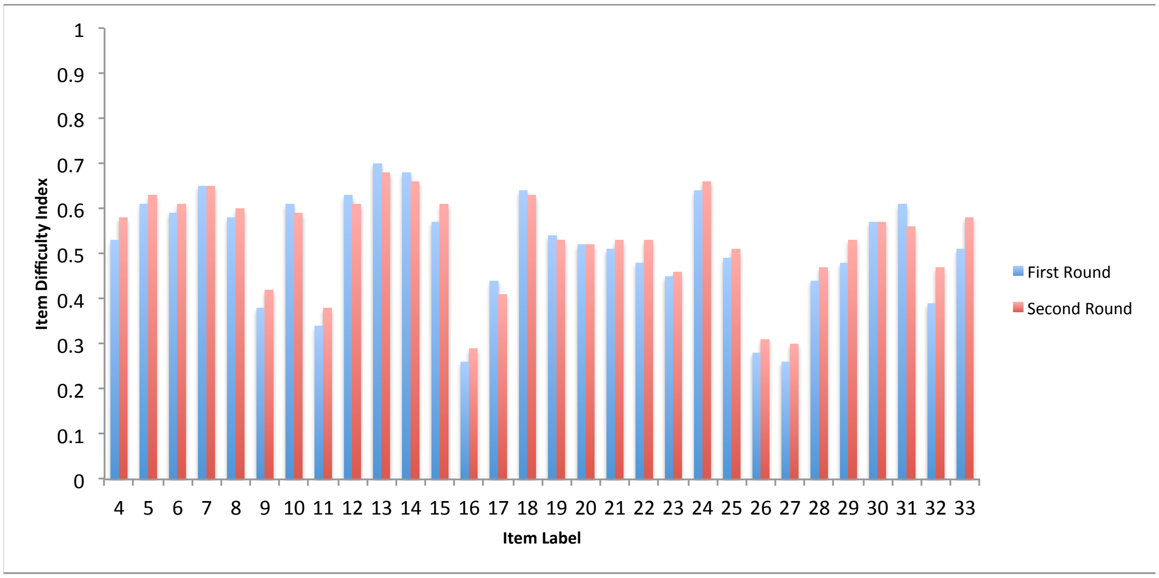


Figure 2. Item difficulty indices for students coming from Chem A.

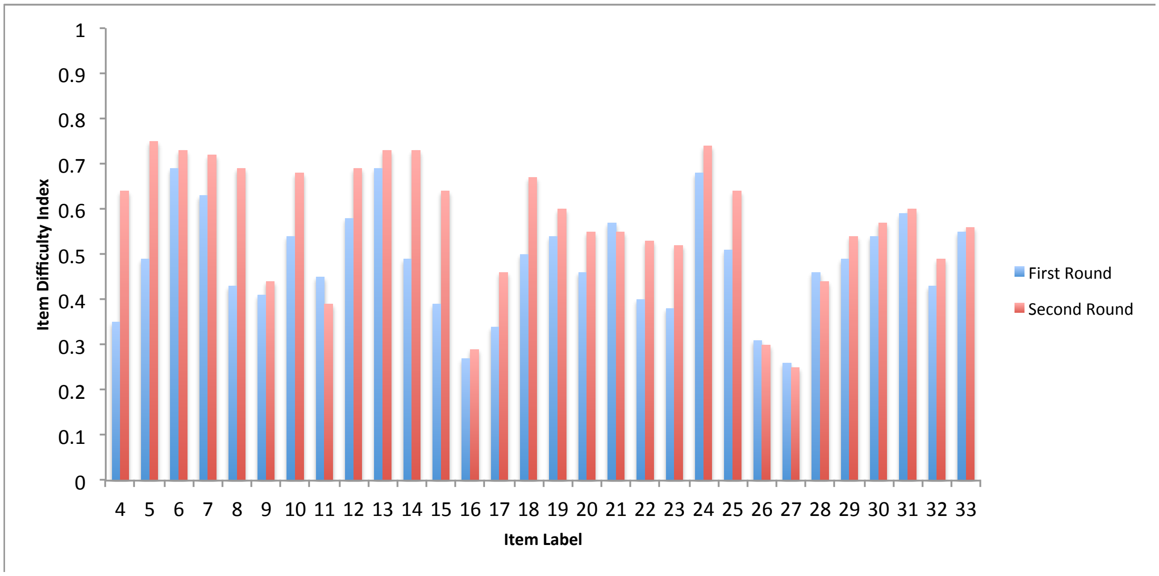


Figure 3. Item difficulty indices for students coming from Chem D.

For students coming from Chem E, item difficulties ranged from 0.21 to 0.82 after the first run, and from 0.27 to 0.94 after the second run (Figure 4). In particular among Chem E students, student performances on item 27 suggest it may be too difficult and 21 may be too easy for inclusion in an inverntory-style assessment such as the one used for this study. Items 5, 7, 8, 10, 13, 14, 15, 18, and 24 (Appendix D) were too easy for Chem E students during the second run of the instrument. This lends more validity to the instrument since students from Chem E are generally assumed to have a better grasp of the concepts covered by the instrument than others do.

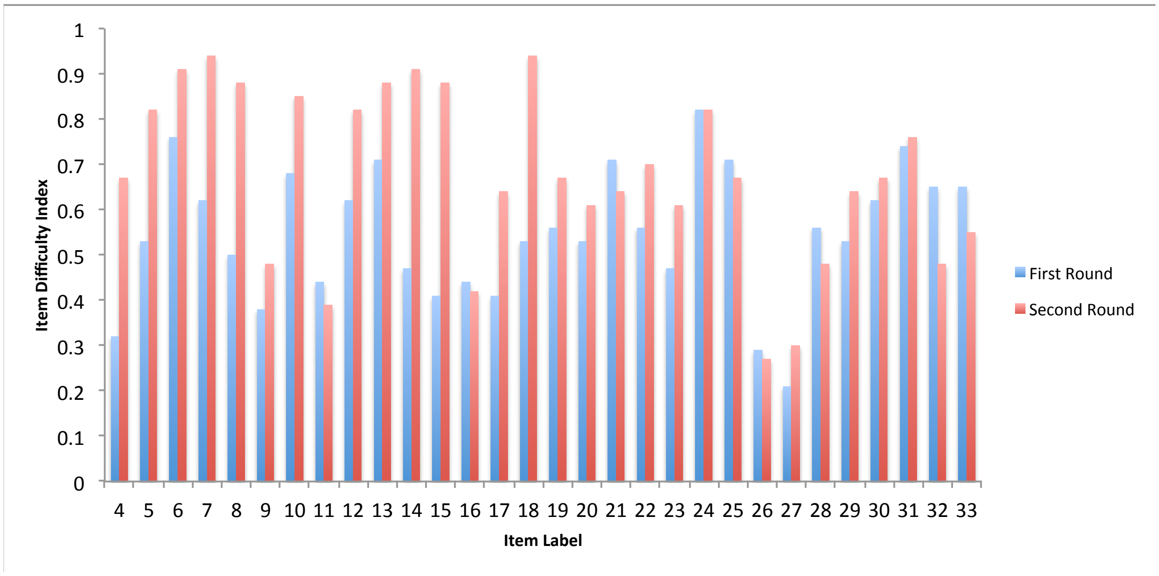


Figure 4. Item difficulty indices for students coming from Chem E.

Item discrimination indices

Aside from item difficulty, the item discrimination indices were calculated for each item. This index is a measure of the extent to which an item distinguishes between the upper 27% and the lower 27% of students based on total score on the instrument (Feldt, 1963). A higher value for this index indicates better discrimination so that a greater fraction of students with higher total scores on an instrument are getting a specific item correct, while more students getting lower total scores are giving incorrect answers to the same item. A value of at least 0.30 is considered to be adequate to distinguish between high- and low-performing groups of students (Ding and Beichner, 2009).

For students coming from Chem A, discrimination indices ranged from a low of 0.12 to a high of 0.64 after the first round of testing, and from a low of 0.07 to a high of 0.69 after the second round (Figure 5). Items 16, 26, 27, 28, 30, 31, and 33 had discrimination indices well below the 0.30 threshold after the first run. This was also true for items 26, 27, 28, 30, and 33 after the second run with Chem A students.

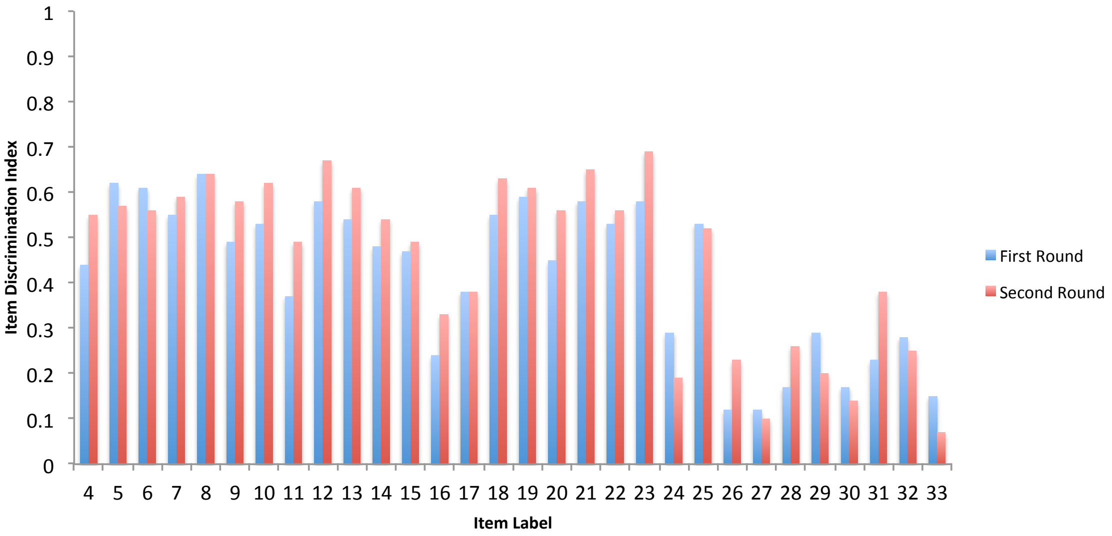
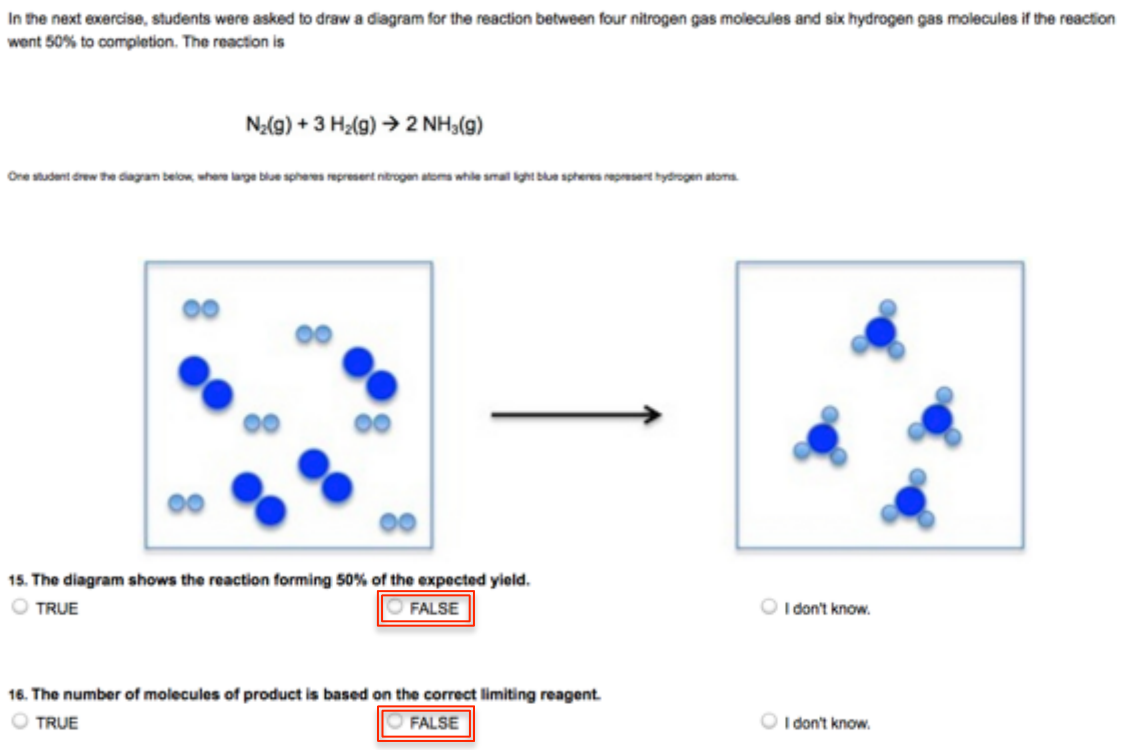


Figure 5. Item discrimination indices for students coming from Chem A.

Table 6 lists responses to items 15 and 16 (Figure 6) from students in Chem A after the first round of testing. These items show how commonly students seem to treat the concepts of limiting reagents and reaction yield independently of each other. A statistically significant Fisher's exact test along with the Cramer's *V* value indicated moderate correlation between how students responded to these questions. Among students who correctly indicated that the diagram did not reflect a reaction at 50% yield, only 57% of them could identify hydrogen gas as the limiting reagent. In fact, more students from Chem A, including some of those who overall performed better on the instrument, were choosing nitrogen as their limiting reagent. In any case, because only 26% of students got item 16 correctly during the first run of the instrument, the low discrimination index for the same item was not surprising.

Table 6. Item response distribution and Fisher exact test between items 15 and 16 using Chem A student responses after the first round of testing.

	Item 16 Response			
Item 15 Response	True	False	No Answer	Row Total
True	217	105	37	359
False	74	120	17	211
No Answer	17	7	17	41
Column Total	308	232	71	611
Cramer's *V*				0.2617
Fisher's exact test *p* value				< 0.001



In the next exercise, students were asked to draw a diagram for the reaction between four nitrogen gas molecules and six hydrogen gas molecules if the reaction went 50% to completion. The reaction is

$$N_2(g) + 3\ H_2(g) \rightarrow 2\ NH_3(g)$$

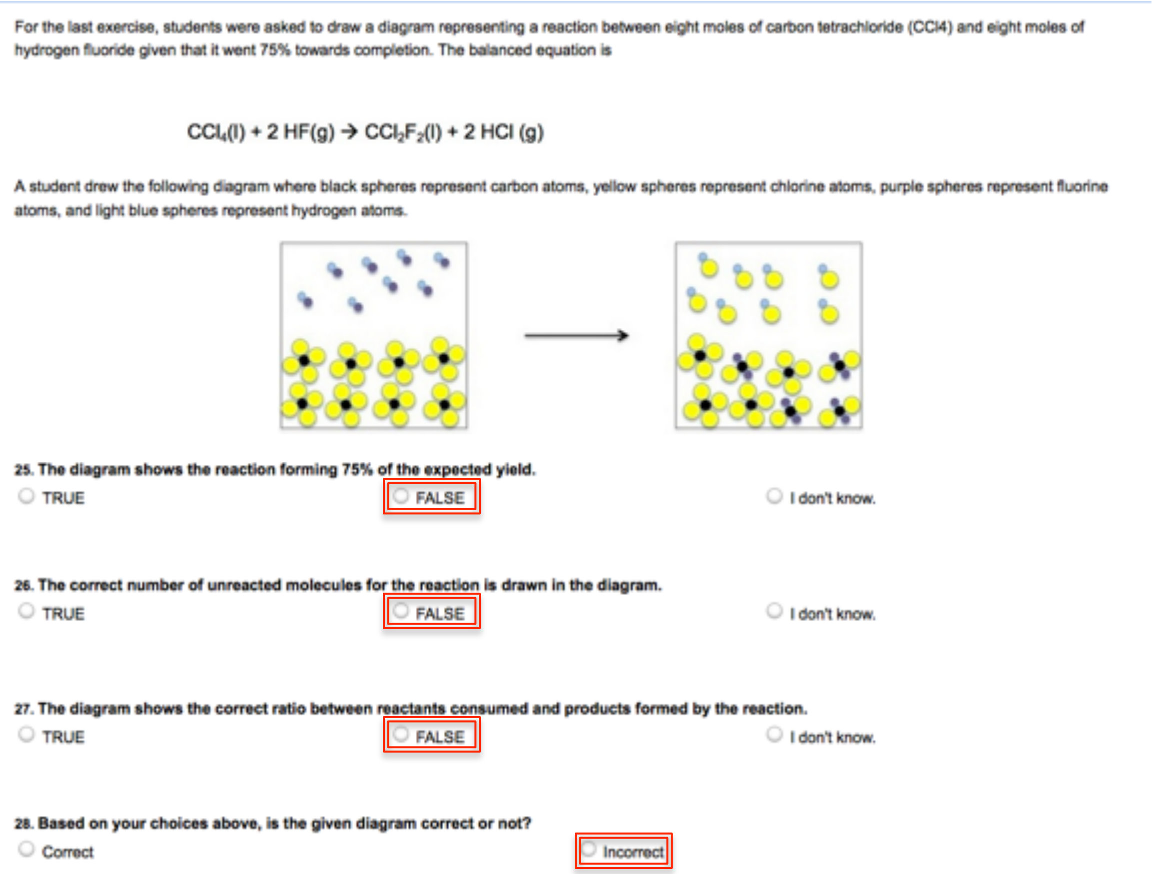
One student drew the diagram below, where large blue spheres represent nitrogen atoms while small light blue spheres represent hydrogen atoms.

15. The diagram shows the reaction forming 50% of the expected yield.
○ TRUE ○ FALSE ○ I don't know.

16. The number of molecules of product is based on the correct limiting reagent.
○ TRUE ○ FALSE ○ I don't know.

Figure 6. Items 15 and 16.

Items 26 and 27 (Figure 7) both behaved similarly as item 16 did with Chem A students. Both questions barely exceeded the 0.25 threshold for item difficulty but had discrimination indices less than 0.30. Poor discrimination indices for items with low difficulty indices are not unusual, because many high-performing students might have given incorrect answers to these items. Among those who correctly indicated that the diagram does not reflect the given percent yield in item 25, 56% incorrectly thought that the diagram showed the correct number of unreacted molecules in item 26 (Table 7). This last group of students were not making the connection between percent yield and unused reactants. With respect to this diagram, students seemed to be treating these items independently of each other. The same pattern of responses was seen between items 25 and 27 (Table 8).



For the last exercise, students were asked to draw a diagram representing a reaction between eight moles of carbon tetrachloride (CCl4) and eight moles of hydrogen fluoride given that it went 75% towards completion. The balanced equation is

$$CCl_4(l) + 2\ HF(g) \rightarrow CCl_2F_2(l) + 2\ HCl\ (g)$$

A student drew the following diagram where black spheres represent carbon atoms, yellow spheres represent chlorine atoms, purple spheres represent fluorine atoms, and light blue spheres represent hydrogen atoms.

25. The diagram shows the reaction forming 75% of the expected yield.

○ TRUE [✓ ○ FALSE] ○ I don't know.

26. The correct number of unreacted molecules for the reaction is drawn in the diagram.

○ TRUE [✓ ○ FALSE] ○ I don't know.

27. The diagram shows the correct ratio between reactants consumed and products formed by the reaction.

○ TRUE [✓ ○ FALSE] ○ I don't know.

28. Based on your choices above, is the given diagram correct or not?

○ Correct [✓ ○ Incorrect]

Figure 7. Items 25 through 28.

Table 7. Item response distribution and Fisher exact test between item 25 and 26 using Chem A student responses after the both rounds of testing.

First Round				
	Item 26 Response			
Item 25 Response	True	False	No Answer	Row Total
True	171	40	8	219
False	168	125	8	301
No Answer	32	9	50	91
Column Total	371	174	66	611
	Cramer’s *V*			0.4540
	Fisher’s exact test *p* value			< 0.001
Second Round				
	Item 26 Response			
Item 25 Response	True	False	No Answer	Row Total
True	129	37	7	173
False	126	102	8	236
No Answer	22	9	23	54
Column Total	277	148	38	463
	Cramer’s *V*			0.3580
	Fisher’s exact test *p* value			< 0.001

Table 8. Item response distribution and Fisher exact test between items 25 and 27 using Chem A student responses after both rounds of testing.

First Round				
	Item 27 Response			
Item 25 Response	True	False	No Answer	Row Total
True	171	37	9	217
False	180	108	12	300
No Answer	24	15	52	91
Column Total	375	160	73	608
Cramer’s *V*				0.4364
Fisher’s exact test *p* value				< 0.001

Second Round				
	Item 27 Response			
Item 25 Response	True	False	No Answer	Row Total
True	131	36	7	174
False	139	92	7	238
No Answer	19	12	23	54
Column Total	289	140	37	466
Cramer’s *V*				0.3532
Fisher’s exact test *p* value				< 0.001

Students who said that the diagram for items 24-28 correctly illustrated the given percent yield (75%) in the problem overwhelmingly said that the same diagram was overall correct (Table 9). However, students who gave a “False” response to item 25 were split almost 2:3 between those who indicated that the diagram was “Correct” overall and those who picked “Incorrect.” Responses to both items 26 (Table 10) and 27 (Table 11), on the other hand, were observed to be strongly correlated with those answers given by students for item 28. The incorrect depiction of the expected yield by the diagram did not weigh as heavily in Chem A students’ minds as did both the incorrect number of unreacted molecules and the ratio between reactants and products. It is certainly possible that this could have resulted from item ordering effects even though these items were all simultaneously presented to the students during testing. Earlier items may have a priming effect that influences the way students respond to succeeding items (Schroeder, Murphy, & Holme, 2012).

Proximity may or may not have been a factor in explaining students’ response patterns between items 29 and 30 (Figure 8). Chem A students who picked “True” as their response to item 29 often also picked “True” (Table 12) as their response for item 30. However, students who chose “False” for item 29 were almost evenly split as far as item 30 was concerned during the first run of the instrument. This lack of predictability in the way one group of students thought about these items reduces the ability of an item to discriminate between high- and low-performers. It suggests that some students treated different aspects of the same PNOM diagram independently of each other.

Table 9. Item response distribution and Fisher exact test between items 25 and 28 using responses from Chem A students after both rounds of testing.

		First Round		
		Item 28 Response		
Item 25 Response	Correct	Incorrect	No Answer	Row Total
True	184	33	0	217
False	114	188	0	302
No Answer	40	49	1	90
Column Total	338	270	1	609
	Cramer's *V*			0.3190
	Fisher's exact test *p* value			< 0.001
		Second Round		
		Item 28 Response		
Item 25 Response	True	False	No Answer	Row Total
True	135	39	0	174
False	88	150	0	238
No Answer	23	31	1	55
Column Total	246	220	1	467
	Cramer's *V*			0.2865
	Fisher's exact test *p* value			< 0.001

Table 10. Item response distribution and Fisher exact test between items 26 and 28 using responses from Chem A students after both rounds of testing.

First Round

	Item 28 Response			
Item 26 Response	Correct	Incorrect	No Answer	Row Total
True	287	84	0	371
False	25	149	0	241
No Answer	28	42	1	71
Column Total	340	275	1	616
Cramer's *V*				0.4084
Fisher's exact test *p* value				< 0.001

Second Round

	Item 28 Response			
Item 26 Response	Correct	Incorrect	No Answer	Row Total
True	215	63	0	278
False	20	128	0	148
No Answer	11	28	1	40
Column Total	246	219	1	466
Cramer's *V*				0.4384
Fisher's exact test *p* value				< 0.001

Table 11. Item response distribution and Fisher exact test between items 27 and 28 using responses from Chem A students after both rounds of testing.

		First Round		
		Item 28 Response		
Item 27 Response	True	False	No Answer	Row Total
True	289	85	0	374
False	22	139	0	161
No Answer	25	52	1	78
Column Total	336	276	1	613
	Cramer’s *V*			0.4127
	Fisher’s exact test *p* value			< 0.001
		Second Round		
		Item 28 Response		
Item 27 Response	True	False	No Answer	Row Total
True	218	72	0	290
False	17	123	0	140
No Answer	13	24	1	38
Column Total	248	219	1	468
	Cramer’s *V*			0.4229
	Fisher’s exact test *p* value			< 0.001

Table 12. Item response distribution and Fisher exact test between items 29 and 30 using responses from Chem A students after both rounds of testing.

First Round				
	Item 30 Response			
Item 29 Response	True	False	No Answer	Row Total
True	224	67	6	297
False	99	106	5	210
No Answer	31	23	49	103
Column Total	354	196	60	610
	Cramer's *V*			0.4496
	Fisher's exact test *p* value			< 0.001
Second Round				
	Item 30 Response			
Item 29 Response	True	False	No Answer	Row Total
True	186	60	5	251
False	65	100	2	167
No Answer	14	16	17	47
Column Total	265	176	24	465
	Cramer's *V*			0.4155
	Fisher's exact test *p* value			< 0.001

Responses to item 33 seem to have been mostly affected by the way students thought about item 29 (Table 13). Of the four questions immediately preceding item 33, the strongest correlation was found with item 32. It is likely that students who chose “False” for item 32 regarded the diagram to be incorrect because of what they may have seen as an error in the diagram. Item order may also have had some role in this response pattern. Item 29, on the other hand, was found to have the weakest relationship with item 33.

Table 13. Effect sizes of Fisher exact tests between Item 33 and each of the other four items related to the last diagram.

False		Item 29	Item 30	Item 31	Item 32
First Round	Cramer’s *V*	0.2583	0.4074	0.3285	0.4796
	p value	< 0.001	< 0.001	< 0.001	< 0.001
Second Round	Cramer’s *V*	0.2856	0.4234	0.3590	0.5035
	p value	< 0.001	< 0.001	< 0.001	< 0.001

Figure 8. Items 29 through 33.

Among students coming from Chem D, item discrimination indices went from 0.03 to 0.66 as a result of the first round of testing, and went from 0.03 to 0.59 the second time (Figure 9). Items behaved similarly using Chem D students’ responses as they did with responses from Chem A students in general. Exceptions to this were items 9 and 25 after the first round of testing, and item 11 after the second round.

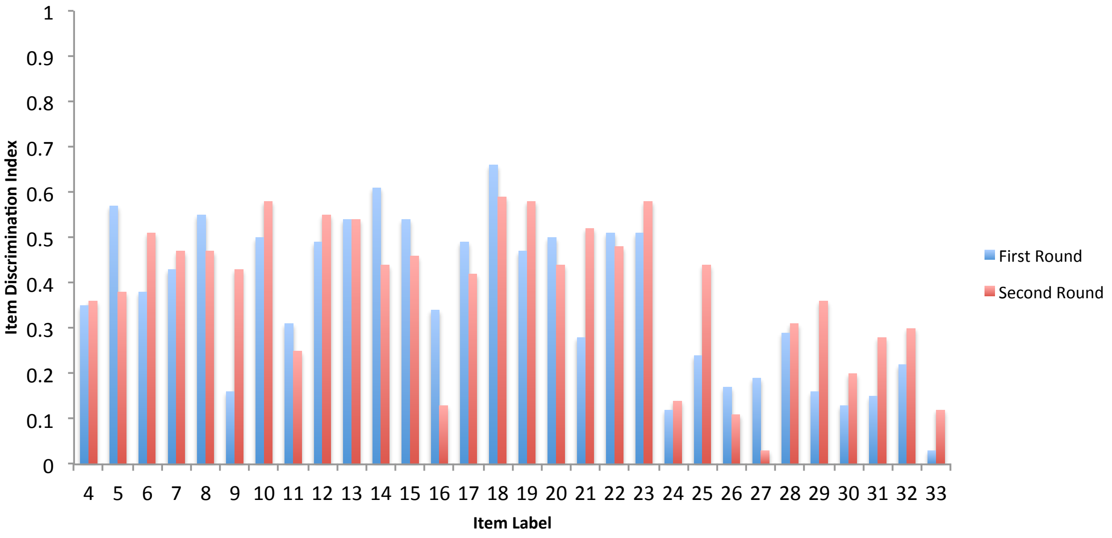
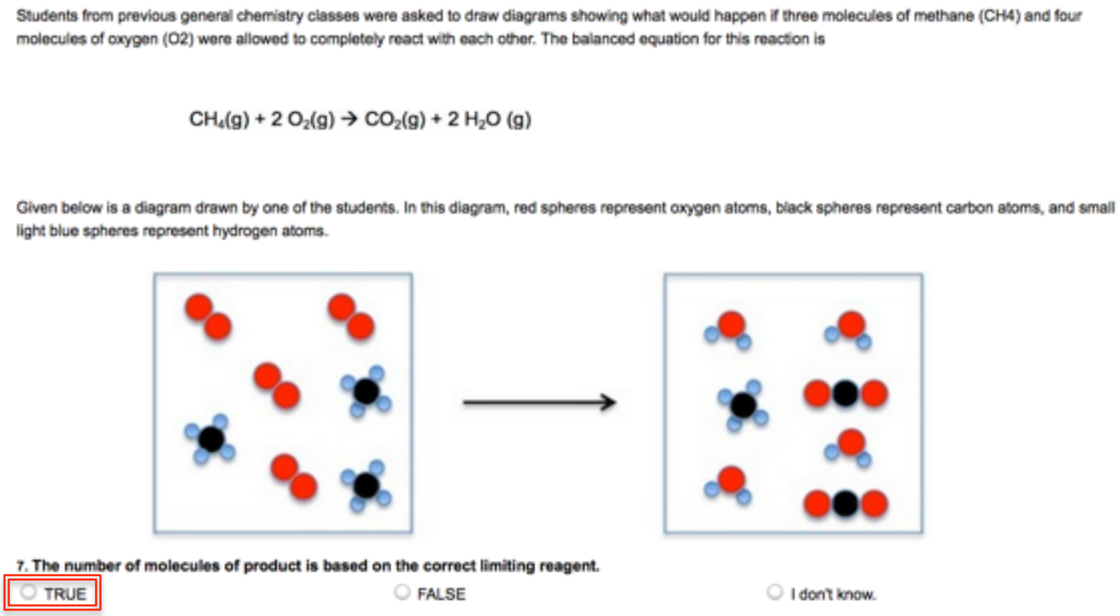


Figure 9. Item discrimination indices for students coming from Chem D.

Table 14 lists Chem D students' responses to items 7 (Figure 10) and 9 (Figure 11) after both rounds of testing. Item 9 had a discrimination index of 0.16 for these students after the first round. Responses after the first round showed that students were not connecting their thought processes between these two items even if they both dealt with identifying the limiting reagent for the same chemical reaction. Even students who responded "True" to item 7 were roughly evenly split with respect to their responses to item 9. Having decided that the diagram for item 7 was associated with the correct choice of the limiting reagent should have prompted students to pick "False" for item 9 since there can only be one correct diagram a specific chemical situation. The small effect size after the first round of testing indicated very little correlation between responses from Chem D students for these two items.

Table 14. Item response distribution and Fisher exact test between items 7 and 9 using Chem D student responses after both rounds of testing.

First Round				
	Item 9 Response			
Item 7 Response	True	False	No Answer	Row Total
True	159	128	13	300
False	84	64	4	152
No Answer	12	4	4	20
Column Total	255	196	21	472
	Cramer's *V*			0.1243
	Fisher's exact test *p* value			0.024
Second Round				
	Item 9 Response			
Item 7 Response	True	False	No Answer	Row Total
True	129	131	7	267
False	55	31	3	89
No Answer	8	3	5	16
Column Total	192	165	15	372
	Cramer's *V*			0.2280
	Fisher's exact test *p* value			< 0.001



Students from previous general chemistry classes were asked to draw diagrams showing what would happen if three molecules of methane (CH4) and four molecules of oxygen (O2) were allowed to completely react with each other. The balanced equation for this reaction is

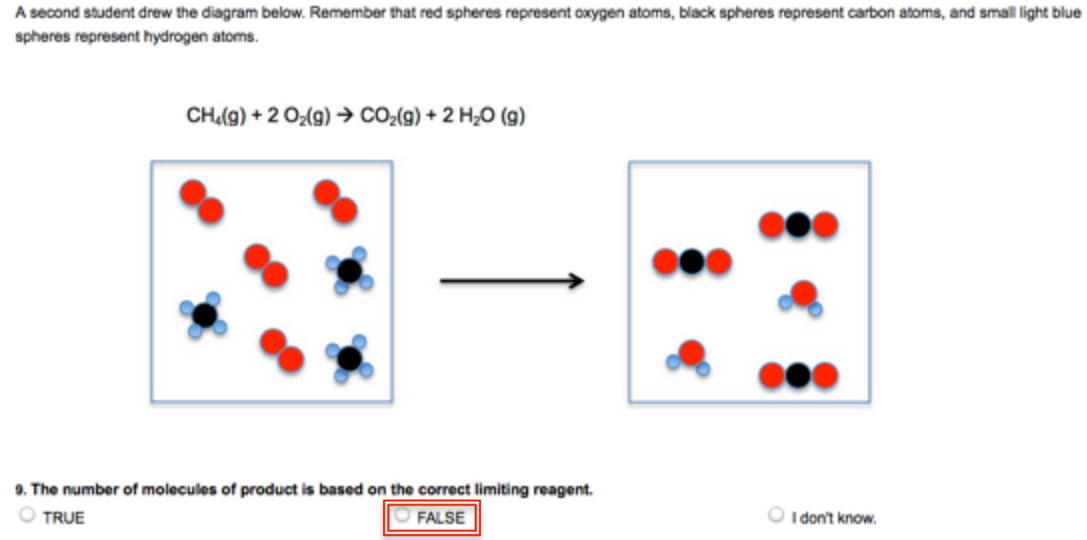
$$CH_4(g) + 2\ O_2(g) \rightarrow CO_2(g) + 2\ H_2O\ (g)$$

Given below is a diagram drawn by one of the students. In this diagram, red spheres represent oxygen atoms, black spheres represent carbon atoms, and small light blue spheres represent hydrogen atoms.

7. The number of molecules of product is based on the correct limiting reagent.

[✓ ○ TRUE] ○ FALSE ○ I don't know.

Figure 10. Item 7.



A second student drew the diagram below. Remember that red spheres represent oxygen atoms, black spheres represent carbon atoms, and small light blue spheres represent hydrogen atoms.

$$CH_4(g) + 2\ O_2(g) \rightarrow CO_2(g) + 2\ H_2O\ (g)$$

9. The number of molecules of product is based on the correct limiting reagent.

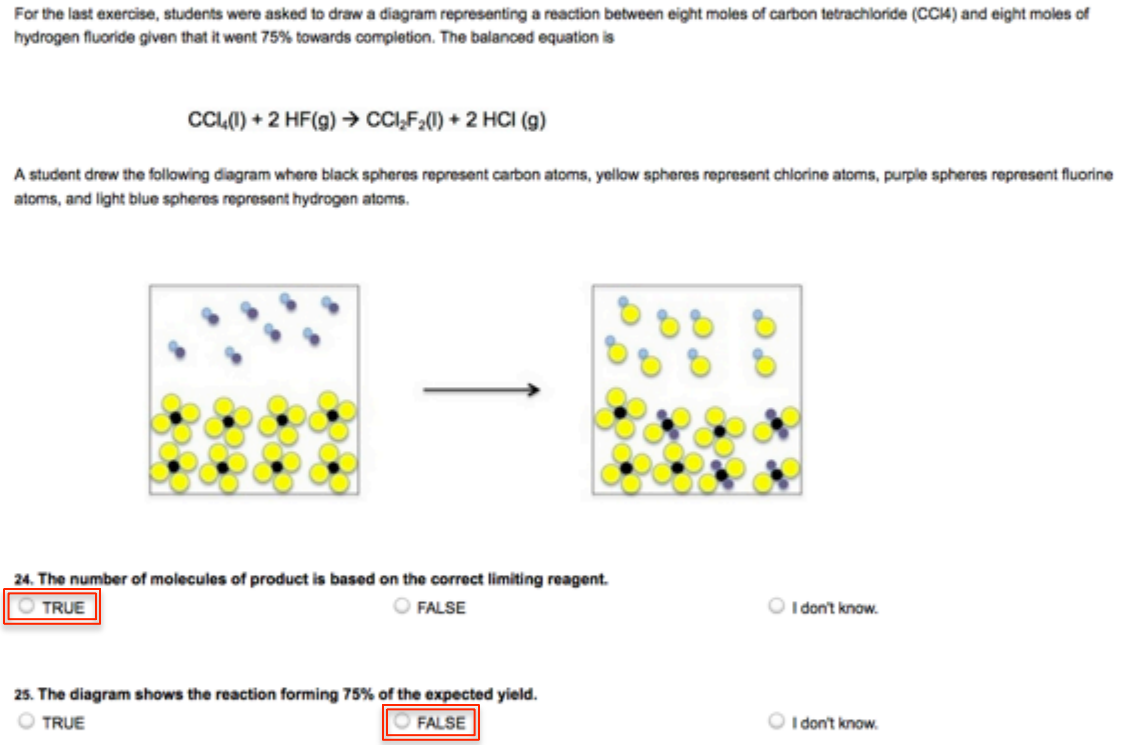
○ TRUE [✓ ○ FALSE] ○ I don't know.

Figure 11. Item 9.

Item 25 (Figure 12) had a discrimination index of 0.24 with Chem D students after the first round of testing. Students who correctly picked “True” for item 24 were distributed roughly 3:4 with respect to their responses to item 25 (Table 15). It is not really clear here how these students thought about the yield of the reaction as given in the diagram, especially since 68% of them got item 24, which deals with determining the limiting reagent, correctly.

Table 15. Item response distribution and Fisher exact test between items 24 and 25 using Chem D student responses after the first round of testing.

	Item 25 Response			
Item 24 Response	True	False	No Answer	Row Total
True	127	166	30	323
False	27	66	13	106
No Answer	5	9	24	38
Column Total	159	241	67	38
	Cramer’s *V*			0.3053
	Fisher’s exact test *p* value			< 0.001



For the last exercise, students were asked to draw a diagram representing a reaction between eight moles of carbon tetrachloride (CCl4) and eight moles of hydrogen fluoride given that it went 75% towards completion. The balanced equation is

$$CCl_4(l) + 2\ HF(g) \rightarrow CCl_2F_2(l) + 2\ HCl\ (g)$$

A student drew the following diagram where black spheres represent carbon atoms, yellow spheres represent chlorine atoms, purple spheres represent fluorine atoms, and light blue spheres represent hydrogen atoms.

24. The number of molecules of product is based on the correct limiting reagent.

[✓ ○ TRUE] ○ FALSE ○ I don't know.

25. The diagram shows the reaction forming 75% of the expected yield.

○ TRUE [✓ ○ FALSE] ○ I don't know.

Figure 12. Items 24 and 25.

Item discrimination indices for students from Chem E went from a low of -0.14 to a high of 0.82 after the instrument was run once and from a low of -0.09 to a high of 0.91 after being run twice (Figure 13). Item 24 had a negative index after the first round most likely due to a very low difficulty index. This is not necessarily a reason for concern when classes are small. This negative value indicates that a large fraction of the high-performing students on the instrument were picking “True” as their response to this item. This probably based on the way the diagram used all of the limiting reagent, hydrogen fluoride, to form products even though the problem requires a 75% yield. Items 5, 6, 7, 14, 18, and 24 also had low discrimination indices among these students in the second round that may be attributed to high difficulty indices as well. Item 27

(Figure 7) had a discrimination index of 0.05 due to a very low difficulty index among Chem E students after both rounds of testing.

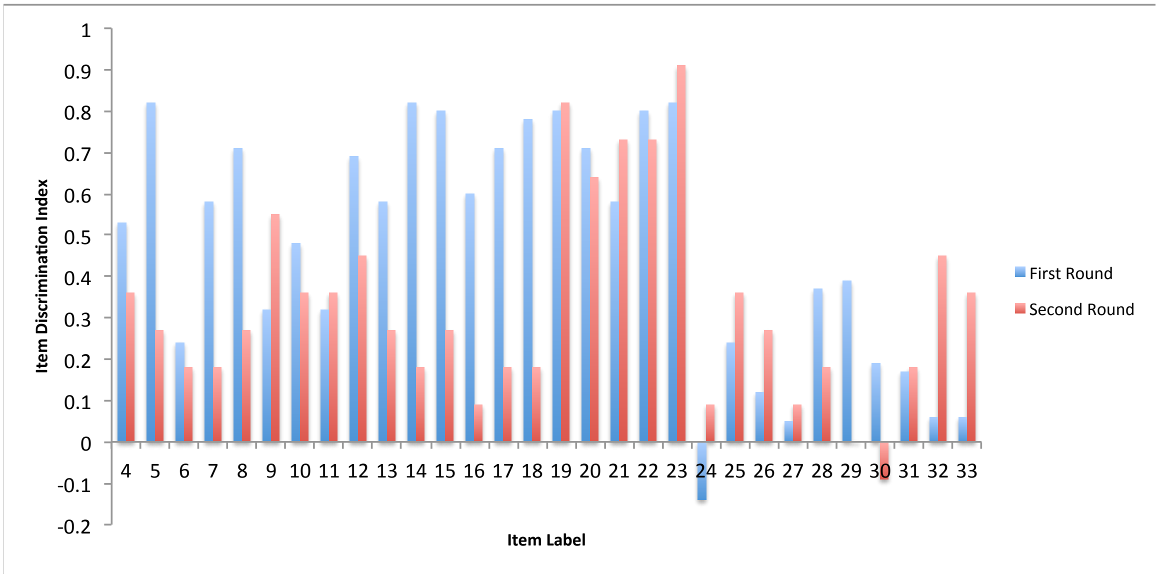


Figure 13. Item discrimination indices for students from Chem E.

Item 30 had a discrimination index of -0.09 despite a difficulty index of 0.67; generally, a low discrimination index would suggest a very easy or very difficult item, but this was not the case. Comparison with responses for item 29 showed that most Chem E students who picked “True” for item 29 also picked “True” for item 30 (Table 16). However, students who picked “False” for item 29 were almost evenly split with respect to their responses for item 30. Students who picked “False” for item 30 may have been referring to the missing HCl molecules in the product side of the diagram to justify their responses. In any case, the large effect size points to the fact that the vast majority of students in this group agreed that the answers to both items 29 and 30 must be “True.”

Table 16. Item response distribution and Fisher exact test between items 29 and 30 using responses from Chem E students after the second round of testing.

	Item 30 Response			
Item 29 Response	True	False	No Answer	Row Total
True	18	4	0	22
False	5	6	0	11
No Answer	0	0	1	1
Column Total	23	10	1	34
Cramer's *V*				0.7547
Fisher's exact test *p* value				< 0.003

Instrument Reliability

The internal consistency of an instrument is measured by Cronbach's α, which indicates how closely the items measure the same construct (Cronbach, 1951). Cronbach's α was determined for the instrument using the responses obtained by each group of students. Cronbach's α values are given in Table 17. Acceptable values for α are greater than or equal to 0.7 (Kline, 2013). The values determined for each group of students were above the 0.7 threshold value, and therefore, indicated reasonable internal consistency among the items included in the instrument.

Table 17. Cronbach's α values for each group of students after each round of testing.

Student Group	Cronbach's α	
	First Round	Second Round
Chem A	0.716	0.723
Chem D	0.707	0.707
Chem E	0.751	0.748

Another way to determine reliability is by evaluating the consistency of the data produced by an instrument across two time periods and calculating a stability coefficient (Adams & Wieman, 2011). Since the instrument was administered twice with each group of students during the same semester, a Wilcoxon signed ranks test was performed to compare scores obtained by students who participated in both rounds of administration of the instrument. While this test revealed no statistically significant differences between scores obtained by Chem A students after each round of testing ($p = 0.225$), significant differences were seen in the performances of Chem D ($p = 0.019$) and Chem E ($p = 0.040$) students after each run, with students from each group showing some improvement. These results likely point to intervention effects coming from instruction that took place during the eight weeks between testing times.

Spearman's rank correlation coefficients (Table 18) were also determined for each set of item analyses. Item analysis parameters determined after each run of the instrument were found to be strongly correlated with each other for students coming from Chem A as shown in Figure 14. This is consistent with the nonsignificant Wilcoxon signed ranks test on total scores (Table 18) as this group of students showed the least differences on instrument performance between each round of testing.

Table 18. Tests of performance stabilities on the instrument for each group of students.

Student Group	Measure of instrument reliability	
	Wilcoxon signed ranks test *p* value	Spearman's rank correlation coefficient, ρ
Chem A	0.225	0.437
Chem D	0.019	0.421
Chem E	0.040	0.366

Corresponding pairs of item analysis parameters were seen to be strongly correlated with each other between both rounds of testing for Chem A students (Figure 14). In particular, item difficulty indices from both rounds of testing were seen to be most strongly correlated with each other. This is consistent with nonsignificant changes in total scores among these students. This also indicates absence of significant changes in scores on individual items on the instrument.

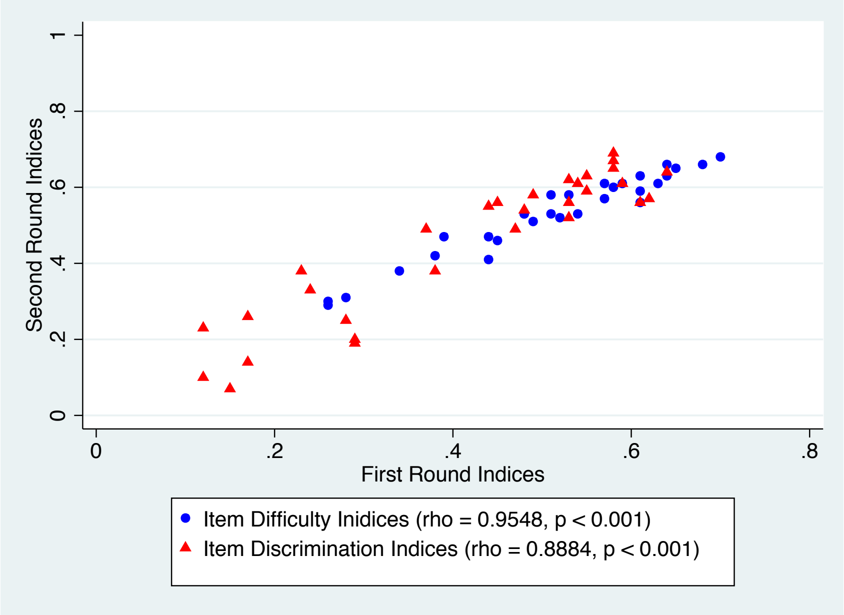


Figure 14. Correlations between item analysis parameters from each run of the instrument for each of the 30 items with students coming from Chem A.

Increases in the difficulty indices of about a third of the items led to less strongly correlated pairs of item analysis parameters resulting from both rounds of testing among Chem D students (Figure 15). This also led to lower correlations between corresponding discrimination indices and point biserial correlation coefficients for each round of testing. The items lost some ability to distinguish between high- and low-performers as the differences between these two groups become narrower.

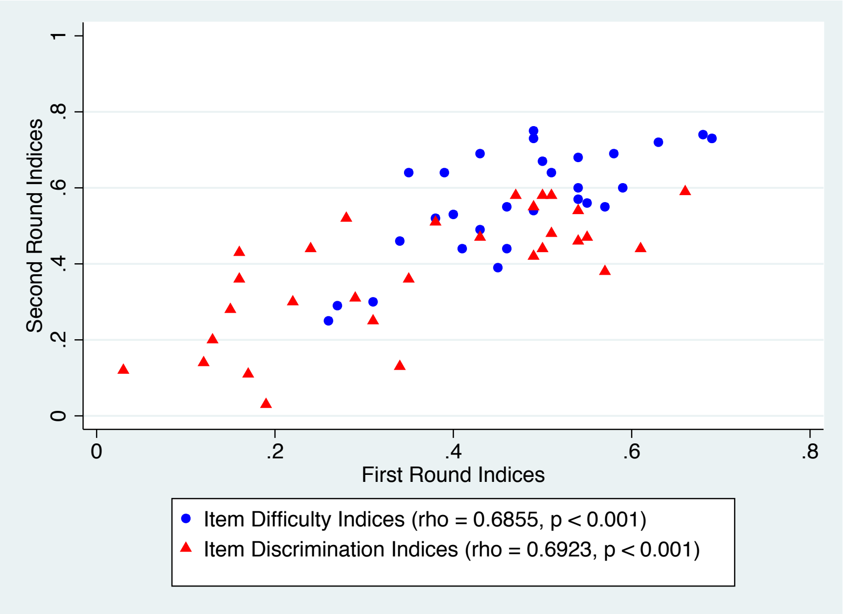


Figure 15. Correlations between item analysis parameters from each run of the instrument for each of the 30 items with students coming from Chem D.

Students from Chem E probably comprise the most homogeneous group among those included in this study in terms of their preparation for college-level chemistry. As a result, "high-" and "low-" performers from these students are expected to have the narrowest differences in terms of total score on the instrument. This was observed specifically after the second round of testing. As a result, many items had low discriminating ability. This could have led to the weak correlations between pairs of item analysis parameters determined for Chem E students after two testing rounds.

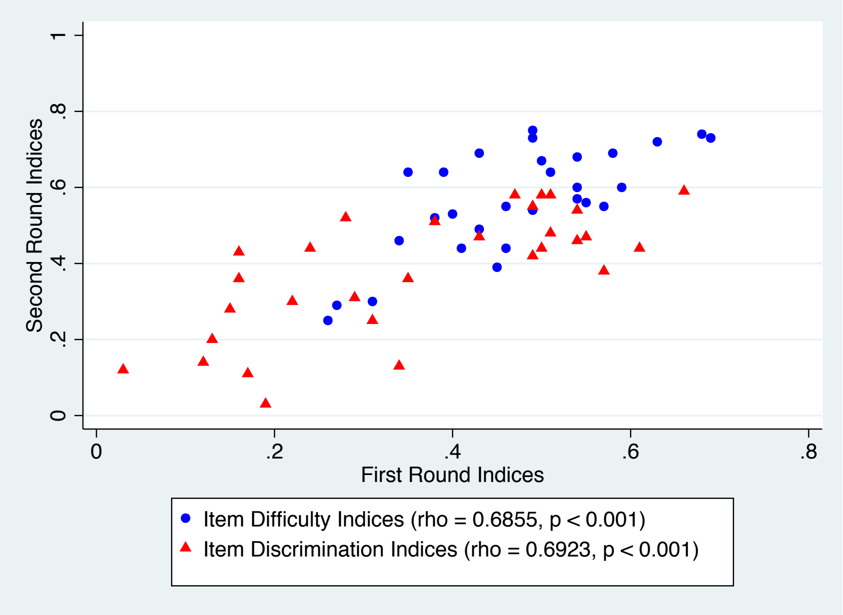


Figure 16. Correlations between item analysis parameters from each run of the instrument for each of the 30 items with students coming from Chem E.

All of these tests indicate good reliability of the instrument as far as testing student understanding of limiting and excess reagents, and yield using PNOM diagrams is concerned. Internal consistency was seen to be good using students from all courses with all Cronbach's α values at or above the 0.7 threshold for low-stakes testing. Correlations among item analysis parameters were seen to be stronger among Chem A and Chem D students compared to those coming from Chem E. That item analysis parameters from Chem E did not correlate as well was likely because these students represent a much smaller and more homogeneous segment of the student populations in terms of their mastery of the concepts covered by the instrument.

Stoichiometry Misconceptions and Student Response Patterns

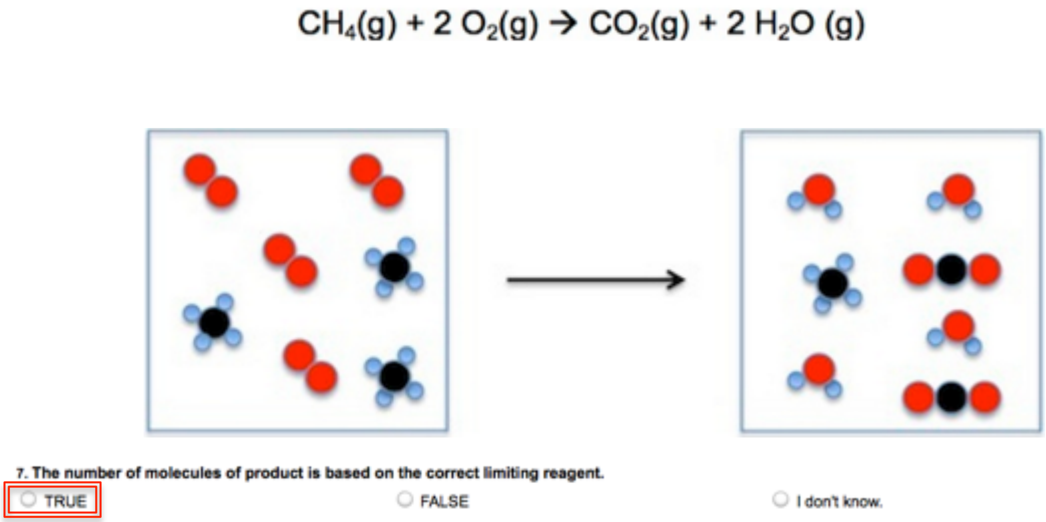
Responses to items on limiting reagents

Items 7, 9, 16, 20, 24, and 31 asked students to determine whether the number of product molecules in each diagram was based on the correct limiting reagent. Difficulty indices for each of these items after both rounds of testing with each course are listed in Table 19. Diagrams corresponding to items 7 and 20 were drawn correctly, while the rest were all incorrect.

Table 19. Difficulty indices for items pertaining to the choice of the limiting reagent for each diagram after both rounds of testing with students from each course.

Group	Round	Item Difficulty Index					
		Item 7	Item 9	Item 16	Item 20	Item 24	Item 31
Chem A	1	0.63	0.41	0.27	0.46	0.68	0.59
	2	0.72	0.44	0.29	0.55	0.74	0.60
Chem D	1	0.65	0.38	0.26	0.52	0.64	0.61
	2	0.65	0.42	0.29	0.52	0.66	0.56
Chem E	1	0.62	0.38	0.44	0.53	0.82	0.74
	2	0.94	0.48	0.42	0.61	0.82	0.76

Based on responses for items 7 (Figure 17) and 9 (Figure 18), the largest groups of students, aside from those who had given the correct response combination to these items, were those who gave correct answers with respect to the correctly drawn diagram but had more difficulty with respect to the incorrect diagram (TT in Figure 19). At least 30% from each student group appeared to have chosen methane as their limiting reagent for item 9, even after they had already correctly picked oxygen gas in item 7. This points to the resilience of the *least amount* misconception. It also points to how dual processing can sometimes fail to lead students to correctly view the diagrams. These students analyzed diagrams independently even though both of these diagrams pertained to the same initial gas mixture for the combustion of methane. Determining that the diagram associated with item 7 was correct should have led students to think that the diagram for item 9 did not use oxygen gas as its limiting reagent. It is very possible, however, that students who picked “True” for item 9 relied more on having seen less methane molecules than there were of oxygen gas in the diagram, and then seeing all of the methane being transformed into carbon dioxide. This they did even after they went through analyzing, probably even more closely, the diagram given for item 7 to come up with their “True” response for this item as well.



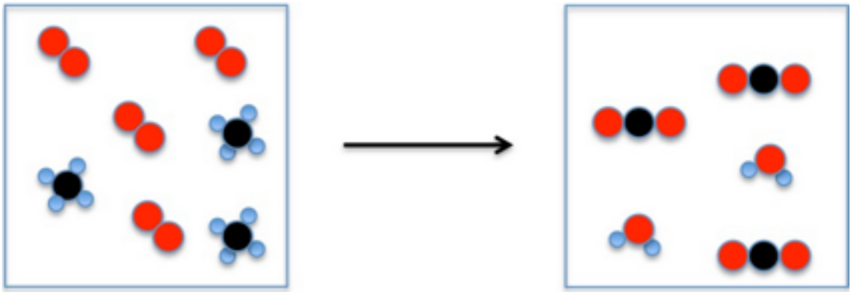
$CH_4(g) + 2\ O_2(g) \rightarrow CO_2(g) + 2\ H_2O\ (g)$

7. The number of molecules of product is based on the correct limiting reagent.

○ TRUE ○ FALSE ○ I don't know.

Figure 17. Item 7.

$CH_4(g) + 2\ O_2(g) \rightarrow CO_2(g) + 2\ H_2O\ (g)$



9. The number of molecules of product is based on the correct limiting reagent.

○ TRUE [✓ ○ FALSE] ○ I don't know.

Figure 18. Item 9.

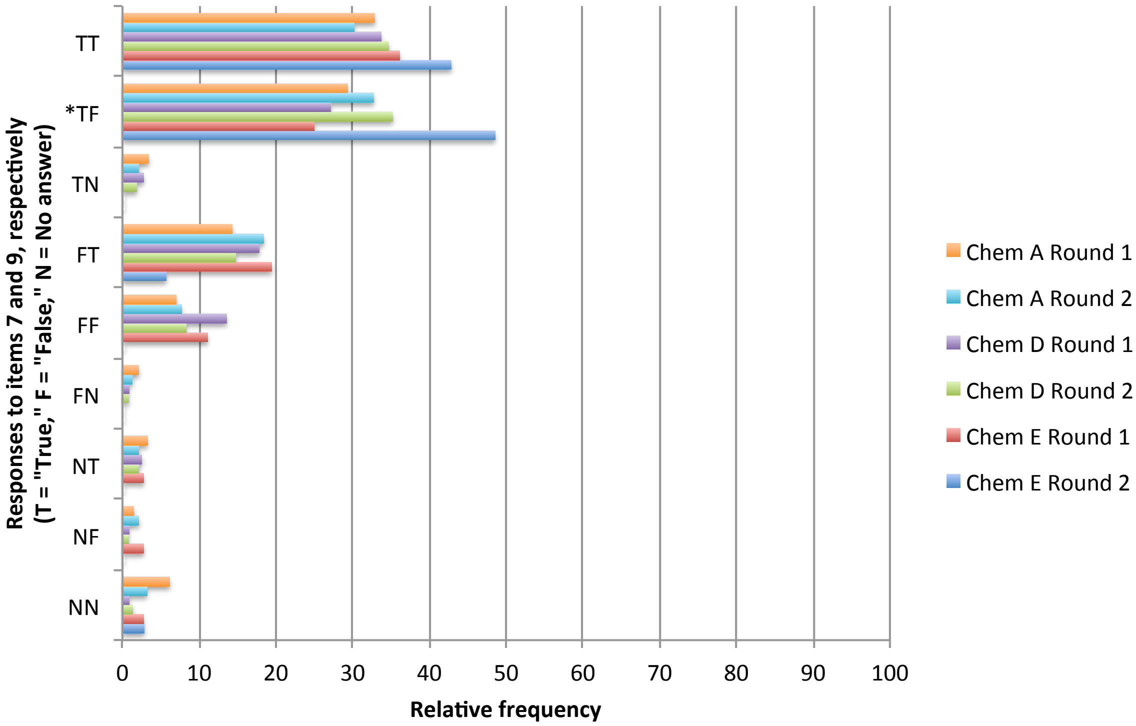
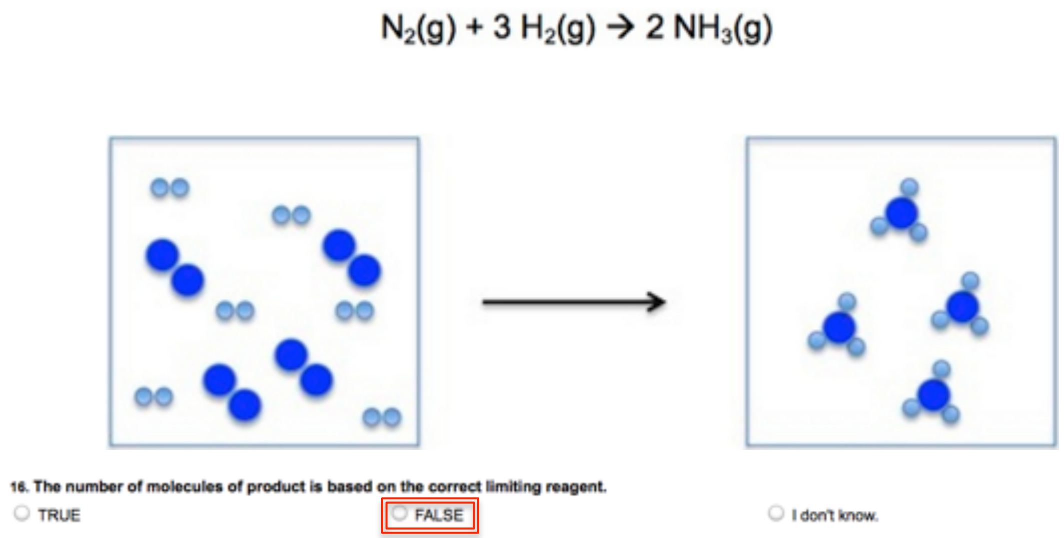


Figure 19. Response distribution between items 7 and 9 for students from all courses after both rounds of testing.

Items 16 (Figure 20) and 20 (Figure 21) produced slightly different response patterns from those between items 7 and 9. More students chose nitrogen as the limiting reagent for both items. The ‘least amount’ misconception and item order may have been factors here. A smaller proportion of students from all groups responded correctly to item 16 possibly because the incorrect diagram was presented first for the ammonia reaction. For example, about 60% of Chem D students correctly chose item 7 to be true about the correct methane combustion diagram, but only 25% could tell that the diagram for item 16 used the wrong limiting reagent during the first round of testing.

Students who picked “True” for item 16 split into subgroups of similar sizes in selecting hydrogen gas to be the limiting reagent in item 20. This gives evidence to the robustness of the dual thinking process theory so that it can also be observed when response patterns are summed over groups of students. All of these students relied on the least amount heuristic in choosing nitrogen as the limiting reagent for item 16. Yet half of them went on to slowly analyze the diagram for item 20 and ended up choosing hydrogen as limiting. It is quite possible that these students treated the two diagrams for the ammonia reaction independently of each other. This, however, was not the case among students who picked “False” for item 16. Among this group, more students picked “True” for item 20 regardless of the course they came from.

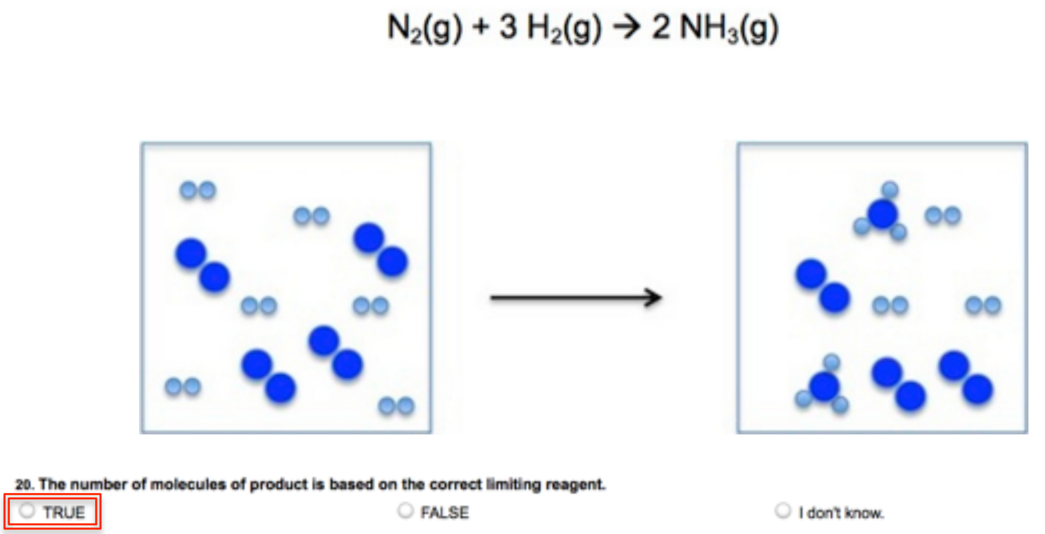


$N_2(g) + 3\ H_2(g) \rightarrow 2\ NH_3(g)$

16. The number of molecules of product is based on the correct limiting reagent.

○ TRUE ○ FALSE ○ I don't know.

Figure 20. Item 16.



$N_2(g) + 3\ H_2(g) \rightarrow 2\ NH_3(g)$

20. The number of molecules of product is based on the correct limiting reagent.

○ TRUE ○ FALSE ○ I don't know.

Figure 21. Item 20.

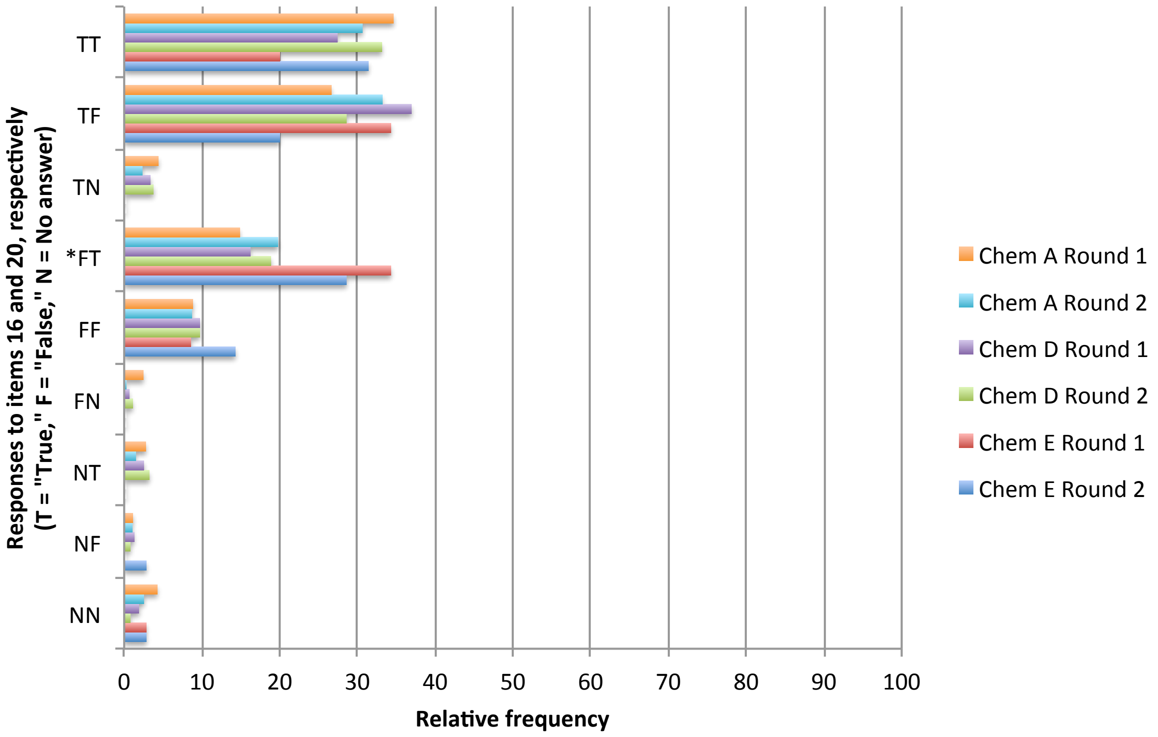
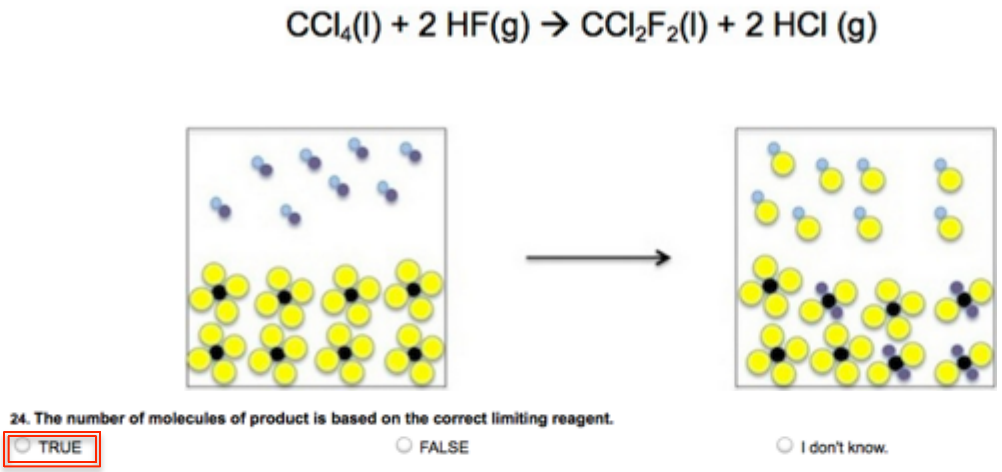


Figure 22. Response distribution between items 16 and 20 from each group of students after both rounds of testing.

With items 24 (Figure 23) and 31 (Figure 24), the *least amount* misconception could not be used since equal numbers of reactant molecules were drawn in the diagram. Most students from all of the groups picked the correct answers (Figure 25) to both items, TT. This seems to be an effective way of persuading students to actually count off the reactant molecules and determine how many product molecules could be obtained from each. The usual even split between responses for the second diagram in the pair in choosing the limiting reagent did not occur for these two items. There were definitely more students who picked “True” than “False” for item 31 after they have also picked “True” for item 24.

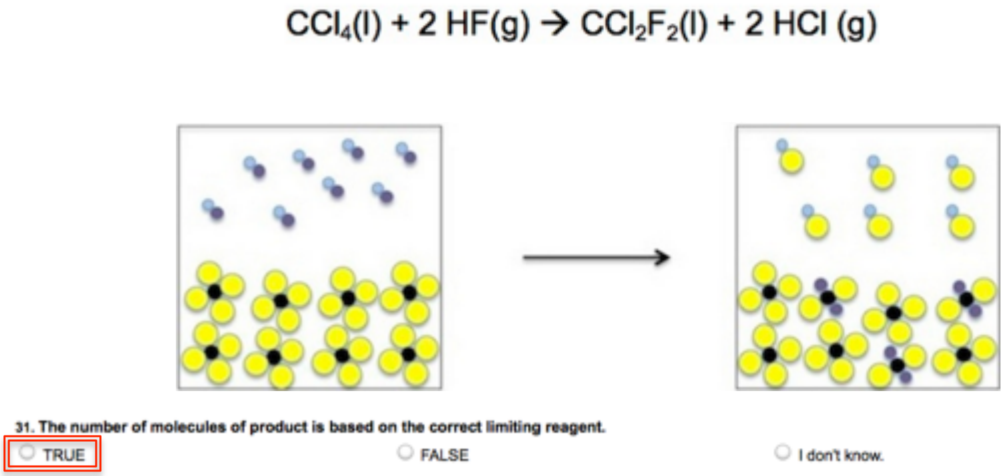


$CCl_4(l) + 2\ HF(g) \rightarrow CCl_2F_2(l) + 2\ HCl\ (g)$

24. The number of molecules of product is based on the correct limiting reagent.

○ TRUE ○ FALSE ○ I don't know.

Figure 23. Item 24.



$CCl_4(l) + 2\ HF(g) \rightarrow CCl_2F_2(l) + 2\ HCl\ (g)$

31. The number of molecules of product is based on the correct limiting reagent.

○ TRUE ○ FALSE ○ I don't know.

Figure 24. Item 31.

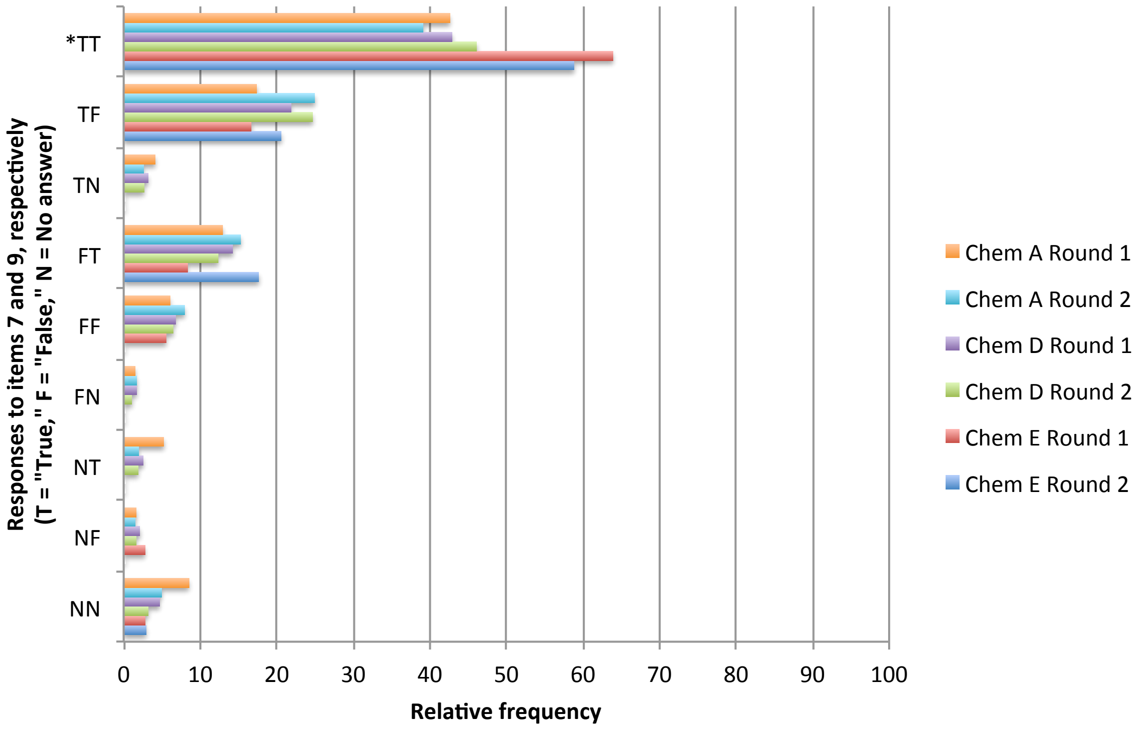


Figure 25. Response distribution between items 24 and 31 for all groups of students after both rounds of testing.

Responses to items on unreacted molecules

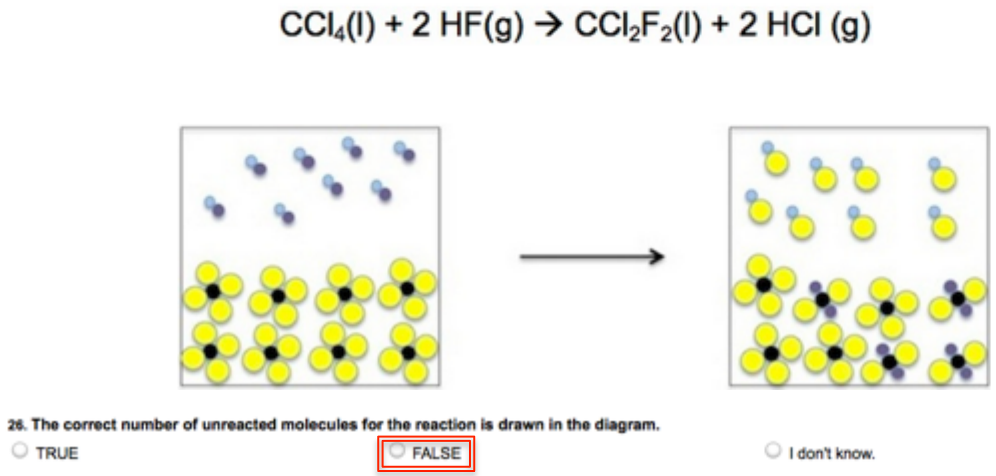
Items 5, 12, 14, 21, 26, and 32 all asked students to determine whether the correct number of unreacted molecules was drawn for each diagram. Difficulty indices for each of these items after both rounds of testing with each course are listed Table 20. Items associated with each of the first four diagrams generally had majority of students in each group getting them correctly. Performances on these items did not appear to have been affected by the accuracy of the associated diagram. The low difficulty indices for items 26 and 32 across all groups of students are quite interesting given that diagrams associated with these items did not show any unreacted molecules.

Table 20. Difficulty indices for items pertaining to the number of unreacted molecules drawn for each diagram after both rounds of testing with students from each course.

Group	Round	Item Difficulty Index					
		Item 5	Item 12	Item 14	Item 21	Item 26	Item 32
Chem A	1	0.61	0.63	0.68	0.51	0.28	0.39
	2	0.63	0.61	0.66	0.53	0.31	0.47
Chem D	1	0.49	0.58	0.49	0.57	0.31	0.43
	2	0.75	0.69	0.73	0.55	0.30	0.49
Chem E	1	0.53	0.62	0.47	0.71	0.29	0.65
	2	0.82	0.82	0.91	0.64	0.27	0.48

It is possible that many students neglected to take into account the percent yield given in the problem in selecting a response for item 26 (Figure 26). Students most likely saw that all atoms of each element in the reactant side of the diagram had been accounted for in the product side and used this conservation of atoms to justify their answers. Item 32 (Figure 27), however, was slightly different. The product side of the diagram took into account the 75% yield given in the problem but lacked unreacted hydrogen fluoride molecules. This seems to have caught some students' attention based on this item's higher difficulty indices compared to those of item 26. Still, at least half of students failed to notice the missing molecules after the second round of testing.

Students from all groups seem to have aligned their responses to items on unreacted molecules with those they gave about limiting reagents. For example, for items 5 and 7, "True" answers for item 7 were mostly matched up with "True" answers for item 5. "False" responses were also mostly aligned with each other. For students from Chem A, this was observed for about 62% of students after the first round of testing and for about 72% after the second round (Figure 28). This indicates that most students were probably treating limiting and excess reagents as parts of the same stoichiometry concept. This response pattern was observed among all groups of students and across all chemical contexts.

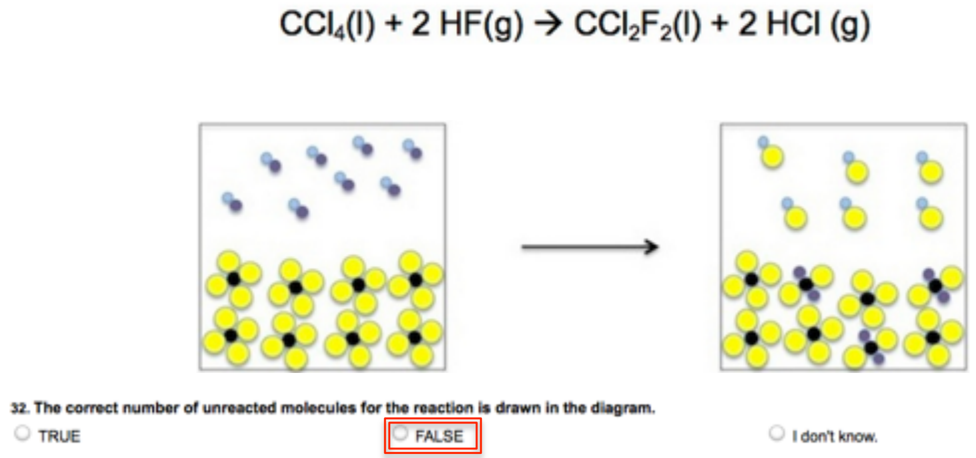


$CCl_4(l) + 2\ HF(g) \rightarrow CCl_2F_2(l) + 2\ HCl\ (g)$

26. The correct number of unreacted molecules for the reaction is drawn in the diagram.

○ TRUE ○ FALSE ○ I don't know.

Figure 26. Item 26.



$CCl_4(l) + 2\ HF(g) \rightarrow CCl_2F_2(l) + 2\ HCl\ (g)$

32. The correct number of unreacted molecules for the reaction is drawn in the diagram.

○ TRUE ○ FALSE ○ I don't know.

Figure 27. Item 32.

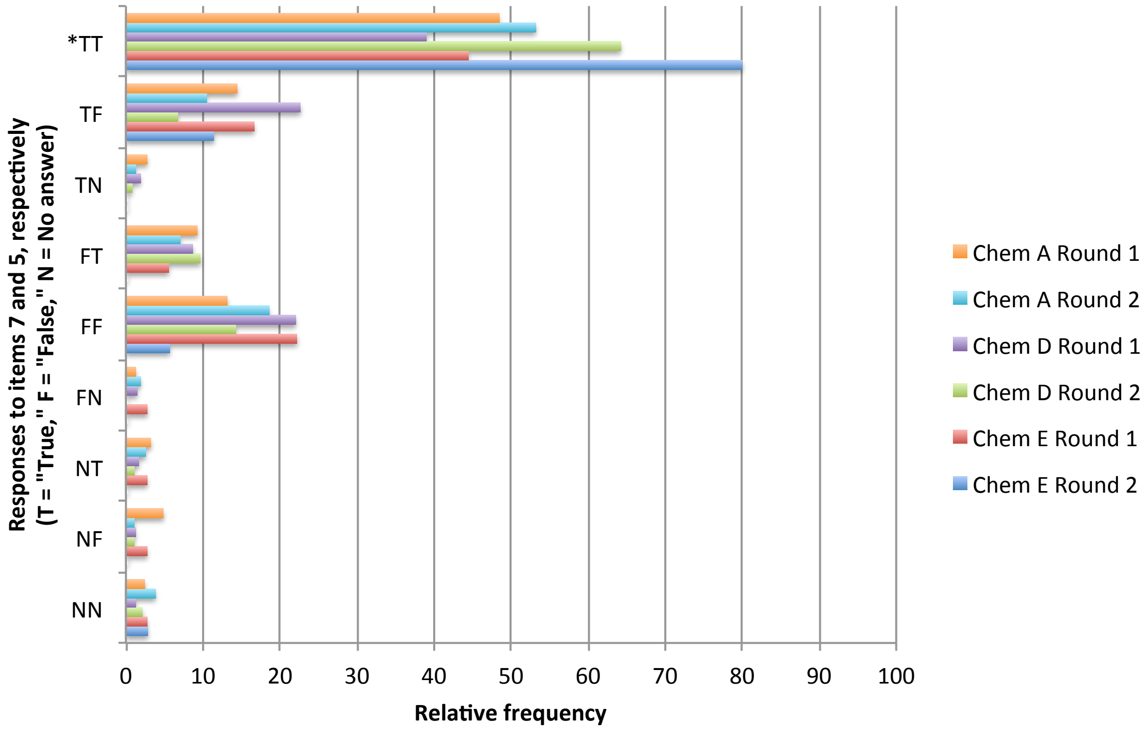


Figure 28. Response distribution between items 7 and 5 from each group of students after both rounds of testing.

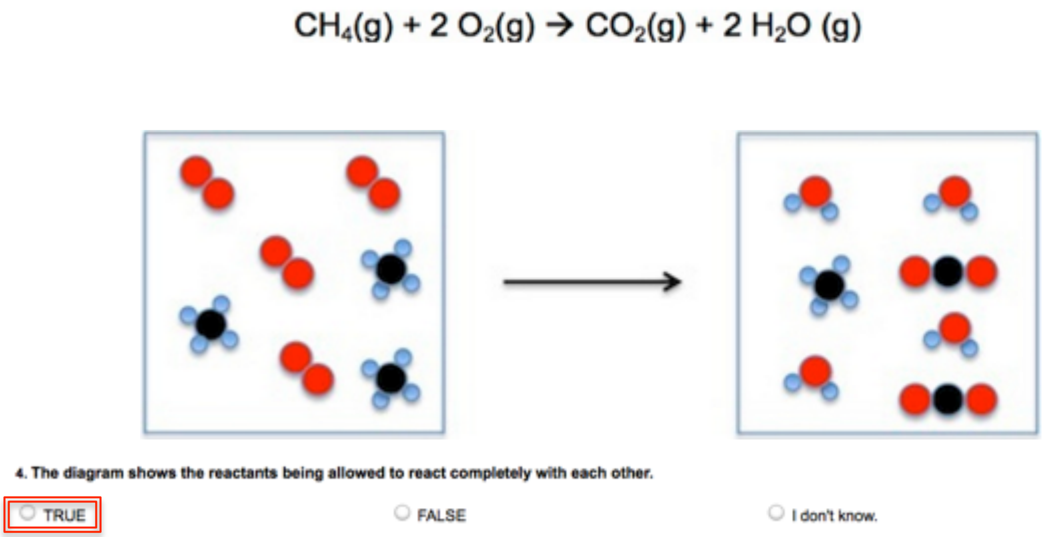
Responses to items on percent yield

Items 4, 11, 15, 22, 25, and 29 all pertained to determining whether each diagram reflected the given percent completion for the problem. Difficulty indices for these items are listed Table 21. Difficulty indices of items corresponding to correctly drawn diagrams (items 4 and 22) had greater increases between rounds of testing, especially among students from Chem D and Chem E.

Table 21. Difficulty indices for items pertaining to whether each diagram reflected the given percent completion in the problem after both rounds of testing with students from each course.

Group	Round	Item Difficulty Index					
		Item 4	Item 11	Item 15	Item 22	Item 25	Item 29
Chem A	1	0.53	0.34	0.57	0.48	0.49	0.48
	2	0.58	0.38	0.61	0.53	0.51	0.53
Chem D	1	0.35	0.45	0.39	0.40	0.51	0.49
	2	0.64	0.39	0.64	0.53	0.64	0.54
Chem E	1	0.32	0.44	0.41	0.56	0.71	0.53
	2	0.67	0.39	0.88	0.70	0.67	0.64

Majority of responses to item 4 (Figure 29) went in the same direction as those of item 7. “True” answers on completeness of the reaction mostly came with “True” answers on the limiting reagent. The same could be said about “False” answers. Between these two items, combining TT and FF responses accounts for at least half of each of the students groups, with second round responses aligning around 70% of the time. This indicates the strength with which students associated the relationship between what they determined to be the limiting reagent and the accuracy of the depiction of the extent of the reaction by each diagram. Students who picked “False” for item 4 used the unreacted methane molecule as a cue that the reaction did not go to completion. These response patterns were repeated between items 9 and 11. This time many students who saw that all of the methane had been converted to carbon dioxide thought that the diagram illustrated completeness of the reaction. For many students then the completeness of a reaction should be tied with the amount of the limiting reagent that proceeds to react. This strong connection between limiting reagents and reaction yield holds even among students who picked the wrong limiting reagent.

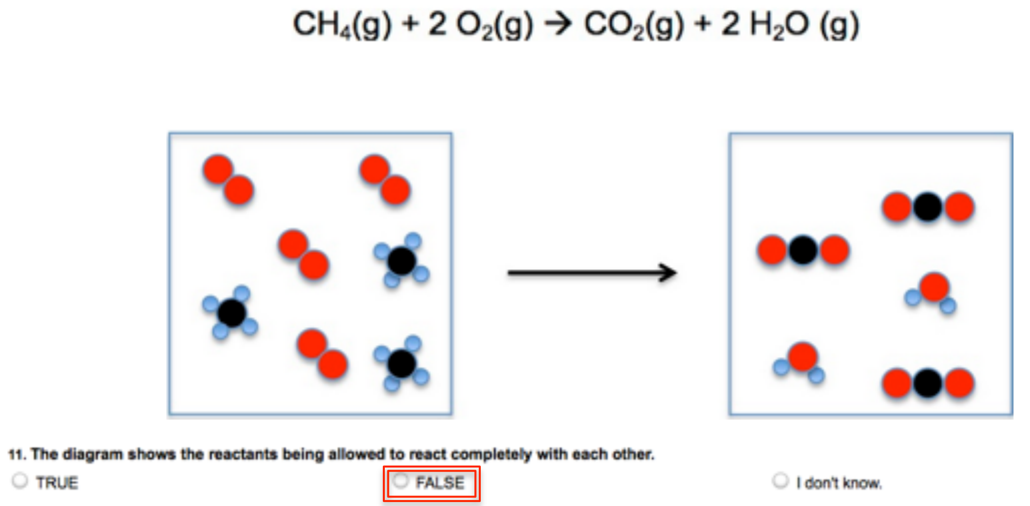


$CH_4(g) + 2\ O_2(g) \rightarrow CO_2(g) + 2\ H_2O\ (g)$

4. The diagram shows the reactants being allowed to react completely with each other.

[✓ ○ TRUE] ○ FALSE ○ I don't know.

Figure 29. Item 4.



$CH_4(g) + 2\ O_2(g) \rightarrow CO_2(g) + 2\ H_2O\ (g)$

11. The diagram shows the reactants being allowed to react completely with each other.

○ TRUE [✓ ○ FALSE] ○ I don't know.

Figure 30. Item 11.

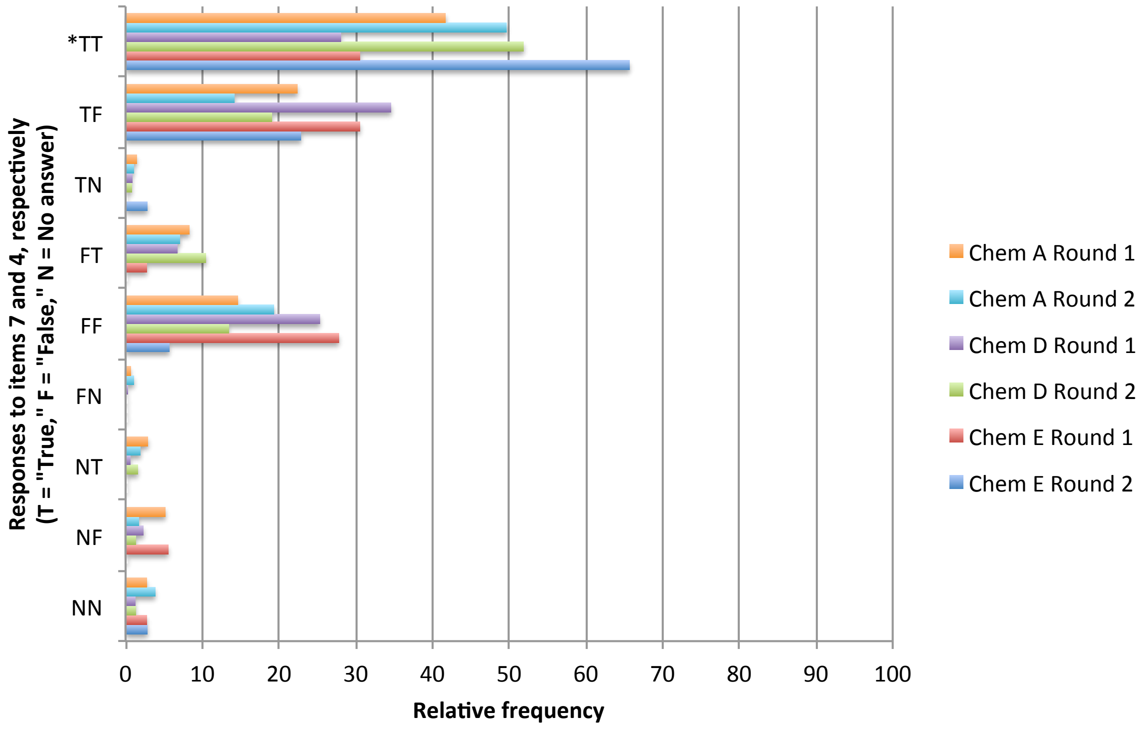


Figure 31. Response distribution between items 7 and 4 for all groups of students after both rounds of testing.

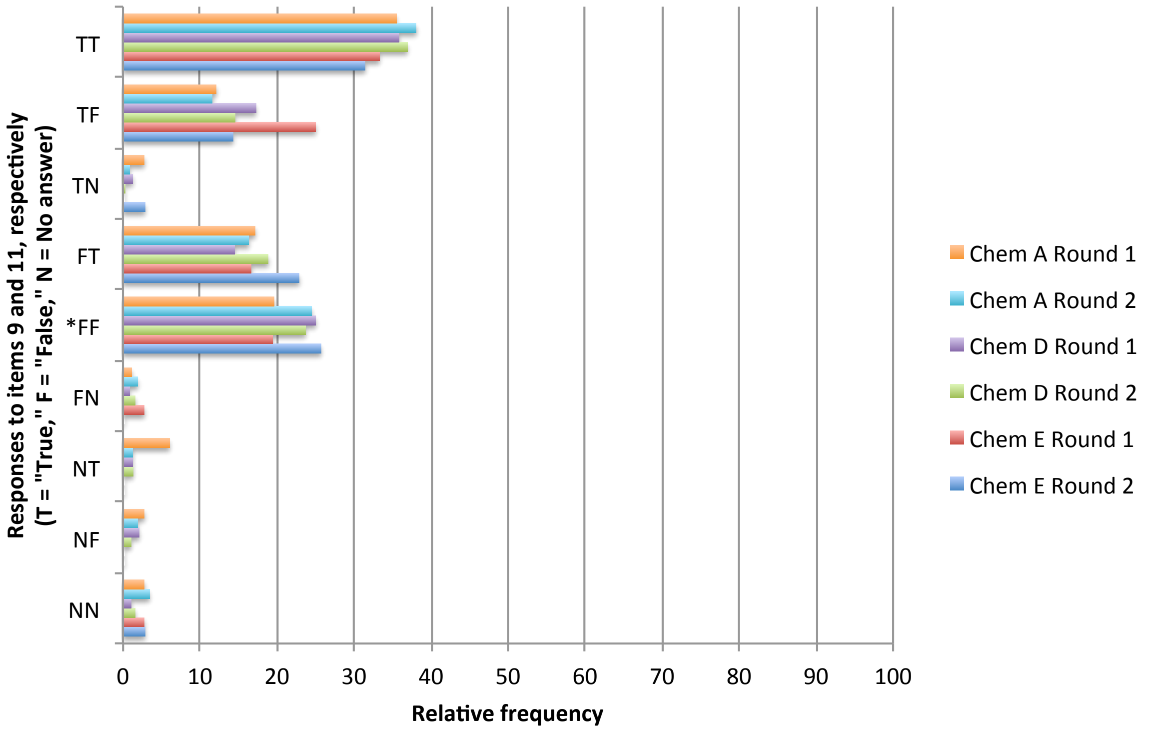
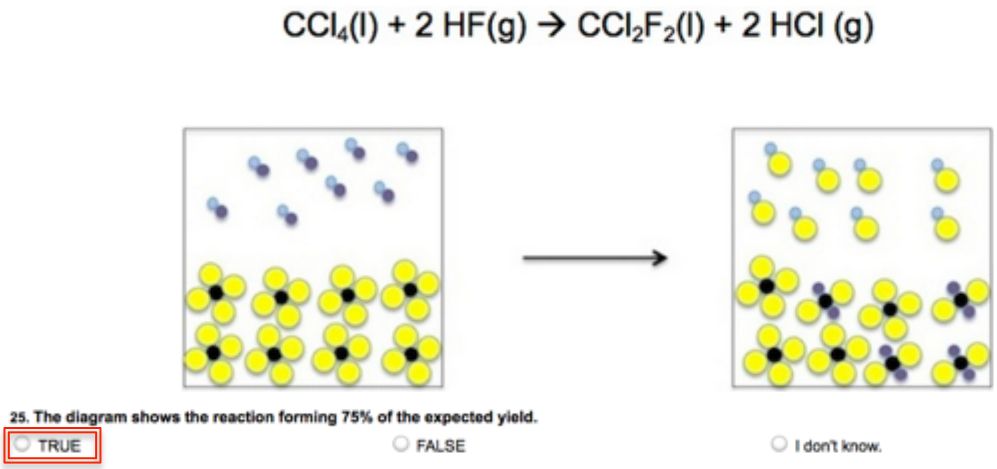


Figure 32. Response distribution between items 9 and 11 for all groups of students after both rounds of testing.

Item 25 (Figure 33) presents a unique situation among items dealing with percent yield. Students from Chem A were almost evenly split between giving aligned (TT or FF) or opposing answers (TF or FT) to items 25 and 24 (the limiting reagent question) during both testing rounds. Chem D students went from being evenly divided during the first round to mostly having opposing answers after the second round. Among Chem D students who saw this diagram as being drawn based on the correct limiting reagent (hydrogen fluoride), the fraction of students who said that it incorrectly illustrated the percent yield given in the problem grew from about 36 to 49 percent between rounds of testing. Students from Chem E were more decisive in giving the TF response combination for items 24 and 25, respectively. Item 24 has actually been coded as a false statement because the diagram illustrates a reaction that went to completion even though the problem requires 75% yield. Students might, however, argue that by defining the limiting reagent as the substance that is completely consumed by a reaction and thus, item 24 may be accepted to be true. This probably explains why depending on the group of students and round of testing, 60-76% of students chose “True” for item 24. On item 25 students are reminded to think about the given percent yield in the problem. By thinking about the question of yield separately from that of the limiting reagent with respect to this diagram, students from Chem A become split in answering item 25. This does not seem to be the case among Chem E students.



$CCl_4(l) + 2\ HF(g) \rightarrow CCl_2F_2(l) + 2\ HCl\ (g)$

25. The diagram shows the reaction forming 75% of the expected yield.

○ TRUE ○ FALSE ○ I don't know.

Figure 33. Item 25.

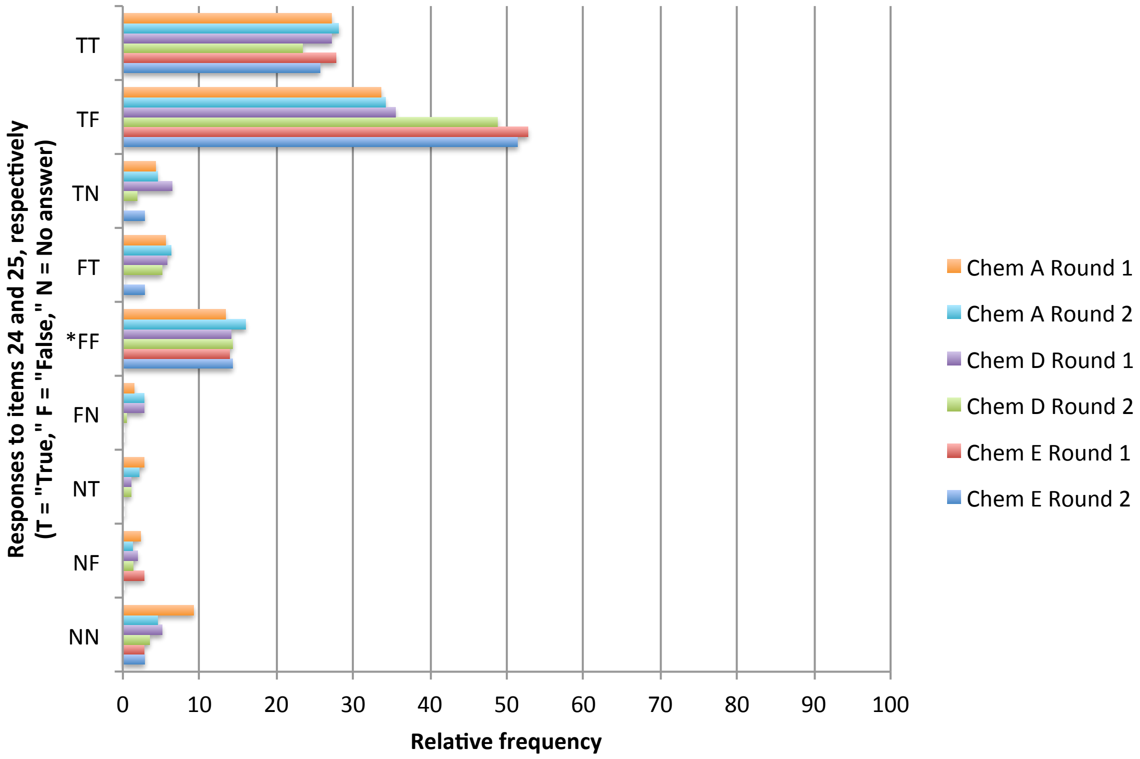


Figure 34. Response distribution between items 24 and 25 for all groups of students after both rounds of testing.

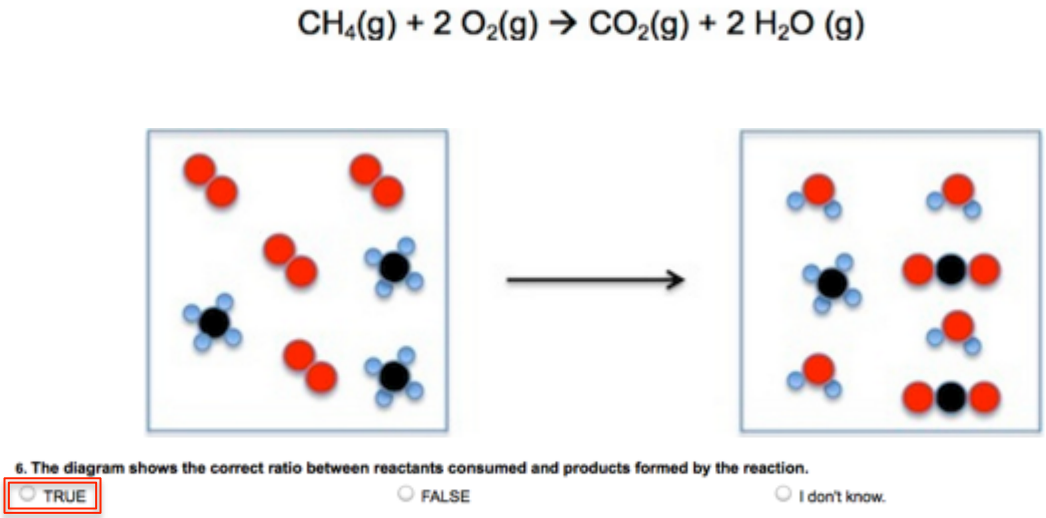
Responses to items on ratios between reactants consumed and products formed

Items 6, 10, 17, 19, 27, and 30 asked students to determine whether each diagram reflected the correct ratio between reactant molecules that have been consumed by the reaction and the product molecules formed given the percent yield. Difficulty indices for these items are listed in Table 22. It appears from these numbers that student performance on these items were affected by factors such as the correctness of the diagram associated with the problem and whether the reaction went to completion or not. Higher difficulty indices were obtained from items associated with correct diagrams (items 6 and 19) than with incorrect diagrams. Items that required students to consider percent yields other than 100 had lower difficulty indices, especially after the second round of testing.

Table 22. Difficulty indices for items pertaining to the ratio between the number of reactant molecules consumed by each reaction and the number of product molecules formed.

Group	Round	Item Difficulty Index					
		Item 6	Item 10	Item 17	Item 19	Item 27	Item 30
Chem A	1	0.59	0.61	0.44	0.54	0.26	0.57
	2	0.61	0.59	0.41	0.53	0.30	0.57
Chem D	1	0.69	0.54	0.34	0.54	0.26	0.54
	2	0.73	0.68	0.46	0.60	0.25	0.57
Chem E	1	0.76	0.68	0.41	0.56	0.21	0.62
	2	0.91	0.85	0.64	0.67	0.30	0.67

The highest difficulty indices on these items came from item 6 (Figure 35), which had a correct and students did not have to think about percent yields since the reaction went to completion. Looking at item response distributions shows that among students from Chem A, TT and FF response combinations between items 6 and 7 added up to about 73.5% after the second round of testing (Figure 36) while only 65.5% did the same between items 6 and 4 (Figure 37). This means that these students more strongly tied their response in item 6 to their choice of the limiting reagent, item 7, than to whether the reaction went to completion, item 4. . These response patterns were observed among other groups of students as well.



$CH_4(g) + 2\ O_2(g) \rightarrow CO_2(g) + 2\ H_2O\ (g)$

6. The diagram shows the correct ratio between reactants consumed and products formed by the reaction.

○ TRUE ○ FALSE ○ I don't know.

Figure 35. Item 6.

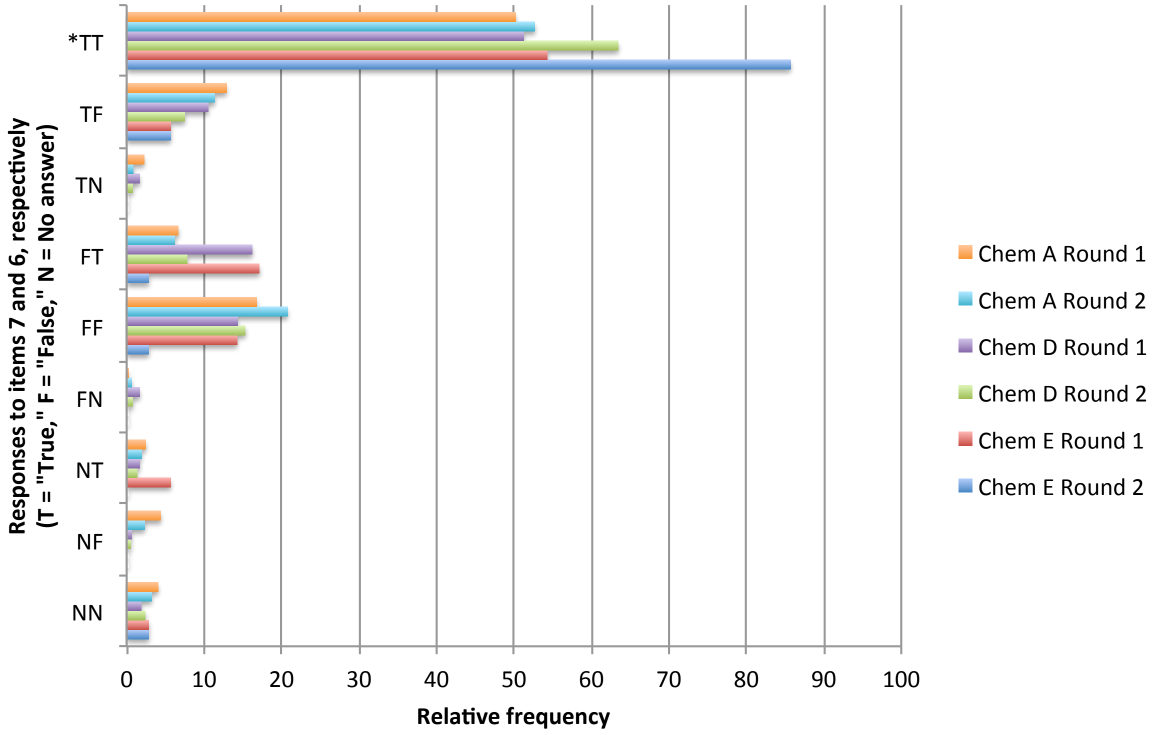


Figure 36. Response distribution between items 7 and 6 for all groups of students after both rounds of testing.

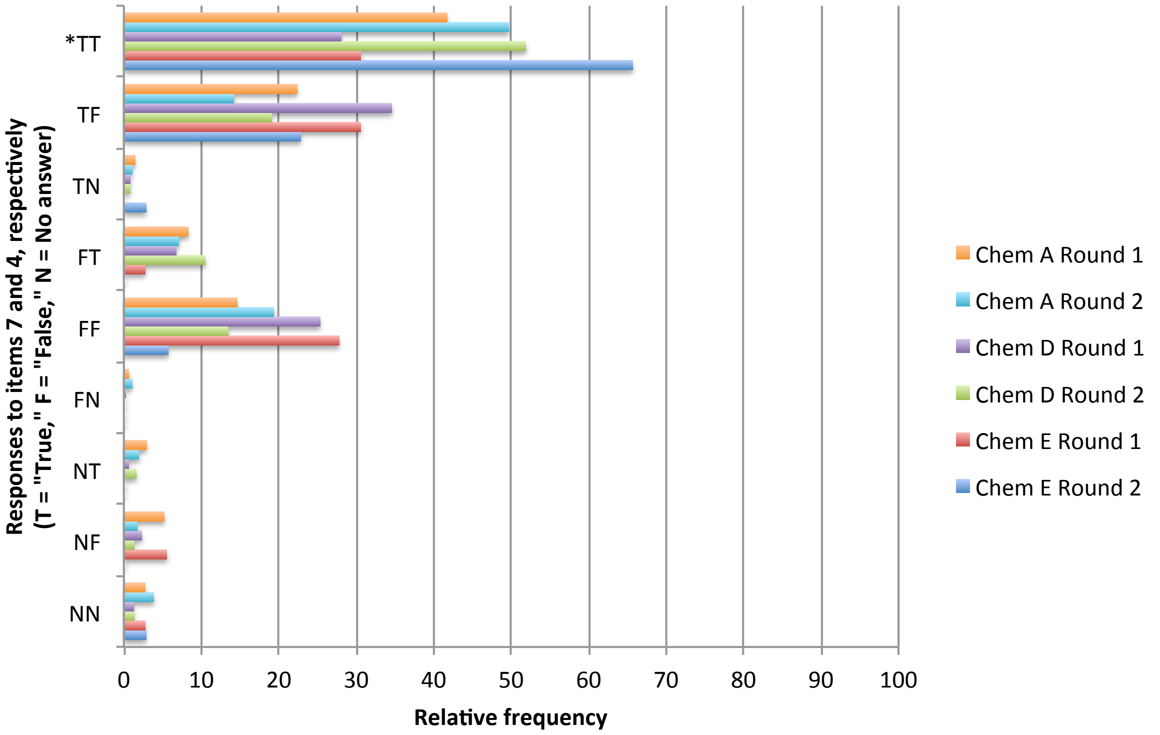
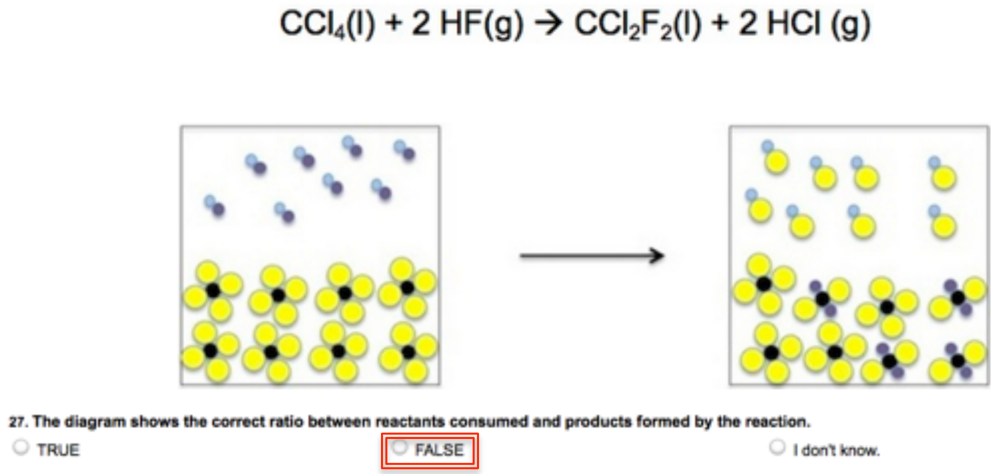


Figure 37. Response distribution between items 4 and 6 for all groups of students after both rounds of testing.

Students from all groups found item 27 (Figure 38) to be the most difficult among items that dealt with ratios between reactants used and products formed. This item dealt with an incorrect diagram that also had the highest number of molecules and highest number of different elements among those included in the instrument, and the problem asked students to think about a 75% yield for the reaction. Responses to this item were strongly aligned with answers given to item 24 (Figure 39), which was the limiting reagent item for this diagram. Alignment with responses to item 24 ranged from 72.1 to 77.2 percent after the second round of testing. On the other hand, response alignment with the percent yield item (Figure 40), item 25, went only from 42.9 to 47.9 percent.



$CCl_4(l) + 2\,HF(g) \rightarrow CCl_2F_2(l) + 2\,HCl\,(g)$

27. The diagram shows the correct ratio between reactants consumed and products formed by the reaction.

○ TRUE [✓ ○ FALSE] ○ I don't know.

Figure 38. Item 27.

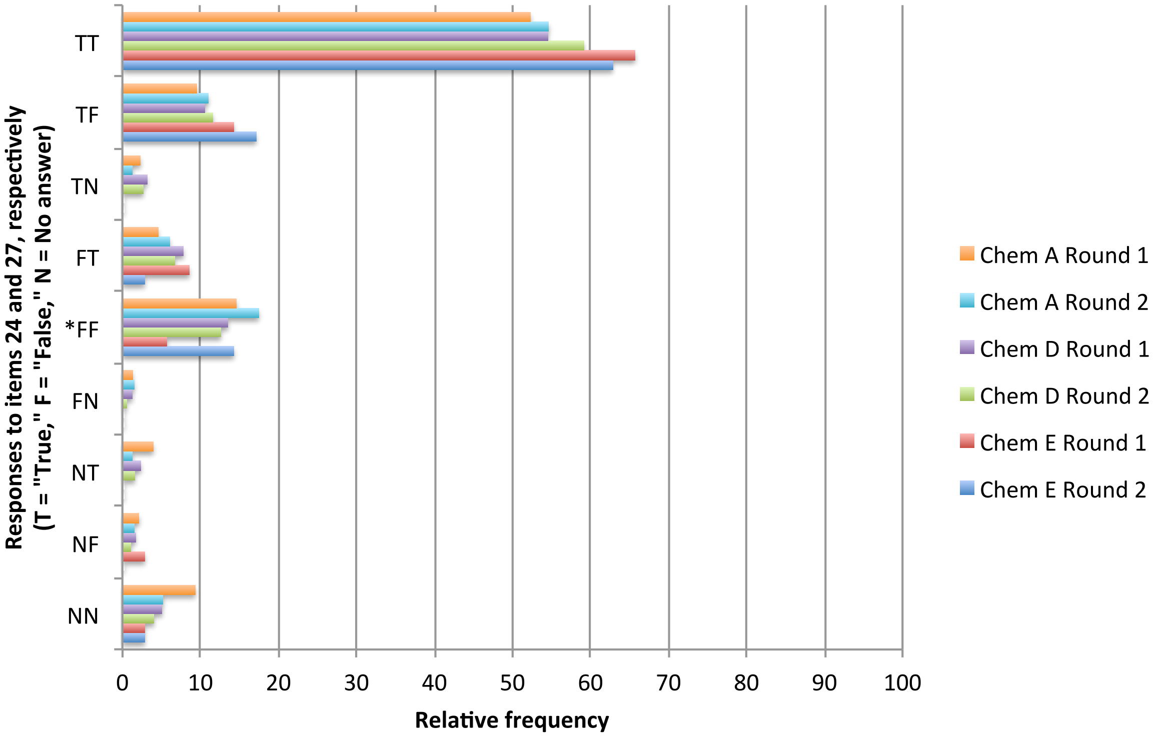


Figure 39. Response distribution between items 24 and 27 for all groups of students after both rounds of testing.

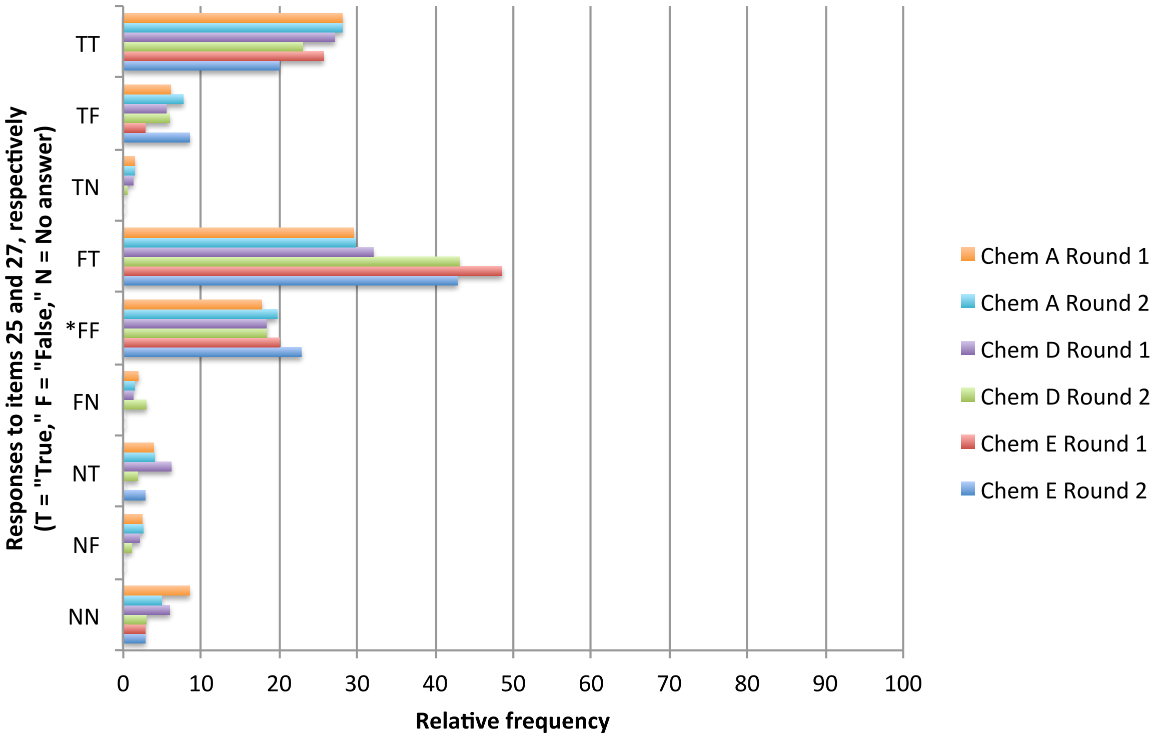


Figure 40. Response distribution between items 25 and 27 for all groups of students after both rounds of testing.

It appears students were consistently relying more strongly on their choice of the limiting reagent than on the percent yield for the reaction in checking the ratio between reactants and products. Even among students who correctly picked “False” for item 25, students were split 1:2 in favor of those picked “True” for item 27. On the other hand, most students who picked “False” for item 24 also ended up picking “False” for item 27. While the concept of limiting reagents is important in general chemistry, these results give a good example of how students can sometimes misplace the priorities they assign to different aspects of a PNOM diagram.

Summary and Implications

This chapter described patterns of responses from three different groups of students to conceptual questions that used PNOM diagrams in relation to the concepts of excess and limiting reagents, and yield. It also identified misconceptions that can be inferred from those response patterns and described how prevalent among general chemistry students these misconceptions are.

Validity of the 30-item instrument was established through consultation with faculty members in charge of courses from which the student participants for this study were taken. Furthermore, reliability of the instrument was established using Cronbach’s α, correlations between item analysis parameters determined from two rounds of testing, and stabilities of student scores between testing and retesting rounds for different groups of students.

Response patterns observed with the use of the instrument revealed common misconceptions among students from all groups. Misconceptions identified included the selection of the limiting reagent in a chemical reaction as the substance present in the least amount, failure to connect responses to items on limiting and excess reagents to those that asked about percent yields correctly, failure to account for unreacted molecules correctly, and neglect of the given percent yield to determine the correct number of product molecules formed. In some instances, response patterns that indicate how some students mistakenly evaluated items pertaining to the same chemical context independently of each other were identified.

Since four different versions of the instrument were implemented with different arrangements of the contexts, ordering effects for items within specific chemical contexts may need to be explored. For instance, in contexts where items on limiting reagents came ahead of those on percent yields, responses to these questions seemed to have been less aligned with each other than those contexts in which the order of these items were reversed, that is, TT and FF response combinations appeared less frequently than TF or FT responses to these items did. On the other hand, having the item on the limiting reagent ahead of that for unreacted molecules seems to result in greater alignment between responses to these items. Future research should include validation interviews to help further understand students’ response patterns.

Instructors might help students overcome the ‘least amount’ misconception in determining the limiting reactant by explicitly illustrating how balanced equations may be used in conjunction with PNOM diagrams to determine exactly how much of each product is formed, how much of each reactant is used, and what would be left behind. For instance, for gaseous reactions such those included in the instrument used for this study, reactant molecules may be counted off in sets defined by the coefficients in the balanced equation until students run out of one of the reactant molecules. These sets of reactant molecules may then be used to form products. Such an approach would also probably help students connect the microscopic and symbolic levels of representations.

Instructors also need to point out that PNOM diagrams used to illustrate chemical reactions must be balanced in terms of the numbers of atoms of each element in the diagram. In most instances, the associated mistake in neglecting to check atom balance in a PNOM diagram is the failure to account for unreacted molecules. In the same manner that chemical equations need to be balanced, PNOM diagrams must also be balanced to explicitly show observation of the law of conservation of mass between both sides of the diagram.

The percent yield of a reaction is just as important as the choice of the limiting reagent in determining the accuracy of a PNOM diagram. Some of the errors students made on the instrument might have been avoided if students had a better understanding of percent yield and how it relates to PNOM diagrams. Students need to keep in mind that (1) the percent yield is determined based on the initial amount of the limiting reagent (either in terms of how much of the limiting reagent was actually used or how much of the product was formed), (2) the amount of unreacted material is also an indication of the percent yield of the reaction. Instructors also need to state that even though items pertaining to a specific PNOM diagram on an assessment instrument are listed separately, students must address these questions in as integrative an approach as possible. This means that responses to items presented earlier must somehow be correctly related with answers to succeeding items, especially if these are all in reference to the same PNOM diagram.

In some instances, the correctness and visual complexity of the diagrams did contribute to the difficulties of some of the items included in the instrument. In teaching students how to interpret PNOM diagrams dealing with stoichiometry concepts, instructors might wish to start with reactions that involve fewer different kinds of atoms. PNOM diagrams that accurately depict specific chemical reactions and have relatively small numbers of molecules in them may also be a good place to start before progressing to more complex scenarios.

REFERENCES

Adams, W. K., & Wieman, C. E. (2011). Development and validation of instruments to measure learning of expert-like thinking. *International Journal of Science Education, 33*(9), 1289-1312.

Agung, S., & Schwartz, M. S. (2007). Students' understanding of conservation of matter, stoichiometry and balancing equations in Indonesia. *International Journal of Science Education, 29*(13), 1679-1702.

Ben-Zvi, R., Eylon, B., & Silberstein, J. (1987). Students' visualisation of a chemical reaction. *Educ. Chem., 24*, 117-120.

Ben-Zvi, R., Eylon, B., & Silberstein, J. (1988). Theories, principles and laws. *Education in chemistry, 25*(3), 89-92.

Bodner, G. M., & Herron, J. D. (2003). Problem-solving in chemistry *Chemical education: Towards research-based practice*.

BouJaoude, S., & Barakat, H. (2000). Secondary school students' difficulties with stoichiometry. *School Science Review, 81*(296), 91-98.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *psychometrika, 16*(3), 297-334.

Davidowitz, B., Chittleborough, G., & Murray, E. (2010). Student-generated submicro diagrams: A useful tool for teaching and learning chemical equations and stoichiometry. *Chemistry Education Research and Practice, 11*(3), 154-164.

de Astudillo, L. R., & Niaz, M. (1996). Reasoning strategies used by students to solve stoichiometry problems and its relationship to alternative conceptions, prior knowledge, and cognitive variables. *Journal of Science Education and Technology, 5*(2), 131-140.

Ding, L., & Beichner, R. (2009). Approaches to data analysis of multiple-choice questions. *Physical Review Special Topics-Physics Education Research, 5*(2), 020103.

Evans, J. S. B. T. (1996). Deciding before you think: Relevance and reasoning in the selection task. *British Journal of Psychology, 87*(2), 223-240.

Evans, J. S. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annu. Rev. Psychol., 59*, 255-278.

Feldt, L. S. (1963). Note on use of extreme criterion groups in item discrimination analysis. *Psychometrika, 28*(1), 97-104.

Gabel, D. (1998). The complexity of chemistry and implications for teaching. *International handbook of science education, 1*, 233-248.

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

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

Kern, A. L., Wood, N. B., Roehrig, G. H., & Nyachwaya, J. (2010). A qualitative report of the ways high school chemistry students attempt to represent a chemical reaction at the atomic/molecular level. *Chemistry Education Research and Practice, 11*(3), 165-172.

Kline, P. (2013). *Handbook of psychological testing*: Routledge.

Luxford, C. J., Linenberger, K. J., Raker, J. R., Baluyut, J. Y., Reed, J. J., De Silva, C., & Holme, T. A. (2014). Building a Database for the Historical Analysis of the General Chemistry Curriculum Using ACS General Chemistry Exams as Artifacts. *Journal of Chemical Education, 92*(2), 230-236.

Maeyer, J., & Talanquer, V. (2010). The role of intuitive heuristics in students' thinking: Ranking chemical substances. *Science Education, 94*(6), 963-984.

Nakhleh, M. B. (1993). Are our students conceptual thinkers or algorithmic problem solvers? Identifying conceptual students in general chemistry. *Journal of Chemical Education, 70*(1), 52.

Nakhleh, M. B., Lowrey, K. A., & Mitchell, R. C. (1996). Narrowing the gap between concepts and algorithms in freshman chemistry. *Journal of Chemical Education, 73*(8), 758.

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

Olmsted, J. (1999). Amounts tables as a diagnostic tool for flawed stoichiometric reasoning. *Journal of Chemical Education, 76*(1), 52.

Sanger, M. J. (2005). Evaluating students' conceptual understanding of balanced equations and stoichiometric ratios using a particulate drawing. *Journal of Chemical Education, 82*(1), 131.

Schroeder, J., Murphy, K. L., & Holme, T. A. (2012). Investigating factors that influence item performance on ACS exams. *Journal of Chemical Education, 89*(3), 346-350.

Treagust, D. F., & Chittleborough, G. (2001). Chemistry: A matter of understanding representations. *Advances in research on teaching, 8*, 239-268.

Upahi, J. E., & Olorundare, A. S. (2012). Difficulties faced by Nigerian senior school chemistry students in solving stoihiometric problems. *Journal of Education and Practice, 3*(12), 181-189.

CHAPTER 4. EYE TRACKING STUDENTS' VISUAL BEHAVIOR AS THEY SOLVE STOICHIOMETRY PROBLEMS USING PNOM DIAGRAMS

The coordination of symbolic and microscopic representations is an important skill to learn in chemistry. Microscopic representations, such as particulate nature of matter (PNOM) diagrams, are commonly used by chemists to explain many chemical phenomena. In the general chemistry classroom, instructors emphasize that an in-depth understanding of chemistry topics, such as stoichiometry, requires not only the ability to follow an algorithm, but also the skill to interpret symbols and explain phenomena at the microscopic level (Ben-Zvi, Eylon, & Silberstein, 1987). Previous studies, for example, have shown that while a majority of students in a first-year chemistry course can write balanced chemical equations accurately, barely more than one out of five can translate chemical reactions represented by PNOM diagrams into the corresponding chemical equation (Davidowitz, Chittleborough, & Murray, 2010). It appears that even though there has been great emphasis on the visual approach to learn chemistry using PNOM diagrams, many students still find it challenging to understand chemical stoichiometry at the microscopic level (Ben-Zvi et al., 1987; Sanger, 2005).

As mentioned in previous chapters, text and diagram representations are coded in different cognitive systems due to their different physical forms (Clark & Paivio, 1991). Thus, information displayed with both text and diagrams are expected to allow learners more cognitive elaborations than solely using text or pictures. Task performance has also been found to be better when learning occurs using both text and diagrams than learning from text or diagrams only (Hegarty, Carpenter, & Just, 1991). Students have been found to inspect diagrams after reading sections of the text as they attempted to select relevant aspects of words or images during task performance. They also tried to build coherent visual and verbal mental models from each representation and then integrated both mental models using their own prior knowledge to generate learning.

Subjects that have different problem-solving abilities usually exhibit different visual behaviors as manifested by differences in eye movement patterns (Rosengrant, 2010; Tang, Topczewski, Topczewski, & Pienta, 2012; Van Gog, Jarodzka, Scheiter, Gerjets, & Paas, 2009). The visualization difference between students with high and low prior knowledge most likely comes from long-term memory (Ericsson & Kintsch, 1995). Students with more prior knowledge have a large number of domain-specific schemas and, thus, may sometimes bypass working memory capacity limits since more of their schemas may have become automated (Kalyuga, Ayres, Chandler, & Sweller, 2003; Kalyuga, Chandler, & Sweller, 1998). Students with low prior knowledge, meanwhile, do not possess the same relevant schematic knowledge to perform tasks more accurately and more efficiently.

The decreased demand from working memory on the part of students with high prior knowledge often results in a greater ability to attend to domain-relevant information, faster processing of such information, and overall improvement of performance (Liu, Gale, & Song, 2007; Weber & Brewer, 2003). Eye tracking has been used to examine how students of different abilities allocated their attention to different parts of visual stimuli and to show that students' reading abilities, for example, may be inferred based on the different lengths of time they spent comprehending text materials (Schmidt-Weigand, Kohnert, & Glowalla, 2010). Students with more limited working memory capacities and less refined schemas, on the other hand, were found to process and integrate information across different representations less accurately and less efficiently (Kozma, 2003). In chemistry, experts have been found to coordinate information within and across representations while most students had difficulties (Kozma, 2003). Previous studies have shown that the number of transitions between regions of visual stimulus corresponding to text and visual representations may be considered as indicators of integrative effort on the part of the learner (Mason, Pluchino, Tornatora, & Ariasi, 2013; Schwonke, Berthold, & Renkl, 2009). Students with greater prior knowledge were found to have more frequent transitions between text and diagrams as they invested more mental effort to integrate information from both sources.

Chemistry education is a especially important domain as far as the visual coordination of representations, as influenced by prior knowledge, is concerned. Reasoning in chemistry often deals with unobservable concepts and processes, which is why the use of visualizations to relate information cannot be overemphasized. This study focused specifically on students' visual behaviors when presented problems on limiting and excess reagents, and yield that included PNOM diagrams. In particular, this study aimed to answer the following research questions:

1. How do students with different levels of prior knowledge divide their attention to text, symbolic and microscopic representations when solving conceptual problems dealing with limiting and excess reagents, and reaction yield? How are these manifested in terms of fixation times, fixation frequencies, and transitions between areas of interest?
2. How do students with different levels of prior knowledge integrate information from symbolic and microscopic representations in terms of frequencies of AOI transitions?

Participant Recruitment

Participants were recruited during the spring semester of 2014 from two different chemistry courses. One group of participants was recruited from those who were currently registered in Chem B, the second of a one-year course in general chemistry offered to physical and biological science majors, chemical engineering majors and other students who intend to register for 300-level courses. This course covers solution properties, kinetics, thermodynamics, electrochemistry, chemical equilibrium, and nuclear chemistry. Students from this course were chosen from among those who have taken Chem A, the first semester of the same one-year general chemistry course, during the previous fall semester and have participated in both round of the instrument survey described in Chapter 3. Participants from Chem B were invited to the study during the fourth and fifth weeks of the semester, and were given a free one-time access to the online general chemistry practice test of the American Chemical Society Examinations Institute. A second group of students were recruited from Chem C, a one-semester survey of chemical principles for those who are not physical and biological science or engineering majors. Topics discussed in Chem C usually include nomenclature, chemical reactions, stoichiometry, atomic structure, periodic properties, chemical bonding, states of matter, solutions, thermochemistry, acid-base theory, oxidation-reduction reactions, basic chemical kinetics, and chemical equilibrium. Participants from Chem C were recruited during the two weeks immediately following their in-course examination on stoichiometry, and were given extra course credit by their instructor. Students were recruited from these specific courses with the intent of building in some expert-novice differentiation among the study’s participants. It was originally hypothesized that because of their extent of their previous exposure to instrument used in this study and the more depth with which chemistry concepts are dealt with in the one-year course, students who came from Chem B will exhibit visual behaviors that are quite different from those coming from Chem C students.

A total of 15 students from Chem B and 14 students from Chem C participated in this study. Data from one student from each of Chem B and Chem C were eliminated due to failure of the eye tracker to capture their visual behavior during parts of their sessions, leaving data from 14 and 13 students, respectively, for analysis.

Collection of Eye Tracking Data

An SMI Red eye tracker with BeGaze 3.3 software was used to collect all experimental data. The eye tracker was place directly below a 23-inch LCD monitor, on which the questions for the study were displayed. All visual stimuli were maximized to occupy the full area of the monitor. As each student came in for their eye tracking session, they were given a brief set of instructions and asked to sign a consent form, indicating that they came to the study voluntarily and that they were at least 18 years of age. Each student was seated about 60 to 70 cm from the front of the monitor. A nine-point eye calibration with the eye tracker, followed by a five-point validation, was performed with each student. This calibration and validation routine was repeated for each student until eye tracking resolution came to within 0.5^{o} along both dimensions of the monitor.

The 30-item instrument administered online among a much larger group of students from different general chemistry courses during the previous semester was modified so that participants for the eye tracking study saw only one item each time along with the relevant balanced chemical equation and PNOM diagram. This was done so that eye tracking data captured for each participant pertained only to the specific item currently displayed on the monitor. Students from Chem C were also given a preview page for each chemical context that consisted only of the stem of the problem, the chemical equation, the description of the color scheme used in the diagram, and the diagram itself. This was done to isolate the time spent by participants tying together the correct colored sphere in each diagram with the correct element. In the process of initially evaluating data collected from Chem B students, it was determined that a large fraction of the time spent on the first item went to matching the correct colored sphere in the diagram with each element in the chemical equation. Thus, participants from Chem B saw five online pages for each chemical context given in the instrument, while those from Chem C saw six. No time limit was imposed on students in responding to the instrument, although it was observed that participants took between five and 25 minutes to go through all items.

After completing the instrument, each student was shown a playback of their gaze video, and asked to describe specific steps associated with how they looked at the different parts of the instrument for each item. For example, they were asked to describe what they were trying to do as they looked back and forth between the two parts of the chemical equation or the diagram, or as their eye fixations went up and down between the chemical equation and the diagram. They were also asked to use the gaze video to remind themselves about thought processes they went through as they responded to each item. These *retrospective think-aloud* (Holmqvist et al., 2011) sessions lasted between 35 and 50 minutes each.

Encoding of Visual Behavior Data

Visual behavior from each participant was encoded in terms of sequences of eye fixation data known as scan paths. Eye fixations were collected as participants viewed across different sections of the visual stimulus known as areas of interest (AOIs). AOIs are generally defined based on the specific type of information about the subjects' visual behaviors the researchers might be curious about (Holmqvist et al., 2011). Usually, AOIs are defined around some specific feature region of the image being studied. For example, AOIs on a facial image might include each eye, the nose, and the mouth of the person in the photo.

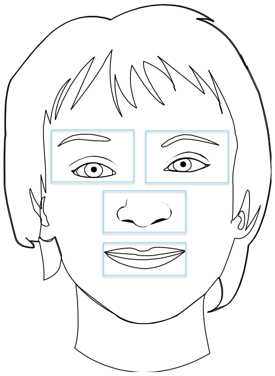


Figure 1. Facial areas of interest as may be defined by an eye tracking researcher ("Face Outline Templates," 2015).

In this specific study, AOIs were defined around each side of the balanced chemical equation, each side of the PNOM diagram given in each item, and on the question for each item. Strings consisting of characters denoting the different AOIs were written out to represent the sequence with which subjects viewed each AOI. Temporal binning was also incorporated into the AOI strings by repeating characters corresponding to each 25 ms of fixation on each AOI (Cristino, Mathôt, Theeuwes, & Gilchrist, 2010). Thus, the scan paths generated from eye tracking data included location, sequence, and durations of eye fixations on visual stimuli.

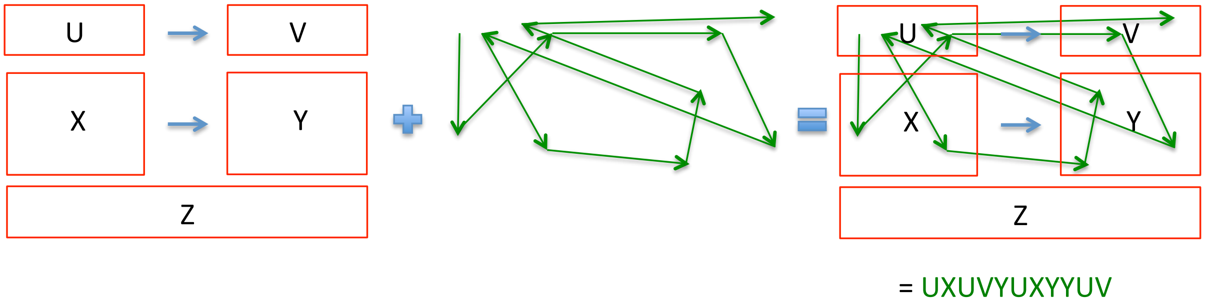


Figure 2. Schematic diagram of the coding procedure used to produce AOI strings from scan paths.

Sequence Alignment and the Needleman-Wunsch Algorithm

AOI strings were compared with each other pairwise to identify similarities and differences among the visual behaviors of subjects with respect to visual stimuli. The simplest of these string comparison methods is the *string edit method* (Levenshtein, 1966). The method defines the edit distance (also known as the Levenshtein distance) between two strings as the minimum cost of transforming one string into another in terms of numbers of insertions, deletions, and substitutions. Weights may be assigned to each operation (Okuda, Tanaka, & Kasai, 1976) or additional operations such as transpositions may be added (Wagner & Lowrance, 1975) in more advanced versions of the calculation. The simplicity of this method allows fast computations of the edit distance but fails to account for the nature of the relationship among AOIs. It cannot distinguish between areas of interest based on, for example, relative locations in the visual stimulus. Close and distant AOIs are treated in exactly the same way. The Levenshtein method also fails to consider physical similarities and cognitive relationships among AOIs as well.

The Needleman-Wunsch algorithm (Needleman & Wunsch, 1970) has been used for decades in bioinformatics to analyze DNA or protein sequences. Just as sequences from related genes may be classified as similar if the number of matching residues from both genes reaches or exceeds a certain threshold, so too can sequences of eye fixations be compared to find the extent of similarities (or differences) between the visual behaviors of two subjects (Cristino, Mathôt, Theeuwes, & Gilchrist, 2010). The Needleman-Wunsch algorithm uses dynamic programming to determine the best alignment between two AOI strings. Dynamic programming refers to the breakdown of a complex problem into simpler subproblems using recursion (Cormen, 2009). The best alignment between two strings is determined by iteratively taking the first i AOIs of string A_1 and the first j AOIs of string A_2, and then obtaining a *similarity score* for the best alignment between these two substrings. Thus, the time to run the Needleman-Wunsch algorithm is proportional to the product of the lengths of the two AOI strings, pq. It is slower than other sequence alignment algorithms such as BLAST (Altschul et al., 1997), but the Needleman-Wunsch algorithm guarantees finding an optimal solution. It was for this reason that this algorithm was specifically chosen to generate similarity scores between AOI strings. The Needleman-Wunsch algorithm has been incorporated in the ScanMatch application to analyze similarities in eye movements (Cristino et al., 2010).

Components of the Needleman-Wunsch Algorithm

The main advantage of the Needleman-Wunsch algorithm over the Levenshtein method comes mostly from the way the former breaks down the similarity score between two strings into two components. The first of these components is determined using a *substitution table* that gives a score to the alignment of every possible pair of characters between the two strings. For example, using the substitution table shown in Table 1, aligning U with V gives a score of 1, while aligning U with Z gives a score of -4.

Table 1. Substitution table used to score the alignment and obtain a similarity score between two AOI strings.

	U	**V**	**X**	**Y**	**Z**
U	7	1	-2	-4	-4
V	1	7	-3	-3	-5
X	-2	-3	5	-3	-3
Y	-4	-3	-3	5	-3
Z	-4	-5	-3	-3	1

The substitution table is used to encode the relationships among AOIs. AOI relationships may be based on distance, some cognitive relationship among AOIs of the visual stimulus, or some other similarity of characteristics among them. In this study, the substitution table was built in three steps. First, when identical AOIs from two AOI strings are matched with each other, a positive score is assigned. The magnitude of this score is normalized against the length of the diagonal of the AOI in question. Matches between AOIs that have shorter diagonals were assigned more points than those coming from AOIs that have longer diagonals. This way, a pair of AOI strings does not get an unnecessarily high similarity score from having more matched fixations on large AOIs (which are anyway more likely to occur) or unnecessarily low scores from having more matched fixations on small AOIs (which are less likely to happen). Next, when fixations landing on two different AOIs are forced to be paired with each other, a negative score is assigned based on the estimated average distance between the two AOIs. An estimate of the average distance between two non-identical AOIs was determined by using the four corners and the center of each AOI in the mismatched pair as sample points. Each sample point from one AOI was paired with every sample point of the other AOI to yield 25 estimates of the distance between the two AOIs, which were eventually used to arrive at an average. Penalties from AOI mismatches were also normalized. The longer the average distance between two AOIs on which fixations are found, the greater is the penalty. The determination of distances between some sample points in AOIs U and Y is shown in Figure 3. AOIs U and V represent the reactant and product sides of the chemical equation, respectively. The corresponding sides of the PNOM diagram are denoted by X and Y, while the question for each item at the bottom of the page seen by participants designated as AOI Z. The succeeding sections of this chapter will sometimes refer to each of the AOIs using these letter designations. All measurements were done on the same monitor used for the actual eye tracking sessions. All pages of the instrument were maximized to cover the entire area of the monitor. Finally, the penalty for mismatches between AOIs corresponding to opposite sides of the chemical equation or opposite sides of the reaction diagram was reduced by two points. This was done to reflect the strength of the chemical relationship between the reactant and product sides of either a chemical equation or PNOM diagram.

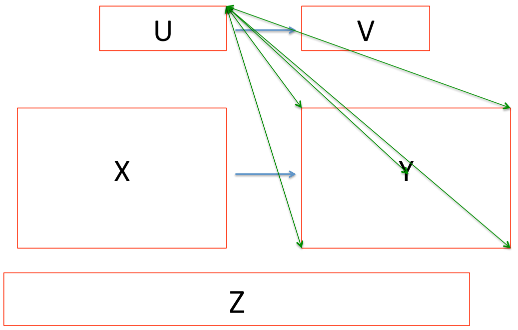


Figure 3. Estimation of the average distance between AOIs U and Y.

The second component of the similarity score between two AOI strings is the *gap penalty*. In some instances, a character on an AOI string might be matched with a blank space inserted within a second AOI string. Gaps may be introduced during the sequence alignment at a certain cost (e.g., -1) if, overall, this results in a better alignment (and thus, a higher similarity score between the two strings) between the different parts of the two AOI strings being compared. Alignments between gaps, however, are not allowed because they are redundant. Small gap penalties favor the inclusion of more gaps resulting in a decrease in the probability of aligning less related AOIs. Meanwhile, large gap penalties discourage the introduction of gaps and may force the alignment of loosely related AOIs. In this study, a penalty of one point per gap was applied.

Implementation of the algorithm

Once the components of the similarity score have been defined, these may be used with the Needleman-Wunsch algorithm to determine the similarity score between two AOI strings. The algorithm starts by building a matrix $\boldsymbol{M}$ with $p + 1$ columns and $q + 1$ rows, where p is the length of one AOI string, A_1, and q is the length of another AOI string, A_2. Columns in $\boldsymbol{M}$ are numbered $i = 0, 1, 2, \dots, p$ and the rows are designated by $j = 0, 1, 2, \dots, q$. Let $\boldsymbol{M}_{i,j}$ be the cell at the ith row and jth column of $\boldsymbol{M}$ and let $\boldsymbol{M}_{0,0} = 0$. The rest of the matrix cells are filled using a recurrence relation:

$$\boldsymbol{M}_{i,j} = max \begin{bmatrix} \boldsymbol{M}_{i-1,j-1} + \sigma\left(S_i(i), S_j(j)\right) \\ \boldsymbol{M}_{i-1,j} + g \\ \boldsymbol{M}_{i,j-1} + g \end{bmatrix} \tag{1}$$

where $\sigma\left(S_i(i), S_j(j)\right)$ is the score corresponding to the alignment of AOIs $S_i(i)$ and $S_j(j)$, read from the substitution table, and g is the gap penalty. The matrix is filled in from left to right, and from top to bottom. Thus, values $\boldsymbol{M}_{i,j}$ for row 0 of the matrix are calculated, first, left to right, followed by those in row 1, and so on. $\boldsymbol{M}_{i,j}$ is obtained as the maximum of three possible values based on the recurrence relation (1). $\boldsymbol{M}_{i-1,j-1}$ is the value of the cell in the previous column and previous row, $\boldsymbol{M}_{i-1,j}$ is the value of the cell in the previous column in the same row as $\boldsymbol{M}_{i,j}$, and $\boldsymbol{M}_{i,j-1}$is the value of the cell in the previous row in the same column as $\boldsymbol{M}_{i,j}$. The values of $\sigma\left(S_i(i), S_j(j)\right)$ are defined by a substitution table for matches and mismatches between every possible pair of AOIs. Once all cells of matrix $\boldsymbol{M}$ have been filled up, it is used to trace back the best possible alignment between A_1 and A_2. This traceback step starts from the bottom right of $\boldsymbol{M}$ and moves to the previous cell used to determine the best value of each current cell. The best alignment between strings A_1 and $A_2$is then obtained as the reverse of the string from the traceback step.

The alignment score obtained for two AOI strings is, in effect, the sum of its components. This sum is then normalized with respect to the longer string. This step is necessary because longer identical strings may be assigned higher alignment scores than if the two strings were shorter. Thus, the normalized similarity score between two AOI strings is given by

$$simscore_{norm} = \frac{alignment\ score}{Max(substitution\ matrix) * Length\ of\ the\ longer\ sequence} \quad (2)$$

With normalization, the match between two identical AOI strings is assigned a score of 1 (Josephson & Holmes, 2002). In this study, similarity scores were determined for every possible pair of AOI strings. All similarity scores were calculated using Mathematica 10.0 (Wolfram Research, 2014). Similarity scores for every pair of subjects were obtained for each item on the instrument.

Permutation Test

With two groups of subjects (which yield two sets of scan paths), a set of between-group similarity scores and two sets of within-group scores were obtained. The between-group similarity scores came from comparing AOI strings corresponding to subjects coming from two different groups. These give insight about how similar (or different) subjects from two groups are with each other. On the other hand, within-group similarity scores are calculated for pairs of subjects that come from the same group. Within-group similarity scores may be taken to describe the extent of similarity between scan paths from two subjects coming from the same treatment group. Similarity scores, thus, are associated with pairs of, and not individual, scan paths. Similarity scores, therefore, do not provide direct numerical measures of individual scan paths. Hence, one cannot use more common statistical methods for comparing similarities of scan paths from two groups of subjects, such as a *t*-test or a Wilcoxon signed-rank test.

The permutation test is a nonparametric test (Feusner & Lukoff, 2008) that has been proposed to compare similarities and differences between two groups of scan paths. Given two groups of subjects with sizes *n* and *m*, respectively, the threshold value $simdiff_{expt}$is defined as the difference between within-group and between-group average similarities and is given by

$$simdiff_{expt} = \overline{sim}_{(within,expt)} - \overline{sim}_{(between,expt)} \tag{3}$$

where $\overline{sim}_{(within,expt)}$ is the average similarity score of scan paths of pairs of subjects from the same group (within-group similarity) and $\overline{sim}_{(between,expt)}$ is that of scan paths of pairs of subjects from different groups (between-group similarity). If subjects from both groups are then randomly re-assigned to two new groups that have the same sizes as the experimental groups, one can calculate

$$simdiff_{test} = \overline{sim}_{(within,test)} - \overline{sim}_{(between,test)}. \tag{4}$$

The value of $simdiff_{test}$is expected to be close to zero and equally likely to be positive or negative. With random regrouping, there is no reason to expect subjects in the same group to have scan paths that are more or less similar to each other than subjects coming from different groups. Based on the sizes of the original groups, there are $\frac{(n+m)!}{n!m!}$ possible random groupings of subjects, so that one group has *n* subjects and the other has *m*. When $simdiff_{test} \geq simdiff_{expt}$ for a specific regrouping of subjects, this means that the random grouping being tested gives a better way of distinguishing among the subjects than the one used in the experiment itself, and that the particular grouping of subjects leads to two groups that each cluster together. The null hypothesis for the permutation test is, therefore, that the members of the groups used in the experiment are interchangeable, and that any differences observed among the subjects cannot be attributed to the original basis of the experimental groups. The *p* value for the permutation test then is the fraction of random regroupings that give $simdiff_{test} \geq simdiff_{expt}$. A schematic representation of this procedure is shown in Figure 4.

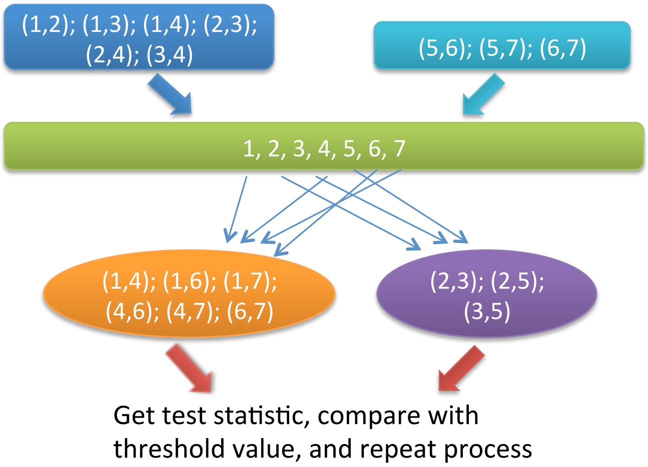


Figure 4. Schematic diagram of the permutation test done on similarity score obtained between pairs of participants.

Running all possible regroupings for a permutation test is quite challenging, if not impossible, due to the large numbers of test groupings one may have even with a small number of subjects in each experimental group. For example, with 14 subjects in one group and 13 in the other, there are approximately $2.0\ x\ 10^7$ possible random regroupings of all the subjects. To alleviate this issue, one may choose to run a randomly selected sample of regroupings using a Monte Carlo strategy to come up with a reasonable estimate of the *p* value associated with the test (Feusner & Lukoff, 2008). Different methods have been proposed to determine the optimum number of regroupings to be used in arriving at a *p* value of low uncertainty, but this remains a topic of debate (Knijnenburg, Wessels, Reinders, & Shmulevich, 2009; Lai, 2007). One approach that has been used is to simply observe fluctuations in the p value as the number of permutations tested is changed over a wide range, typically from as low as 20 to as high as 100,000 (Tang et al., 2012). The script used to run the permutation tests with AOI strings obtained from this study was written using Python 3.4.3 (Python Software Foundation, 2014) and is listed in Appendix _.

Descriptive Statistics

Participants from Chem B responded correctly to 14 to 29 items out of 30 on the instrument with a mean of 20.2 ± 2.44 at 95% confidence, while those among from Chem C ranged from 9 to 18 with a mean of 12.9 ± 1.14. These means were determined to be significantly different from each with $p < 0.001$ at $\alpha = 0.05$. Additionally, participants were also sorted between those who obtained at least 15 items correctly (high-performing group) and those who did not (low-performing group) regardless of which course they came from. Of the 27 participants, 18 were classified as high performers. Among these 18, five of them were from Chem C. One student from Chem B, on the other hand, got a score of 14 points.

Permutation Test Results

One important issue in running permutation tests is the number of regrouping samples necessary to obtain reliable p values. The permutation test in this study was optimized by varying the number of regrouping samples from 10 to 20000. Figure 5 shows the initial fluctuations and eventual stabilization for item 16, which yielded a significant difference between the visual behaviors of high- and low-performing subjects. The *p* values are seen to fluctuate at 2000 or less regrouping samples, but tend to stabilize at about 0.007-0.008 beginning at 3000 samples. It appears from these data that using 3000 regrouping samples to run the permutation test for each item would be reasonable.

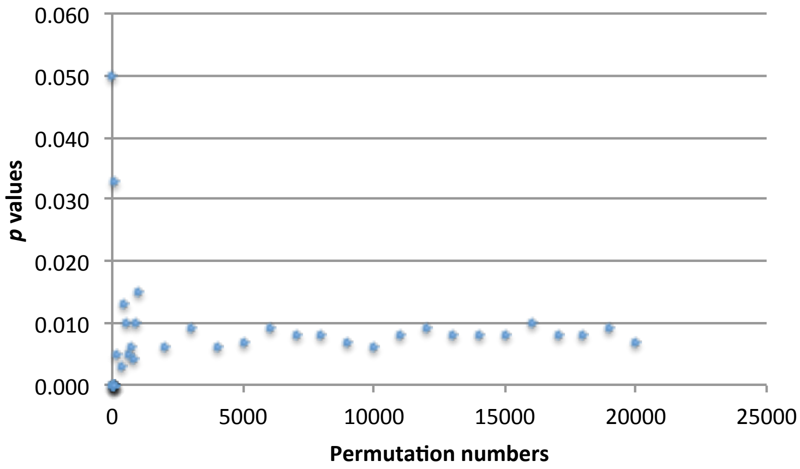


Figure 5. *p* values obtained from varying the number of regrouping samples for item 16.

Sets of similarity scores for each of the second through the fifth items for every chemical context were obtained using the Needleman-Wunsch algorithm as described earlier. The first item of each context was excluded from analysis due to differences in the way these have been presented to students coming from each course. Similarity scores for each item analyzed were then grouped together, first, based on the course from which student came. Permutation tests were then run on the similarity scores based on this grouping of participants using 3000 regrouping samples for each test. A second set of permutation tests were also run using groups of the same participants based on their performance on the instrument. The *p* value obtained from each test is given in Table 2.

Based on the *p* values determined, none of the permutation tests resulted in statistically significant differences between students from Chem B and Chem C based on the course from where they came except for item 33. This may be attributed to instrument fatigue on the part of one group of participants, although this was not verified.

What the *p* values do suggest, though, is that the visual behaviors of the participants for this study were not so much a function of their courses, but rather were more correlated with how they performed on the instrument. Thus, subjects were regrouped based on the number of correct responses they had on the instrument. Subjects who responded correctly to at least 15 out of the 30 items were reclassified as high performers. Eighteen students belonged to this group. The remaining nine participants comprised the low-performing group. Specifically, the permutation tests suggest that significant differences between how high- and low-performing students occurred in seven (items 6, 7, 16, 20, 21, 31, and 32) of the 24 items tested from the instrument. It appears that although most of the participants who got lower scores on the instrument came from Chem C, those who did well from Chem C looked at different parts of the instrument in ways similar to what most students from Chem B did. By extension, the lone student from Chem B who got less than half of the items on the instrument correctly seems to have visually behaved in the same way as students from Chem C who fared just as poorly.

Table 2. Permutation test *p* values based on comparison of similarity scores based on grouping participants by course and by performance on the instrument.

Item number	Permutation test *p* value

	By course	By performance
5	0.285	0.205
6	0.095	0.001*
7	0.052	< 0.001*
8	0.608	0.962
10	0.497	0.876
11	0.582	0.144
12	0.092	0.380
13	0.334	0.720
15	0.275	0.925
16	0.079	0.007*
17	0.252	0.225
18	0.135	0.069
20	0.498	0.036*
21	0.218	0.045*
22	0.613	0.470
23	0.322	0.303
25	0.396	0.095
26	0.565	0.357
27	0.356	0.373
28	0.171	0.597
30	0.857	0.148
31	0.574	0.021*
32	0.105	0.024*
33	0.019*	0.290

*significant at $\alpha = 0.05$

Visual Behavioral Patterns of High- and Low-Performing Participants

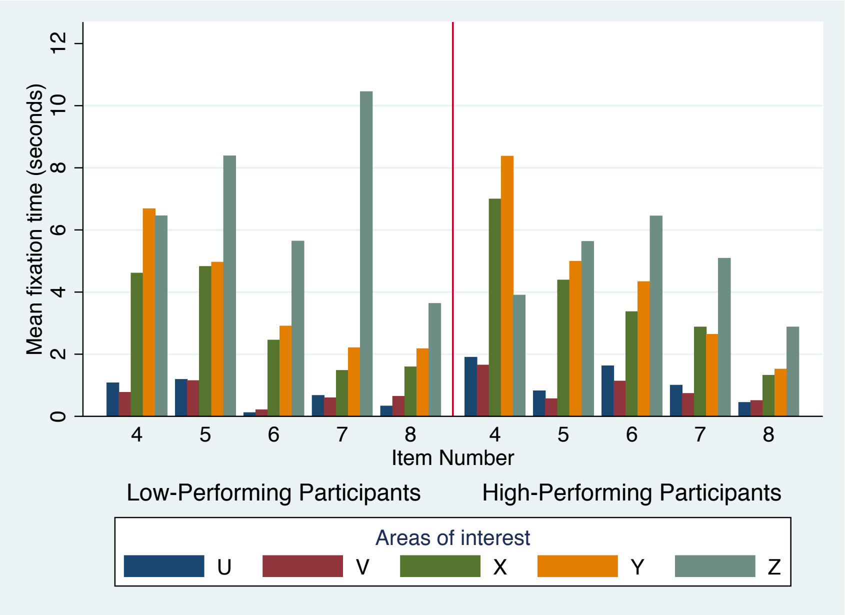
General visual behavioral patterns

The mean fixation times spent by groups of participants (based on their performance on the instrument) analyzing the different AOIs for each item are plotted in Figures 6 through 8. In general, students from both groups were observed to have spent less time responding to the items for the second diagram associated with each chemical reaction. The decrease in time spent on the second diagram of each pair may be attributed to a reliance on the part of the participants on their *episodic memory* in retaining some of the information obtained while responding to questions about the first diagram of each pair (Tulving & Thomson, 1973). Episodic memory refers to the storage and retrieval of temporally dated, spatially located, and personally experienced events or episodes. The appearance of a balanced chemical equation or the colors associated with each element in a given context might be considered to be an episode. Several students referred to having noticed that certain components of the problems, namely the chemical equation, the reactant side of the PNOM diagram, and the color scheme used to represent each element in the PNOM diagram, did not change in going from the first to the second diagram of each pair: "*I was aware of the diagram from the first picture so I didn't spend as much time looking back at it.*"

Across all sets of questions, participants spent the least amount of time fixating on the chemical equation, especially after they have already figured out which colored sphere was associated with each element in the equation. On the other hand, participants from both groups generally spent the most time fixating on the question for each item which was found at the bottom of each page. Exceptions to this were the first item for each of three different contexts among the low performers and two contexts for the high performers. In these instances, students spent time figuring out the color scheme used in the PNOM diagrams. For these items, participants usually spent the most time fixating on one or both sides of the PNOM diagram.

One key difference among high and low performers across all items is seen in the way fixation times on parts of the same PNOM diagram varied as subjects went from one item to the next. For each set of items pertaining to the same diagram, high performers generally spent monotonically decreasing times looking at diagrams as they went from one item to item. No such trend with fixation times on diagrams can be described among the low-performers. The monotonic decrease in fixation times on the diagram across items for the same context was most distinctly observed for the methane combustion (Figure 6) and the carbon tetrachloride disubstitution reaction contexts (Figure 8). It is possible that high performers were more efficiently using their memories by taking into account different aspects of the diagrams such as which reactant was limiting, how many of each product should be formed, how many leftover molecules should be in the diagram, even as they were still trying to figure out their responses to the first item for each diagram. As a result, in moving on to the next item for the same diagram, the high performers were observed to be making their decisions much more quickly than they had for the earlier items. Low performers, on the other hand, may have applied a more compartmentalized approach to their diagram analysis. This means that at least some of the low-performing participants treated items pertaining to the same diagram independently of each other. For example, among high performers who saw a correctly drawn diagram chose to indicate that it was both based on the correct choice of the limiting reagent and that the correct number of leftover molecules were drawn. This was not generally true among the low performers. Several of them picked “True” for one item and chose “False” for the other, even though both items pertained to the same diagram. Fixation times on diagrams for low performers usually decreased between ends of a series of items, but went through upswings of varying extents somewhere in the middle.

High performers also seemed to have been more affected by the visual complexities of the PNOM diagrams than low performers were, especially when responding to the first item for each diagram. Figure 8 shows mean fixation times lasting more than nine and ten seconds, respectively, on the reactant and product sides of the first diagram for the carbon tetrachloride reaction context among high performers. The same group only spent about seven seconds looking at the reactant side and slightly more than eight seconds studying the product side of first diagram for methane combustion (Figure 6). The methane combustion context diagrams included a smaller number of different types of atoms and a smaller number of molecules, than those drawn in the carbon tetrachloride reaction diagrams. Fixation times on the diagrams among low performers, on the other hand, were not as affected by visual complexities. For the first methane combustion diagram, low performers spent a little over four seconds, and more than six seconds, respectively, looking at the reactant and product sides. As the low performers studied the first diagram for the carbon tetrachloride reaction, they spent less than three seconds and less than six seconds looking at the left and right sides, respectively. It appears that high performers were more deliberate in analyzing changes between diagrams than low performers were. This, again, points to the high performers' more integrative approach as they tried to figure out things well beyond which colored sphere represented which element.



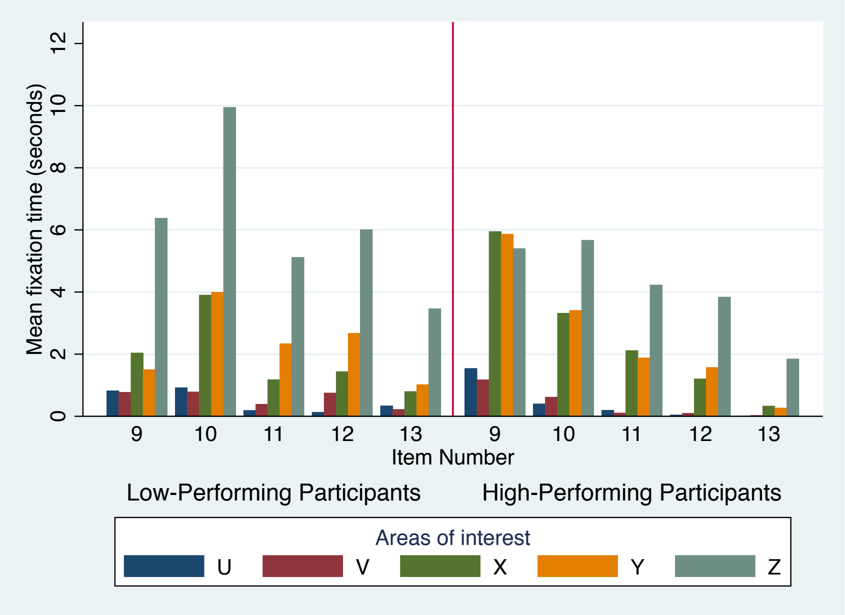


Figure 6. Mean fixation times spent by low- and high-performing students on the different areas of interest for items 4 through 13. (U and V are the reactant and product sides of the balanced chemical equation, respectively; X and Y are the reactant and product sides of the PNOM diagram, respectively; and Z is the question for each item.)

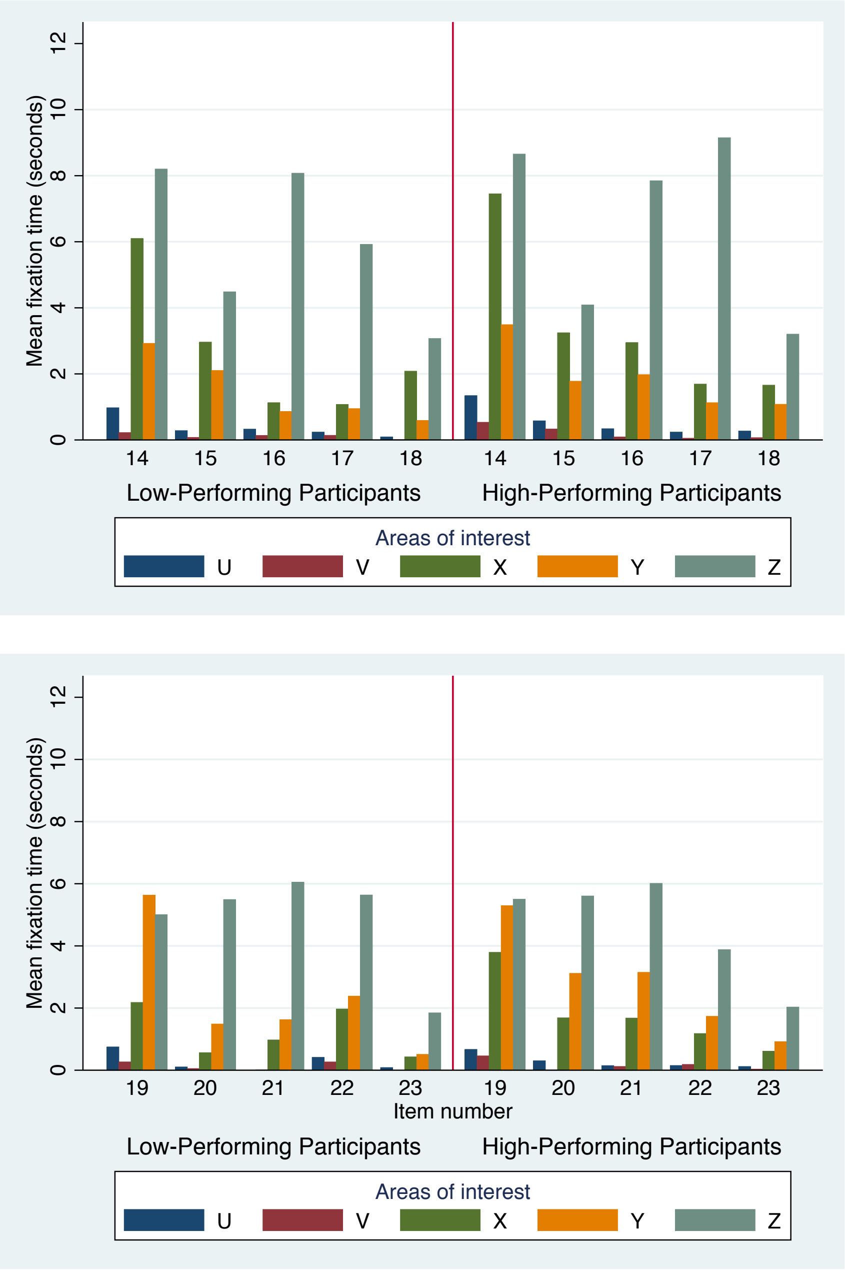


Figure 7. Mean fixation times spent by low- and high-performing students on the different areas of interest for items 14 through 23. (U and V are the reactant and product sides of the balanced chemical equation, respectively; X and Y are the reactant and product sides of the PNOM diagram, respectively; and Z is the question for each item.)

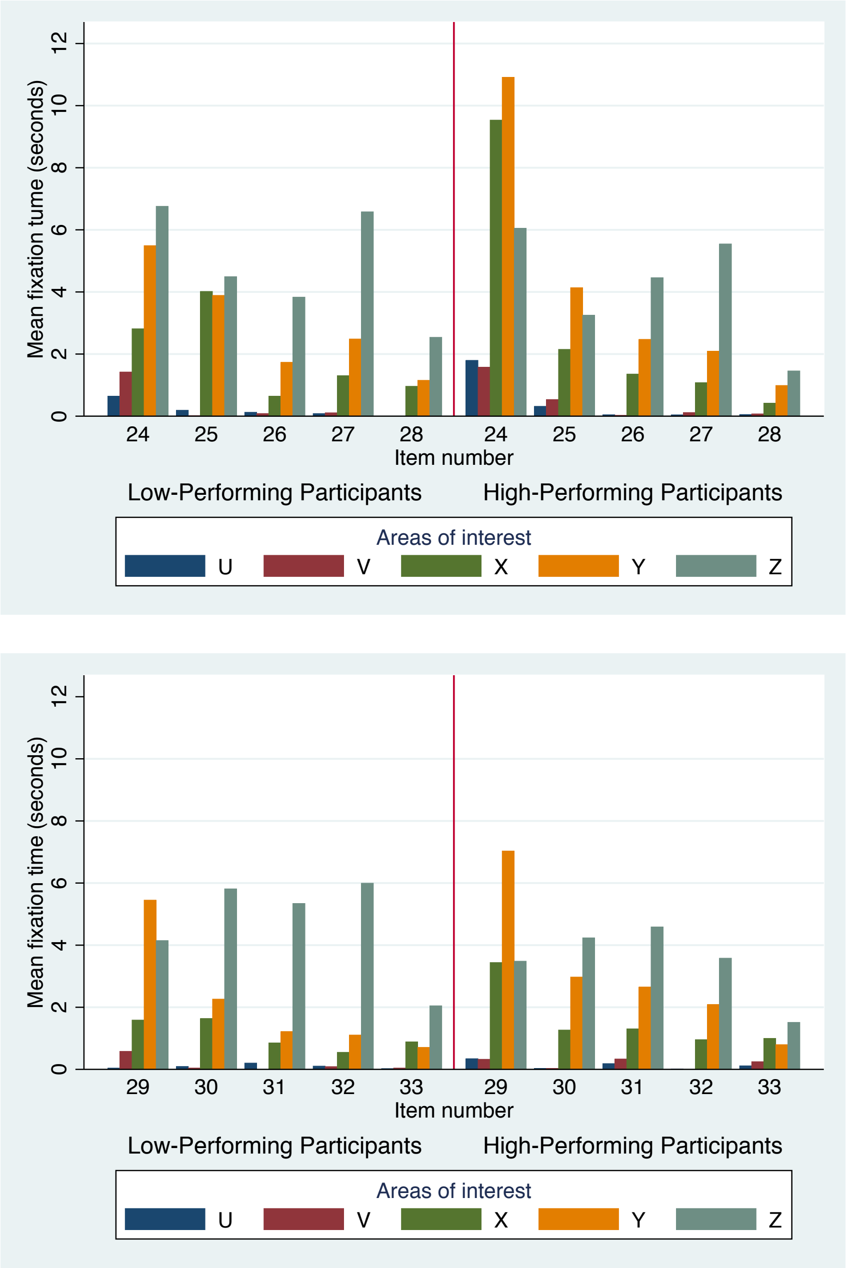


Figure 8. Mean fixation times spent by low- and high-performing students on the different areas of interest for items 24 through 33. (U and V are the reactant and product sides of the balanced chemical equation, respectively; X and Y are the reactant and product sides of the PNOM diagram, respectively; and Z is the question for each item.)

Heat maps may be used to visualize the durations of fixations on different AOIs across stimuli. Color gradients are used to illustrate the mean fixation times across participants from the same group. More intense colors (usually red/orange) indicate the longest mean fixations on an AOI while less intense colors (blue/green) indicate shorter fixations. As an example of how heat maps may be used to compare the way different groups of students looked at visual stimuli, consider Figure 9, which is on item 30. For this specific item, one should see that high-performing participants did not pay as much attention to the chemical equation as the low-performing students did, this being the second item for the same diagram. The lack of red/orange blots across the PNOM diagram AOIs for high-performers also indicate that at this point they were mostly doing cursory checks on the diagram. The larger and more intensely colored blots on the heat map for low performers meanwhile indicate that these students were still very much involved with carefully examining each side of the diagram even though they have seen exactly the same thing just in the previous item. Both high- and low-performers, however, spent a large fraction of their time for this item carefully reading the question at the bottom of the page. High performers were just as concerned with understanding the question correctly as the low performers were.

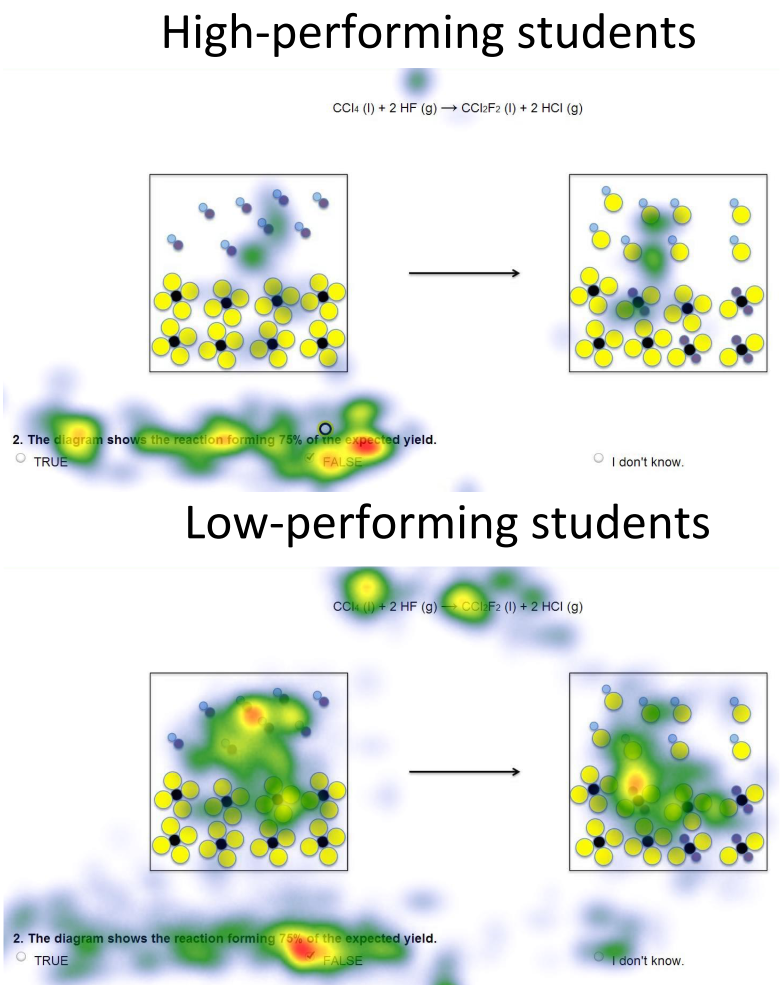
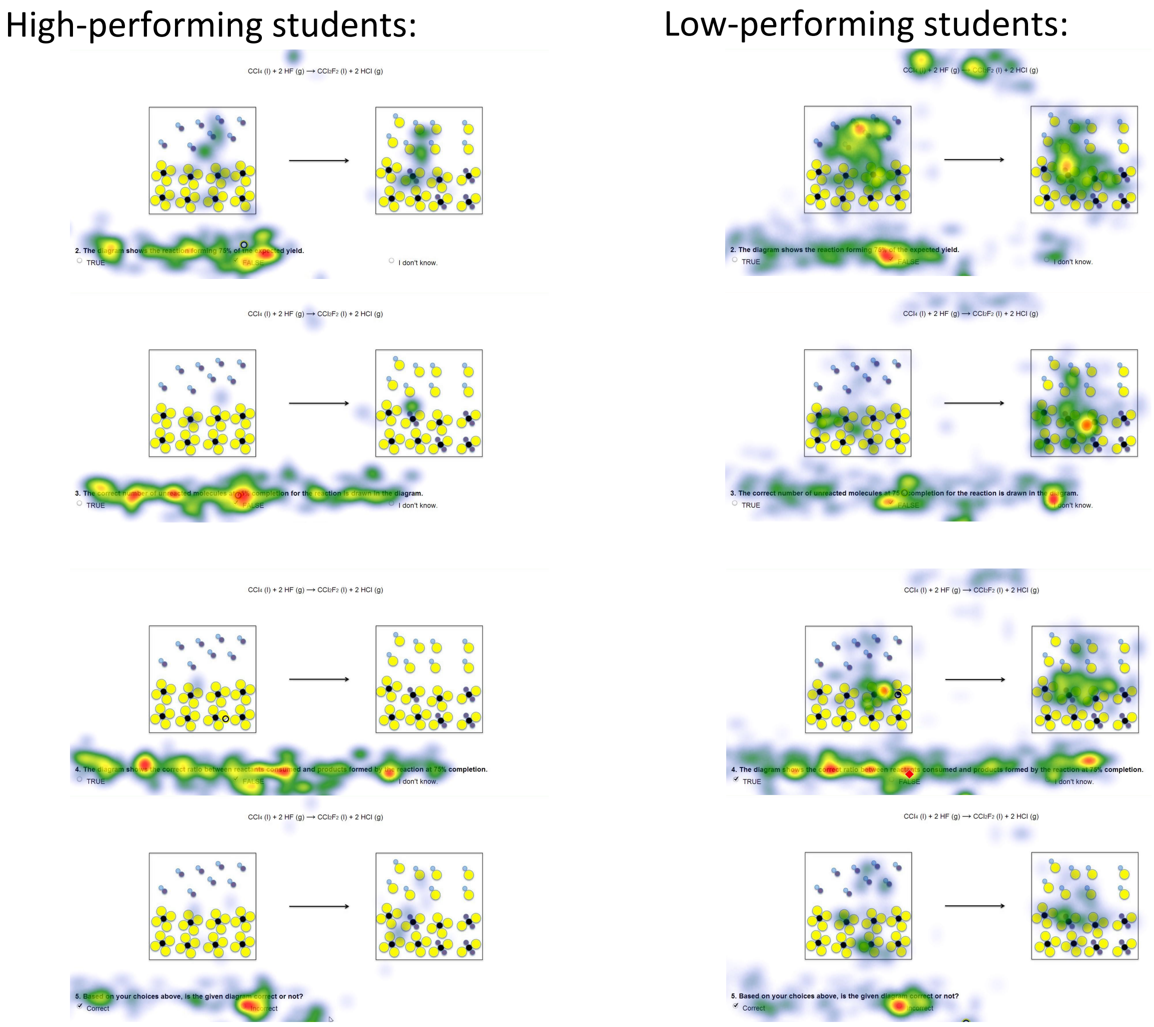


Figure 9. Item 26 heat maps for high- and low-performing students.

Heat maps may also be used to compare the visual behaviors of high- and low-performing participants across series of items. Figure 10 shows heat maps across items 25 through 28. Notice the intensity of colors across the PNOM diagrams among low-performing participants especially early in this series. It is pretty clear that this group of students needed to constantly attend to the diagram and not just rely on their memory to come up with a response to each item, even though the same diagram was shown through this series. On the other hand, high-performing participants are seen to have only attended to the diagram at the top of their series of heat maps.


High-performing students:

Low-performing students:

Figure 10. Heat maps for items 25 through 28.

Heat maps for another series of items indicate that differences between the viewing patterns of high- and low-performers may sometimes be more subtle (Figure 11). Items 5 through 8 show both groups paying more or less equal attention to the chemical equation while responding to earlier items in this series. However, a gradual shortening of fixation times on the reaction is shown by the decrease in intensities of blots across the chemical equation for high performers. This decrease in fixation time on the reaction did not occur for low-performing students. This again indicates the greater trust placed by high-performing students on their memories as they went from item to item. In contrast, low-performing students seem to have paid much more attention to both sides of the diagram even as the moved on to the later items. The intense red/orange colors of blots, especially on the reactant side of the diagram did not fade until the last item among high-performers. This lends evidence to what high-performers assigned greater priority to as they went from one item to the next. Low performers, in contrast, spent more time on the diagram early in the series and were more concerned with understanding the question correctly.

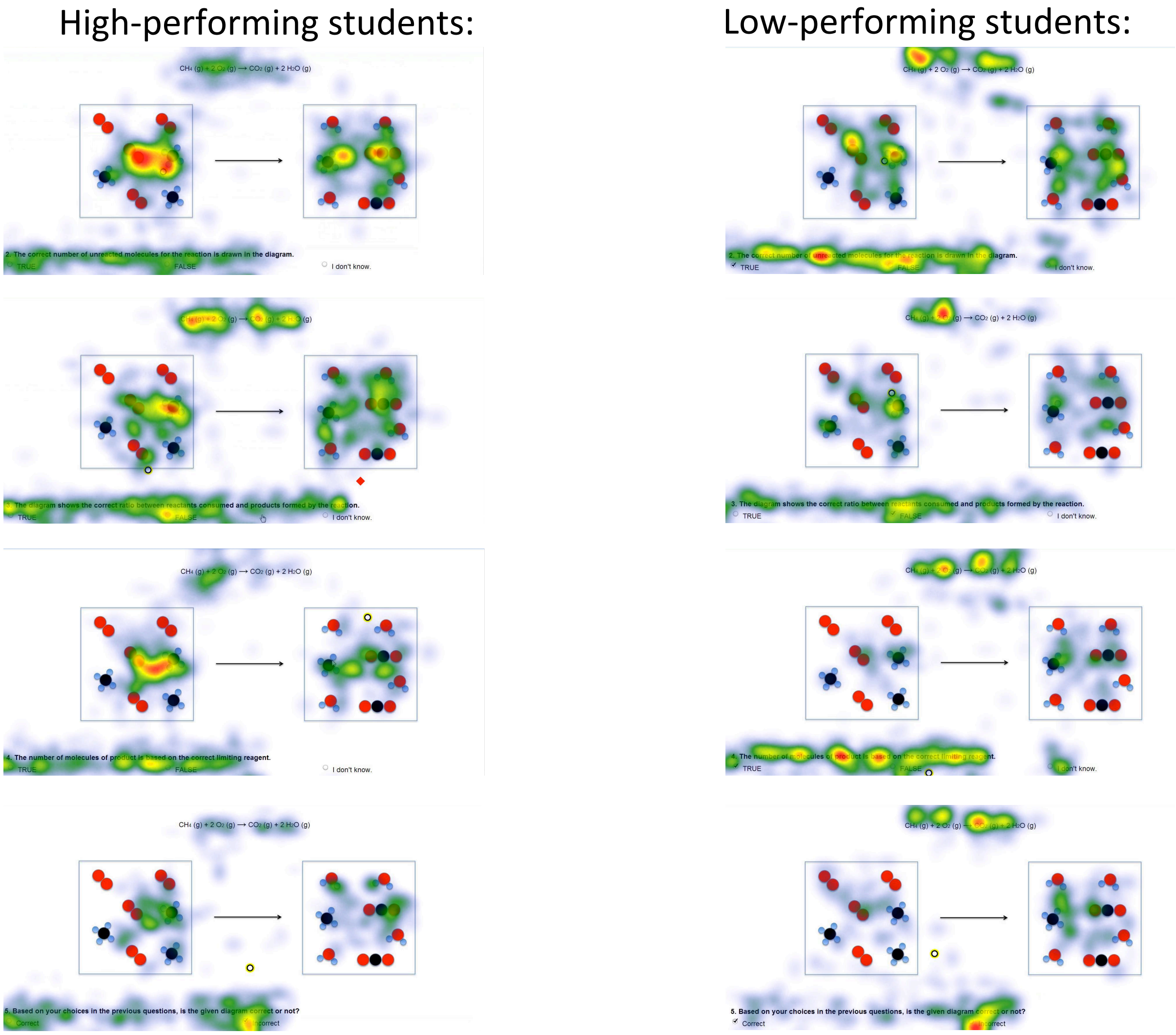


Figure 11. Heat maps for items 5 through 8.

Differences between high- and low-performers on specific items

Seven items showed significant differences among the scan paths taken by high- and low-performing participants as indicated by *p* values coming from the permutation tests on the similarity scores. Each of these items were analyzed in terms of mean fixation times and fixation counts participants from both groups allocated to each of the AOIs, and in terms of mean numbers of transitions between pairs of AOIs for each item. Examining AOI fixation times and counts give a quantitative indication of which AOI each group paid the most attention to while responding to each item. Looking at AOI transitions gives ideas about which pair(s) of AOIs were most commonly coordinated by participants. In a few instances, the direction of coordination between AOIs in a pair was also deemed significant.

Differences in AOI fixation times between high- and low-performing participants for items yielding significant permutation tests were determined with point biserial correlation coefficients (Tate, 1954). This correlation coefficient, r_{pb}, is used when one of the variables for which a relationship is being examined is dichotomous, in this case, whether participants belong to either the high- or low-performing group, while the other is continuous, such as the fixation time on each AOI. Assigning the value 1 to members of one group and 0 to the other, r_{pb} is calculated as

$$r_{pb} = \frac{M_1 - M_0}{s_{n-1}} \sqrt{\frac{n_1 n_0}{n^2}} \tag{5}$$

where s_{n-1} is the standard deviation, M_1 is the mean value of the continuous variable for all participants in group 1, M_0 is that for group 0, n_1 and n_0 are the number of participants in their respective groups, and n is the total sample size. The sign of this coefficient gives the direction of the effect a change in the dichotomous variable has on the continuous variable. Values between 0.3 and 0.7 (or -0.3 and -0.7) indicate moderate correlation between the variables, while those exceeding 0.7 (or less than -0.7) indicate strong correlation.

For correlations that involve a dichotomous variable and an ordinal variable (such as fixation counts and number of transitions between two AOIs), the rank biserial correlation coefficient, r_{rb}, is more appropriate (Cureton, 1956). This coefficient is calculated as

$$r_{rb} = \frac{2(M_1 - M_0)}{n} \tag{6}$$

The magnitude and sign of this correlation coefficient are interpreted in the same way as those of the point biserial.

Item 6

Item 6 (Figure 12) pertains to the ratio between the number of reactant molecules used and the number of product molecules formed by the combustion of methane. The permutation test *p* value for this was determined to be 0.001, indicating a statistically significant difference between the sequences in which high- and low-performing participants looked at AOIs for this item. Mean fixation times and mean fixation counts for each group of participants as well as the correlation coefficients associated with these are listed in Table 3. High-performing participants were observed to have spent more time examining each of the different AOIs for this item than low-performing participants did, with moderate correlations between grouping and fixation times on the AOIs for the two sides of the chemical equation (U and V, respectively). High-performing participants also fixated on each of the AOIs more frequently than the low-performing participants did, with moderate correlations for AOIs V, Y (product side of the PNOM diagram), and Z (question for the item), and a strong correlation for AOI U. These suggest that at this point, participants from the high-performing group were still spending a large fraction of their time trying to put together the information they can obtain from each of the AOIs for this item and may still be comparing ratios obtained from the equation to those between the numbers of molecules in the diagram.

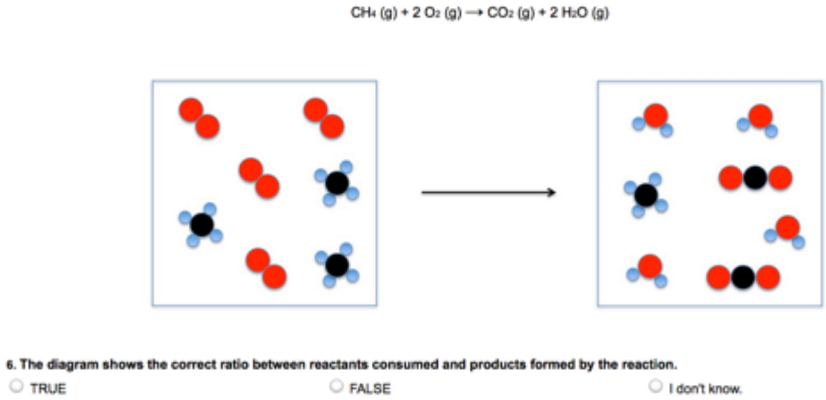
Table 3. Means and correlation coefficients for fixation times and fixation counts associated with areas of interest for item 6 resulting from grouping participants into high- and low-performers based on total score on the instruments (r_{pb} = point biserial correlation coefficient; r_{rb} = rank biserial correlation coefficient).

Mean	Fixation Time (seconds)			Fixation Count		
AOI	High Group	Low Group	r_{pb}	High Group	Low Group	r_{rb}
U	1.46	0.12	0.389*	7.1	0.7	0.701**
V	1.11	0.21	0.332*	6.0	1.4	0.635*
X	3.31	2.48	0.172	14.8	10.2	0.272
Y	4.33	2.93	0.221	19.9	12.9	0.333*
Z	6.59	5.76	0.086	27.4	24.0	0.354*

* moderate correlation; ** strong correlation

To gain insights about how participants tried to put information derived from AOIs together, transitions between pair of AOIs for item 6 were also analyzed (Table 4). Specifically, AOI transitions were examined in terms of means based on performance group and the direction of each AOI transition. Moderately strong correlations were found for the U → V, U → X, V → Y, and Y → V AOI transitions when the specific direction of AOI transitions are taken into account. Total number of transitions between

AOIs U and V, U and X, and V and Y were also seen to exhibit moderately strong correlation with performance group. The fact that the total number of transitions between U and V are greater for high performers than for low performers may be an indication of the value high performing participants assigned to the balance of the equation for the chemical situation they were dealing with as they responded to this item. That the U → V direction had moderately strong correlation while the V → U direction did not is probably a result of the left-to-right direction in which subjects read the equation. The moderate correlations for the U ↔ X and V ↔ Y transitions indicate the emphasis placed by high performers on the relationships between corresponding sides of the chemical equation and the diagrams. Part of the emphasis here may simply be the need to match each of the chemical formulas in the equation with the correct representation in the PNOM diagram. High-performing subjects may also have been concerned with checking whether ratios between species on each side of the equation were the same as the corresponding side in the diagram since this item asked subjects to check the ratios between reactants used and products formed by the reaction.



CH_4 (g) + 2 O_2 (g) → CO_2 (g) + 2 H_2O (g)

6. The diagram shows the correct ratio between reactants consumed and products formed by the reaction.

○ TRUE ○ FALSE ○ I don't know.

Figure 12. Item 6. The correct response is boxed in red.

Table 4. Means and correlation coefficients for transition counts between each AOI pair for item 6 for high- and low-performers based on total score (r_{rb} = rank biserial correlation coefficient). Arrows indicate the direction of each transition.

From-To AOI Pair	Mean AOI Transition Count		r_{rb}
	High Group	Low Group	
U → V	0.6	0.0	0.410*
U → X	1.6	0.1	0.609*
U → Y	0.2	0.0	0.112
U → Z	0.1	0.0	0.056
V → U	0.6	0.0	0.285
V → X	0.1	0.1	-0.056
V → Y	1.2	0.1	0.517*
V → Z	0.2	0.1	0.056
X → U	1.6	0.6	0.299
X → V	0.0	0.0	0.000
X → Y	1.7	1.7	-0.074
X → Z	1.0	1.0	0.062
Y → U	0.3	0.0	0.226
Y → V	0.9	0.0	0.347*
Y → X	1.7	1.4	-0.062
Y → Z	0.9	0.7	0.168
Z → U	0.1	0.0	0.112
Z → V	0.1	0.0	0.112
Z → X	0.8	1.2	-0.031
Z → Y	1.2	0.7	0.259
U ⇔ V	1.1	0.0	0.478*
U ⇔ X	3.1	0.7	0.509*
U ⇔ Y	0.5	0.0	0.285
U ⇔ Z	0.2	0.0	0.168
V ⇔ X	0.1	0.1	-0.056
V ⇔ Y	2.1	0.1	0.533*
V ⇔ Z	0.3	0.1	0.168
X ⇔ Y	3.3	3.1	-0.099
X ⇔ Z	1.8	2.2	0.049
Y ⇔ Z	2.2	1.3	0.246

* moderate correlation; ** strong correlation

Items pertaining to the correct choice of the limiting reagent

Four of the six items pertaining to the correct choice of the limiting reagent were determined to have produced statistically significant differences between how high- and low-performing participants viewed the different AOIs for these items. These were items 7 (Figure 13), 16 (Figure 14), 20 (Figure 15), and 31 (Figure 16).

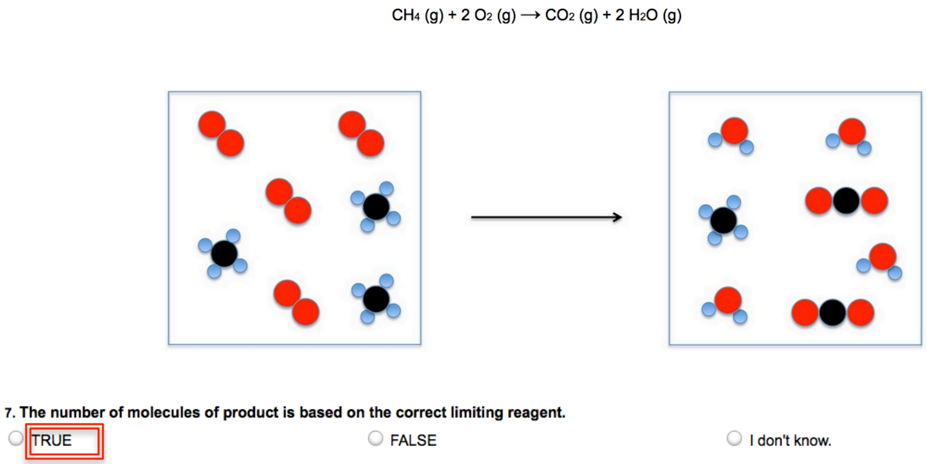
Item 7 is the fourth of the series of items for this diagram. Fixation counts on AOIs X and Z were determined to be moderately correlated with performance on the instrument (Table 5). High performers focused their attention on the left side of the diagram, but more in relation to the reactants in the equation. There were more transitions between AOIs U and X for high performers, with emphasis on the U → X direction (Table 6). The U → X emphasis may indicate efforts to use the relationship between reactants in the equation to check the ratio between the numbers of molecules in the reactant side of the diagram. As in the case of the low performers, some high performers may also have been prompted by the item’s emphasis on the limiting reagent to look more at the reactant side of the equation. This is indicated by the somewhat greater number of U ↔ Z transitions for the high performers. In particular, this time it was the low-performing participants who looked more at these AOIs than the high-performers did. The emphasis of this item on the limiting reagent may have prompted participants to look more in the direction of the reactant side of the diagram. An examination of the heat maps for this item (third row of Figure 11) also shows that low performers tended to focus their attention of the words “limiting reagent” in AOI Z, indicating that this may have been a source of difficulty for these participants, resulting in the much longer fixation times and much more frequent fixations on this AOI by the low performers.

High performers focused their attention on the left side of the diagram, but more in relation to the reactants in the equation. There were more transitions between AOIs U and X for high performers, with emphasis on the U → X direction (Table 6). This directional emphasis may indicate efforts to use the relationship between reactants in the equation to check the ratio between the numbers of molecules in the reactant side of the diagram. As in the case of the low performers, some high performers may also have been prompted by the item's emphasis on the limiting reagent to look more at the reactant side of the equation. This is indicated by the somewhat greater number of U ↔ Z transitions for the high performers.

Table 5. Means and correlation coefficients for fixation times and fixation counts associated with areas of interest for item 7 resulting from grouping participants into high- and low-performers based on total score on the instruments (r_{pb} = point biserial correlation coefficient; r_{rb} = rank biserial correlation coefficient).

	Fixation Time (seconds)			Fixation Count		
AOI	High Group	Low Group	r_{pb}	High Group	Low Group	r_{rb}
U	0.98	0.73	0.085	2.2	4.2	0.200
V	0.68	0.60	0.023	2.2	3.4	-0.068
X	2.93	1.51	0.296	1.6	6.2	0.368*
Y	2.55	2.21	0.055	4.6	9.2	0.239
Z	5.13	10.84	-0.509*	9.5	42.0	-0.626*

* moderate correlation



CH_4 (g) + 2 O_2 (g) → CO_2 (g) + 2 H_2O (g)

7. The number of molecules of product is based on the correct limiting reagent.

○ TRUE ○ FALSE ○ I don't know.

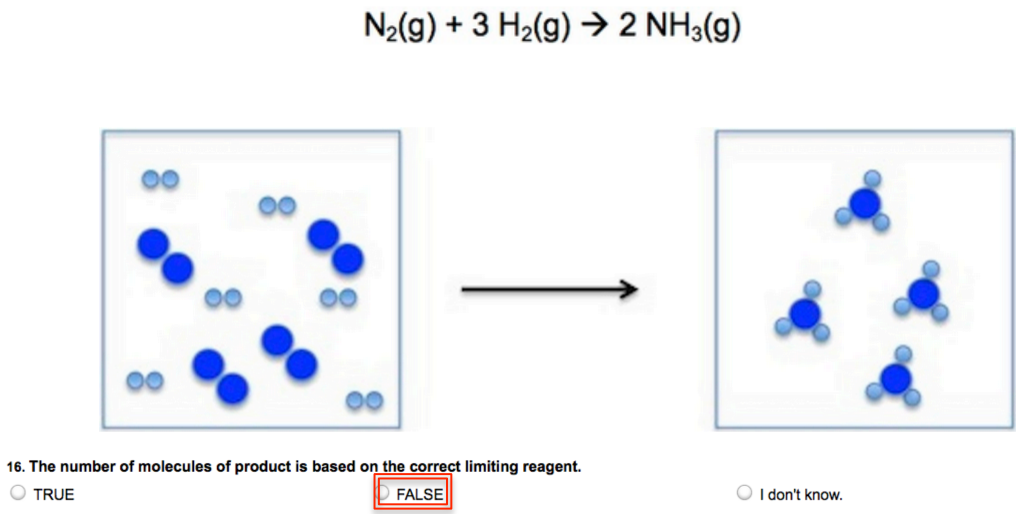
Figure 13. Item 7. The correct response is boxed in red.

Table 6. Means and correlation coefficients for transition counts between each AOI pair for item 7 resulting for high- and low-performers based on total score (r_{rb} = rank biserial correlation coefficient). Arrows indicate the direction of each transition.

From-To AOI Pair	Mean AOI Transition Count		r_{rb}
	High Group	Low Group	
U → V	0.2	0.1	0.056
U → X	1.3	0.1	0.525*
U → Y	0.1	0.0	0.112
U → Z	0.3	0.0	0.226
V → U	0.2	0.2	0.037
V → X	0.3	0.1	0.118
V → Y	0.6	0.1	0.187
V → Z	0.0	0.2	-0.223
X → U	0.9	0.6	0.137
X → V	0.0	0.0	0.000
X → Y	1.6	1.7	-0.056
X → Z	0.5	1.1	-0.226
Y → U	0.3	0.0	0.168
Y → V	0.6	0.0	0.226
Y → X	0.7	1.3	-0.340*
Y → Z	0.8	0.7	0.056
Z → U	0.2	0.0	0.168
Z → V	0.1	0.0	0.056
Z → X	0.8	1.1	0.062
Z → Y	0.8	0.7	0.087
U ⇔ V	0.3	0.3	0.137
U ⇔ X	2.3	0.7	0.368*
U ⇔ Y	0.4	0.0	0.226
U ⇔ Z	0.4	0.0	0.347*
V ⇔ X	0.3	0.1	0.118
V ⇔ Y	1.2	0.1	0.252
V ⇔ Z	0.1	0.2	-0.168
X ⇔ Y	2.3	3.0	-0.181
X ⇔ Z	1.3	2.2	-0.043
Y ⇔ Z	1.6	1.3	0.112

* moderate correlation

Item 16 (Figure 14) pertains to the incorrectly drawn diagram for the production of ammonia from the elements at 50% completion. Moderately strong correlations were found between the performance of participants on the instrument and the both the duration of fixations on AOIs X and Y, as well as the frequencies of these fixations (Table 7). For each of these AOIs, the high-performing participants were observed to have focused more attention to both sides of the diagram than the low-performing participants did. In particular, high-performing participants generally looked both ways at the diagram as they tried to establish the relationships between the numbers of reactant and product molecules (Table 8).



$N_2(g) + 3\ H_2(g) \rightarrow 2\ NH_3(g)$

16. The number of molecules of product is based on the correct limiting reagent.

○ TRUE

○ FALSE

○ I don't know.

Figure 14. Item 16. The correct response is boxed in red.

The large number of fixations on the reactant side of the diagram may have been prompted by the emphasis on the limiting reagent the question for this item has. Longer fixation times on X may have come from its greater visual complexity compared to Y, given that X has two different kinds of molecules and overall a greater number of molecules than Y does. To determine the correct answer for this question, participants needed to see that all of the hydrogen gas molecules were used used as opposed to the nitrogen gas molecules, which should have led them to see that the reaction did not reflect 50% completion as required by the problem. They also needed to see there were missing nitrogen atoms from the product side of the diagram. The greater number of fixations on the reactant side of the diagram may be from counting more molecules compared to the product side. All subjects who answered this item correctly came from the high-performing group.

Table 7. Means and correlation coefficients for fixation times and fixation counts associated with areas of interest for item 16 resulting from grouping participants into high- and low-performers based on total score on the instruments (r_{pb} = point biserial correlation coefficient; r_{rb} = rank biserial correlation coefficient).

	Fixation Time (seconds)			Fixation Count		
AOI	High Group	Low Group	r_{pb}	High Group	Low Group	r_{rb}
U	0.34	0.33	0.011	2.1	2.0	0.043
V	0.09	0.14	-0.084	0.6	0.4	0.012
X	2.98	1.12	0.363*	12.5	5.3	0.478*
Y	2.09	0.85	0.379*	8.7	4.0	0.447*
Z	8.04	8.22	-0.018	32.8	30.4	0.012

* moderate correlation

Table 8. Means and correlation coefficients for transition counts between each AOI pair for item 16 resulting for high- and low-performers based on total score (r_{rb} = rank biserial correlation coefficient). Arrows indicate the direction of each transition.

From-To AOI Pair	Mean AOI Transition Count		r_{rb}
	High Group	Low Group	
U → V	0.1	0.0	0.112
U → X	0.1	0.2	-0.156
U → Y	0.1	0.0	0.112
U → Z	0.1	0.1	0.000
V → U	0.0	0.0	0.000
V → X	0.1	0.0	0.056
V → Y	0.1	0.1	0.000
V → Z	0.0	0.0	0.000
X → U	0.2	0.1	0.056
X → V	0.1	0.0	0.056
X → Y	1.6	0.6	0.440*
X → Z	1.1	0.7	0.137
Y → U	0.1	0.1	0.000
Y → V	0.1	0.0	0.056
Y → X	1.2	0.7	0.207
Y → Z	1.0	0.6	0.292
Z → U	0.1	0.2	-0.112
Z → V	0.1	0.0	0.056
Z → X	1.1	0.6	0.175
Z → Y	1.1	0.8	0.118
U ⇔ V	0.1	0.0	0.112
U ⇔ X	0.3	0.3	-0.056
U ⇔ Y	0.2	0.1	0.062
U ⇔ Z	0.2	0.3	-0.062
V ⇔ X	0.1	0.0	0.056
V ⇔ Y	0.2	0.1	0.006
V ⇔ Z	0.1	0.0	0.056
X ⇔ Y	2.7	1.2	0.361*
X ⇔ Z	2.1	1.2	0.187
Y ⇔ Z	2.1	1.3	0.194

* moderate correlation

With item 20, the product side of the diagram becomes more visually complex with the inclusion of all the unreacted molecules of hydrogen and nitrogen gases even as the number of ammonia product molecules drawn gets reduced by two. As a result, there were three different types of molecules to account for and the product side of the diagram loses also the roughly symmetrical arrangement among the molecules. It appears then that the high-performing participants were affected more by these changes in the visual complexity on this side of the diagram based on their more frequent and longer fixations on Y (Table 9). High-performing participants also tended to compare the numbers of molecules drawn on Y with those on the reactant side of the diagram (X) more frequently (Table 10) . On the other hand, low-performing participants were more concerned with relating the number of molecules on Y with the question (Z). It is not clear how low performers were trying to relate the limiting reagent with the numbers of different kinds of molecules shown in AOI Y for this item.

Table 9. Means and correlation coefficients for fixation times and fixation counts associated with areas of interest for item 20 resulting from grouping participants into high- and low-performers based on total score on the instruments (r_{pb} = point biserial correlation coefficient; r_{rb} = rank biserial correlation coefficient).

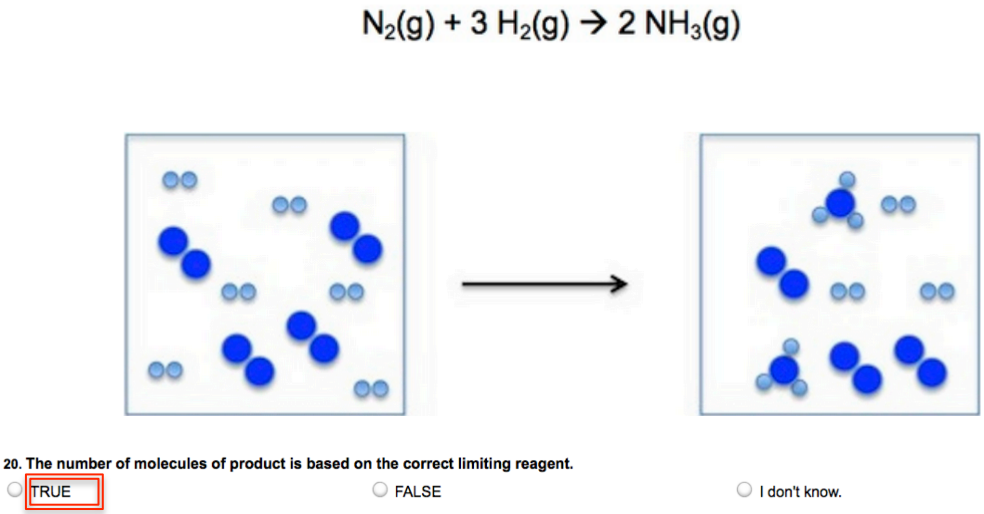
	Fixation Time (seconds)			Fixation Count		
AOI	High Group	Low Group	r_{pb}	High Group	Low Group	r_{rb}
U	0.23	0.11	0.143	1.2	0.7	0.006
V	0.01	0.05	-0.302*	0.1	0.3	-0.175
X	1.77	0.57	0.302*	7.4	3.0	0.259
Y	3.08	1.46	0.252	13.9	6.7	0.319*
Z	5.70	5.98	-0.039	24.2	22.2	0.000

* moderate correlation

Table 10. Means and correlation coefficients for transition counts between each AOI pair for item 20 for high- and low-performers based on total score (r_{rb} = rank biserial correlation coefficient). Arrows indicate the direction of each transition.

From-To AOI Pair	Mean AOI Transition Count		r_{rb}
	High Group	Low Group	
U → V	0.0	0.1	-0.112
U → X	0.1	0.0	0.112
U → Y	0.2	0.0	0.167
U → Z	0.1	0.0	0.112
V → U	0.0	0.0	0.000
V → X	0.0	0.0	0.000
V → Y	0.0	0.0	0.000
V → Z	0.1	0.1	-0.056
X → U	0.2	0.0	0.112
X → V	0.0	0.0	0.000
X → Y	0.9	1.0	-0.207
X → Z	0.7	0.6	0.012
Y → U	0.1	0.0	0.112
Y → V	0.1	0.0	0.056
Y → X	1.2	0.4	0.382*
Y → Z	0.8	1.6	-0.485*
Z → U	0.0	0.1	-0.112
Z → V	0.0	0.0	0.000
Z → X	0.5	0.7	-0.093
Z → Y	1.1	0.9	0.080
U ⇔ V	0.0	0.1	-0.112
U ⇔ X	0.3	0.0	0.112
U ⇔ Y	0.3	0.0	0.163
U ⇔ Z	0.1	0.1	0.000
V ⇔ X	0.0	0.0	0.000
V ⇔ Y	0.1	0.0	0.056
V ⇔ Z	0.1	0.1	-0.056
X ⇔ Y	2.1	1.4	0.087
X ⇔ Z	1.2	1.2	0.025
Y ⇔ Z	1.8	2.4	-0.207

* moderate correlation



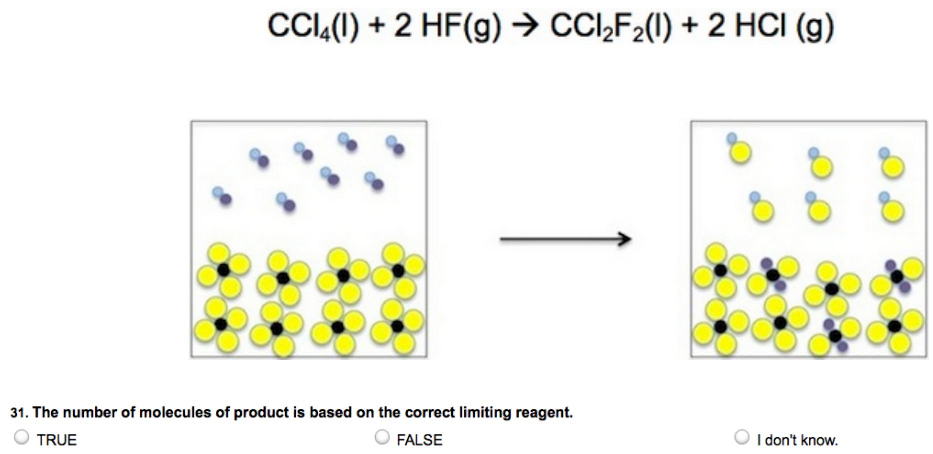
$N_2(g) + 3\ H_2(g) \rightarrow 2\ NH_3(g)$

20. The number of molecules of product is based on the correct limiting reagent.

○ TRUE ○ FALSE ○ I don't know.

Figure 15. Item 20.

AOI Y in item 31 (Figure 16) was also more closely examined by high-performing participants than those with lower scores (Table 11) . Participants were being asked to consider this diagram given that the reaction went 75% towards completion. This item required participants to figure out that with 75% completion for the reaction, there should be three HCl, three CCl_2F_2, five CCl_4, and two HF molecules drawn on the product side of the diagram. High-performing participants looked across the two sides of the diagram more often than low-performing students did, with a tendency to go left to right rather than right to left (Table 12). High performers also spent more time and looked more frequently at Y than low performers did.



$CCl_4(l) + 2\ HF(g) \rightarrow CCl_2F_2(l) + 2\ HCl\ (g)$

31. The number of molecules of product is based on the correct limiting reagent.

○ TRUE ○ FALSE ○ I don't know.

Figure 16. Item 31. The correct response is boxed in red.

Overall, with regards to items asking about the correctness of the PNOM diagrams with respect to the choice of the limiting reagent for each situation, high performers tended to focus on establishing the relationships between molecules found on both sides of PNOM diagram, X and Y. The attention from high performers on X and Y seemed to have been divided evenly between the two sides. Low performers, on the other hand, seem to have focus more of their attention to the question AOIs, Z.

Table 11. Means and correlation coefficients for fixation times and fixation counts associated with areas of interest for item 31 resulting from grouping participants into high- and low-performers based on total score on the instruments (r_{pb} = point biserial correlation coefficient; r_{rb} = rank biserial correlation coefficient).

	Fixation Time (seconds)			Fixation Count		
AOI	High Group	Low Group	r_{pb}	High Group	Low Group	r_{rb}
U	0.19	0.23	-0.026	0.6	0.8	0.049
V	0.34	0.00	0.136	0.4	0.0	0.056
X	1.33	0.85	0.107	5.3	4.3	0.130
Y	2.68	1.22	0.227	9.2	3.9	0.375*
Z	4.65	5.37	-0.086	18.3	22.7	-0.207

*moderate correlation

Table 11. Means and correlation coefficients for transition counts between each AOI pair for item 31 for high- and low-performers based on total score (r_{rb} = rank biserial correlation coefficient). Arrows indicate the direction of each transition.

From-To AOI Pair	Mean AOI Transition Count		r_{rb}
	High Group	Low Group	
U → V	0.1	0.0	0.056
U → X	0.1	0.0	0.056
U → Y	0.1	0.0	0.112
U → Z	0.0	0.1	-0.112
V → U	0.0	0.0	0.000
V → X	0.1	0.0	0.056
V → Y	0.0	0.0	0.000
V → Z	0.1	0.0	0.056
X → U	0.1	0.0	0.112
X → V	0.0	0.0	0.000
X → Y	1.2	0.4	0.340*
X → Z	0.5	0.7	-0.226
Y → U	0.1	0.0	0.112
Y → V	0.0	0.0	0.000
Y → X	1.2	0.8	0.200
Y → Z	0.7	0.8	-0.068
Z → U	0.0	0.1	-0.112
Z → V	0.0	0.0	0.000
Z → X	0.4	0.2	0.080
Z → Y	0.8	0.8	-0.056
U ↔ V	0.1	0.0	0.056
U ↔ X	0.2	0.0	0.168
U ↔ Y	0.2	0.0	0.168
U ↔ Z	0.0	0.2	-0.112
V ↔ X	0.1	0.0	0.056
V ↔ Y	0.0	0.0	0.000
V ↔ Z	0.1	0.0	0.056
X ↔ Y	2.4	1.2	0.305*
X ↔ Z	0.9	0.9	-0.137
Y ↔ Z	1.5	1.6	-0.037

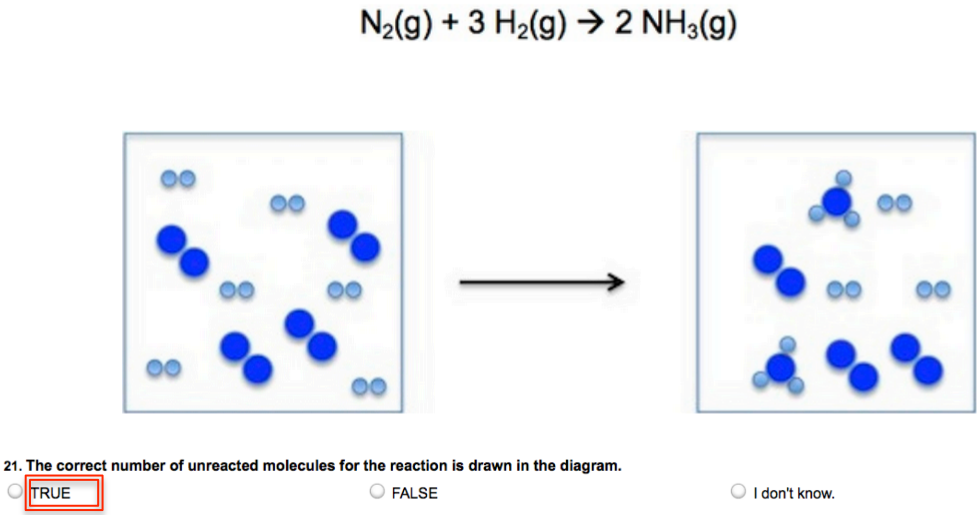
* moderate correlation

Items pertaining to unreacted molecules

Two of the six items pertaining to unreacted molecules produced significant differences between the visual behaviors of high- and low-performing groups. Fixations on individual AOIs for item 21 (Figure 17) do not reach the threshold for any of them to be identified as being at least moderately strong, although there were some indications that high-performing participants tended to look more at AOI Y. This probably is more a result of prompting from the question than anything else, especially since high performers were also observed to have gone between Y and Z just slightly more often than low performers did. This lack of any practical significance of any difference between visual behaviors of high and low performers questions of this type was repeated with item 32 (Figure 18). Among the individual AOIs for this item, only Z showed a statistical significance that may have some practical importance. Low performers read the question for item 32 more than two seconds longer than high performers did. However, the greater focus on Z by the low performers was not tied to attention with any of the other AOIs as indicated by a lack of any significant AOI transitions. This could very have been simply a sign of instrument fatigue among some of the participants.

Table 12.Means and correlation coefficients for fixation times and fixation counts associated with areas of interest for item 21 resulting from grouping participants into high- and low-performers based on total score on the instruments (r_{pb} = point biserial correlation coefficient; r_{rb} = rank biserial correlation coefficient).

	Fixation Time (seconds)			Fixation Count		
AOI	High Group	Low Group	r_{pb}	High Group	Low Group	r_{rb}
U	0.14	0.01	0.219	0.6	0.0	0.226
V	0.12	1.00	0.179	0.3	0.1	0.012
X	1.72	1.78	0.185	6.4	5.6	0.000
Y	3.17	6.62	0.280	12.7	8.1	0.285
Z	5.93	5.98	-0.082	23.9	26.4	-0.168



$N_2(g) + 3\,H_2(g) \rightarrow 2\,NH_3(g)$

21. The correct number of unreacted molecules for the reaction is drawn in the diagram.

○ TRUE ○ FALSE ○ I don't know.

Figure 17. Item 21. The correct response is boxed in red.

Table 13.Means and correlation coefficients for transition counts between each AOI pair for item 21 for high- and low-performers based on total score (r_{rb} = rank biserial correlation coefficient). Arrows indicate the direction of each transition.

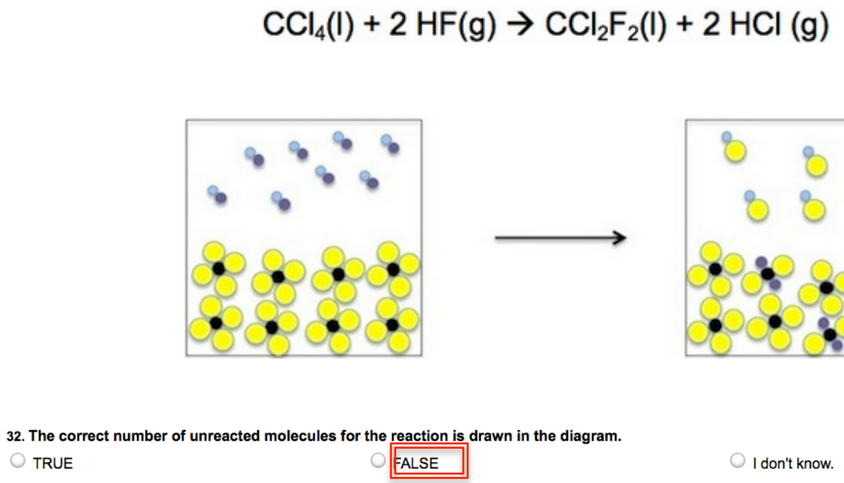
From-To AOI Pair	Mean AOI Transition Count		r_{rb}
	High Group	Low Group	
U → V	0.0	0.0	0.000
U → X	0.0	0.0	0.000
U → Y	0.1	0.0	0.056
U → Z	0.0	0.0	0.000
V → U	0.0	0.0	0.000
V → X	0.1	0.0	0.056
V → Y	0.1	0.0	0.056
V → Z	0.1	0.0	0.056
X → U	0.1	0.0	0.056
X → V	0.0	0.0	0.000
X → Y	1.1	0.7	0.124
X → Z	0.4	0.3	0.074
Y → U	0.2	0.0	0.112
Y → V	0.0	0.0	0.000
Y → X	1.0	0.4	0.168
Y → Z	1.1	0.7	0.354*
Z → U	0.0	0.0	0.000
Z → V	0.0	0.0	0.000
Z → X	0.6	0.7	0.068
Z → Y	1.2	0.8	0.279
U ⇔ V	0.0	0.0	0.000
U ⇔ X	0.1	0.0	0.056
U ⇔ Y	0.3	0.0	0.112
U ⇔ Z	0.0	0.0	0.000
V ⇔ X	0.1	0.0	0.056
V ⇔ Y	0.1	0.0	0.056
V ⇔ Z	0.1	0.0	0.056
X ⇔ Y	2.1	1.1	0.130
X ⇔ Z	1.1	1.0	-0.031
Y ⇔ Z	2.3	1.4	0.432*

* moderate correlation

Table 14.Means and correlation coefficients for fixation times and fixation counts associated with areas of interest for item 32 resulting from grouping participants into high- and low-performers based on total score on the instruments (r_{pb} = point biserial correlation coefficient; r_{rb} = rank biserial correlation coefficient).

	Fixation Time (seconds)			Fixation Count		
AOI	High Group	Low Group	r_{pb}	High Group	Low Group	r_{rb}
U	0.01	0.11	-0.308*	0.1	0.6	-0.168
V	0.00	0.09	-0.377*	0.0	0.4	-0.226
X	0.94	0.55	0.166	3.8	2.8	0.118
Y	2.16	1.09	0.228	8.4	4.1	0.265
Z	3.82	6.12	-0.318*	18.3	24.3	-0.292

* moderate correlation



$CCl_4(l) + 2\ HF(g) \rightarrow CCl_2F_2(l) + 2\ HCl\ (g)$

32. The correct number of unreacted molecules for the reaction is drawn in the diagram.

○ TRUE ○ FALSE ○ I don't know.

Figure 18. Item 32. The correct response is boxed in red.

Table 15.Means and correlation coefficients for transition counts between each AOI pair for item 32 for high- and low-performers based on total score (r_{rb} = rank biserial correlation coefficient). Arrows indicate the direction of each transition.

From-To AOI Pair	Mean AOI Transition Count		r_{rb}
	High Group	Low Group	
U → V	0.0	0.0	0.000
U → X	0.0	0.0	0.000
U → Y	0.0	0.0	0.000
U → Z	0.0	0.0	0.000
V → U	0.0	0.0	0.000
V → X	0.0	0.0	0.000
V → Y	0.0	0.1	-0.112
V → Z	0.0	0.0	0.000
X → U	0.0	0.1	-0.112
X → V	0.0	0.0	0.000
X → Y	0.6	0.2	0.252
X → Z	0.6	0.6	0.049
Y → U	0.0	0.0	0.000
Y → V	0.0	0.1	-0.112
Y → X	0.4	0.6	-0.037
Y → Z	0.7	0.8	0.087
Z → U	0.0	0.1	-0.112
Z → V	0.0	0.0	0.000
Z → X	0.4	0.1	0.292
Z → Y	0.8	1.0	-0.068
U ⇔ V	0.0	0.0	0.000
U ⇔ X	0.0	0.1	-0.111
U ⇔ Y	0.0	0.0	0.000
U ⇔ Z	0.0	0.1	-0.112
V ⇔ X	0.0	0.0	0.000
V ⇔ Y	0.0	0.2	-0.112
V ⇔ Z	0.0	0.0	0.000
X ⇔ Y	1.1	0.8	0.137
X ⇔ Z	1.1	0.7	0.226
Y ⇔ Z	1.4	1.8	0.000

Visual Steps Taken By Participants

Retrospective think-alouds

To understand how participants coordinated the different AOIs for each item, they were each shown a playback of their gaze video and asked to *think aloud* during the playback. This procedure is known as *cued retrospective think-aloud* (RTA) during which a participant uses the playback of his or her gaze video to be reminded of thought processes that occurred as each participant pondered each item on the instrument. The use of a think-aloud protocol with eye tracking allows the triangulation of data about cognitive processing that occurred as each subject examined visual stimuli (Jarodzka, Scheiter, Gerjets, & Van Gog, 2010). The alternative to an RTA is to do what is known as a *concurrent think-aloud* (CTA), in which a participant verbalizes his or her thought processes as items are responded to and eye movements are being recorded (Holmqvist et al., 2011).

One drawback of the CTA is that verbal utterances have long been suspected to affect eye movements during the execution of a task (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). As visual stimuli are described, the planning of speech is a process in itself, which thus, requires more time on the part of the subject and modifies eye behavior (Holsanova, 2006). The think-aloud process takes away some of the resources from the cognitive system so that although thinking aloud may not change the way in which a task is performed, doing a CTA may slow down not just the eye movement but the general processes and learning as well (Nielsen, Clemmensen, & Yssing, 2002). Advantages of the CTA, on the other hand, include being able to record two data sources at the same time, which yields data that are very closely linked to each other. CTA also provides the momentous perspective, that is, being able to explore what thought processes occur in light of all other processes taking place at the very moment of interest (Holmqvist et al., 2011).

The separation of eye movement recording during the performance of a task from verbalizations in an RTA risks losing some detail from memory. How much detail is lost remains unclear since for the most part researchers agree that participants' memories about what they did earlier remain intact. Participants will often look at objects during the gaze replay in roughly the same order they did during the actual task performance (Guan, Lee, Cuddihy, & Ramey, 2006). RTAs also generally result in more detailed descriptions as the gaze video is played back to the subject compared to an uncued verbalization (Van Gog, Paas, van Merriënboer, & Witte, 2005). They give more details about actions done and how each step was performed. Cued RTAs have also been observed to yield more comments about a subject's cognitive processes while CTAs are more focused on manipulations (Hyrskykari, Ovaska, Majaranta, Räihä, & Lehtinen, 2008). Verbalizations from CTAs usually include more action and outcome statements, while RTAs are more about strategies and reasons for actions.

Participants were shown gaze replays of how they visually examined each item in the same order the items were given to them during testing. Each RTA lasted between 20 and 35 minutes and was mainly focused on asking participants to explain what they were attempting to do as they visually coordinated the different AOIs on each stimulus: *“What were you doing as your eyes moved up and down between _______ and ______?”*

Interviews were then transcribed and open-coded in terms of thought processes attributed by participants to sets of eye movements they made during their analysis of each item.

Atom-to-formula matching

When participants saw the first item for each context, one of their first concerns was to match each of the different circles in the PNOM diagram with the correct element. Corollary to this is their need to match each of the groups of circles in the diagram with corresponding formulas of substances involved in the given balanced equation. Participants needed to do this to make sure that they were thinking of the correct element or molecule as they examine each of the circles or groups of circles, respectively, in the diagram. Consider, for example, student 18 from Chem D, as she tried to put together which circles were which for the first ammonia context (emphasis added):

> Participant: *I was trying to visualize the reaction equation with the description of having four nitrogen gas molecules and six hydrogen gas molecules and 50% completion of the reaction.*
>
> Interviewer: *Okay.*
>
> Participant: *All of that.* ***All those words compared to the visual diagrams that are trying to line that up.***
>
> Interviewer: *Okay.* ***So I see some up and down movements usually on the same part of the diagram and the reaction*** *like if you were looking at the left side of*

the diagram, you were also looking at the left side of the reaction. Were you trying to do something there?

Participant: *Between looking at the picture and looking at the reaction?*

Interviewer: *Yes.*

Participant: *Yeah.* ***I was just trying to match up like where it's says N_2, where it was trying to say, okay, this is N_2 or this is H_2*** *and just counting the molecules and trying to compare that to what was stated in the diagram.*

Interviewer: *When you say you match up N_2 with N_2, what does that mean?*

Participant: *Just like as opposed to they are just shapes with no labels trying to like label it in my head have like, okay, this is –* ***like this is a nitrogen or that's a hydrogen, just trying to put the two together in my mind, the text information with the visual information.***

Interviewer: *When did you finally say, okay, I am ready to move on to the first question?*

Participant: *Kind of just* ***when I put all of that together when I was sure that I was looking at the molecules that I was sure nitrogen and what I was sure was the hydrogen and then looking and trying to make sense on my head of the diagram on the right****, but that was what was the completed, like those were the products I'm just trying to –* ***when I had it all together in my head that I was -- at least semi-confident that I knew what I was looking at without having to constantly glance at the reaction*** *which I kind of did anyway. Then I just thought like I was okay with it, I guess.*

It can be inferred from the above excerpt that the matching between circles in the diagrams and formulas in the equation occurred as participants looked up and down between these AOIs. Often, the vertical eye movements would go from one side of the equation to the same side of the diagram. Usually, this process would occur in two steps, with participants first matching each differently colored circle with the correct element, and then they would move on to match groups of circles with specific formulas in the equation: "*Oh, these red are oxygens, this grouping of black with the blue is the* CH_4." Specifically for the for first page for each context, there would also be fixations on the text at the middle, where the correspondences between colored circles and elements were explicitly stated. Participants generally stated that these were all necessary steps for them to take before they were able to move on to responding to the first item for each context. The check with the diagram and the equation though was not limited to each of the first items. In moving on to the succeeding items, participants were observed to initially look back at both the equation and the diagram:

> Interviewer: *Okay. Moving on to the second question, what are you doing here?*
>
> Participant: *Kind of the same thing. Looking at the question and then looking at the equation to see what I would think in my mind I should be able to expect in the diagram and then looking at the diagram, again at what we started with and what we ended with.*

Periodic checks with the equation and the diagram were performed by participants generally to remind themselves of the correct representations between formulas and groups of colored circles. A schematic diagram of the directions in which eye movements associated with this process is given in Figure 19.

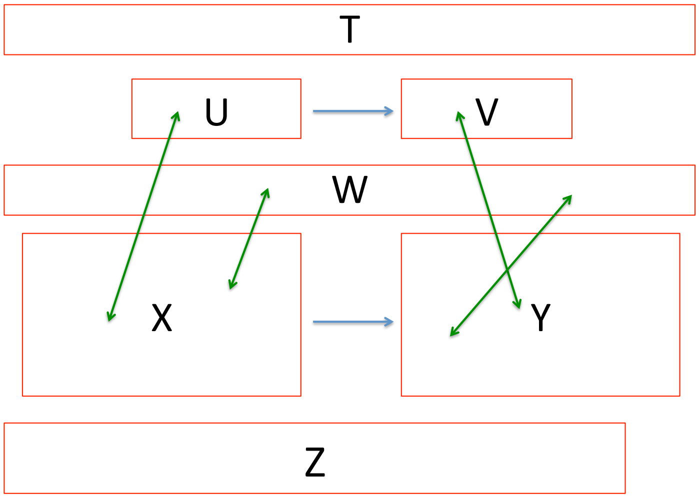


Figure 19. Schematic diagram of eye movements associated with matching between circles drawn in the PNOM diagram with the correct species in the balanced equation.

Atom/molecule counting

Another step that accompanied many eye movements among participants in this study was the counting of atoms and/or molecules (Figure 20). At least initially, most of the counting of the atoms and molecules occurred with participants moving their eyes from one atom (or molecule) to the next within the same box:

> Interviewer: *Okay. All right, moving on to the next question. Tell me what you're doing. What are you doing there looking at those atoms individually on the left side?*
>
> Participant: *Counting them, so then I can go to the next box and count them*.

Counting usually occurred first on one side of the diagram, and then moved to the other side. Participants then usually compared the numbers of atoms of each kind on the two sides of the diagram: "*Counting the ones in the first box, so I can go to the second box* *and see if they're the same number represented*." Counting was more often done on a molecule-by-molecule basis although there were some instances where individual atoms were counted off as well. Among the purposes of counting atoms according to participants were to determine:

- whether there were equal numbers of each atom on both sides of the diagram ("*I was counting how many of the dots were on the left but then were also on the right, so and like a perfect equation they should all be showing up on the other side…*");
- ratios between numbers of molecules shown in the diagram ("…*and I was counting the molecules, like finding out kind of what the ratios of them were."*);
- limiting reagents based on the amounts of reagents ("*Probably to decide what would be the limiting reagent, because by counting the least amount.*");
- numbers of unreacted reactants ("… *looking at what we started with for the products and just counting and trying to see what I think should react with what and how much I could expect to have left over.*"); and
- percent yields illustrated by each diagram ("*Well, the 50% completion. So I figured that, of the nitrogens there's eight on the left and four on the right. So 50% of them were being used in this completion and only four of them would be shown.*").

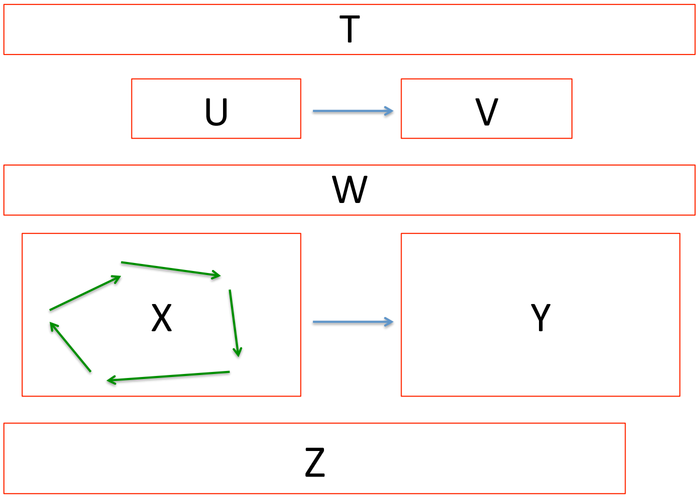


Figure 20. Schematic diagram of eye movements associated with counting atoms and/or molecules in a PNOM diagram.

Atom-to-atom matching

Unlike atom-to-formula matching, which occurs between the PNOM diagram and the balanced equation, atom-to-atom matching (Figure 21) occurred between the two sides of the diagram. Generally, the main purpose of participants in going from one side of the diagram to the other was to figure out just where among the products could an atom from a specific reactant have ended up: *"I looked at the diagram and then I'd be like* CH_4*, find it, and then I looked at the other side and I tried to find the same* CH_4*, see if it's there and then check back with myself on the color and dots and what they're matched with."* The reverse may also be true, that is, participants may have also been interested in where each atom on the product side of the diagram could have possibly come from among the reactants: *"I'm trying to, I guess, figure out where the different elements and molecules on the right came from the left."* Often this step was part of making sure that there were the same number of atoms of each kind on both sides of the diagram. This step probably resulted more from the fact that the instrument was administered online. Participants were, of course, not allowed to mark individual atoms or molecules on the monitor to make sure they had accounted the different species on the diagram correctly, although pen and paper were available for them to use: “*I think, just when I kind of had read through the equation, then I looked back and again trying to mentally group them and see if I saw that grouping on the right side.*”

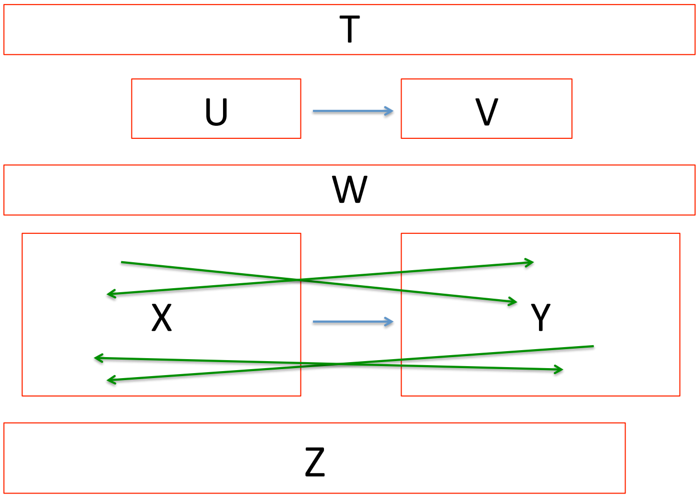


Figure 21. Schematic diagram of eye movements associated with atom-atom matching across both sides of the PNOM diagram.

Balancing atoms across the PNOM diagram

Balancing atoms across the diagram (Figure 22) can probably be viewed as a composite of the atom-to-atom matching and counting steps described earlier. Participants who have described going through this step often referred to checking whether the number of each type of atoms on the two sides of the PNOM diagram were identical or not:

> *“By looking at the numbers and the corresponding molecule color, I was going back and forth and counting up there and counting down there to make sure they were the same.”*

Often participants would count molecules on the left side of the diagram first, and then switch over to counting individual atoms on the right side as a result of the reconfiguration of the atoms drawn to represent the product molecules that have been formed:

> *"On the left, I was going molecule-by-molecule and then on right, atom-by-atom because they're all split up, you couldn't – like the red's I did, I count two, four, six, eight and then on the other side, I probably started with the black because that was all like, well, no, I started with the red and then I went to the black because there was only three blacks, and so it was kind of easy for me to see, but on the left, it was easier for me to see the pairs rather than them all split up."*

Checking the balance between atoms in the diagram most commonly occurred when participants saw a new context for the first time, either with the first item for students from Chem B or the page before the first item for those who came from Chem C. It was not uncommon, however, for participants to do a final rebalancing of the atoms when they got to the last item for a context (emphasis added):

> *"I wanted to check one last time. I guess the main thing would be that there is the correct number of molecules in both the reactant side and the product side and that **nothing is lost in the diagram**."*

It was also clear from what some participants said that adherence to the law of conservation of mass as far as the diagram was concerned was important to them: "*I saw all the dots there.*" There were times when participants anchored their decisions on the correctness (or lack of it) of the diagram on whether the same number of atoms for each element was seen on both sides of the diagram (emphasis added):

> Participant: … *the diagram appeared to not have been done correctly because we were left without unreacted uh, substances. It should have been included in the diagram there.*
>
> Interviewer: *So, when you say you were left without unreacted substances…*
>
> Participant: *Right, yeah*
>
> Interviewer: *that should have been included, where, in which part of the diagram…*
>
> Participant: *In the right part of the diagram, there should have,* ***you see that there's only four nitrogen molecules (ammonia) in the right part of the diagram where in beginning we have eight total nitrogen molecules (atoms) so some of them have disappeared.***

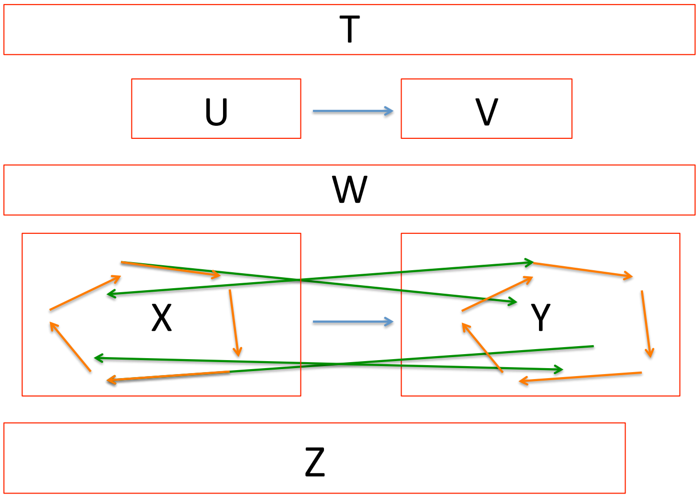


Figure 22. Schematic diagram of eye movements associated with balancing the numbers of atoms across both sides of the PNOM diagram.

Determination of ratios among molecules

Ratios between numbers of molecules in the diagrams appeared to have helped some participants decide whether diagrams were drawn correctly or not as well:

> *"For this one, I guess I was looking at again like the ratios between water and carbon dioxide. There were two water for every carbon dioxide and there were four water and two carbon dioxide, so I thought was okay."*

Often this was done by going up and down between the diagram and the equation after numbers of molecules on each side of the diagram have been determined (Figure 33). Sometimes checks were done with the other side of the equation and the diagram were done as well:

> *"So I guess I'm just paying attention to the coefficients of one of the molecules and just comparing those to what I see in the picture."*

Ratios between molecules in the diagram were sometimes compared by participants with ratios between the coefficients in the balanced equation to determine which reactant was limiting:

> Interviewer: *How did you decide that H_2 was the limiting um, reactant?*
>
> Participant: *Because it requires more moles of H_2 compared to moles of nitrogen or N_2 to create NH_3 and seeing as we had six moles of H_2 and four moles of N_2 even though it shows one of N_2 to three of H_2 you can see that we have less H_2 than would be required to react completely with all of the nitrogen.*

Sometimes the ratios were used to identify excess reagents as well:

> "*Because when I, um, tried to combine them in the reactant picture, I just followed three* H_2*, which they provide, then one* N_2 *so there's six there, that should be combined with the two* N_2*s and there would be left over two of the* N_2*.*"

However, these ratios did not always help participants make their decisions correctly. This was especially true when they sometimes saw that the ratios between molecules drawn in the diagram were not identical with those given by the coefficients in the balanced equation:

> "*I was trying to see if the ratio* (*sic*) *were correct. But then the whole thing that confused me was, uh there was like three* CH_4 (*in the diagram*) *but the reaction just shows one*."

Especially among the reactants, there were participants who thought the ratios between numbers of molecules drawn in the diagram should reflect ratios between coefficients in the balanced equation. This was observed when one participant was asked about how he came up with the decision that the first diagram drawn for the methane combustion reaction was incorrect:

> Interviewer: *Okay. What about on deciding whether the diagram is correct or not?*
>
> Participant: *Probably I think I just went back on the ratios and thought, the ratio on the left side is just not matching up* (*with what is in the equation*)*, so therefore I concluded it must be incorrect.*

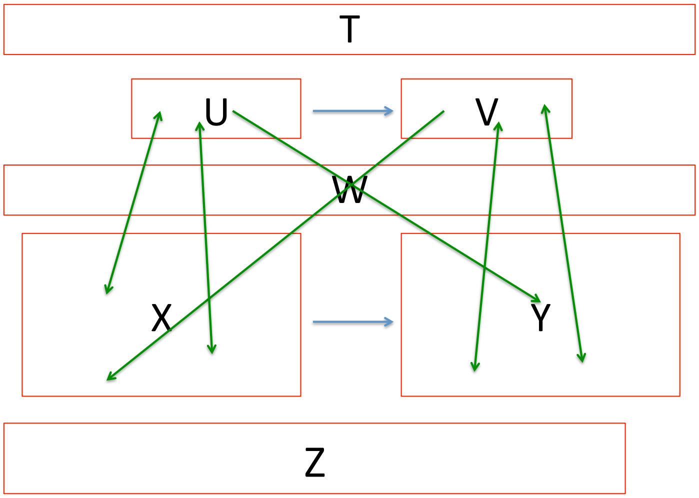


Figure 23. Schematic diagram of eye movements associated with comparisons of ratios between substances as indicated by coefficients in the balanced equation and numbers of molecules in the PNOM diagram.

Summary and Implications

This study explored the differences in fixation durations and frequencies on different areas of interest between high- and low-performing students asked to respond to questions that dealt with stoichiometry concepts with the use of PNOM diagrams. High-performing participants were observed to allocate monotonically decreasing lengths of time fixating on the two sides of the PNOM diagram given for each context as they moved from one item to the next in each series. No such pattern was observed among the low-performing students. One way to explain this observation is to attribute a better working memory capacities to the high-performing participants than their low-performing counterparts. High-performing students also appeared to have examined diagram AOIs longer and more frequently than the low-performing students did. This was most especially true on items that dealt with the correct choice of the limiting reagent on which each diagram was based. This served to indicate how high-performing students selected information from the PNOM diagrams to be the most important for each item. Diagram AOIs were also observed to be most commonly involved in AOI transitions that yielded statistically significant differences between high- and low-performing students. High-performing students again seem to have chosen to compare information from the diagram AOIs most frequently with those obtained from other AOIs. Low-performing students, on the other hand, tended to focus their visual attention on the question.

One similarity between high- and low-performing students was the short amount of fixation times with reaction AOIs. This probably came from greater familiarity on the part of both groups of students with obtaining information from text rather than from a PNOM diagram. Two of the three chemical reactions used in the instrument were probably familiar to many students (methane combustion and production of ammonia). There is also the fact that a sentential (left-to-right) reading of a chemical equation may at least give the student a rudimentary understanding of which elements are involved in the reaction, what reactants are mixed up and which products are formed, as well as stoichiometric ratios between them. Left-to-right viewing of PNOM does not ordinarily yield the same information. Students have had to first determine the specific key to the diagram being used in terms of color and size of the spheres, count atoms and/or molecules of different kinds on each side, determine how each atom on the reactant side of the diagram might have ended on the product side, and compare ratios both between the two sides of the diagram, and between the diagram and the chemical equation. These cognitive processes were described by students during retrospective think-alouds that took place immediately after each student’s eye tracking activity.

An important caveat with regards to all of the interpretations above is how the use of qualitative interviews with eye tracking imposes a practical limit to the number of participants for this study. The use of retrospective think-aloud to determine exactly how students were putting together information from the different AOIs limited the number of participants to the minimum required to obtain saturation. In some ways, differences observed between high- and low-performing students may not be generalized to the entire general chemistry student population since other factors that would have been observed with a larger number of participants were probably missed. On the other hand, the fact that the instrument made use of the same set of questions through the different chemical contexts may have allowed the identification of repeating visual behavioral patterns that are probably worth exploring with a bigger sample size.

REFERENCES

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research, 25*(17), 3389-3402.

Ben-Zvi, R., Eylon, B., & Silberstein, J. (1987). Students' visualisation of a chemical reaction. *Educ. Chem., 24*, 117-120.

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational psychology review, 3*(3), 149-210.

Cormen, T. H. (2009). *Introduction to algorithms*: MIT press.

Cristino, F., Mathôt, S., Theeuwes, J., & Gilchrist, I. D. (2010). ScanMatch: A novel method for comparing fixation sequences. *Behavior Research Methods, 42*(3), 692-700.

Cureton, E. E. (1956). Rank-biserial correlation. *Psychometrika, 21*(3), 287-290.

Davidowitz, B., Chittleborough, G., & Murray, E. (2010). Student-generated submicro diagrams: A useful tool for teaching and learning chemical equations and stoichiometry. *Chemistry Education Research and Practice, 11*(3), 154-164.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological review, 102*(2), 211.

Face Outline Templates. (2015). Retrieved from http://www.presentationmagazine.com/face-outlines-template-3019.htm

Feusner, M., & Lukoff, B. (2008). Testing for statistically significant differences between groups of scan patterns. *Proceedings of the Symposium on Eye-tracking Research & Applications* 43-46.

Guan, Z., Lee, S., Cuddihy, E., & Ramey, J. (2006). *The validity of stimulated restrospective think-aloud method as measured by eye tracking.* Paper presented at the CHI 2006.

Hegarty, M., Carpenter, P. A., & Just, M. A. (1991). Diagrams in the comprehension of scientific texts.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*: Oxford University Press.

Holsanova, J. (2006). Dynamics of picture viewing and picture description. In L. Albertazzi (Ed.), *Visual Thought. The Depictive Space of the Mind* (pp. 233-254). Amsterdam, Philadelphia: John Benjamins Publishing Company.

Hyrskykari, A., Ovaska, S., Majaranta, P., Räihä, K.-J., & Lehtinen, M. (2008). Gaze path stimulation in retrospective think-aloud. *Journal of Eye Movement Research, 2*(4), 1-18.

Jarodzka, H., Scheiter, K., Gerjets, P., & Van Gog, T. (2010). In the eyes of the beholder: How experts and novices interpret dynamic stimuli. *Learning and Instruction, 20*(2), 146-154.

Josephson, S., & Holmes, M. E. (2002). Visual attention to repeated internet images: testing the scanpath theory on the world wide web. *Proceedings of the Symposium on Eye- tracking Research & Applications*, 43-49.

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational psychologist, 38*(1), 23-31.

Kalyuga, S., Chandler, P., & Sweller, J. (1998). Levels of expertise and instructional design. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 40*(1), 1-17.

Knijnenburg, T. A., Wessels, L. F. A., Reinders, M. J. T., & Shmulevich, I. (2009). Fewer permutations, more accurate P-values. *Bioinformatics, 25*(12), i161-i168.

Kozma, R. (2003). The material features of multiple representations and their cognitive and social affordances for science understanding. *Learning and Instruction, 13*(2), 205-226.

Lai, Y. (2007). Conservative adjustment of permutation p-values when the number of permutations is limited. *International journal of bioinformatics research and applications, 3*(4), 536-546.

Levenshtein, V. I. (1966). *Binary codes capable of correcting deletions, insertions, and reversals*.

Liu, X., Gale, A., & Song, T. (2007, 2007). *Detection of terrorist threats in air passenger luggage: Expertise development*.

Mason, L., Pluchino, P., Tornatora, M. C., & Ariasi, N. (2013). An eye-tracking study of learning from science text with concrete and abstract illustrations. *The Journal of Experimental Education, 81*(3), 356-384.

Needleman, S. B., & Wunsch, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology, 48*(3), 443-453.

Nielsen, J., Clemmensen, T., & Yssing, C. (2002). *Getting access to what goes on in people's heads? Reflections on the think-aloud technique.* Paper presented at the Second Nordic Conference on Human-Computer Interaction, New York.

Okuda, T., Tanaka, E., & Kasai, T. (1976). A method for the correction of garbled words based on the Levenshtein metric. *Computers, IEEE Transactions on, 100*(2), 172-178.

Python Software Foundation. (2014). Python (Version 3.4.3). Beaverton, OR: Python Software Foundation.

Rosengrant, D. (2010). Gaze scribing in physics problem solving. 45-48.

Sanger, M. J. (2005). Evaluating students' conceptual understanding of balanced equations and stoichiometric ratios using a particulate drawing. *Journal of Chemical Education, 82*(1), 131.

Schmidt-Weigand, F., Kohnert, A., & Glowalla, U. (2010). A closer look at split visual attention in system-and self-paced instruction in multimedia learning. *Learning and Instruction, 20*(2), 100-110.

Schwonke, R., Berthold, K., & Renkl, A. (2009). How multiple external representations are used and how they can be made more useful. *Applied Cognitive Psychology, 23*(9), 1227-1243.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*(5217), 1632-1634.

Tang, H., Topczewski, J. J., Topczewski, A. M., & Pienta, N. J. (2012). Permutation test for groups of scanpaths using normalized Levenshtein distances and application in NMR questions. *Proceedings of the Symposium on Eye Tracking Research and Applications*, 169-172.

Tate, R. F. (1954). Correlation between a discrete and a continuous variable. Point-biserial correlation. *The Annals of mathematical statistics*, 603-607.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological review, 80*(5), 352.

Van Gog, T., Jarodzka, H., Scheiter, K., Gerjets, P., & Paas, F. (2009). Attention guidance during example study via the model’s eye movements. *Computers in Human Behavior, 25*(3), 785-791.

Van Gog, T., Paas, F., van Merriënboer, J. J. G., & Witte, P. (2005). Uncovering the problem-solving process: cued retrospective reporting versus concurrent and retrospective reporting. *Journal of Experimental Psychology: Applied, 11*(4), 237.

Wagner, R. A., & Lowrance, R. (1975). An extension of the string-to-string correction problem. *Journal of the ACM (JACM), 22*(2), 177-183.

Weber, N., & Brewer, N. (2003). Expert memory: The interaction of stimulus structure, attention, and expertise. *Applied Cognitive Psychology, 17*(3), 295-308.

Wolfram Research, I. (2014). Mathematica. Champaign, Illinois: Wolfram Research, Inc.

CHAPTER 5. CONCLUSION AND IMPLICATIONS

This chapter summarizes the overall conclusions determined from all findings present in Chapters 2 through 4 from the perspective of the underlying theories discussed in Chapter 1. Implications of these conclusions in terms of chemistry education research and classroom instruction in a general chemistry setting will be discussed. Suggestion for future work based on the findings given in this dissertation will also be listed.

Summary of Research Findings

The goals of this research project were to:

1. Identify and explain students' understandings of the concepts of excess and limiting reagents, and yield based on how information from symbolic and microscopic representations are selected, coordinated and integrated by students;
2. Determine the extent to which misconceptions on these concepts occur among first-year chemistry students through a large-scale administration of an online instrument;
3. Distinguish between the visual behaviors of high- and low-performing students tasked with solving problems in stoichiometry that make use of PNOM diagrams; and

4. Identify cognitive processes students used in association with specific types of visual behavior as students coordinated PNOM diagrams to respond to questions on limiting reagents and reaction yield.

Chapter 2 discussed the role of the dual processing theory of reasoning in how students came up (or failed to come up) with PNOM diagrams to illustrate specific contexts on limiting reagents and reaction yield. It was observed that in executing both tasks assigned to subjects, several subjects used a heuristic-analytic sequence, where students would first use a heuristic to select information on which to focus their attention, and then analyze this selected information to come up with a final solution to the problem. Several students were, for example, observed to have used the factor-label method to identify which of two reactants would be limiting the amount of products formed and determine the number of product molecules formed. They then drew their diagrams guided by these calculations, in some instances, carefully accounting for each atom drawn to make sure the same number of each type of atoms were drawn on each side of the PNOM diagram. Several students who used this approach were also observed to have failed to follow up the heuristic stage with an analytic stage as indicated by their failure to come up with an appropriate diagram for the given chemical context. In fact, a few students simply failed to draw diagrams completely. Others incorrectly chose the reactant present in the smaller amount as their limiting reagent during the heuristic stage, which ultimately led them to draw incorrect diagrams later. A few failed to account for leftover reactant molecules to maintain the balance of atoms between the two sides of their diagrams. Some students, for instance, who started their solutions for Task 2 by using the given 50% reaction yield did so by taking half of the numbers of reactant molecules given, allowed these to react, and then completely neglected the other half of the molecules.

The subjects who had most success in drawing appropriate diagrams for tasks discussed in Chapter 2 were those who went directly into the analytic stage, or went through at least some cognitive dissonance as they started their sketching their illustrations following the application of a heuristic. Several subjects immediately broke down each reactant molecule into its component atoms and then recombined those to form sets of product molecules based on ratios indicated by the balanced equation. A few others broke the entire ensemble of reactant molecules drawn in the diagram into smaller sets whose compositions were based on ratios determined from the given balanced equation. In the case of Task 1, this always led to an accurate determination of the excess molecule. For Task 2, the successful subjects went ahead to apply the reaction yield to determine just how many molecules of the limiting reactant will react or to figure out the number of product molecules formed. Everything else was treated as leftover.

The above differences in the actions between subjects who had successfully drawn diagrams for each task and those who failed to do so are all consistent with ideas espoused by the dual processing theory. In most cases, a heuristic was used to select information for further processing during the analytic stage of problem solving. Those who successfully came up with diagrams either applied the correct heuristics to choose information relevant to each problem and then correctly analyzed these information, or showed signs of completely skipping the heuristic stage in favor of analysis. Unsuccessful students, on the other hand, either used the wrong heuristic (and, thus, chose irrelevant information) or failed to follow through with a correct analysis of the information obtained from the heuristic stage of the solution.

Chapter 3 demonstrated the robustness of the dual processing theory when looking at the prevalence of response patterns among large samples of student populations. Large proportions of student samples seemed to have applied one kind of reasoning in responding to a question for the first of a pair of diagrams pertaining to the same chemical context and then another line of thinking to respond to the same question for the second diagram. This was seen mostly with questions on whether the diagrams were drawn using the correct limiting reactant. It was not unusual to see that, for instance, students who chose "True" as their response with respect to the first diagram of a pair, were split into almost equal proportions in terms of their responses to the same question with respect to the second diagram. Students seemed to have treated diagrams for the same chemical context independently of each other. That large fractions of the student samples thought two different PNOM diagrams for the exactly the same chemical context could both have been based on the correct choice of the limiting reactant strongly suggests the use of different thought processing systems for each diagram. This points to questions about what visual cues from PNOM diagrams influenced students' selection to use one thought processing system over the other. Response patterns that repeated over large enough segments of the student samples also pointed to the resilience and spread of some common errors such the least amount misconception in selecting the limiting reagent and failing to adhere to the law of conservation of mass when working with PNOM diagrams. These observations were certainly not unexpected in light of principles described by the dual processing theory. The design of the instrument, which is rather different from what is usually given in assessments written by most instructors, allowed the illustration of how students may have been mostly heuristics to respond to one question pertaining to a specific diagram and then switch to a more analytical mode in dealing with the same question for another diagram.

The third phase of this project was an eye tracking study that explored the visual behaviors that came with how students analyzed PNOM diagrams used to illustrate concepts of limiting reagents and reaction yields. Retrospective think-alouds (RTAs) were then used to describe cognitive processes underlying specific viewing patterns as students studied the PNOM diagrams and selected their responses to each item on the instrument.

The eye tracking study also illustrated differences between the visual behaviors of high- and low-performing subjects. High performers demonstrated indications of using more integrative approaches in studying the diagrams given for each context. These subjects tended to take as much information as they can from a diagram and then tried to decide how accurately the diagram reflects the chemical context described in the text as soon as they first saw the diagram. This kind of cognitive behavior was manifested in terms of monotonically decreasing fixation times on diagram areas of interest as high performing subjects went from one item to the next for each context.

Low performers, meanwhile, tended to treat items for the same diagram in a more compartmentalized manner. This meant that low performers viewed diagrams most likely only from the perspective specifically addressed by the item these subjects were responding to.

Items that asked whether the diagrams drawn were based on the correct limiting reagent showed statistically significant differences between high and low performers' visual behaviors more often than other types of items did. All items on the limiting reagents that resulted in a difference between high and low performers showed high performers looking longer and more frequently at the one or both sides of the PNOM diagram given for the item than low performers did. This is consistent with high performers' greater tendency to count the atoms and molecules drawn and check the balance of atoms between the two sides of the diagram.

Aside from items related to the selection of the correct limiting reagent for a diagram, the visual behavior of high performers differed significantly from that of low performers for an item pertaining to the ratio between number of reactant molecules consumed and number of product molecules formed. This observation was made only for the first diagram near the very beginning of the instrument. High performers spent significantly longer times fixating on both sides of the balanced equation and the question at the bottom of the page, and they also looked more frequently at the product side of the diagram than low performers did. Transitions between the same sides of the equation and the diagram were also more frequently observed with the high performers. It might be an indication of high-performing subjects' greater willingness to spend more mental effort to study the different parts of the visual stimulus and become familiar with what the question was asking.

The only times low-performing subjects were observed to have spent significantly longer times at any area of interest were in two items. In both instances, low performers took more time on the question AOI than high performers did. One item for which this difference was observed occurred very early in the instrument, probably indicating mental effort being spent on comprehending the question. The other instance occurred towards the end, which may be a sign of the onset of instrument fatigue or a result of increased visual complexity of the diagram.

Implications

This study demonstrated the different cognitive processes used by students with different levels of prior knowledge when constructing, interpreting, and using information from PNOM diagrams as they solved problems on limiting reagents and reaction yields. These are important findings for general chemistry instructors to keep in mind as they design materials for instruction, assessment, and simulation of chemical reactions dealing with fundamental stoichiometry concepts such as limiting reagents and yields. Instructors need to remember that most student behavior can be understood in terms of dual thinking processes used either in the construction of PNOM diagrams to illustrate chemical reactions or in evaluating the accuracy of diagrams provided in terms the way stoichiometric concepts are illustrated.

When asked to draw PNOM diagrams, students, for the most part, used heuristics such as factor-label method to select information on which to focus their attention, then shifted to analysis to either come up with their illustrations or make final decisions about PNOM diagrams. While this approach led some students to correct solutions, most subjects who started with heuristics were not quite as successful. This should not be surprising since the heuristics employed by most students are not really associated with the kind of conceptual understanding instructors hope for students to have when using PNOM diagrams.

What was found to work more consistently was for students to go directly into analytical mode. Students showed that this could be done by breaking up the given reactant molecules either (1) into sets whose compositions were derived from the balanced equation and then using those sets to form corresponding sets of product molecules, or (2) into the component atoms and then recombining these atoms again as guided by the product side of the balanced equation. It might be helpful to make these counting processes more explicit for students to follow as instructors illustrate specific concepts such as using up all of the molecules of the limiting reagent as product molecules are formed, adherence to the law of conservation of mass even as PNOM diagrams are used by accounting for leftover molecules in the product mixture, and how the percent yield of a reaction may be determined either as a fraction of the number of limiting reagent molecules that reacted or of the number of product molecules expected to be formed.

It was also observed that the type of representations used by students in drawing diagrams was not related to the success with which students came up with diagrams. Several students who have used chemical symbols to represent atoms of the elements were observed to have arrived at conceptually sound diagrams to pretty much the same extent as those who used spheres did. More importantly, the fact that more students actually drew Lewis-type structures to represent molecules in their diagrams point to their greater familiarity with this representation style than with the spheres commonly seen in textbook illustrations and assessment items.

Results from the large-scale administration of the instrument clearly show the extent to which dual-thinking processes occurred using large samples of students. Instructors need to constantly remind students that an accurately drawn PNOM diagram for a specific chemical situation needs to correctly show each and every aspect of the such a situation. For a PNOM diagram to be correct, it must: be based on the correct choice of the limiting reagent; have ratios between reactants used and products formed that are the same as those given by the balanced equation; reflect the required yield based on the fraction of limiting reagent molecules that have been transformed into products; and adhere to the law of conservation of mass by having identical numbers of atoms of each element on both sides. While these may be asked using separate items on an instrument, instructors need to tell students that these questions should be responded to in an integrated manner, and not independently of each other. It is also important to tell students that, although real-life situations are never as precise as the ones instructors use to demonstrate chemical principles, the rigor with which these principles have been established must reflected by PNOM diagrams so that the diagrams may be used to model chemical situations as close to reality as possible. This means that for the purposes of what students deal with in a typical general chemistry course, PNOM diagrams should only be completely correct or incorrect in illustrating a specific chemical context.

The eye tracking phase of the project pointed to the concept of limiting reagents as the most likely source of significant differences between the visual behaviors of high- and low-performing students. This is not surprising at all. Data presented in the earlier chapters have repeatedly demonstrated the resilience of heuristics and misconceptions students used in relation to this principle. The longer lengths of time spent and greater frequencies with which high-performing students fixated on the PNOM diagrams as they responded to limiting-reagent questions were probably good measures of these items' level of difficulty. High-performing subjects described how they counted and recounted atoms and molecules to come up with their responses. Instructors need to think about how these behaviors can be made more explicit for other students to follow.

High-performing students demonstrated their greater capacity to integrate information into their working memories as they tended to look less and less at the same AOIs in going from one item to the next when dealing with the same chemical context. This was particularly true about times spent on and frequencies looking at diagram AOIs. While it may be unreasonable to expect every student to have exactly the same abilities in incorporating visual information into their working memory, instructors can probably explore ways of helping students organize these details in a more explicit manner. The explicit demonstration, for instance, of counting techniques to account for atoms and molecules on both sides of the diagram might be useful.

Suggestions for Future Work

The use of RTAs to explore cognitive processes underlying visual behaviors of students severely limited the sizes of student samples used for the eye tracking phase of the study. It was apparent from the permutation test *p* values that some differences between the visual behaviors of high- and low-performing students could have turned out to be statistically significant had the sample sizes used been even slightly bigger. Of the 24 items analyzed, about two to four more items showed seemed worth exploring further with larger sample sizes.

It was observed during the first phase of this research that some students drew out diagrams using sets of molecules that reflected ratios given by the balanced equation. The examples used in the instruments for this study all dealt with gaseous and liquid molecules, which are both supposed to have relatively random arrangements in space. It is definitely important for students to eventually understand nuances such as these with respect to the behavior of gases and liquids. However, many students have referred to the visual complexity of the diagrams, especially as the numbers of molecules increased in later items. If the goal is to help students understand stoichiometry concepts with using diagrams, one has to wonder whether it would help to temporarily cast other theories aside, and draw gas and liquid molecules in a manner that would facilitate a better organization of students' thoughts. A pilot study on the effects of spatial arrangements among molecular representations on cognitive processes might be useful.

Many high performers from Chem B described how they tried to obtain as much information as they could even from their initial view of each diagram. Among the issues they immediately tried to resolve were which reagent was limiting, were there the correct numbers of product and leftover molecules, and were all atoms accounted for. This was done even before students started responding to the first item associated with each diagram. However, the original design of the instrument, where the first item for each diagram was given at the same time the diagram was shown, did not allow the isolation of time spent on this initial examination of the diagram for Chem B students. Even time spent on determining which colored sphere was which atom could not be determined from the original design of the instrument. On the other hand, many students from Chem C described how they used the initial view of each diagram simply to check correspondences between colored spheres and the atoms' identities. It was only after they have convinced themselves that they knew which atom was which did Chem C students move on to responding to the questions for each diagram. As a result, these differences in the way students conducted their initial examination of each diagram could not be compared. A more carefully designed study using exactly the same presentation format of the instrument for all subjects would probably lead to some hints about pieces of information subjects were actually chunking together, what kinds of visual behavior they were using, and how these influenced their behavior on the succeeding pages of the instrument.

The instrument used "true" or "false" questions with diagrams students were asked to look at from specific perspectives. Actual questions that appear in the general chemistry exams released by the ACS Examinations Institute, which many instructors use in some way, are formatted slightly differently. For instance, the PNOM diagrams used in this study's instrument made use of colored spheres, while those that appear on the ACS General Chemistry exams use different shades of black, white, and grey. For this study, students were asked to evaluate a single PNOM diagram for each question. This was done in an attempt to identify which specific aspect(s) of stoichiometry students experienced the least or most difficulty with, as well as to see how cognitive processes and visual behaviors were modified when students were asked to look at specific aspects of the diagrams. On the other hand, most items on the ACS exam ask students to select a PNOM diagram that best illustrates a specific chemical context among three or four distractors. This was not explored at all in this study. Exploring how students' visual behaviors are affected by such an item format for questions on stoichiometry that use PNOM diagrams should prove to be interesting.

Finally, the incorporation of a longitudinal aspect on the development of students' competencies with the use of PNOM diagrams as they gain more expertise on stoichiometric principles might also be worth exploring. This study has shown that, indeed, there are differences in the cognitive processes and visual behaviors with which diagrams were analyzed by students with different levels of prior knowledge. Still, a look at how these processes and behaviors change as students gain some maturity as they learn more chemical concepts would be interesting.

APPENDIX A: HIGH SCHOOL BACKGROUND AND DEMOGRAPHIC SURVEY

This survey gathers your high school background as well as demographic information as part of your participation in our study to determine problem-solving strategies used by students in solving problems encountered in different General Chemistry courses. Please answer these questions as honestly and as accurately as you can. **If you feel uncomfortable answering any or all of these questions, you may leave them blank.** When you are done with the survey, please return it to your General Chemistry laboratory TA.

High School Math and Science Coursework

For questions 1 through 10, please indicate which mathematics and science courses you took in high school by encircling the letter corresponding to the highest level at which you took each course.

	Courses	**Not taken**	**Regular**	**Honors**	**Advanced Placement**	**Post-Secondary**
1	Biology	A	B	C	D	E
2	Chemistry	A	B	C	D	E
3	Physics	A	B	C	D	E
4	Physical Science	A	B	C	D	E
5	Earth Science	A	B	C	D	E
6	Algebra	A	B	C	D	E
7	Geometry	A	B	C	D	E
8	Trigonometry	A	B	C	D	E
9	Pre-Calculus	A	B	C	D	E
10	Calculus	A	B	C	D	E

11. Please select the choice that best describes the number of hours per week spent in HIGH SCHOOL chemistry lab.

A. None
B. one to two
C. three to four
D. five to six
E. more than six

12. Please select the choice that best completes the following statement: “In HIGH SCHOOL, I was ranked academically in the _______.”

A. top 5% of my class
B. top 10% of my class
C. top 25% of my class
D. outside of the top 25% of my class
E. I do not know

Demographics

14. Please select your gender:

A. Male
B. Female

15. Please indicate your age at the time this survey is conducted: __________ years

16. Please select the choice that best completes the statement: “I consider myself as _____.”

A. White/Caucasian
B. Black/African-American
C. Native American
D. Hispanic American
E. Asian American or Pacific Islander
F. Others (including mixed ethnicities)

17. My intended major is ____.

18. I am currently registered in Chem ___________ .

ISU ID Number: ____

Email address: ____

APPENDIX B: PILOT INTERVIEW PARTICIPANT DESCRIPTIONS

Student	Course	Ethnicity	Gender	Major
Beyonce	Chem A	African-American	Female	Biochemistry
Justin	Chem E	Caucasian	Male	Chemical Engineering
Taylor	Chem E	Caucasian	Female	Chemistry
Rihanna	Chem E	African-American	Female	Chemistry
Eminem	Chem E	Caucasian	Male	Chemical Engineering
Charice	Chem E	Asian-American	Female	Biochemistry
Kelly	Chem E	Caucasian	Female	Chemical Engineering
Adam	Chem E	Caucasian	Male	Chemical Engineering
Jason	Chem A	Caucasian	Male	Undeclared Engineering
BJ	Chem A	Caucasian	Male	Microbiology
Clark	Chem E	Caucasian	Male	Chemistry
Miley	Chem E	Caucasian	Female	Chemical Engineering
Avril	Chem A	Caucasian	Female	Biochemistry
Psy	Chem E	Asian-American	Male	Chemical Engineering
Calvin	Chem A	Caucasian	Male	Materials Science
Philip	Chem A	Caucasian	Male	Animal Science
Billy	Chem A	Caucasian	Male	Undeclared Engineering
Austin	Chem A	Caucasian	Male	Biochemistry

APPENDIX C: PILOT INTERVIEW GUIDE

Greeting

Hello! My name is (Interviewer's name). I am a graduate student here in the Department of Chemistry at Iowa State University. Thank you so much for your help. Today you will be participating in a study on solving problems in general chemistry. Its goal is to find ways students solve stoichiometry problems. If at any point you decide you do not want to continue that is your choice and you are free to stop and we will end the interview. Do you have any questions before I begin the instructions?

Subject: (Responds.)

Interviewer: So I see that you are a (state Subject's major). How is that going so far?

Subject: (Responds.)

Interviewer: How did you come about registering in Chem ___?

Subject: (Responds.)

Interviewer: How are you doing in chemistry class? How do you find being in it so far?

Subject: (Responds.)

Informed Consent Document

Interviewer: Before we proceed, I need you to sign this Informed Consent Form, which states that: a) you have voluntarily agreed to come here and be interviewed for the purposes of this study, b) that what happens here will in no way affect your grade in chemistry, and c) that all the information I collect here will not be linked to your identity, but will be used in combination with information I collect from other students.

If you want, you may go through the document, and if you agree to continue with the interview, I need you to write your name and sign on the indicated space. Do you have any questions? (Interviewer hands the Informed Consent Form to the Subject. Subject signs the form. If the Subject signs the consent form, the Interviewer shall start recording the session with both the Livescribe pen and the back-up recorder.)

Interviewer: Today is (states the **day, date, and time** of interview). This is the interview with (states Subject's pseudonym) on solving stoichiometry problems.

Study Description

Interviewer: I am going to ask you to solve a few chemistry problems for me. I need you to describe things that you are thinking about so that I can listen to you as you go along solving these problems. I am not interested in the answer you come up with as I am with how you think about the different tasks needed to solve each problem. I am going to hand you each task, which you would then read aloud to me. Then you would go ahead and try to solve each problem, writing and saying things out loud as they come to your mind. From time to time, I will ask you something like "What are you thinking?" Other times I might ask you to clarify things that you say by saying "I'm not sure I understand what you are saying." Do you have any questions about this procedure?

Subject: (Responds.)

Equipment Description

Interviewer: This is a pen that records everything you say as you write things down into this special notebook that comes with the pen. (Interviewer hands over Livescribe pen and notebook to Subject.) It is important that you try to speak loudly enough so that the pen can record everything you say and that you write as much of your thoughts as they come to your mind into the notebook. Is all of this clear so far?

Subject: (Responds.)

Interviewer: We have a copy of the textbook used in your class, which you may use to help you solve problems. Here's a calculator for you to use, too, if you see the need for it. (Interviewer hands over textbook and calculator to Subject.) Any questions at this point?

Subject: (Responds.)

Practice Task

Interviewer: Okay, now that we've cleared everything, here's the first problem for you to solve. Can you please read it out to me? (Hands over a copy of the practice task to the subject.)

Subject: (Reads the practice task.) Consider the reaction: $CCl_4(l) + 2\ HF(g) \rightarrow CCl_2F_2(l) + 2\ HCl(g)$. When 4.0 mol of CCl_4 reacts with an excess of HF, about 1.5 mol of CCl_2F_2 is obtained. What is percent yield for the reaction?

Interviewer: Go ahead, write and tell me what you are thinking about.

(Subject writes and describes a solution to the Practice Task.)

Task 1

Interviewer: That was good. Let's move on to the next problem. Can you please read it out to me?

(Hands over a copy of Task 1 to the Subject.)

Subject: (Reads Task 1.) Given the equation: $CH_4(g) + 2\ O_2(g) \rightarrow CO_2(g) + 2\ H_2O(g)$. Draw a diagram representing what would happen if 3 molecules of CH_4 and 4 molecules of O_2 were allowed to react completely.

Probing Questions

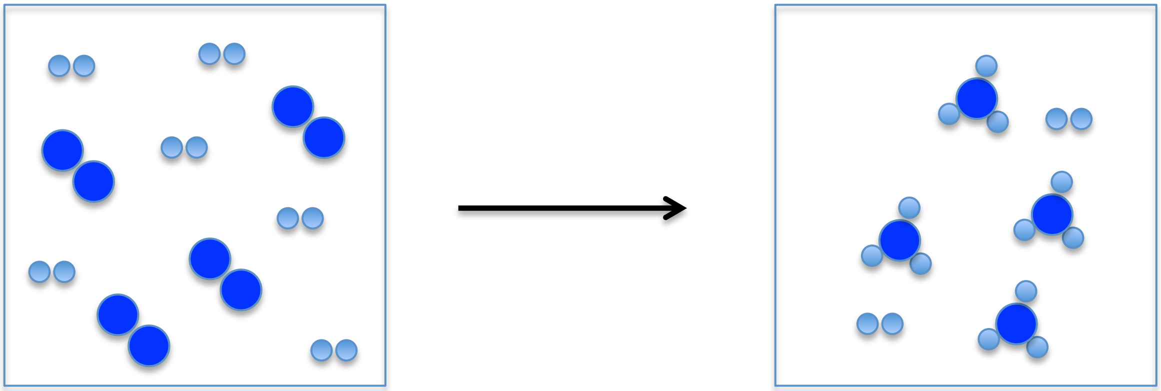
What does each of the symbols you have drawn represent?

Describe to me how you came up with your diagram.

Task 2

Interviewer: Let's move on to the next problem. Can you please read it out to me? (Hands over a copy of Task 2 to the Subject.)

Subject: (Reads Task 2.) Consider the following diagram, where each shaded circle is a hydrogen atom and each unshaded circle is a nitrogen atom. If the reaction has a 50% yield, how would you change the diagram?



Probing Question

Describe to me exactly how you came up with the changes to the diagram.

Closing

Interviewer: Thank you very much for the help you have given me today. You have been good about saying out loud what you were thinking as you solved these problems. Would you be willing to participate in a similar interview in the future?

Subject: (Responds “Yes” or “No.”)

APPENDIX D: ONLINE INSTRUMENT PAGES

INTRODUCTION

The purpose of this study is to determine misconceptions students have in solving problems in General Chemistry.

DESCRIPTION OF PROCEDURES

If you agree to participate, you will be asked to complete an online questionnaire that will be used to identify misconceptions students may have when dealing with problems in general chemistry. You will then be asked to give your demographic information. We shall also ask the office of the registrar to give us the ACT scores of students who agree to participate in this study. There will be clinical interviews with a few students who will be given invites. Your responses and information will be combined with those coming from other students. The investigators will not focus on the results from any individual student.

CONFIDENTIALITY

Records identifying participants will be kept confidential to the extent permitted by applicable laws and regulations and will not be made publicly available. However, federal government regulatory agencies (e.g., NIH), auditing departments of Iowa State University, and the Institutional Review Board (a committee that reviews and approves human subject research studies) may inspect and/or copy your records for quality assurance and data analysis. These records may contain private information.

To ensure confidentiality to the extent permitted by law, the following measures will be taken: all records will be kept in a secure room and any electronic information will be stored on a password-protected computer. If the results are published, your identity will not be associated with the publication.

RISKS

There are no foreseeable risks associated with your participation in this study. Your responses will be combined, and your name will not be attached to any information generated by this study. Your instructors in Chem 167 will not be aware of your decision to allow (or not allow) your data to be used until after grades for the class have been submitted.

BENEFITS

If you decide to participate in this study there may be no direct benefit to you. There may be extra credit assigned to completing one or more components of the study. It is hoped that the information gained in this study will benefit the chemistry learning community by informing instructors about the strategies students use in solving problems in general chemistry, as well as the misconceptions such students may have.

COSTS AND COMPENSATION

You will not have any costs from participating in this study. You will not be financially compensated for participating in this study. You may receive free food.

PARTICIPANT RIGHTS

Your participation in this study is completely voluntary and you may refuse to participate or decide to leave the study at any time. If you decide not to participate in the study or to leave the study early, it will not result in any penalty or loss of benefits to which you are otherwise entitled.

QUESTIONS OR PROBLEMS

You are encouraged to ask questions at any time during this study. For further information contact:

Mr. John Baluyut
(515) 294-6905
jbaluyut@iastate.edu, or

Prof. Thomas A. Holme
(515) 294- 9025
taholme@iastate.edu.

If you have any questions about the rights of research subjects or research-related injury, please contact the IRB Administrator, (515) 294-4566, IRB@iastate.edu, or Director, (515) 294-3115, Office for Responsible Research, Iowa State University, Ames, Iowa 50011.

*** 1. At the time of this survey, are you 18 years of age or older?**

☑ Yes ○ No

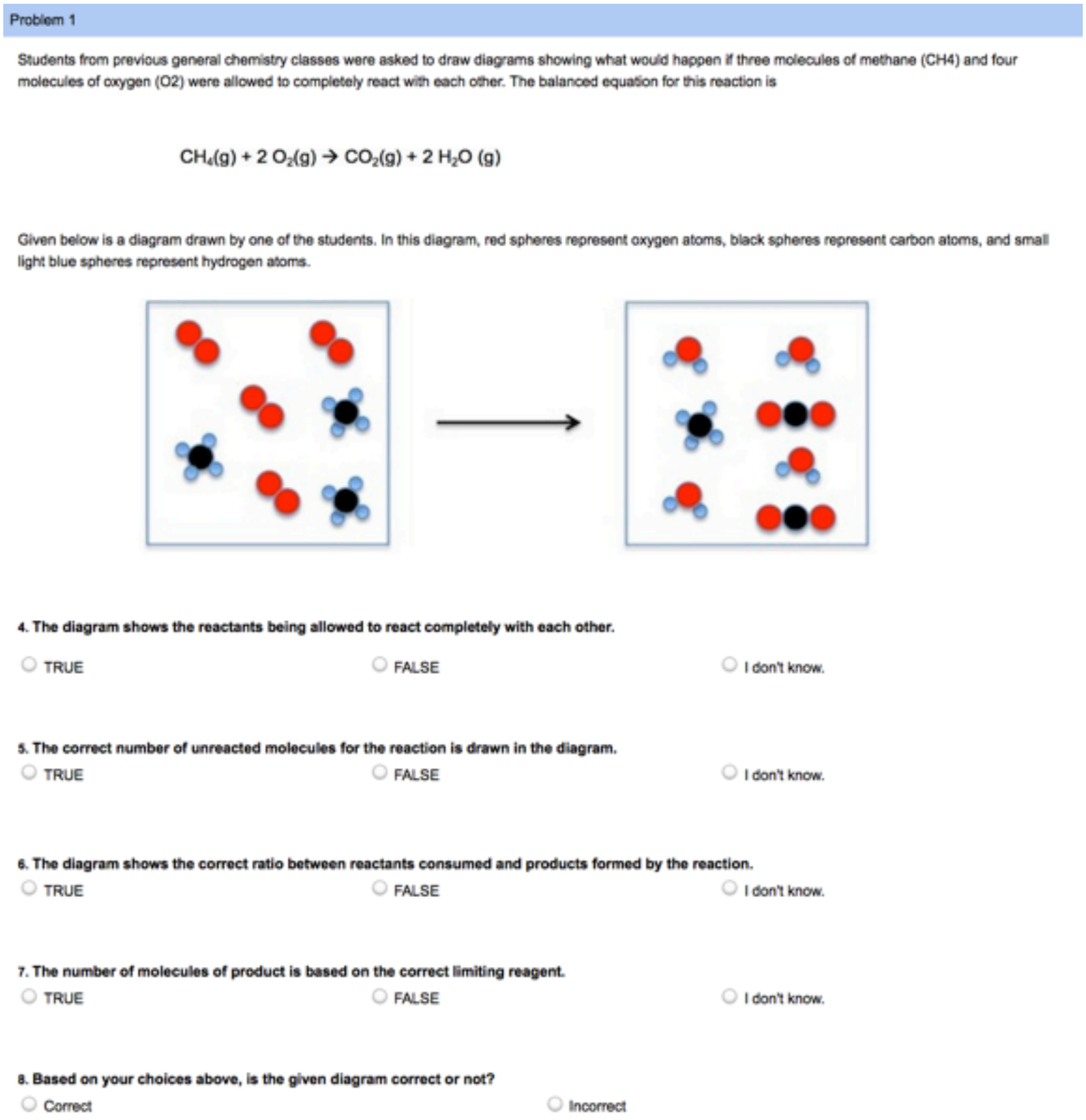
*** 2. Do you voluntarily agree to participate in this study?**

☑ Yes ○ No

*** 3. PARTICIPANT SIGNATURE**

By entering your "Iowa State University" EMAIL address below, you are indicating that the study has been explained to you, that you have been given the time to read the informed consent document, that your questions have been satisfactorily answered, and that you are the student agreeing to participate in the study. Your email address will also be used to acknowledge your participation in the study AND to verify your status as an Iowa State University student.

Please enter your ISU email address on the space below:



Problem 1

Students from previous general chemistry classes were asked to draw diagrams showing what would happen if three molecules of methane (CH4) and four molecules of oxygen (O2) were allowed to completely react with each other. The balanced equation for this reaction is

$$CH_4(g) + 2\ O_2(g) \rightarrow CO_2(g) + 2\ H_2O\ (g)$$

Given below is a diagram drawn by one of the students. In this diagram, red spheres represent oxygen atoms, black spheres represent carbon atoms, and small light blue spheres represent hydrogen atoms.

4. The diagram shows the reactants being allowed to react completely with each other.

○ TRUE ○ FALSE ○ I don't know.

5. The correct number of unreacted molecules for the reaction is drawn in the diagram.

○ TRUE ○ FALSE ○ I don't know.

6. The diagram shows the correct ratio between reactants consumed and products formed by the reaction.

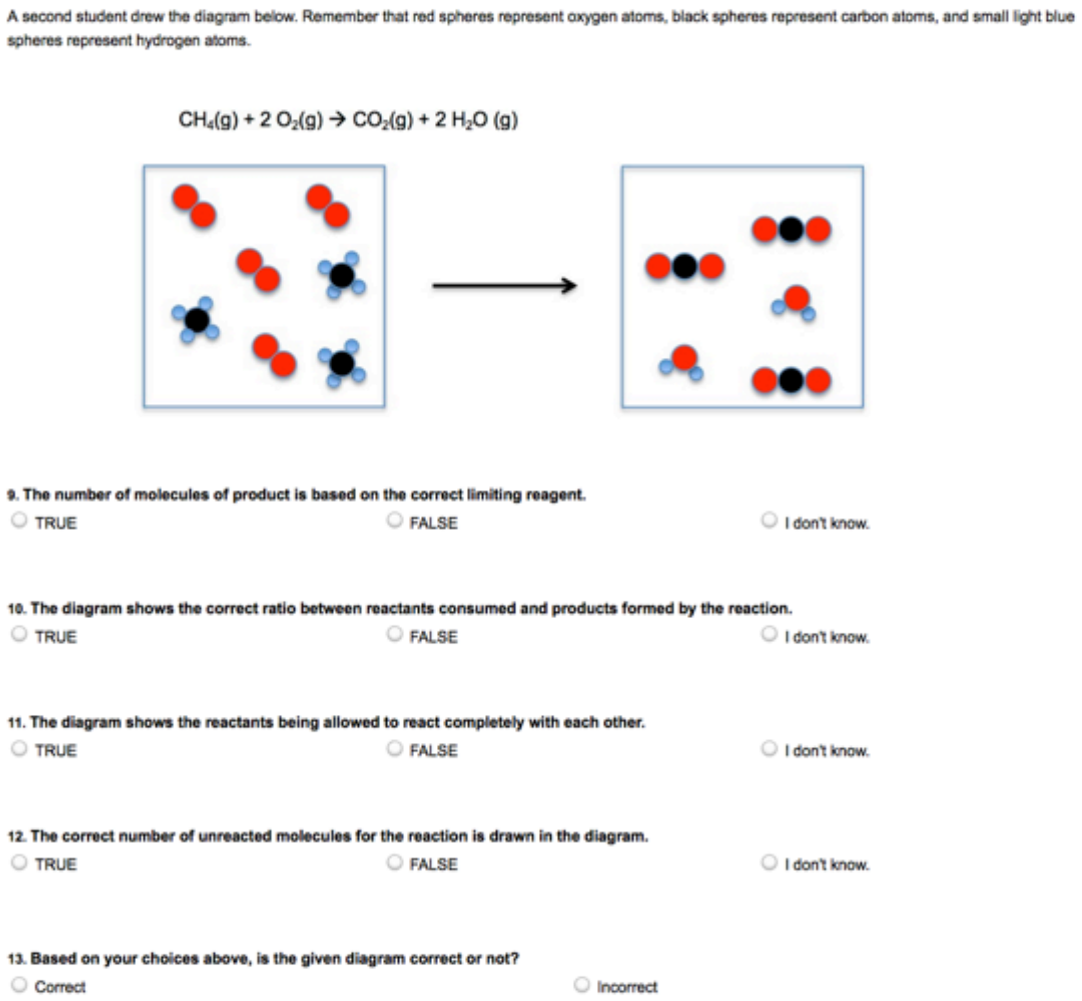
○ TRUE ○ FALSE ○ I don't know.

7. The number of molecules of product is based on the correct limiting reagent.

○ TRUE ○ FALSE ○ I don't know.

8. Based on your choices above, is the given diagram correct or not?

○ Correct ○ Incorrect



A second student drew the diagram below. Remember that red spheres represent oxygen atoms, black spheres represent carbon atoms, and small light blue spheres represent hydrogen atoms.

$CH_4(g) + 2 O_2(g) \rightarrow CO_2(g) + 2 H_2O (g)$

9. The number of molecules of product is based on the correct limiting reagent.

○ TRUE ○ FALSE ○ I don't know.

10. The diagram shows the correct ratio between reactants consumed and products formed by the reaction.

○ TRUE ○ FALSE ○ I don't know.

11. The diagram shows the reactants being allowed to react completely with each other.

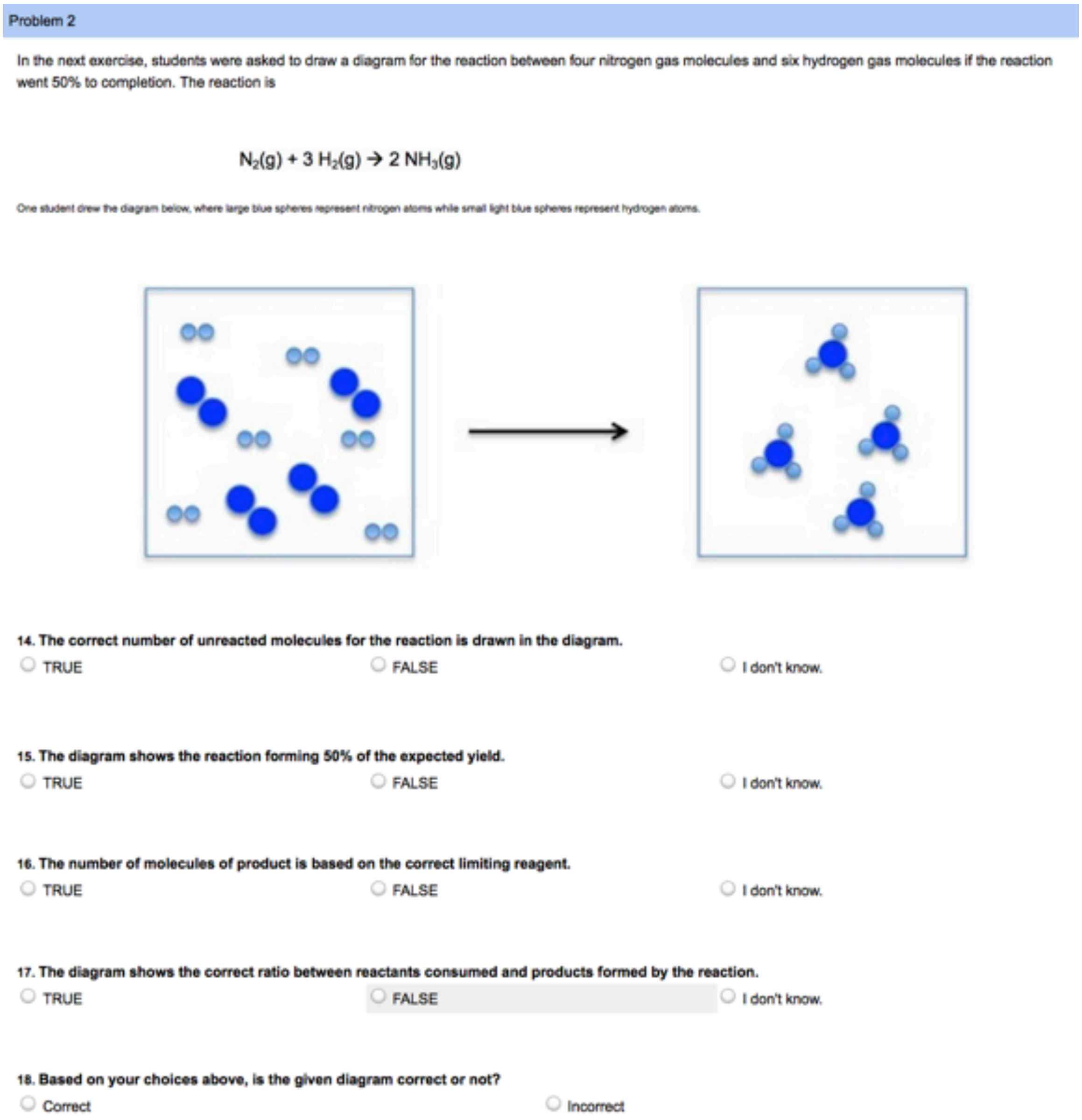
○ TRUE ○ FALSE ○ I don't know.

12. The correct number of unreacted molecules for the reaction is drawn in the diagram.

○ TRUE ○ FALSE ○ I don't know.

13. Based on your choices above, is the given diagram correct or not?

○ Correct ○ Incorrect



Problem 2

In the next exercise, students were asked to draw a diagram for the reaction between four nitrogen gas molecules and six hydrogen gas molecules if the reaction went 50% to completion. The reaction is

$$N_2(g) + 3\ H_2(g) \rightarrow 2\ NH_3(g)$$

One student drew the diagram below, where large blue spheres represent nitrogen atoms while small light blue spheres represent hydrogen atoms.

14. The correct number of unreacted molecules for the reaction is drawn in the diagram.

○ TRUE ○ FALSE ○ I don't know.

15. The diagram shows the reaction forming 50% of the expected yield.

○ TRUE ○ FALSE ○ I don't know.

16. The number of molecules of product is based on the correct limiting reagent.

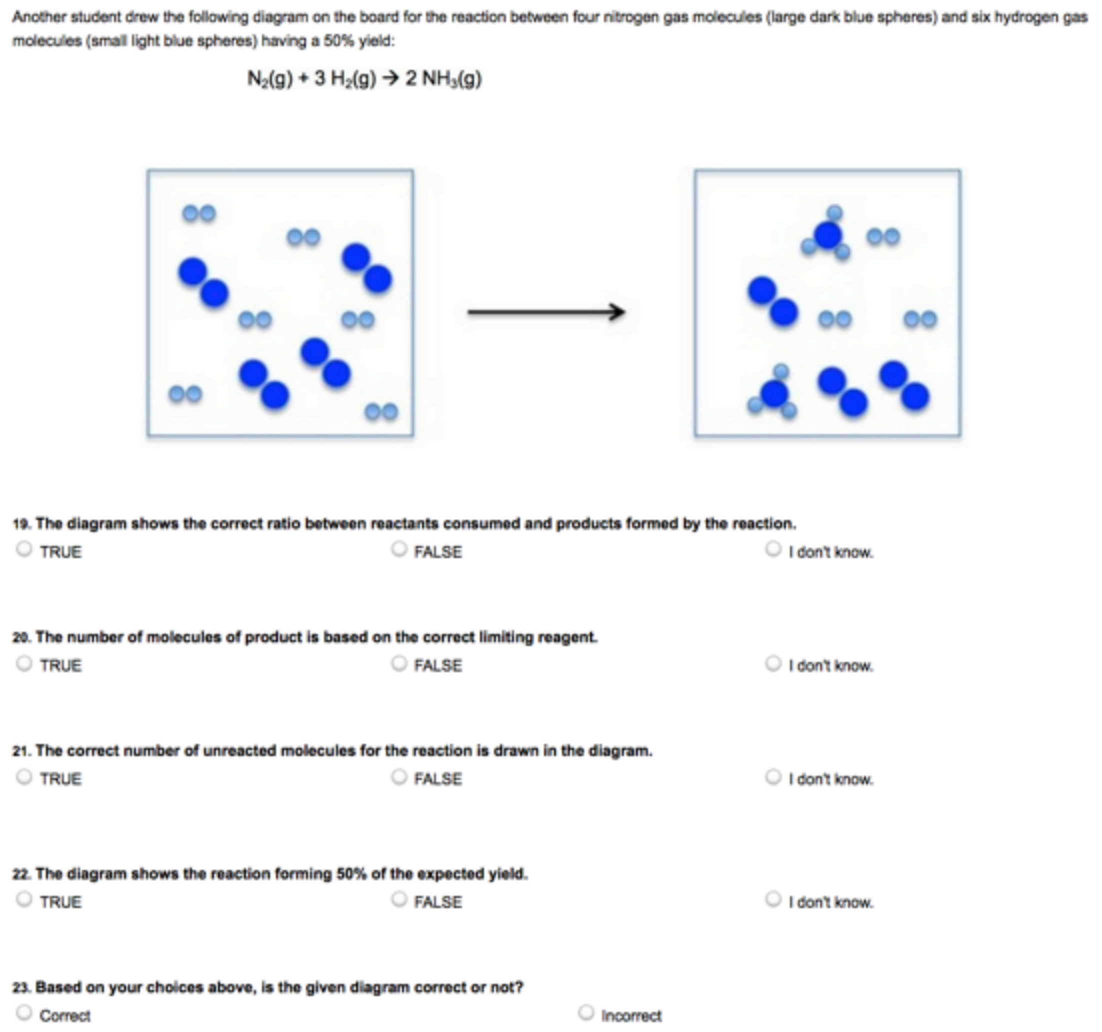
○ TRUE ○ FALSE ○ I don't know.

17. The diagram shows the correct ratio between reactants consumed and products formed by the reaction.

○ TRUE ○ FALSE ○ I don't know.

18. Based on your choices above, is the given diagram correct or not?

○ Correct ○ Incorrect



Another student drew the following diagram on the board for the reaction between four nitrogen gas molecules (large dark blue spheres) and six hydrogen gas molecules (small light blue spheres) having a 50% yield:

$N_2(g) + 3 H_2(g) \rightarrow 2 NH_3(g)$

19. The diagram shows the correct ratio between reactants consumed and products formed by the reaction.

○ TRUE ○ FALSE ○ I don't know.

20. The number of molecules of product is based on the correct limiting reagent.

○ TRUE ○ FALSE ○ I don't know.

21. The correct number of unreacted molecules for the reaction is drawn in the diagram.

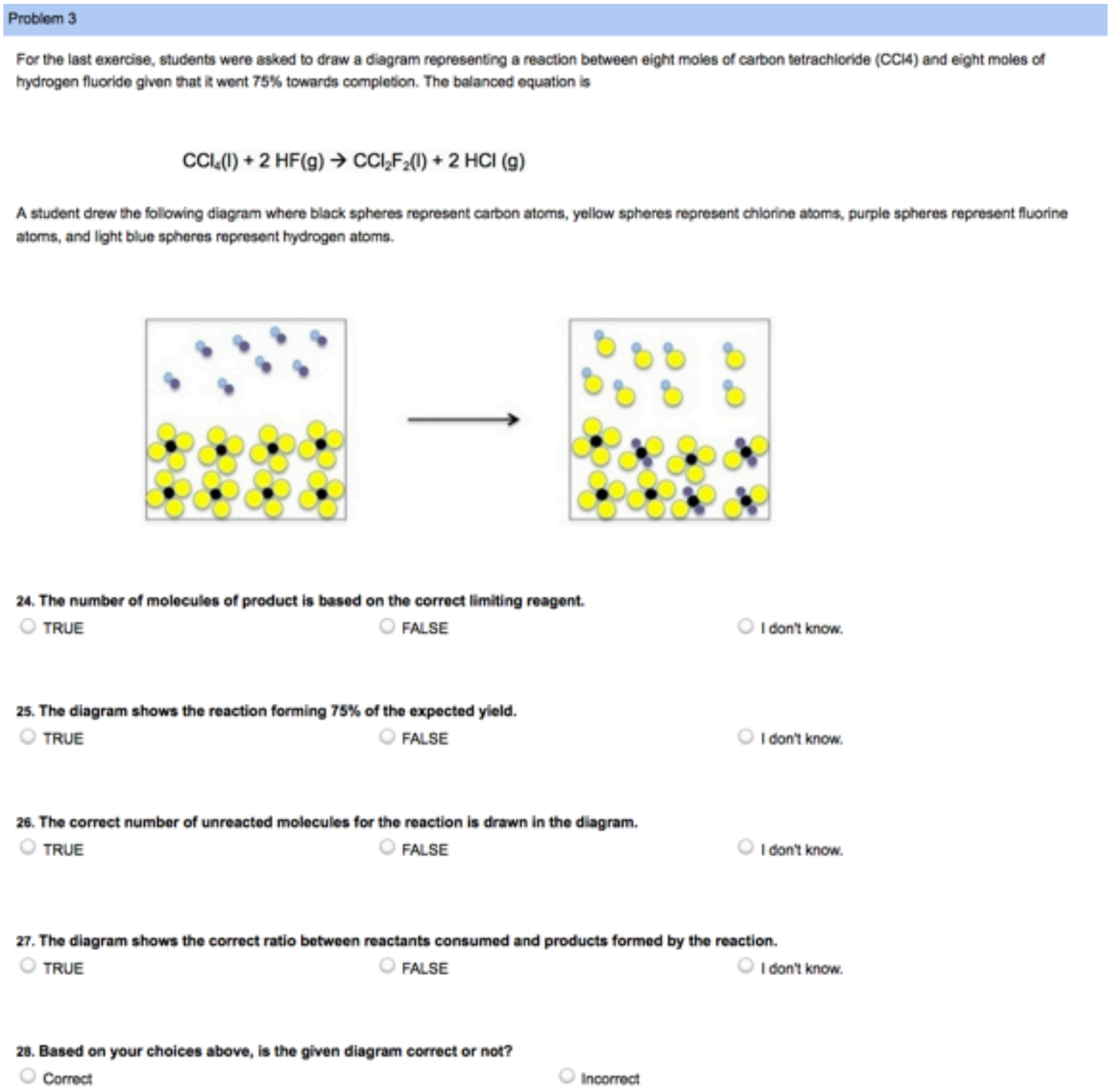
○ TRUE ○ FALSE ○ I don't know.

22. The diagram shows the reaction forming 50% of the expected yield.

○ TRUE ○ FALSE ○ I don't know.

23. Based on your choices above, is the given diagram correct or not?

○ Correct ○ Incorrect



Problem 3

For the last exercise, students were asked to draw a diagram representing a reaction between eight moles of carbon tetrachloride (CCl4) and eight moles of hydrogen fluoride given that it went 75% towards completion. The balanced equation is

$$CCl_4(l) + 2\,HF(g) \rightarrow CCl_2F_2(l) + 2\,HCl\,(g)$$

A student drew the following diagram where black spheres represent carbon atoms, yellow spheres represent chlorine atoms, purple spheres represent fluorine atoms, and light blue spheres represent hydrogen atoms.

24. The number of molecules of product is based on the correct limiting reagent.

○ TRUE ○ FALSE ○ I don't know.

25. The diagram shows the reaction forming 75% of the expected yield.

○ TRUE ○ FALSE ○ I don't know.

26. The correct number of unreacted molecules for the reaction is drawn in the diagram.

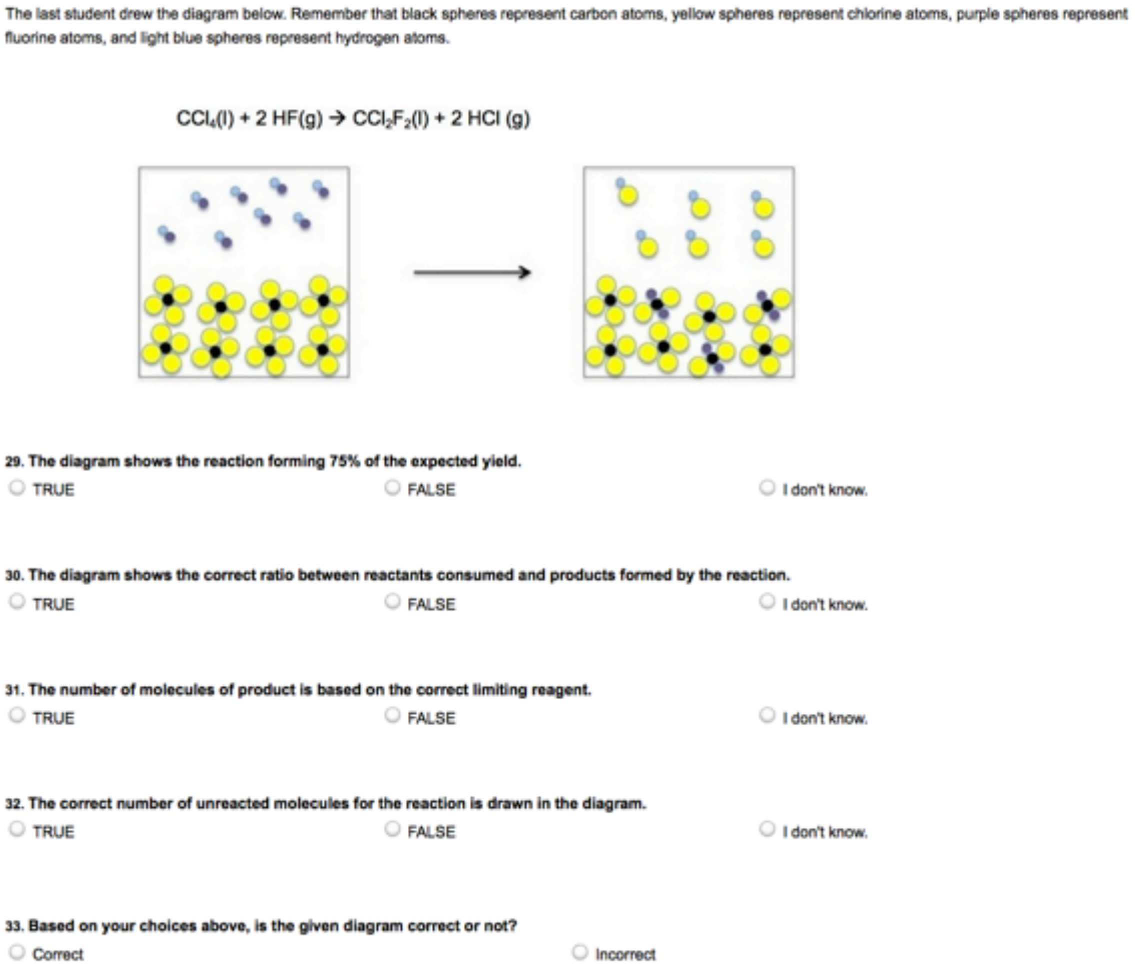
○ TRUE ○ FALSE ○ I don't know.

27. The diagram shows the correct ratio between reactants consumed and products formed by the reaction.

○ TRUE ○ FALSE ○ I don't know.

28. Based on your choices above, is the given diagram correct or not?

○ Correct ○ Incorrect



The last student drew the diagram below. Remember that black spheres represent carbon atoms, yellow spheres represent chlorine atoms, purple spheres represent fluorine atoms, and light blue spheres represent hydrogen atoms.

$CCl_4(l) + 2\ HF(g) \rightarrow CCl_2F_2(l) + 2\ HCl\ (g)$

29. The diagram shows the reaction forming 75% of the expected yield.

○ TRUE ○ FALSE ○ I don't know.

30. The diagram shows the correct ratio between reactants consumed and products formed by the reaction.

○ TRUE ○ FALSE ○ I don't know.

31. The number of molecules of product is based on the correct limiting reagent.

○ TRUE ○ FALSE ○ I don't know.

32. The correct number of unreacted molecules for the reaction is drawn in the diagram.

○ TRUE ○ FALSE ○ I don't know.

33. Based on your choices above, is the given diagram correct or not?

○ Correct ○ Incorrect

APPENDIX E: PYTHON 3.0 SCRIPT FOR PERMUTATION TEST

```
# Import packages
import datetime
import csv
import os.path

# Define Functions
def importCSV():
    csvLocation = input("Where is the CSV file of subjects? (Example:
C:/Users/Bob/test.csv)")
    fileExists = os.path.exists(csvLocation)
    if not fileExists:
        print("File can't be found.")
        importCSV()
    else:
        return parseCSV(csvLocation)

def parseCSV(csvLocation):
    masterDictionary = {}
    reader = csv.DictReader(open(csvLocation))
    for row in reader:
        key = row["Subject A"] + '-' + row["Subject B"]
        masterDictionary[key] = dict(SimscoreA=row['SimscoreA'],
SimscoreB=row['SimscoreB'])
    return masterDictionary

# Main Script
# Turn CSV into array of scores
scoreDictionary = importCSV()

#Use Simscore A or Simscore B?
whatScore = int(input("What score do you want to use? (Simscore A = 0, Simscore B =
1)"))
scoreKey = ''
if whatScore == 0:
    scoreKey = "SimscoreA"
else:
    scoreKey = "SimscoreB"

#print("Simscore for subjects 1 and 2 is", scoreDictionary["1-2"][scoreKey])

# Sample Group Sizes
totalsamplesize=int(input ('Total number of all subjects: '))
```

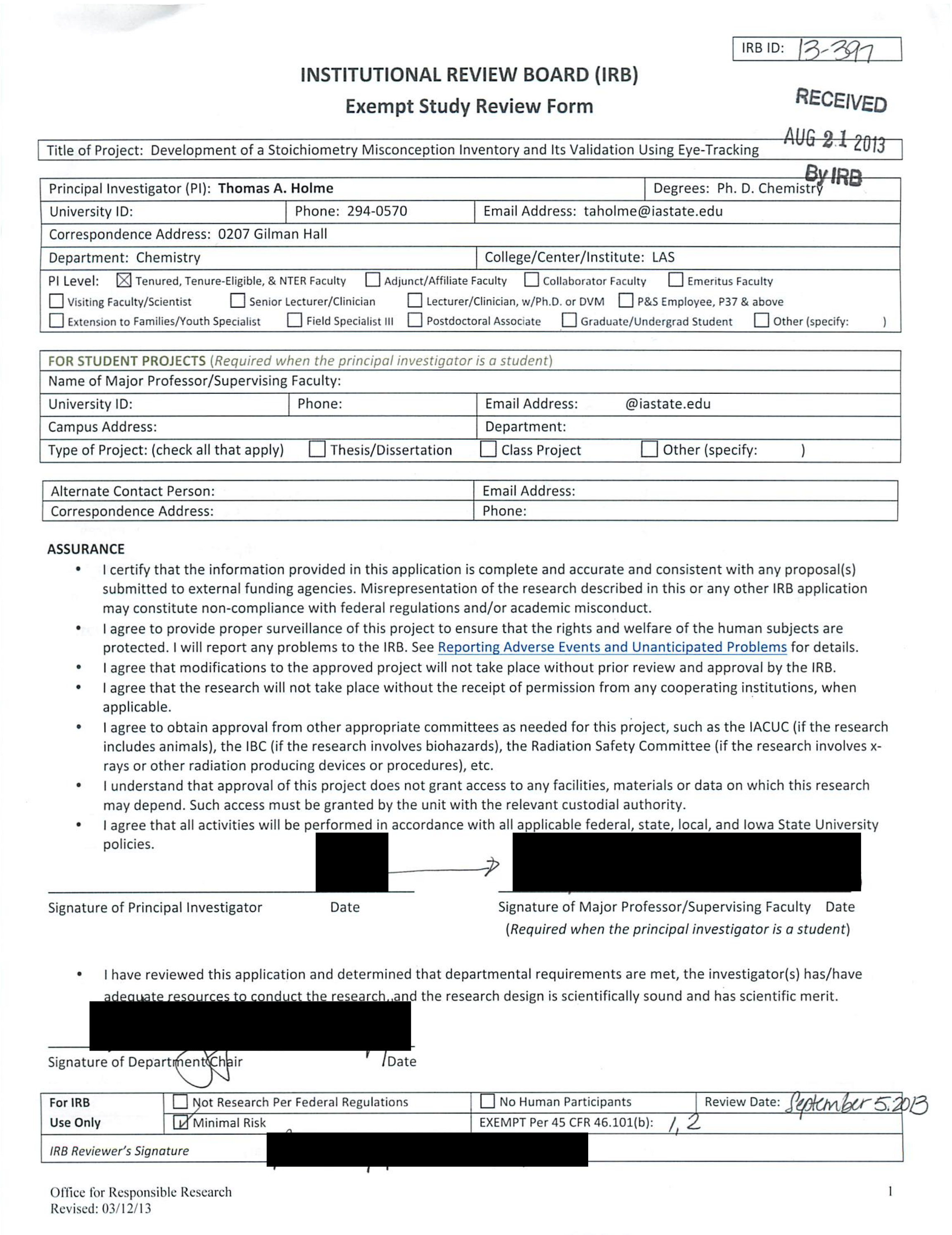
```
print ("\n")
# Input similarity scores between pairs of subjects
scorematrix = [[0 for i in range(totalsamplesize)] for i in range(totalsamplesize)]
print ("starting...")
for i in range(totalsamplesize - 1):
        for j in range(i + 1,totalsamplesize):
                subjectKey = str(i + 1) + "-" + str(j + 1)
                scorematrix[i][j] = float(scoreDictionary[subjectKey][scoreKey])

print ("ending...")
# Input original grouping of subjects
groupmatrix = [0 for k in range(totalsamplesize)]
test = 1
while test > 0:
    keep = 0
    keep = int(input ("Do you want to make new group assignments? (yes = 1, no = 0)"))
    if keep > 0:
        firstgroupsize=int(input ('Number of subjects in the first group: '))
        remaingroupsize=totalsamplesize-firstgroupsize # number of subjects in the second
group
        print ('Number of subjects in the second group: ',remaingroupsize)
        print ("\n")
        for k in range(totalsamplesize):
            print ("subject ",k+1)
            groupmatrix[k] = int(input("initial group assignment:"))
            print ("\n")
    # Timestamp for start of threshold value calculation
    threshold_start = datetime.datetime.now()
    print ("start of threshold calculations: %s" % threshold_start)
    # Calculate threshold value
    withinscoresum = 0
    betweenscoresum= 0
    withinscoreave = 0
    betweenscoreave = 0
    withincounter = 0
    betweencounter = 0
    threshold = 0
    for i in range(0,totalsamplesize-1):
        for j in range(i + 1,totalsamplesize):
            if groupmatrix[i] == groupmatrix[j]:
                withinscoresum += scorematrix[i][j]
                withincounter += 1
            else:
                betweenscoresum += scorematrix [i][j]
```

```
        betweencounter += 1
  withinscoreave = withinscoresum/withincounter
  print (betweenscoresum, betweencounter)
  betweenscoreave = betweenscoresum/betweencounter
  threshold = withinscoreave - betweenscoreave
  print ("\n")
  print ("threshold value = ", "%8.4e"% betweenscoreave, "-", "%8.4e"%
withinscoreave, "=","%8.4e"% threshold,"\n")
  # Timestamp for end of threshold value calculation
  threshold_end = datetime.datetime.now()
  print ("end of threshold calculations: %s" % threshold_end)
  threshold_time = threshold_end - threshold_start
  print ("time for threshold calculations:",threshold_time,"\n")
  print ("\n")
  pvaluecounter = 0
  pvalue = 0
  run = 0
  testvalue = 0
  import random
  numberofruns = int(input ('Number of regrouping samples:'))
  # Timestamp for start of sampling calculations
  sampling_start = datetime.datetime.now()
  print ("start of regrouping calculations: %s" % sampling_start)
  print ("\n")
  testdict = {}
  print ("\n")
  for trial in range(numberofruns):
      check = 0
      while check == 0:
          testgroup = random.sample(range(1,totalsamplesize + 1),firstgroupsize)
          testgroup.sort()
          testtuple = tuple(testgroup)
          if testtuple not in testdict:
              check += 1
      testdict[testtuple] = trial
      withinscoresum = 0
      betweenscoresum = 0
      withinscoreave = 0
      betweenscoreave = 0
      withincounter = 0
      betweencounter = 0
      score = 0
      run += 1
      for i in range(0,totalsamplesize-1):
```

```
      for j in range(i + 1,totalsamplesize):
          if i in testtuple and j in testtuple:
              withinscoresum += scorematrix [i][j]
              withincounter += 1
          elif i not in testtuple and j not in testtuple:
              withinscoresum += scorematrix [i][j]
              withincounter += 1
          else:
              betweenscoresum += scorematrix [i][j]
              betweencounter += 1
    withinscoreave = withinscoresum/withincounter
    betweenscoreave = betweenscoresum/betweencounter
    testvalue = withinscoreave - betweenscoreave
    if testvalue > threshold:
      pvaluecounter += 1
pvalue = pvaluecounter/numberofruns
print ("\n")
print ("P-Value:","%.3f"% pvalue,"\n")
if pvalue < 0.05:
    print ('Group memberships among subjects are NOT interchangeable.')
else:
    print ('Group memberships among subjects are interchangeable.')
sampling_end = datetime.datetime.now()
print ("end of regrouping calculations: %s" % sampling_end)
sampling_time = sampling_end - sampling_start
print ("time for regrouping calculations:",sampling_time,"\n")
test = int(input ("Do you want to generate another p-value? (yes = 1, no = 0)"))
```

APPENDIX F: INSTITUTIONAL REVIEW BOARD APPROVAL FORM



IRB ID: 13-397

RECEIVED
AUG 21 2013
By IRB

INSTITUTIONAL REVIEW BOARD (IRB)
Exempt Study Review Form

Title of Project: Development of a Stoichiometry Misconception Inventory and Its Validation Using Eye-Tracking

Principal Investigator (PI): **Thomas A. Holme** | Degrees: Ph. D. Chemistry
University ID: | Phone: 294-0570 | Email Address: taholme@iastate.edu
Correspondence Address: 0207 Gilman Hall
Department: Chemistry | College/Center/Institute: LAS
PI Level: ☒ Tenured, Tenure-Eligible, & NTER Faculty ☐ Adjunct/Affiliate Faculty ☐ Collaborator Faculty ☐ Emeritus Faculty
☐ Visiting Faculty/Scientist ☐ Senior Lecturer/Clinician ☐ Lecturer/Clinician, w/Ph.D. or DVM ☐ P&S Employee, P37 & above
☐ Extension to Families/Youth Specialist ☐ Field Specialist III ☐ Postdoctoral Associate ☐ Graduate/Undergrad Student ☐ Other (specify:)

FOR STUDENT PROJECTS *(Required when the principal investigator is a student)*
Name of Major Professor/Supervising Faculty:
University ID: | Phone: | Email Address: @iastate.edu
Campus Address: | Department:
Type of Project: (check all that apply) ☐ Thesis/Dissertation ☐ Class Project ☐ Other (specify:)

Alternate Contact Person: | Email Address:
Correspondence Address: | Phone:

ASSURANCE

- I certify that the information provided in this application is complete and accurate and consistent with any proposal(s) submitted to external funding agencies. Misrepresentation of the research described in this or any other IRB application may constitute non-compliance with federal regulations and/or academic misconduct.
- I agree to provide proper surveillance of this project to ensure that the rights and welfare of the human subjects are protected. I will report any problems to the IRB. See Reporting Adverse Events and Unanticipated Problems for details.
- I agree that modifications to the approved project will not take place without prior review and approval by the IRB.
- I agree that the research will not take place without the receipt of permission from any cooperating institutions, when applicable.
- I agree to obtain approval from other appropriate committees as needed for this project, such as the IACUC (if the research includes animals), the IBC (if the research involves biohazards), the Radiation Safety Committee (if the research involves x-rays or other radiation producing devices or procedures), etc.
- I understand that approval of this project does not grant access to any facilities, materials or data on which this research may depend. Such access must be granted by the unit with the relevant custodial authority.
- I agree that all activities will be performed in accordance with all applicable federal, state, local, and Iowa State University policies.

[redacted] [redacted]

Signature of Principal Investigator Date | Signature of Major Professor/Supervising Faculty Date
(Required when the principal investigator is a student)

- I have reviewed this application and determined that departmental requirements are met, the investigator(s) has/have adequate resources to conduct the research, and the research design is scientifically sound and has scientific merit.

[redacted]

Signature of Department Chair Date

For IRB Use Only	☐ Not Research Per Federal Regulations ☑ Minimal Risk	☐ No Human Participants EXEMPT Per 45 CFR 46.101(b): 1, 2	Review Date: September 5, 2013

IRB Reviewer's Signature [redacted]

Office for Responsible Research
Revised: 03/12/13 1